



COMPOSITIO MATHEMATICA

Non-trivial matrix actions preserve normality for continued fractions

Joseph Vandehey

Compositio Math. **153** (2017), 274–293.

[doi:10.1112/S0010437X16007740](https://doi.org/10.1112/S0010437X16007740)



FOUNDATION
COMPOSITIO
MATHEMATICA



LONDON
MATHEMATICAL
SOCIETY
EST. 1865



Non-trivial matrix actions preserve normality for continued fractions

Joseph Vandehey

ABSTRACT

A seminal result due to Wall states that if x is normal to a given base b , then so is $rx + s$ for any rational numbers r, s with $r \neq 0$. We show that a stronger result is true for normality with respect to the continued fraction expansion. In particular, suppose $a, b, c, d \in \mathbb{Z}$ with $ad - bc \neq 0$. Then if x is continued fraction normal, so is $(ax + b)/(cx + d)$.

1. Introduction

A number $x \in [0, 1)$ with base 10 expansion $x = 0.a_1a_2a_3 \dots$ is said to be normal (to base 10) if for any finite string $s = [c_1, c_2, \dots, c_k]$ we have that

$$\lim_{n \rightarrow \infty} \frac{\#\{0 \leq i \leq n : a_{i+j} = c_j, 1 \leq j \leq k\}}{n} = \frac{1}{10^k}.$$

Although almost all real numbers are normal,¹ we still do not know of a single commonly used mathematical constant, such as π , e , or $\sqrt{2}$, that is normal.

In his PhD thesis under Lehmer, Wall [Wal50] proved a series of results on normal numbers which are now considered classical, elementary facts. Among them, Wall proved that if x is normal, then $qx + r$ is normal for any rational numbers q, r with $q \neq 0$. Chang [Cha76] appears to have discovered this result independently, while Doty *et al.* [DLN07] knew of Wall's result and reproved it by a different method. Aistleitner [Ais11] has given the only significant extension of Wall's result the author is aware of, showing that if x is normal and $y \in \mathbb{R}$ is a number with almost all of its digits equal to 0, then $x + ry$ is normal for any rational r . (See also [Bug12, p. 97].)

Although the definition of normality is easily extended to many other digital systems, questions about which operations preserve normality are still unanswered in most cases. Recently, the present author, with Airey and Mance [AMV15], studied how rational multiplication and addition act for Q -Cantor series expansions.

However, in this paper, we shall be interested in normality for continued fraction expansions, which we shall abbreviate as CF-normality and define explicitly in a moment. Mendès France first asked the question of which operations preserve CF-normality [Mau00, pp. 17–18]. He actually

Received 18 June 2015, accepted in final form 6 July 2016, published online 6 February 2017.

2010 Mathematics Subject Classification 11K16 (primary), 11A55 (secondary).

Keywords: normal number, continued fraction, ergodic theory.

The author acknowledges assistance from the Research and Training Group grant DMS-1344994 funded by the National Science Foundation.

This journal is © Foundation Compositio Mathematica 2017.

¹ For readability's sake, we will no longer say 'to base 10' every time. The results mentioned here would follow if we replaced 10 by any integer base $b \geq 2$.

asked a simpler question, namely if non-zero rational multiplication preserves simple normality² for continued fractions. Yann Bugeaud extended the question to ask whether non-zero rational multiplication preserved CF-normality [Bug12, Problem 10.56, p. 222]. Part of the difficulty of proving such a result comes from the fact that rational multiplication and addition are operations that are not very well understood for continued fractions. Research on these topics appears to have come almost completely from a computational side (‘Given a continued fraction expansion x , how can we quickly compute the continued fraction expansion of $qx + r$?’). Notable works include Gosper [Gos72], Raney [Ran73], and Liardet and Stambul [LS98]. On the theoretical side (‘If x has a continued fraction expansion with property Y , does $qx + r$ have property Z ?’), the author is unaware of any significant result.

We recall some standard definitions for continued fractions. The (regular) continued fraction expansion of a number $x \in \mathbb{R}$ is given by

$$x = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \dots}}, \quad \text{with } a_0 \in \mathbb{Z} \text{ and } a_i \in \mathbb{N} \text{ for } i \in \mathbb{N}.$$

We will denote this expansion by $\langle a_0; a_1, a_2, \dots \rangle$ for typographical simplicity. This expansion is infinite if and only if x is irrational. We will refer to the n th digit of the continued fraction expansion of x by $a_n(x)$ or just a_n if the choice of x is clear. The Gauss map $T : [0, 1) \rightarrow [0, 1)$ given by

$$Tx = \begin{cases} \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

acts as a forward shift on the continued fraction digits, ignoring a_0 . The Gauss measure μ given by

$$\mu(A) = \int_A \frac{1}{(\log 2)(1+x)} dx$$

is a probability measure, preserved by T , and is ergodic with respect to T . Given a string $s = [c_1, c_2, \dots, c_k]$, we define the cylinder set of s to be

$$C_s = \{x \in [0, 1) : a_i(x) = c_i, 1 \leq i \leq k\},$$

and we say this cylinder set has rank k . We shall also need the usual matrix action on real numbers given by

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} x = \frac{\alpha x + \beta}{\gamma x + \delta}.$$

With these definitions in mind, we say that a point $x \in [0, 1)$ is CF-normal if for any string s , we have

$$\lim_{n \rightarrow \infty} \frac{\#\{0 \leq i < n : T^i x \in C_s\}}{n} = \mu(C_s). \tag{1}$$

Since $T^i x \in C_s$ if and only if the string s appears in the continued fraction expansion of x starting at the $(i + 1)$ th position, the left-hand side of (1) represents the limiting frequency with which s appears in the continued fraction expansion of x . By the pointwise ergodic theorem, almost all

² Simple normality asks that all of the one-digit strings appear with the correct limiting frequency.

numbers $x \in [0, 1)$ are CF-normal. We extend the definition of CF-normal to all $x \in \mathbb{R}$ to say x is CF-normal if $x - a_0(x)$ is CF-normal.

Our main result will be the following, which not only answers Bugeaud’s question in the affirmative, but shows that any non-trivial linear fractional transformation preserves CF-normality. The method of proof, which we shall outline in §1.1, makes use of ideas about extending ergodicity to a skew product (as was done in [JL88]) as well as ideas about finite automata acting on the digits of a normal number (as was done in [BCH15]).

THEOREM 1.1. *Let M be a 2×2 matrix with coefficients in \mathbb{Z} and non-zero determinant. Let $x \in \mathbb{R}$ be CF-normal. Then Mx is also CF-normal.*

In particular, if $x = \langle 0; a_1, a_2, \dots \rangle$ is CF-normal, $c_0 \in \mathbb{Z}$, and $c_1, c_2, \dots, c_k \in \mathbb{N}$, then

$$\begin{pmatrix} 1 & c_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & c_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & c_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & c_k \end{pmatrix} x = \langle c_0; c_1, c_2, \dots, c_k, a_1, a_2, \dots \rangle.$$

It is easy to see that this preserves normality. However, all of the matrices on the left here have determinant ± 1 , and thus so does their product. In fact, for any matrix with determinant ± 1 , the action of the matrix on an irrational number x will alter the head of the expansion and leave the tail unchanged (see [BvdPSZ14, Theorem 2.37]), thus preserving CF-normality.

We emphasize that Theorem 1.1 works for any matrix that does not have determinant zero, not just those with determinant ± 1 . The reason why we exclude matrices with determinant zero is that in such a case Mx will always be the same *rational* number, regardless of which x is chosen.

Due to earlier work of Kraaikamp and Nakada [KN00] and the author [Van14], we know that normality for regular continued fraction expansions, the kind we are studying in this paper, is equivalent to normality for nearest-integer continued fractions and continued fractions with odd partial quotients. Thus, Theorem 1.1 holds for these expansions as well.

1.1 The idea and outline of the proof

Let us return to the question of normality to base 10 and give a glimpse into why Wall’s result is true.

Given a number x that is base 10 normal, how often do we expect to see the digit 7 appear in the base 10 expansion of $2x$? We should see a 7 appear in the n th position of $2x$ whenever we see one of the strings [3, 5], [3, 6], [3, 7], [3, 8], [3, 9], [8, 5], [8, 6], [8, 7], [8, 8], or [8, 9] appear starting in the n th position of x . We call these strings *trigger strings* for the string 7. But since x is normal, each of these strings appears with limiting frequency $1/100$ and there are 10 of them, so we expect to see 7 appear with limiting frequency $1/10$. In this case, understanding how often trigger strings occur relies on knowing how the sequence $(10^n x)$ is distributed modulo 1.

A slightly harder problem: if x is base 10 normal, then how often do we expect to see 7 in the base 10 expansion of $x/3$? Here, to determine what the n th digit of $x/3$ is, we must not only know what the n th digit of x is, we must know something about all of the first $n - 1$ digits. In particular, the n th digit of $x/3$ is 7 if the n th digit of x is 1, 2, or 3, and the sum of the first $n - 1$ digits of x is 2 modulo 3. If one could show that each of these options appeared with limiting frequency $1/30$,³ that would give the desired limiting frequency for the string 7 in the base 10 expansion of $x/3$.

³This is possible using the techniques of [Van14].

This suggests that to show that division by 3 preserves normality to base 10, we want to understand how the pairs

$$\left(10^n x \pmod{1}, \sum_{i=1}^{n-1} a_i(x) \pmod{3} \right) \tag{2}$$

distribute in the set $[0, 1) \times \{0, 1, 2\}$. (Here $a_i(x)$ is referring to the i th digit of the base 10 expansion, not the continued fraction expansion, of x .)

We will run into a similar difficulty with continued fractions. What we will want is some way to examine the tail of the expansion of Mx , and to do this, we want to be able to push the matrix M through the first n digits of the continued fraction expansion for x like so:

$$M \left(\frac{1}{a_1 + \frac{1}{a_2 + \dots + \frac{1}{a_n + T^n x}}} \right) = b_0 + \frac{1}{b_1 + \frac{1}{b_2 + \dots + \frac{1}{b_m + M_n(T^n x)}}}, \tag{3}$$

with M_n belonging to some set \mathcal{M} . Our analogy to (2) will be the sequence $(T^n x, M_n)$ and we want to show this distributes nicely in the space $[0, 1) \times \mathcal{M}$.

As an explicit example, suppose $M = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Then

$$Mx = \frac{1}{2}x = \frac{1}{2} \frac{1}{a_1 + Tx} = \frac{1}{2a_1 + 2(Tx)}.$$

By reinterpreting this in terms of matrices, we get that

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 2a_1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

If instead, $M = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$, then there are more possibilities, depending on the value of $a_1(x)$:

$$\begin{aligned} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 2n \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \\ \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} &= \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}, \\ \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 2n + 1 \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}. \end{aligned}$$

In §2, we state some of the results of Liardet and Stambul mentioned earlier, which will essentially say that if we started with a ‘nice’ matrix M , then we can in fact choose the matrices M_n in (3) to always belong to a particular finite set \mathcal{M} ; and in Lemma 2.3, we show that it suffices to prove Theorem 1.1 when $M \in \mathcal{M}$. Theorem 3.1 then says that the sequence $(T^n x, M_n)$ is nicely distributed with respect to some measure, provided M starts in a subset of \mathcal{M} with good properties. In §4, we again show that it suffices to prove Theorem 1.1 when M is in such a subset. Finally, we complete the proof of Theorem 1.1 in §5, using a new definition of trigger strings as part of a key lemma.

1.2 Notation and definitions

We will denote the continued fraction expansion of x by $\langle a_0; a_1, a_2, \dots \rangle$ and the continued fraction expansion of Mx by $\langle b_0; b_1, b_2, \dots \rangle$. We will commonly use n to denote an index for x (so we might talk about the n th digit of x) while we will use m to denote an index for Mx . The length of a string $s = [c_0; c_1, \dots, c_n]$, denoted by $|s|$, is n , regardless of the value of c_0 ; and when $c_0 = 0$, we will often not write it or the semi-colon, so the string $[c]$ with no semi-colon will always denote the string $[0; c]$. When we wish to distinguish between strings of digits in x and strings of digits in Mx , we will use $s = [c_0; c_1, c_2, \dots, c_n]$ for strings in x and $t = [d_0; d_1, d_2, \dots, d_m]$ for strings in Mx . We will also make reference to a string r that is a substring in the expansion of Mx . We will use M to denote an arbitrary matrix and \mathcal{M} to denote a collection of matrices. Some further definitions related to strings and matrices will be given in § 2.

We shall make use of standard asymptotic notation as well. We will say that $f(n) = O(g(n))$ if there exists a constant C (called an implicit constant) such that $|f(n)| \leq C \cdot g(n)$. We will say that $f(n) \asymp g(n)$ (with implicit constant C) if $f(n) = O(g(n))$ and $g(n) = O(f(n))$ (both with implicit constant C). If we have two $k \times k$ matrices K_1, K_2 , then we say that $K_1 \asymp K_2$ (with implicit constant C) if $(K_1)_{i,j} \asymp (K_2)_{i,j}$ for $1 \leq i, j \leq k$ (uniformly with implicit constant C). We will say $f(n) = o(g(n))$ if $f(n)/g(n) \rightarrow 0$ as $n \rightarrow \infty$. If a variable appears in a subscript of a big-O or little-o, this denotes that the implicit constant or rate of decay is dependent on this variable.

We will say a vector or matrix is non-negative (or positive) if all its coordinates are non-negative (or positive). We will call a vector a probability vector if it is non-negative and the sum of its coordinates is 1.

2. Matrices and resultant strings

In this section, we will make use of several results of Liardet and Stambul [LS98]. Since we rely so heavily on their results, we will also borrow a lot of their notation. We make some critical changes, however, and will point out where our use differs from theirs.

We will define \mathcal{S} to be the set of all strings $[c_0; c_1, c_2, \dots, c_k]$ with $c_0 \in \mathbb{N}_{\geq 0}$ and $c_i \in \mathbb{N}$ for $1 \leq i \leq k$. We will equate the string $[0;]$ with the empty string, denoted by \wedge . We will let \mathcal{S}^* denote the subset of \mathcal{S} of strings with $c_0 = 0$. Given two strings $s = [c_0; c_1, c_2, \dots, c_k]$ and $s' = [c'_0; c'_1, c'_2, \dots, c'_{k'}]$ both in \mathcal{S} , we define the concatenation $s.s' \in \mathcal{S}$ in the following non-standard way:

$$s.s' := [c_0; c_1, c_2, \dots, c_{k-1}, c_k + c'_0, c'_1, c'_2, \dots, c'_{k'}].$$

We also have $s.\wedge = \wedge.s = s$. The definition of concatenation extends naturally to \mathcal{S}^* .

Given a string $s = [c_0; c_1, c_2, \dots, c_k] \in \mathcal{S}$, we let

$$\Pi_s = \begin{pmatrix} 1 & c_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & c_1 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & c_k \end{pmatrix}.$$

(If we let

$$J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

then, in Liardet and Stambul’s notation, this would be $\Pi_{c_0 c_1 c_2 \dots c_k}$ multiplied by J on the right.) We always have that $|\det(\Pi_s)| = 1$.

Given a string $s \in \mathcal{S}$, we will define $s.x$ for $x \in \mathbb{R}$ by $\Pi_s x$. By the way that concatenation is defined, we can quickly see that $\Pi_{s.s'} = \Pi_s \Pi_{s'}$.

We shall let \mathcal{E} denote the set of matrices

$$\mathcal{E} := \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \mid \alpha, \beta, \gamma, \delta \in \mathbb{N}_{\geq 0}, \alpha\delta - \beta\gamma \neq 0, (\alpha - \gamma)(\beta - \delta) < 0 \right\}.$$

(This is what Liardet and Stambul refer to as \mathcal{E}_2 . We note that \mathcal{E} is invariant under the action of multiplying by J on the left or the right, so that our changes to the definition of Π_s from Liardet and Stambul will not have a noticeable impact here.)

Given $D \in \mathbb{N}$, $D > 1$, let \mathcal{M}_D denote the subset of \mathcal{E} of matrices M satisfying $|\det(M)| = D$.

LEMMA 2.1 [LS98, Lemma 5]. *For any $D \in \mathbb{N}$, $D > 1$, we have that \mathcal{M}_D is finite.*

Let $D \in \mathbb{N}$, $D > 1$ be fixed. For $a \in \mathbb{N}$, we wish to define two functions $\phi_a : \mathcal{M}_D \rightarrow \mathcal{M}_D$ and $\psi_a : \mathcal{M}_D \rightarrow \mathcal{S}$ given by the following:

- (i) if $M\Pi_{[a]} = M'$ for some $M' \in \mathcal{M}_D$, then $\phi_a(M) = M'$ and $\psi_a(M) = \wedge$; and
- (ii) if $M\Pi_{[a]} \notin \mathcal{M}_D$, then according to [LS98, Theorem 1], there exists a unique factorization $M\Pi_{[a]} = \Pi_s M'$ with $M' \in \mathcal{M}_D$ and $s \in \mathcal{S}$, and so we let $\phi_a(M) = M'$ and $\psi_a(M) = s$.

With the functions ϕ_a and ψ_a defined in this way, we have that for any $M \in \mathcal{M}_D$ and irrational $x = \langle a_1, a_2, a_3 \dots \rangle \in [0, 1)$, that

$$Mx = M\Pi_{[a_1]}(Tx) = \Pi_{\psi_{a_1}(M)}\phi_{a_1}(M)(Tx) = \psi_{a_1}(M).(\phi_{a_1}(M)(Tx)).$$

Iterating this procedure again gives the following:

$$Mx = \psi_{a_1}(M).\psi_{a_2}(\phi_{a_1}(M)).(\phi_{a_2}(\phi_{a_1}(M))(T^2x)).$$

In order to continue iterating this procedure without devolving into a typographical nightmare, let us define two new functions $\Phi_s : \mathcal{M}_D \rightarrow \mathcal{M}_D$ and $\Psi_s : \mathcal{M}_D \rightarrow \mathcal{S}$ for $s \in \mathcal{S}^*$.⁴ These functions are defined iteratively beginning with the base cases

$$\Phi_{\wedge}(M) = M, \quad \Psi_{\wedge}(M) = \wedge,$$

and the iterative relations

$$\begin{aligned} \Phi_{[c_1, c_2, \dots, c_k]}(M) &= \phi_{c_k}(\Phi_{[c_1, c_2, \dots, c_{k-1}]}(M)), \\ \Psi_{[c_1, c_2, \dots, c_k]}(M) &= \Psi_{[c_1, c_2, \dots, c_{k-1}]}(M).\psi_{c_k}(\Phi_{[c_1, c_2, \dots, c_{k-1}]}(M)). \end{aligned}$$

This function $\Psi_s(M)$, which Liardet and Stambul refer to as the output of the transducer associated to the string s and the initial state M , we shall instead refer to as the *resultant string* of the pair (s, M) .

With these definitions, we see that for any $M \in \mathcal{M}_D$ and any $x = \langle a_1, a_2, a_3 \dots \rangle \in [0, 1)$ the following relation holds

$$Mx = \Psi_{[a_1, a_2, \dots, a_n]}(M).(\Phi_{[a_1, a_2, \dots, a_n]}(M)(T^n x)), \quad n \in \mathbb{N}_{\geq 0}. \tag{4}$$

LEMMA 2.2. *Let $M \in \mathcal{M}_D$ for an integer $D > 1$ and let $s = [c_1, c_2, \dots, c_k] \in \mathcal{S}$. We have that*

$$\frac{k-1}{|\mathcal{M}_D|} - 1 \leq |\Psi_s(M)| \ll k.$$

⁴ Our use of capital Φ and Ψ is substantially different from Liardet and Stambul's. They use Φ and Ψ to denote the set of functions $\{\phi_c\}$ and $\{\psi_c\}$, respectively. Our notation $\Psi_{[c_1, c_2, \dots, c_k]}(M)$ is equivalent to $\mu([\Psi, \Phi]_{c_1, \dots, c_k})$ in their notation.

Proof. The lower inequality is an immediate consequence of [LS98, Theorem 2].

The upper inequality starts with the fact that $|s.s'| = |s| + |s'|$, thus we only need to prove that $|\psi_c(M)| \ll 1$

Consider the matrix $M\Pi_{[c]}$. If this is in \mathcal{M}_D , then $\psi_c(M) = \wedge$ and there is nothing to prove. Suppose otherwise for the remainder of the proof, that $M\Pi_{[c]} \notin \mathcal{M}_D$.

Let us define $\alpha, \alpha', \beta, \beta', \gamma, \gamma', \delta, \delta'$ by

$$M\Pi_{[c]} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & c \end{pmatrix} = \begin{pmatrix} \alpha' & \beta' \\ \gamma' & \delta' \end{pmatrix}. \tag{5}$$

By [LS98, Theorem 1, part (ii)], since the matrices of (5) equal $\Pi_{\psi_c(M)}\phi_c(M)$, we have that if γ' (or δ') is larger than 0, then the fraction α'/γ' (respectively, β'/δ') has at least $|\psi_c(M)|$ digits beyond the zeroth digit. However, any fraction with at least n digits in its continued fraction expansion has a denominator that is at least n . Therefore γ' (respectively, δ') either is zero or is greater than or equal to $|\psi_c(M)|$.

By Lemma 2.1, the set \mathcal{M}_D is finite, so the coefficients of M must be bounded from above, let us say by K . If $\delta \neq 0$, then $\gamma' = \delta \neq 0$ and thus $|\psi_c(M)| \leq \gamma' = \delta \leq K$. Otherwise, if $\delta = 0$, then $\delta' = \gamma \neq 0$, and so $|\psi_c(M)| \leq \delta' = \gamma \leq K$. These two cases complete the proof. \square

We also require the following result.

LEMMA 2.3. *It suffices to prove Theorem 1.1 in the case where $M \in \mathcal{M}_D$ for some $D \in \mathbb{N}$, $D > 1$ and where $x \in [0, 1)$.*

Proof. By sending

$$x \mapsto x - [x] \quad \text{and} \quad M \mapsto M \begin{pmatrix} 1 & [x] \\ 0 & 1 \end{pmatrix},$$

we may assume without loss of generality that $x \in [0, 1)$.

As noted after the statement of Theorem 1.1, if M is a 2×2 matrix with coefficients in \mathbb{Z} and determinant ± 1 , then Mx is normal if and only if x is normal. We will make frequent use of this fact in the remainder of this proof.

Let $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ have coefficients in \mathbb{Z} and non-zero determinant.

Suppose that either $\gamma = 0$ or $\delta = 0$. Since $\det(M) \neq 0$, we must have that $\alpha \neq 0$ or $\beta \neq 0$ respectively. Thus, there exists some integer a such that the bottom row of

$$M' = \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}, \quad M = \begin{pmatrix} \alpha & \beta \\ \alpha a + \gamma & \beta a + \delta \end{pmatrix}$$

consists only of non-zero integers. Since the determinant of $\begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$ is 1, we have that $M'x$ is normal if and only if Mx is normal. Also $|\det(M')| \neq 0$. Thus, we may assume without loss of generality that $\gamma \neq 0$ and $\delta \neq 0$.

Now suppose that $\text{sgn}(\alpha) \neq \text{sgn}(\gamma)$ or that $\text{sgn}(\beta) \neq \text{sgn}(\delta)$. Then there exists some integer a such that

$$M' = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, \quad M = \begin{pmatrix} \alpha + a\gamma & \beta + a\delta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \alpha' & \beta' \\ \gamma' & \delta' \end{pmatrix}$$

satisfies $\text{sgn}(\alpha') = \text{sgn}(\gamma')$ and $\text{sgn}(\beta') = \text{sgn}(\delta')$. Since $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ has determinant 1, we have that $M'x$ is normal if and only if Mx is normal. Thus, we may assume without loss of generality that $\text{sgn}(\alpha) = \text{sgn}(\gamma)$ and $\text{sgn}(\beta) = \text{sgn}(\delta)$.

Now suppose $\text{sgn}(\gamma) = -1$ or $\text{sgn}(\delta) = -1$. Then by appropriately multiplying M on the left by a matrix of the form $\begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$ (which again has determinant ± 1), we may assume without loss of generality that $\text{sgn}(\gamma) = 1$ and $\text{sgn}(\delta) = 1$. This also preserves all previous assumptions.

Thus, in particular, we may assume that all of the coefficients of M are positive. By [LS98, Theorem 1, part (i)], there exists a string $s = [c_0; c_1, \dots, c_n] \in \mathcal{S}$ and $M' \in \mathcal{M}_D$ such that $M = \Pi_s M'$. However, since Π_s has determinant ± 1 , we have that Mx is normal if and only if $M'x$ is normal, and thus we may assume without loss of generality that $M \in \mathcal{M}_D$ all along. \square

3. Normality on a skew product

By our heuristic argument in § 1.1, we want to show that the sequence

$$(T^n x, \Phi_{[a_1, a_2, \dots, a_n]}(M))$$

is nicely distributed in some sense. This will allow us to show that the last term of (4), which gives the tail of the CF expansion of Mx , is likewise nicely distributed.

Let $\Omega \subset [0, 1)$ denote the subset of irrational points and let \mathcal{M} denote some finite set, which we will later take to be a set of matrices. We will let x denote elements in Ω and M denote elements of \mathcal{M} . We will consider cylinder sets of Ω to be the intersection of the usual cylinder sets (for the continued fraction expansion) of $[0, 1)$ with Ω .

We wish to extend the Gauss map T to a transformation \tilde{T} on a larger domain $\tilde{\Omega} = \Omega \times \mathcal{M}$. For any $(x, M) \in \tilde{\Omega}$, we define the skew product

$$\tilde{T}(x, M) = (Tx, f_{a_1(x)}(M))$$

for some function $f_{a_1} : \mathcal{M} \rightarrow \mathcal{M}$ that is indexed by the first digit of x . Since the second coordinate of $\tilde{T}(x, M)$ only depends on M and the first CF-digit of x , we see that this second coordinate is constant for all x in the same rank 1 cylinder. Given a cylinder set C_s for Ω , we call $C_s \times \{M\}$ (for any $M \in \mathcal{M}$) a cylinder set for $\tilde{\Omega}$. Moreover, we define $\tilde{\mu}(E \times \{M\}) = \mu(E)/|\mathcal{M}|$ for any measurable subset E of Ω and $b \in \mathcal{M}$.

For easier readability, we will use (E, M) to denote $E \times \{M\}$ for any measurable set $E \subset \Omega$, with measurability being determined by Lebesgue measure or, equivalently, the Gauss measure. We will also let (E, \mathcal{M}) denote $E \times \mathcal{M}$.

We adapt our definition of normality on this space. We will say that $(x, M) \in \tilde{\Omega}$ is \tilde{T} -normal with respect to a measure ρ on $\tilde{\Omega}$, if for any cylinder set (C_s, M') we have

$$\lim_{n \rightarrow \infty} \frac{\#\{0 \leq i < n : \tilde{T}^i(x, M) \in (C_s, M')\}}{n} = \rho(C_s, M').$$

We say \tilde{T} is transitive if for any $M_1, M_2 \in \mathcal{M}$, there exists a string $s \in \mathcal{S}$ of length n such that

$$T^n(C_s, M_1) = (\Omega, M_2).$$

The goal of this section is to prove the following result, which is similar to a previous result of the author [Van14]; however, as this paper contains significant departures (notably not assuming that the functions f_a are bijective and hence not being able to assume that $\tilde{\mu}$ is \tilde{T} -invariant), we present the proof in full.

THEOREM 3.1. *If \tilde{T} is transitive, then there exists a probability measure ρ on $\tilde{\Omega}$ that is absolutely continuous with respect to $\tilde{\mu}$ and such that \tilde{T} preserves ρ and is ergodic with respect to ρ . Moreover, if x is CF-normal, then for any $M \in \mathcal{M}$, the point (x, M) is \tilde{T} -normal with respect to ρ .*

Since almost all numbers $x \in [0, 1]$ are CF-normal, almost all $(x, M) \in \tilde{\Omega}$ are \tilde{T} -normal with respect to ρ . Therefore, calling these points (x, M) ‘normal’ is reasonable to do.

The proof of Theorem 3.1 breaks into two pieces. Proving that ρ exists and satisfies the desired properties follows very standard ergodic theoretic techniques (such as Knopp’s lemma). Proving that any CF-normal $x \in \Omega$ lifts to \tilde{T} -normal points (x, M) requires much more work. We will need to show that the density of ρ and the density of $\tilde{\mu}$ are within a constant multiple of one another and then apply the Pyatetskiĭ–Shapiro normality criterion (Lemma 3.3).

Remark 3.2. Although we are using T as the Gauss map here, the only properties of the Gauss map that we use are the fact that all cylinders are full (i.e. $T^{|s|}C_s = \Omega$), that the Pyatetskiĭ–Shapiro normality criterion can be applied via the cylinder sets, that the corresponding ergodic, invariant measure μ is finite, and that T satisfies Renyi’s condition (6). Therefore, any map T which also satisfies these conditions will also satisfy Theorem 3.1.

3.1 Necessary lemmata for Theorem 3.1

In order to simplify the readability of the proof of Theorem 3.1, we will include several lemmas here. All of these results make use of the definitions and assumptions at the start of §3.

In order to show that (x, M) is \tilde{T} -normal with respect to ρ , we will need to make use of the Pyatetskiĭ–Shapiro normality criterion in the following form.

LEMMA 3.3. *Let $(x, M) \in \tilde{\Omega}$ and suppose a measure ρ exists satisfying the first part of Theorem 3.1. If for any cylinder set (C_s, M') , we have*

$$\limsup_{n \rightarrow \infty} \frac{\#\{0 \leq i \leq n - 1 : T^i(x, M) \in (C_s, M')\}}{n} \leq \sigma \cdot \rho(C_s, M')$$

for some constant σ independent of our choice of cylinder set, then x is \tilde{T} -normal with respect to ρ .

Proof. This is a simple consequence of [MS03, Theorem 1]. We briefly describe how using the terminology from their paper.

We let the family $\{C_m\}$ denote the family of all cylinder sets on $\tilde{\Omega}$ and also let $\varphi(t) = \sigma \cdot t$. Since the set $A_l(T, \chi_I, \delta)$ is a disjoint union of rank- l cylinder sets, we have that $H_\varphi(A_l(T, \chi_I, \delta))$ equals $\sigma \cdot \mu(A_l(T, \chi_I, \delta))$ and thus goes to 0 as $l \rightarrow \infty$. \square

A well-known consequence of Renyi’s condition for continued fraction expansions (see [Sch95, ch. 9]) states that there exists an absolute constant $\mathcal{C} > 0$ so that for any measurable set E and cylinder C_s of rank n , we have that

$$\frac{1}{\mathcal{C}}\mu(E)\mu(C_s) \leq \mu(T^{-n}E \cap C_s) \leq \mathcal{C}\mu(E)\mu(C_s). \tag{6}$$

It is clear that one could replace C_s by any set that can be expressed as a disjoint union of rank n cylinder sets.

We will also want a similar equality (sans the cylinder set) to hold for \tilde{T} and $\tilde{\mu}$, for which we will require the following results.

LEMMA 3.4. *Let $\{K_n\}_{n=1}^\infty$ be a sequence of $k \times k$ Markov matrices such that $K_{n_1} \asymp K_{n_2}$ uniformly for $n_1, n_2 \in \mathbb{N}$. Assume that there exists a power ℓ such that K_1^ℓ has all positive coordinates, and also assume that $\{K_1\}_{i,i} > 0$ for $1 \leq i \leq k$. Then there exists an integer $n_0 \in \mathbb{N}$ and a constant $c \in (0, 1)$ such that for any $1 \times k$ probability vector \vec{v} and any $n \geq n_0$, we have that all coordinates of $\vec{v}K_1K_2K_3 \dots K_n$ are in the interval $(c, 1 - c)$.*

Proof. This is a special case of [SCZ11, Proposition 2.13]: the restriction that $K_{n_1} \asymp K_{n_2}$ uniformly for $n_1, n_2 \in \mathbb{N}$ implies that $K_{n+1}K_{n+2} \dots K_{n+\ell} \asymp K_1^\ell$ with an implicit constant dependent on ℓ . This allows us to replace the ‘uniform irreducibility’ assumption with our simpler condition that there exists a power ℓ such that K_1^ℓ has all positive coordinates. \square

LEMMA 3.5. *Suppose \tilde{T} is transitive, then there exists a constant $\mathcal{D} > 0$ and integer $n_0 \in \mathbb{N}$ such that for any measurable set $E \subset \tilde{\Omega}$ and any $n \geq n_0$, we have that*

$$\frac{1}{\mathcal{D}}\tilde{\mu}(E) \leq \tilde{\mu}(\tilde{T}^{-n}E) \leq \mathcal{D}\tilde{\mu}(E). \tag{7}$$

Proof. We will make a few definitions to start. Let $\mathcal{S}_{n,i}$ denote the set of cylinders $(C_s, M) \subset \tilde{\Omega}$ such that $|s| = n$ and $\tilde{T}^n(C_s, M) = (\Omega, M_i)$. We let $\mathcal{A}_{i,j}$ denote the set of $a \in \mathbb{N}$ such that $f_a(M_i) = M_j$. We also consider a sequence of probability vectors $\vec{v}_n = \{v_{n,1}, v_{n,2}, \dots, v_{n,|\mathcal{M}|}\}$, $n \geq 0$, such that

$$v_{n,i} = \tilde{\mu}\left(\bigcup_{(C_s, M) \in \mathcal{S}_{n,i}} (C_s, M)\right).$$

For example, $\vec{v}_0 = \{1/|\mathcal{M}|, 1/|\mathcal{M}|, \dots, 1/|\mathcal{M}|\}$.

Our goal will be to show that there exists a constant $c \in (0, 1)$ and integer $n_0 \in \mathbb{N}$ so that $v_{n,i} \in (c, 1 - c)$ if $n \geq n_0$.

Consider a sequence of $|\mathcal{M}| \times |\mathcal{M}|$ matrices $(K_n)_{n=1}^\infty$ defined by

$$(K_n)_{i,j} = \frac{1}{v_{n,i}}\tilde{\mu}\left(\bigcup_{(C_s, M) \in \mathcal{S}_{n,i}} \bigcup_{a \in \mathcal{A}_{i,j}} (C_{s,[a]}, M)\right).$$

It is clear by construction that these matrices are stochastic and that $\vec{v}_{n+1} = \vec{v}_n K_n$.

We let $K_{\ell,n} = K_{\ell+1}K_{\ell+2} \dots K_n$.

Unfortunately, the matrices K_n are not all the same, so $K_{0,n}$ represents a Markov chain that is time-inhomogeneous. However, they are not far from being time-homogenous. By writing the matrix coefficients in a different way and applying (6), we see that

$$\begin{aligned} (K_n)_{i,j} &= \frac{1}{v_{n,i}}\tilde{\mu}\left(\bigcup_{(C_s, M) \in \mathcal{S}_{n,i}} \left(C_s \cap T^{-n}\left(\bigcup_{a \in \mathcal{A}_{i,j}} C_{[a]}\right), M\right)\right) \\ &\asymp \frac{1}{v_{n,i}}\tilde{\mu}\left(\bigcup_{(C_s, M) \in \mathcal{S}_{n,i}} (C_s, M)\right)\mu\left(\bigcup_{a \in \mathcal{A}_{i,j}} C_{[a]}\right) \\ &= \mu\left(\bigcup_{a \in \mathcal{A}_{i,j}} C_{[a]}\right), \end{aligned}$$

with the same implicit constant \mathcal{C} as in (6). Thus, $K_{n_1} \asymp K_{n_2}$ for any $n_1, n_2 \in \mathbb{N}$ with uniform implicit constant \mathcal{C}^2 . This also implies that for any fixed L , we have that $K_1^L \asymp K_{nL, (n+1)L}$ uniformly for any $n \in \mathbb{N}$ with implicit constant \mathcal{C}^{2L} .

Since \tilde{T} is assumed to be transitive, we know that for any $M_i, M_j \in \mathcal{M}$, there exists a cylinder (C_s, M_i) with $|s| = \ell$ such that $T^\ell(C_s, M_i) = (\Omega, M_j)$. This implies that $(K_{0,\ell})_{i,j} > 0$ and thus that $(K_1^\ell)_{i,j} > 0$. In other words, K_1 is an irreducible matrix. Also, we can find integers $\ell_1, \ell_2, \dots, \ell_{|\mathcal{M}|} \in \mathbb{N}$ so that $(K_1^{\ell_i})_{i,i} > 0$. Since all terms of K_1 are non-negative by construction,

if $(K_1^{\ell_i})_{i,i} > 0$ then we have $(K_1^{m\ell_i})_{i,i} > 0$ for any $m \in \mathbb{N}$. Thus, if we let $L = \text{lcm}(\ell_1, \ell_2, \dots, \ell_{|\mathcal{M}|})$, then we have that K_1^L is strictly positive along its diagonal.

Suppose K_1^L is itself irreducible. Then since K_1^L has non-negative coefficients with a strictly positive diagonal, there is some power of it such that every coefficient is strictly positive (see [Mey00, equation (8.3.5) on p. 672]). We may therefore apply Lemma 3.4 to the sequence of matrices $\{K_{nL, (n+1)L}\}_{n=0}^\infty$. So there exists $c' \in (0, 1)$ and n'_0 such that $v_{nL,i} \in (c', 1 - c')$ for all i and all $n \geq n'_0$.

Suppose K_1^L is not irreducible. By [BR91, Theorem 3.4.5], since K_1 itself is irreducible, there exists a permutation matrix P such that

$$PK_1^L P^T = \begin{pmatrix} C_1 & 0 & \cdots & 0 \\ 0 & C_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_r \end{pmatrix}, \tag{8}$$

where the C_j are irreducible matrices. In this case, we would apply Lemma 3.4 to each C_j and thus show that there exist $c'_j \in (0, 1)$ and n'_j such that $v_{nL,i} \in (c'_j, 1 - c'_j)$ for $n \geq n'_j$ and for indexes i corresponding to the matrix C_j after undoing the permutation. By taking $c' = \min\{c'_j\}$ and $n'_0 = \max\{n'_j\}$, we get the same result as in the previous paragraph.

Regardless of whether K_1^L is irreducible or not, we have shown that there exists $c' \in (0, 1)$ and n'_0 such that $v_{nL,i} \in (c', 1 - c')$ for all $n \geq n'_0$.

No column of K_1 consists of all zeros (otherwise there would be an $M \in \mathcal{M}$ that is never visited, contrary to the transitivity of \tilde{T}), therefore the sum of the coefficients in any column vector of $K_{nL, nL+j}$ for $j \leq L$ is uniformly bounded from below. Thus, we can therefore find a constant c and n_0 such that $v_{n,i} \in (c, 1 - c)$ for all i and all $n \geq n_0$. In particular, $v_{n,i} \asymp 1$.

Now we can prove the desired statement (7). It suffices to show the statement is true for $n \geq n_0$ and for sets of the form (E, M_i) for some measurable subset $E \subset \Omega$ and $M_i \in \mathcal{M}$. In this case, we have that $\tilde{T}^{-n}(E, M_i)$ equals the union of $((T^{-n}E) \cap C_s, M)$ for $(C_s, M) \in \mathcal{S}_{n,i}$ as defined above. Therefore, by applying (6), we have

$$\begin{aligned} \tilde{\mu}(\tilde{T}^{-n}(E, M_i)) &= \sum_{(C_s, M) \in \mathcal{S}_{n,i}} \tilde{\mu}((T^{-n}E) \cap C_s, M) = \sum_{(C_s, M) \in \mathcal{S}_{n,i}} \frac{1}{|\mathcal{M}|} \mu((T^{-n}E) \cap C_s) \\ &\asymp \mu(E) \sum_{(C_s, M) \in \mathcal{S}_{n,i}} \frac{1}{|\mathcal{M}|} \mu(C_s) = \mu(E) v_{n,i} \asymp \mu(E) \\ &= |\mathcal{M}| \cdot \tilde{\mu}(E, M_i) \asymp \tilde{\mu}(E, M_i), \end{aligned}$$

as desired. □

3.2 Proof of Theorem 3.1

First, we will show that \tilde{T} is ergodic (despite not necessarily being invariant) with respect to $\tilde{\mu}$.

Suppose, we have a \tilde{T} -invariant subset of $\tilde{\Omega}$ called E that has non-zero $\tilde{\mu}$ -measure. We define $E^c = \tilde{\Omega} \setminus E$ and define E_M as the set of $x \in \Omega$ such that $(x, M) \in E$, so that $E = \bigcup_{M \in \mathcal{M}} (E_M, M)$.

We claim that

$$\tilde{\mu}(E_M, M) > 0 \tag{9}$$

for all $M \in \mathcal{M}$. Since E has positive $\tilde{\mu}$ -measure, there must exist a set $E_{M'}$ of positive μ -measure. By transitivity, for any $M \in \mathcal{M}$, there exists a string s with $|s| = n$ so that $\tilde{T}^n(C_s, M) = (\Omega, M')$.

Thus,

$$\tilde{\mu}((C_s, M) \cap \tilde{T}^{-n}(E_{M'}, M')) = \tilde{\mu}((C_s \cap T^{-n}E_{M'}, M)) = \frac{1}{|\mathcal{M}|} \mu(C_s \cap T^{-n}E_{M'}).$$

However, by Renyi’s condition, this is within a positive constant multiple of $\mu(C_s)\mu(E_{M'})$, and in particular is positive. Since E is a \tilde{T} -invariant set, $\tilde{T}^{-n}(E_{M'}, M') \subset E$, and thus (9) holds.

Now we wish to show that E has a substantial intersection with every cylinder set on $\tilde{\Omega}$, in particular, by showing that there exists a constant $\epsilon > 0$ so that for all cylinder sets (C_s, M) , we have

$$\tilde{\mu}(E \cap (C_s, M)) \geq \epsilon \cdot \tilde{\mu}(C_s, M). \tag{10}$$

Since there are only finitely many elements in \mathcal{M} , there must exist $\epsilon' > 0$, such that $\mu(E_{M'}) \geq \epsilon'$ for all $M' \in \mathcal{M}$. Let us now fix an arbitrary cylinder (C_s, M) with $n := |s|$, and let M' be such that $\tilde{T}^n(C_s, M) = (\Omega, M')$. By applying (6), we have

$$\begin{aligned} \tilde{\mu}((C_s, M) \cap E) &= \tilde{\mu}((C_s, M) \cap \tilde{T}^{-n}(E_{M'}, M')) = \tilde{\mu}((C_s \cap T^{-n}E_{M'}, M)) \\ &= \frac{1}{|\mathcal{M}|} \mu(T^{-n}E_{M'} \cap C_s) \geq \frac{1}{\mathcal{C}|\mathcal{M}|} \mu(E_{M'})\mu(C_s) \\ &\geq \frac{\epsilon'}{\mathcal{C}|\mathcal{M}|} \mu(C_s) = \frac{\epsilon'}{\mathcal{C}} \tilde{\mu}(C_s, M). \end{aligned}$$

Therefore, letting $\epsilon = \epsilon'/\mathcal{C}$ gives (10).

Since the cylinder sets generate the Borel sets on $\tilde{\Omega}$, we can find, for any $\delta > 0$, a set E_δ such that $\tilde{\mu}(E^c \Delta E_\delta) < \delta$ and E_δ is a disjoint union of a finite number of cylinder sets. Therefore, by applying (10), we have

$$\begin{aligned} \tilde{\mu}(E \cap E^c) &= \tilde{\mu}(E \cap E_\delta) + O(\delta) \geq \epsilon \tilde{\mu}(E) \tilde{\mu}(E_\delta) + O(\delta) \\ &= \epsilon \tilde{\mu}(E) \tilde{\mu}(E^c) + O(\delta). \end{aligned}$$

But $\tilde{\mu}(E \cap E^c) = 0$ and δ was an arbitrary positive number. Thus, either $\tilde{\mu}(E) = 0$ or $\tilde{\mu}(E^c) = 0$. Since we know E has positive measure, this therefore implies that E must have full measure, and \tilde{T} is ergodic with respect to $\tilde{\mu}$.

We will now construct a measure ρ that is absolutely continuous with respect to $\tilde{\mu}$ such that \tilde{T} is not only ergodic but also invariant with respect to ρ .

We define a sequence of measures ρ_n on $\tilde{\Omega}$ by

$$\rho_n(A) = \frac{1}{n} \sum_{k=0}^{n-1} \tilde{\mu}(\tilde{T}^{-k}A) = \int_{\tilde{\Omega}} \left(\frac{1}{n} \sum_{k=0}^{n-1} \chi_A(\tilde{T}^k t) \right) d\tilde{\mu}(t). \tag{11}$$

By Lemma 3.5, we can show that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \tilde{\mu}(\tilde{T}^{-k}E) \leq \mathcal{D}\tilde{\mu}(E),$$

for any measurable set E . Therefore, by a theorem of Ryll–Nardzewski (see [DS58, p. 683]), the integrand of (11) converges pointwise to a L_1 function g_A almost everywhere, and since the integrand is dominated by 1, the integrand must in fact converge uniformly to g_A almost everywhere. Therefore, we may define $\rho(A) = \lim_{n \rightarrow \infty} \rho_n(A)$. The Vitali–Hahn–Saks theorem [Bro69] shows that ρ is in fact a probability measure on $\tilde{\Omega}$. Since $\rho_n(\tilde{T}^{-1}E) = \rho_n(E) + O(1/n)$

for any measurable set E , we have that ρ is preserved by \tilde{T} . Likewise, by Lemma 3.5 again, we can see that

$$\frac{\tilde{\mu}(E)}{\mathcal{D}} \leq \rho_n(E) \leq \mathcal{D}\tilde{\mu}(E), \tag{12}$$

and thus the same is true if we replace ρ_n by ρ .⁵

Thus it remains to show that if $x \in \Omega$ is CF-normal, then (x, M) is \tilde{T} -normal with respect to ρ for any $M \in \mathcal{M}$.

So consider a point $x \in \Omega$ that is CF-normal. Then for every cylinder C_s and every $M \in \mathcal{M}$, we have

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \#\{1 \leq n \leq N \mid \tilde{T}^n(x, M) \in (C_s, \mathcal{M})\} &= \lim_{N \rightarrow \infty} \frac{1}{N} \#\{1 \leq n \leq N \mid T^n x \in C_s\} \\ &= \mu(C_s). \end{aligned}$$

Thus, in particular, we have for any $M, M' \in \mathcal{M}$ that

$$\begin{aligned} &\limsup_{N \rightarrow \infty} \frac{1}{N} \#\{1 \leq n \leq N \mid \tilde{T}^n(x, M) \in (C_s, M')\} \\ &\leq \limsup_{N \rightarrow \infty} \frac{1}{N} \#\{1 \leq n \leq N \mid \tilde{T}^n(x, M) \in (C_s, \mathcal{M})\} \\ &= \mu(C_s) = \frac{1}{|\mathcal{M}|} \tilde{\mu}(C_s, M') \leq \frac{\mathcal{D}}{|\mathcal{M}|} \rho(C_s, M'). \end{aligned}$$

Thus, by Lemma 3.3, the points (x, M) for all $M \in \mathcal{M}$ are \tilde{T} -normal with respect to ρ . □

The referee of this paper proposed an alternative method for studying ρ . We include the outline of this other method in case it proves more useful to those wishing to extend these results than the method given above.

One could consider the map \mathcal{T} which acts by

$$\mathcal{T}(x, y, M) = \left(Tx, \frac{1}{y + a_1(x)}, f_{a_1(x)}(M) \right).$$

Let $X = \overline{\{\mathcal{T}^n(x, 0, M) : n \geq 0, x \in [0, 1], M \in \mathcal{M}\}}$, and let $d\nu$ be $dx dy / (1 + xy)^2$ crossed with the counting measure on \mathcal{M} and appropriately normalized to be a probability measure. Then (X, \mathcal{T}, ν) can be shown to be the natural extension of $(\tilde{\Omega}, \tilde{T}, \rho)$ via the methods of [KSS12]. Moreover ρ can be given explicitly by

$$d\rho = \int_{\{y \in [0, 1] : (x, y, M) \in X\}} d\nu.$$

By picking an appropriate cylinder set in X and applying \mathcal{T} enough times, one can show that for each M there exists a ‘horizontal stripe’ $[0, 1] \times [y_1^M, y_2^M] \times \{M\}$ in X . One can then show that, for any set $E \subset E_M$, one has

$$\mu(E) \ll \nu(E \times [y_1^M, y_2^M] \times \mathcal{M}) \ll \rho(E, M) \leq \mu(E).$$

This provides the crucial relationship between μ and ρ that is used in the final part of Theorem 3.1.

⁵ This part of the proof draws heavily on techniques used in the proof of [IG90, Theorem 5.3.5].

4. Building a dynamical system

As we hinted at the start of §3, we would like to build a dynamical system \tilde{T} from $\Omega \times \mathcal{M}_D$ to itself by $\tilde{T}(x, M) = (Tx, \phi_{a_1}(M))$, with ϕ_a defined as it was in §2. We will use this definition for \tilde{T} throughout the rest of the paper; however, it may turn out that this system is not transitive and thus Theorem 3.1 may not apply. Thus, we will require the following definition to modify our dynamical system.

We will call a subset $\mathcal{M}'_D \subset \mathcal{M}_D$ a transitive component if the following conditions are satisfied:

- (i) for any string $s \in \mathcal{S}$, and any $M \in \mathcal{M}'_D$ we have that $\Phi_s(M) \in \mathcal{M}'_D$;
- (ii) for any $M, M' \in \mathcal{M}'_D$ there exists a string $s \in \mathcal{S}$ such that $\Phi_s(M) = M'$.

Note that any two distinct transitive components of \mathcal{M}_D must have empty intersection.

With this definition the transformation \tilde{T} given by $\tilde{T}(x, M) = (Tx, \phi_{a_1}(M))$ is transitive on $\Omega \times \mathcal{M}'_D$ for any transitive component $\mathcal{M}'_D \subset \mathcal{M}_D$. Thus, Theorem 3.1 applies for this \tilde{T} .

LEMMA 4.1. *There exists at least one transitive component of \mathcal{M}_D . Moreover, there exists a string s such that $\Phi_s(M)$ is in a transitive component for any $M \in \mathcal{M}_D$ (although not necessarily always in the same transitive component for each M).*

Proof. Consider a directed graph G whose vertices are matrices $M \in \mathcal{M}_D$ and which has an edge from M_1 to M_2 if there exists a $j \in \mathbb{N}$ such that $\phi_j(M_1) = M_2$. Note this graph has out-degree always at least 1. A subgraph G' of G is said to be strongly connected if for any M_1, M_2 in $V(G')$, the vertex set of G' , there exists a path from M_1 to M_2 and vice versa. We can partition G into its strongly connected components, which are the maximal strongly connected subgraphs of G . (Note that if there is a vertex M such that there is no path from M to another vertex and back to itself, then M is its own strongly connected component.) Let us call these components G_1, G_2, \dots, G_n . Note that if $M \in V(G_i)$, then $V(G_i)$ consists of all vertices which are strongly connected to M .

Now let us consider another directed graph \mathcal{G} whose vertices are G_1, G_2, \dots, G_n and where there is an edge from G_i to G_j with $i \neq j$ if there exists an edge from some $M \in G_i$ to some $M' \in G_j$ in the directed graph G . We do not let \mathcal{G} contain an edge which goes from a vertex to itself. Note that if there is an edge from G_i to G_j , then by the strong-connectivity of these components, there is a path from any $M \in G_i$ to any $M' \in G_j$. We see that \mathcal{G} cannot have any cycles, as otherwise it would be possible to find matrices in two different strongly connected components that are strongly connected to one another, contradicting the maximality of these components. Thus, \mathcal{G} is acyclic.

Any finite acyclic directed graph must contain at least one sink. We claim that the set of vertices of any sink G_i of \mathcal{G} is a transitive component for \mathcal{M}_D . Let us fix a sink G_i and let \mathcal{M}' denote the matrices in $V(G_i)$. The first condition for being a transitive component is satisfied because if there existed a string $s \in \mathcal{S}$ and $M \in \mathcal{M}'$ such that $\Phi_s(M) \notin \mathcal{M}'$, then there would be at least one path from G_i to another strongly connected component, contradicting the assumption that G_i is a sink. For the second condition, this follows from the fact that within any G_i there is a path from any vertex to any other vertex and also to itself.

Now consider all of the matrices that do not lie in a transitive component, let us call them M_1, M_2, \dots, M_k . Consider M_1 and suppose it is in G_i . As there must be a path from G_i to a sink of \mathcal{G} , there exists a string s_1 such that $\Phi_{s_1}(M_1)$ is in a transitive component. Now consider $\Phi_{s_1}(M_2)$. Regardless of what matrix in \mathcal{M}_D the matrix $\Phi_{s_1}(M_2)$ happens to be, there is, by the

same argument, a string s_2 such that $\Phi_{s_2}(\Phi_{s_1}(M_2)) = \Phi_{s_1.s_2}(M_2)$ is in a transitive component. (The string s_2 could equal \wedge if $\Phi_{s_1}(M_2)$ is already in a transitive component.) Likewise there is a string s_3 such that $\Phi_{s_1.s_2.s_3}(M_3)$ is in a transitive component, and so on. The desired string s is simply $s_1.s_2.s_3 \cdots .s_k$. \square

The second part of Lemma 4.1 allows us to reduce the cases of Theorem 1.1 yet further. By Lemma 2.3, it suffices to prove Theorem 1.1 in the case where $x \in [0, 1)$ is a CF-normal number and $M \in \mathcal{M}_D$ for some $D \in \mathbb{N}$, $D > 1$. Let s be a string satisfying Lemma 4.1. If x is CF-normal, its continued fraction expansion contains s . Thus, there exists some n such that $\tilde{T}^n(x, M) \in \Omega \times \mathcal{M}'_D$, for some transitive component $\mathcal{M}'_D \subset \mathcal{M}_D$, and so

$$Mx = \Psi_{[a_1, a_2, \dots, a_n]}(M) \cdot (\Phi_{[a_1, a_2, \dots, a_n]}(M)(T^n x)),$$

with $\Phi_{[a_1, a_2, \dots, a_n]}(M) \in \mathcal{M}'_D$. Since the action of T and the action of string concatenation both preserve CF-normality and CF-non-normality, it suffices to assume that $x \in [0, 1)$ is CF-normal and that M is in some transitive component \mathcal{M}'_D .

LEMMA 4.2. *Let \tilde{T} and ρ be the transformation and measure corresponding to some transitive component \mathcal{M}'_D , the latter of whose existence is guaranteed by Theorem 3.1. Let $k \in \mathbb{N}$, and let $f : \Omega \times \mathcal{M}'_D \rightarrow \mathbb{R}$ be a bounded function that is constant on rank- k cylinder sets.*

For any CF-normal $x \in [0, 1)$ and any $M \in \mathcal{M}'_D$, we have

$$\sum_{i=0}^{n-1} f(\tilde{T}^i(x, M)) = n \cdot \left(\int_{\tilde{\Omega}} f \, d\rho \right) (1 + o(1)).$$

Proof. Since there are countably many pairs (s, M) with $|s| = k$, let us enumerate them as $(s_i, M_i)_{i=1}^\infty$.

Since f is constant on rank- k cylinder sets, we may write

$$f = \sum_{i=1}^\infty a_i \cdot 1_{(C_{s_i}, M_i)}(\cdot),$$

where $a_i \in \mathbb{R}$ is uniformly bounded and $1_E(\cdot)$ is the standard indicator function of a set E . Let us write $A_+ = \sup_i a_i$ and $A_- = \inf_i a_i$. We then construct functions f_j^+ and f_j^- , $j \in \mathbb{N}$, by

$$f_j^\pm = A_\pm + \sum_{i=1}^j (a_i - A_\pm) 1_{(C_{s_i}, M_i)}(\cdot).$$

By construction we have that $f_j^- \leq f \leq f_j^+$ and, moreover, both f_j^+ and f_j^- converge pointwise to f (and, hence, by dominated convergence, also converge in norm).

Now let $x \in [0, 1)$ be CF-normal and let $M \in \mathcal{M}'_D$. From the second part of Theorem 3.1, we have that for any cylinder set (C_s, M') that

$$\sum_{i=0}^{n-1} 1_{(C_s, M')}(\tilde{T}^i(x, M)) = n \cdot \left(\int_{\tilde{\Omega}} 1_{(C_s, M')} \, d\rho \right) (1 + o(1)).$$

This statement also holds if $1_{(C_s, M')}$ is replaced by any function that can be expressed as a finite sum of indicator functions of cylinder sets, such as f_j^\pm .

Thus, we have, for any $j \geq 0$,

$$\sum_{i=0}^{n-1} f_j^-(\tilde{T}^i(x, M)) \leq \sum_{i=0}^{n-1} f(\tilde{T}^i(x, M)) \leq \sum_{i=0}^{n-1} f_j^+(\tilde{T}^i(x, M)),$$

$$n \cdot \left(\int_{\tilde{\Omega}} f_j^- d\rho \right) (1 + o(1)) \leq \sum_{i=0}^{n-1} f(\tilde{T}^i(x, M)) \leq n \cdot \left(\int_{\tilde{\Omega}} f_j^+ d\rho \right) (1 + o(1)).$$

As both integrals above converge in norm to $\int_{\tilde{\Omega}} f d\rho$ as j tends to infinity, this completes the proof. \square

5. Proof of Theorem 1.1

As we have noted in Lemma 2.3 and §4, it suffices to prove Theorem 1.1 in the case where $x \in [0, 1)$ is CF-normal and $M \in \mathcal{M}'_D$ where \mathcal{M}'_D is some transitive component of \mathcal{M}_D .

Note that Theorem 3.1 applies to \tilde{T} acting on the set $\Omega \times \mathcal{M}'_D$, giving us an ergodic, invariant measure ρ on this space, and (x, M) is normal with respect to \tilde{T} and ρ .

For the first and largest step of the proof, we want to show that for any string $r \in \mathcal{S}^*$, this string appears in Mx with a limiting frequency that does not depend on the CF-normal number x . (However, we will resume throughout the proof that M is a fixed matrix.) In particular, we want a constant ρ_r such that

$$\lim_{m \rightarrow \infty} \frac{\#\{0 \leq i \leq m : T^i(Mx \pmod{1}) \in C_r\}}{m} = \rho_r, \tag{13}$$

for all CF-normal $x \in [0, 1)$.

Let $x = \langle a_1, a_2, a_3, \dots \rangle$, and let $Mx = \langle b_0; b_1, b_2, \dots \rangle$ so that

$$Mx = \Psi_{[a_1, a_2, \dots, a_n]}(M) \cdot (\Phi_{[a_1, a_2, \dots, a_n]}(M)(T^n x)).$$

We let $\ell(n)$ denote the length of $\Psi_{[a_1, a_2, \dots, a_n]}(M)$. The following two lemmas assume that r and \mathcal{M}'_D are fixed and that x is any CF-normal number in $[0, 1)$ and $M \in \mathcal{M}'_D$. We will also let $\tilde{\Omega} = \Omega \times \mathcal{M}'_D$.

LEMMA 5.1. *We have for some constant $c_1 > 0$*

$$\ell(n) = c_1 n(1 + o(1)).$$

Proof. It is clear that $|s.s'| = |s| + |s'|$, so that $\ell(n)$ is just the sum of the lengths of the corresponding resultant strings

$$\sum_{i=1}^n |\psi_{a_i}(\Phi_{[a_1, a_2, \dots, a_{i-1}]}(M))|.$$

As a result, let us consider the function $g : \mathbb{N} \times \mathcal{M}'_D \rightarrow \mathbb{N}$ which acts by taking the pair (a, M) to the length of the resultant string $\psi_a(M)$. Let $G(x, M)$ equal $g(a, M)$ if $x \in C_{[a]}$. Note that

$$G = \sum_{a \in \mathbb{N}, M \in \mathcal{M}_D} g(a, M) \cdot 1_{C_{[a], M}}(\cdot).$$

By Lemma 2.2, we have that $g(a, M) \ll 1$, so that G is a bounded function that is constant on all rank-one cylinder sets.

So we have that $\ell(n) = \sum_{i=0}^{n-1} G(\tilde{T}^i(x, M))$, and by Lemma 4.2, $\ell(n) = n \cdot (\int_{\tilde{\Omega}} G d\rho)(1 + o(1))$. \square

LEMMA 5.2. We have for some constant $c_r > 0$,

$$\#\{0 \leq i \leq \ell(n) : T^i(Mx \pmod{1}) \in C_r\} = c_r n(1 + o(1)).$$

Proof. In order to prove this lemma, we will need to define trigger strings for this system in the same way that we loosely defined trigger strings to help us understand base-10 normality in § 1.1. This will require a few ancillary definitions. Throughout these definitions, we shall allow $s, s', s^* \in \mathcal{S}^*$ and $M, M', M^* \in \mathcal{M}'_D$.

We say that (s, M) is a *subpair* of (s', M') if s' can be written as $s_+ . s . s_-$ for strings $s_+, s_- \in \mathcal{S}^*$ with $M = \Phi_{s_+}(M')$. In particular, $\tilde{T}^{|s_+|}(C_{s'}, M') \subset (C_s, M)$.

We say that r *appears nicely in the resultant* of (s, M) if r is a substring of $\Psi_s(M)$ starting after the zeroth digit of $\Psi_s(M)$ and ending before the last digit. Thus, by our definition of string concatenation, r also appears as a substring of $\Psi_{s'}(M')$ for any pair (s', M') of which (s, M) is a subpair. In fact, if $s' = s_+ . s . s_-$, as in the previous paragraph, and r appears as a substring of $\Psi_s(M)$ starting at the n th digit, then r appears as a substring of $\Psi_{s'}(M')$ starting at the $|\Psi_{s_+}(M')| + n$ th digit. We refer to these two copies of r as being in the same *relative position*.

Finally we will define (s, M) to be a trigger string for r of multiplicity k if there are k copies of r that appear nicely in the resultant of (s, M) each of which does not appear nicely in the same relative position in the resultant of (s^*, M^*) for any pair (s^*, M^*) that is a proper subpair of (s, M) . By the lower inequality in Lemma 2.2, we see that the length of s in a trigger string (s, M) must be bounded, say by $\ell = \ell_r$. In consequence, since $|s|$ is bounded, the upper inequality in Lemma 2.2 implies that the multiplicity k of any trigger string must be bounded as well.

From these definitions, we can see that for each occurrence of r that appears nicely within the resultant of (s', M') , there exists a unique subpair (s, M) of (s', M') for which (s, M) is a trigger string for r and for which r appears nicely in the resultant of (s, M) in the same relative position as it did within the resultant of (s', M') . In particular, the number of occurrences of r that appear nicely within the resultant of (s', M') should be equal to the sum of all of the multiplicities of all the subpairs (s, M) of (s', M') .

The count

$$\#\{0 \leq i \leq \ell(n) : T^i(Mx \pmod{1}) \in C_r\}$$

can be thought of as counting the number of occurrences of r that appear nicely within the resultant of $([a_1, a_2, \dots, a_n], M)$, up to $O(1)$. The $O(1)$ accounts for those appearances of r that start within the first $\ell(n) + 1$ digits of Mx but do not appear nicely within the resultant of $([a_1, a_2, \dots, a_n], M)$.

So let $f(x, M)$ denote the sum of the multiplicities of all trigger strings (s, M) such that $x \in C_s$. Our above work implies that $f(x, M)$ is a bounded function and is constant on cylinder sets of rank ℓ . We can then apply Lemma 4.2 to see that

$$\#\{0 \leq i \leq \ell(n) : T^i(Mx \pmod{1}) \in C_r\} = \sum_{i=0}^{n-1} f(\tilde{T}^i(x, M)) + O(1) = n \left(\int_{\tilde{\Omega}} f \, d\rho \right) (1 + o(1)),$$

as desired. □

From here we are nearly done. First, let $\ell^{-1}(m) := \max\{n : \ell(n) \leq m\}$. By Lemma 5.1, $\ell^{-1}(m) = m/c(1 + o(1))$, so therefore we have that

$$0 \leq m - \ell(\ell^{-1}(m)) \leq o(m).$$

Thus, we have that

$$\begin{aligned} & \lim_{m \rightarrow \infty} \frac{\#\{0 \leq i \leq m : T^i(Mx \pmod{1}) \in C_r\}}{m} \\ &= \lim_{m \rightarrow \infty} \frac{\#\{0 \leq i \leq \ell(\ell^{-1}(m)) : T^i(Mx \pmod{1}) \in C_r\} + o(m)}{\ell(\ell^{-1}(m)) + o(m)} \end{aligned}$$

and so (13) follows from Lemmas 5.1 and 5.2.

Now consider the sets

$$E_M = \{Mx : x \in [0, 1) \text{ is CF-normal}\} \text{ and } E = \{x : x \in \mathbb{R} \text{ is CF-normal}\}.$$

We have shown that for any string $r \in \mathcal{S}^*$ there exists a constant ρ_r such that for all $y \in E_M$, the string r appears in the continued fraction expansion of y with limiting frequency ρ_r , even though we do not know what any of these constants ρ_r equal. On the other hand, for all strings $r \in \mathcal{S}^*$ and all $x \in E$, the limiting frequency of r in the continued fraction expansion of x is $\mu(C_r)$. Thus, either $\rho_r = \mu(C_r)$ for all r and E_M is a subset of E or $\rho_r \neq \mu(C_r)$ for some r and E_M is disjoint from E .

However, E has full Lebesgue measure and E_M , being a non-trivial linear fractional transformation of a positive measure set, has positive measure, so E_M must be a subset of E , and the theorem is proved.

6. Further questions

In one, admittedly peculiar, sense, the generalization that we have proved of Wall’s result is not the natural generalization to make. What makes rational numbers so nice for any base b , is that they are eventually periodic. So one could ask the following.

Suppose x is CF-normal and q and r have eventually periodic continued fraction expansions, that is, they are both quadratic irrationals, with $q \neq 0$. Must it be true that $qx + r$ is CF-normal as well?

Also, while Theorem 1.1 solves Bugeaud’s problem, it does not solve Mendès France’s problem. CF-normality is a much stronger condition than CF-simple normality, and our proof relies crucially on full CF-normality. So we ask, as Mendès France did: does non-zero rational multiplication and rational addition preserve CF-simple normality?

ACKNOWLEDGEMENTS

The author would like to thank Justin Moore for asking a thought-provoking question on mathoverflow.net regarding the effect adding $1/2$ has on a continued fraction expansion, Bill Mance for bringing Bugeaud’s question to his attention, and Cor Kraaikamp for pointing the author to the work of Liardet and Stambul. The author would also like to thank the anonymous referee whose many comments greatly improved and simplified this paper.

REFERENCES

Ais11 C. Aistleitner, *On modifying normal numbers*, Unif. Distrib. Theory **6** (2011), 49–58; [MR 2904038](#).
 AMV15 D. Airey, B. Mance and J. Vandehey, *Normality preserving operations for Cantor series expansions and associated fractals part II*, New York J. Math. **21** (2015), 1311–1326.
 BCH15 V. Becher, O. Carton and P. A. Heiber, *Normality and automata*, J. Comput. System Sci. **81** (2015), 1592–1613; [MR 3389924](#).

- BvdPSZ14 J. Borwein, A. van der Poorten, J. Shallit and W. Zudilin, *Neverending fractions: an introduction to continued fractions*, Australian Mathematical Society Lecture Series, vol. 23 (Cambridge University Press, Cambridge, 2014).
- Bro69 J. K. Brooks, *On the Vitali-Hahn-Saks and Nikodým theorems*, Proc. Natl. Acad. Sci. USA **64** (1969), 468–471.
- BR91 R. A. Brualdi and H. J. Ryser, *Combinatorial matrix theory*, Encyclopedia of Mathematics and its Applications, vol. 39 (Cambridge University Press, Cambridge, 1991); [MR 1130611](#) (93a:05087).
- Bug12 Y. Bugeaud, *Distribution modulo one and Diophantine approximation*, Cambridge Tracts in Mathematics, vol. 193 (Cambridge University Press, Cambridge, 2012); [MR 2953186](#).
- Cha76 K. T. Chang, *A note on normal numbers*, Nanta Math. **9** (1976), 70–72; [MR 0444599](#) (56 #2949).
- DLN07 D. Doty, J. H. Lutz and S. Nandakumar, *Finite-state dimension and real arithmetic*, Inform. and Comput. **205** (2007), 1640–1651; [MR 2368644](#) (2008m:68073).
- DS58 N. Dunford and J. T. Schwartz, *Linear Operators. I. General Theory*, Pure and Applied Mathematics, vol. 7 (Interscience, London, 1958); with the assistance of W. G. Bade and R. G. Bartle; [MR 0117523](#) (22 #8302).
- Gos72 R. W. Gosper, *Continued fraction arithmetic*, in *HAKMEM Item 101B*, MIT Artificial Intelligence Memo 239 (MIT, Cambridge, MA, 1972).
- IG90 M. Iosifescu and Ş. Grigorescu, *Dependence with complete connections and its applications*, Cambridge Tracts in Mathematics, vol. 96 (Cambridge University Press, Cambridge, 1990); [MR 1070097](#) (91j:60098).
- JL88 H. Jager and P. Liardet, *Distributions arithmétiques des dénominateurs de convergents de fractions continues*, Nederl. Akad. Wetensch. Indag. Math. **50** (1988), 181–197; [MR 952514](#).
- KN00 C. Kraaikamp and H. Nakada, *On normal numbers for continued fractions*, Ergodic Theory Dynam. Systems **20** (2000), 1405–1421.
- KSS12 C. Kraaikamp, T. A. Schmidt and W. Steiner, *Natural extensions and entropy of α -continued fractions*, Nonlinearity **25** (2012), 2207–2243; [MR 2946184](#).
- LS98 P. Liardet and P. Stambul, *Algebraic computations with continued fractions*, J. Number Theory **73** (1998), 92–121; [MR 1654886](#) (2000c:11014).
- Mau00 C. Mauduit, *Problem session dedicated to Gérard Rauzy*, Dynamical systems (Luminy–Marseille, 1998), 1–19 (World Scientific, River Edge, NJ, 2000), [MR 1796141](#) (2001i:11001).
- Mey00 C. Meyer, *Matrix analysis and applied linear algebra* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000); with 1 CD-ROM (Windows, Macintosh and UNIX) and a solutions manual; [MR 1777382](#).
- MS03 N. G. Moshchevitin and I. D. Shkredov, *On the Pyatetskii-Shapiro criterion for normality*, Mat. Zametki **73** (2003), 577–589; [MR 1991904](#) (2005e:37013).
- Ran73 G. N. Raney, *On continued fractions and finite automata*, Math. Ann. **206** (1973), 265–283; [MR 0340166](#) (49 #4922).
- SCZ11 L. Saloff-Coste and J. Zúñiga, *Merging and stability for time inhomogeneous finite Markov chains*, Surveys in stochastic processes, EMS Ser. Congr. Rep. (2011), 127–151; [MR 2883857](#) (2012m:60166).
- Sch95 F. Schweiger, *Ergodic theory of fibred systems and metric number theory*, Oxford Science Publications (The Clarendon Press, Oxford University Press, New York, 1995); [MR 1419320](#) (97h:11083).

PRESERVING CONTINUED FRACTION NORMALITY

- Van14 J. Vandehey, *On the joint normality of certain digit expansions*, Preprint (2014), [arXiv:1408.0435](https://arxiv.org/abs/1408.0435).
- Wal50 D. D. Wall, *Normal numbers*, ProQuest LLC, Ann Arbor, MI, 1950, PhD thesis, University of California, Berkeley, CA; [MR 2937990](https://www.jstor.org/stable/2937990).

Joseph Vandehey vandehey.1@osu.edu
Ohio State University, 100 Math Tower, 231 West 18th Avenue,
Columbus OH, 43210-1174, USA