#A13 INTEGERS 9 (2009), 141-162

THE 3x+1 CONJUGACY MAP OVER A STURMIAN WORD

Josefina López
Santiago Yosondúa, Oaxaca, México
josefinapedro@hotmail.com

Peter Stoll
Santiago Yosondúa, Oaxaca, México
josefinapedro@hotmail.com

Received: 12/1/08, Revised: 2/25/09, Accepted: 3/5/09

Abstract

The $3x+1$ map T is defined on the 2-adic integers $\mathbb{Z}_2$ by $T(x) = x/2$ for even x and $T(x) = (3x+1)/2$ for odd x. Under iteration of T, the sequence $(T^k(x) \mod 2)_{k=0}^{\infty}$, called the *parity vector* of $x \in \mathbb{Z}_2$, can be interpreted as an infinite word over the alphabet $\{0, 1\}$ or as the digits of the 2-adic integer $\Phi^{-1}(x) = \sum_{k=0}^{\infty}(T^k(x) \bmod 2) \cdot 2^k$. For any $v \in \mathbb{Z}_2$ (or equivalently for the infinite word v), the inverse map Φ (called the $3x+1$ *conjugacy map*) yields the unique $x \in \mathbb{Z}_2$ with parity vector v. It is unknown if there exists any aperiodic v with an eventually periodic $\Phi(v)$. In this paper we compute $\Phi(v)$ for a class of aperiodic infinite words v of minimal complexity, the mechanical words with irrational slope and intercept 0. Our main result is a generalized continued fraction expansion of $-1/\Phi(x)$, convergent under the 2-adic metric of $\mathbb{Z}_2$. The given examples suggest that Φ always maps Sturmian words to infinite words of full complexity.

1. Introduction

Let $\mathbb{Z}_2$ denote the ring of 2-adic integers. Each $x \in \mathbb{Z}_2$ can be expressed uniquely as an infinite string $x_0x_1x_2\cdots$ of 1's and 0's, called the *binary representation* of x. The x_k are the digits of x, written from left to right. For instance, $-1 = 1111\cdots$ and $1 = 1000\cdots$. The 2-adic norm $|\cdot|_2$ in $\mathbb{Z}_2$ is given by $|x|_2 := 2^{-n}$ if $x \neq 0$ and $|x|_2 := 0$ if $x = 0$, where x_n is the first nonzero digit of x. The distance is defined by $d(x, y) = |x - y|_2$ for all $x, y \in \mathbb{Z}_2$. With this metric $\mathbb{Z}_2$ is a compact and complete topological space. Let $0 \leq d_0 < d_1 < d_2 < \cdots$ be a finite or infinite sequence of nonnegative integers defined by $d_i := k$ whenever $x_k = 1$ for a 2-adic integer $x = x_0x_1\cdots x_k\cdots$. Then x can be written as the finite or infinite sum $x = 2^{d_0} + 2^{d_1} + 2^{d_2} + \cdots$.

The $3x+1$ map T is defined on the 2-adic integers $\mathbb{Z}_2$ by $T(x) = x/2$ for even x and $T(x) = (3x+1)/2$ for odd x. The 2-adic shift map S is defined on the 2-adic integers $\mathbb{Z}_2$ by $S(x) = x/2$ for even x and $S(x) = (x-1)/2$ for odd x.

The maps T and S are conjugates (Bernstein, Lagarias [2]). There exists a unique homeomorphism $\Phi : \mathbb{Z}_2 \to \mathbb{Z}_2$ (the $3x+1$ conjugacy map) with $\Phi(0) = 0$ and

$$\Phi \circ S \circ \Phi^{-1} = T. \tag{1}$$

There are explicit formulas for Φ and Φ^{-1}:

$$\Phi(2^{d_0} + 2^{d_1} + 2^{d_2} + \cdots) = -\sum_i \frac{1}{3^{i+1}} 2^{d_i} \tag{2}$$

(see [1]) and

$$\Phi^{-1}(x) = \sum_{k=0}^{\infty} (T^k(x) \mod 2) \cdot 2^k \tag{3}$$

(see ([2]), where $T^k(x)$ denotes the k-th iterate of T and $T^0(x) := x$.

Also, Φ is a 2-adic isometry ([2]):

$$|\Phi(x) - \Phi(y)|_2 = |x - y|_2 \qquad \text{for all } x, y \in \mathbb{Z}_2. \tag{4}$$

Moreover ([2]),

$$\Phi(x) \equiv x \pmod 2 \qquad \text{for all } x \in \mathbb{Z}_2. \tag{5}$$

The sequence

$$(T^k(x) \mod 2)_{k=0}^{\infty} \tag{6}$$

is called the *parity vector* of $x \in \mathbb{Z}_2$ (Lagarias [6]). Its elements are the digits of the 2-adic integer $\Phi^{-1}(x)$. The concatenation of these digits is an infinite word v over the alphabet $\{0, 1\}$, written from left to right:

$$v = x_0 x_1 x_2 \cdots, \qquad \text{where } x_k \equiv T^k(x) \pmod 2.$$

We define $d_i := k$ whenever $x_k = 1$. Then we can define $\Phi(v) := \Phi(2^{d_0} + 2^{d_1} + 2^{d_2} + \cdots)$ and refer to Φ, whether it is a 2-adic integer or an infinite word.

Let $\mathbb{A}^{\mathbb{N}_0}$ be the set of (right) infinite words over the alphabet $\mathbb{A} = \{0, 1\}$:

$$\mathbb{A}^{\mathbb{N}_0} := \{x_0 x_1 x_2 \cdots : x_k \in \mathbb{A},\ k = 0, 1, 2, \ldots\}.$$

The set $\mathbb{A}^{\mathbb{N}_0}$ is equipped with a distance defined as follows:

$$\begin{aligned} &\text{for } x, y \in \mathbb{A}^{\mathbb{N}_0}, \quad d(x, y) := 2^{-n} \\ &\text{with } n = \min\{k \geq 0 : x_k \neq y_k\}, \end{aligned}$$

and the convention that $d(x, y) := 0$ if and only if $x = y$. With this metric (essentially the same as the metric of $\mathbb{Z}_2$) $\mathbb{A}^{\mathbb{N}_0}$ is a compact and complete topological space (Lothaire [8], Chapter 1).

Note that $\mathbb{A}^{\mathbb{N}_0}$ and $\mathbb{Z}_2$ are homeomorphic spaces: both are homeomorphic to the Cantor space. A sequence of words in $\mathbb{A}^{\mathbb{N}_0}$ converges to a limit $x \in \mathbb{A}^{\mathbb{N}_0}$ if and only if the corresponding sequence of 2-adic integers converges to $y \in \mathbb{Z}_2$, and such that y is the 2-adic integer corresponding to the word x. This fact simplifies the redaction and the notation: If our interest is the digits structure of an $x \in \mathbb{Z}_2$, we use the language of words. We even use the same symbols for words and 2-adic integers when there is no confusion in the context.

Let $\mathbb{Q}_{\text{odd}}$[1] denote the ring of rational numbers, having an odd denominator in reduced fraction form. We know that $\mathbb{Q}_{\text{odd}}$ is isomorphic to the subring $\mathbb{Q}_2 \subset \mathbb{Z}_2$ of eventually periodic 2-adic integers (i.e., their word of digits is eventually periodic). This isomorphism enables us to do arithmetic within $\mathbb{Q}_{\text{odd}}$ instead of struggling with the cumbersome elements of $\mathbb{Q}_2$. We take care that no fractions with even denominator arise. The elements $a/b \in \mathbb{Q}_{\text{odd}}$ (and also $a/b \in \mathbb{Q}$) have a 2-adic norm given by

$$\left|\frac{a}{b}\right|_2 := 2^{-\text{val}_2(\frac{a}{b})} \text{ where } \text{val}_2\left(\frac{a}{b}\right) := \max\left\{r \,:\, 2^r \text{ divides } \frac{a}{b}\right\} \geq 0;\ \text{val}_2(0) := \infty.$$

If the parity vector (6) of some $x \in \mathbb{Z}_2$ is an eventually periodic infinite word v, then $x = \Phi(v)$ is also eventually periodic. This follows from (2) by computing the corresponding geometric series. Thus $\Phi(\mathbb{Q}_{\text{odd}}) \subset \mathbb{Q}_{\text{odd}}$. The periodicity conjecture, concerning the famous $3x+1$ problem, states that $\Phi(\mathbb{Q}_{\text{odd}}) = \mathbb{Q}_{\text{odd}}$ (Bernstein, Lagarias [2]). If this conjecture is true, Φ maps each aperiodic parity vector v onto an aperiodic 2-adic integer: $\Phi(v) \notin \mathbb{Q}_{\text{odd}}$.

The "nicest" aperiodic infinite words are the Sturmian words: infinite words over the alphabet $\{0,1\}$ which have exactly $(n+1)$ different *factors*[2] of length n for each $n \geq 0$. Indeed, Sturmian words are aperiodic infinite words of minimal complexity (see [8]). They can even be described explicitly in arithmetic form, known as lower and upper mechanical words (see [8]):

$$\lfloor (j+1)\alpha + \rho \rfloor - \lfloor j\alpha + \rho \rfloor \quad \text{or} \quad \lceil (j+1)\alpha + \rho \rceil - \lceil j\alpha + \rho \rceil \quad \text{for } j = 0,1,2,\ldots$$

where $\alpha \in (0,1)$ is an irrational number (the *slope*) and $\rho \in [0,1)$ (the *intercept*).

In this paper, we compute $\Phi(v)$ for mechanical words v with intercept 0. A special word

$$c_\alpha := \lfloor (j+1)\alpha \rfloor - \lfloor j\alpha \rfloor = \lceil (j+1)\alpha \rceil - \lceil j\alpha \rceil \quad \text{for } j = 1,2,3,\ldots$$

is called the *characteristic word*. Note that here $j \neq 0$. Then

$$0c_\alpha = \lfloor (j+1)\alpha \rfloor - \lfloor j\alpha \rfloor \quad \text{and } 1c_\alpha = \lceil (j+1)\alpha \rceil - \lceil j\alpha \rceil \quad \text{for } j = 0,1,2,\ldots.$$

Let $\alpha = [0; a_1, a_2, \ldots]$ be the simple continued fraction expansion of the irrational number α with partial denominators $(a_k)_{k\geq 0}$,

$$\alpha = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

[1]Monks, Yazinski [9]. Halbeisen, Hungerbűhler [5] use $\mathbb{Q}[(2)]$.

[2]A finite word $w = w_0 w_1 \cdots w_{n-1}$ is a factor of an infinite word v if $w = v_i v_{i+1} \cdots v_{i+n-1}$ for some integer i.

and convergents $(p_k/q_k)_{k\geq 0}$ defined by

$$\begin{aligned} p_{-2} &:= 0, & q_{-2} &:= 1, \\ p_{-1} &:= 1, & q_{-1} &:= 0, \\ p_k &= a_k p_{k-1} + p_{k-2}, \\ q_k &= a_k q_{k-1} + q_{k-2}. \end{aligned}$$

Then always

$$\begin{aligned} p_0 &= 0, & q_0 &= 1, \\ p_1 &= 1, & q_1 &= a_1. \end{aligned}$$

For any irrational $\alpha \in (0,1)$ given as a simple continued fraction, we obtain a 2-adic convergent series expansion in terms of p_k's and q_k's for $\Phi(1c_\alpha) \in \mathbb{Z}_2$ (Theorem 1). From $\Phi(1c_\alpha)$ one easily gets by (1) and (5):

$$\Phi(c_\alpha) = \frac{3\Phi(1c_\alpha)+1}{2} \quad \text{and} \quad \Phi(0c_\alpha) = 3\Phi(1c_\alpha)+1.$$

As main result we get a convergent generalized continued fraction expansion of $-1/\Phi(1c_\alpha)$ in $\mathbb{Z}_2$, formally with rational integers as partial denominators and numerators (Corollary 2).

In Section 2 we summarize our results without proof. We give several examples of $\Phi(1c_\alpha)$'s for different α-values. The examples suggest the full complexity of the infinite words, i.e., they have 2^n different factors of length n for every $n \geq 0$. We show the exact number of digits, necessary for checking the claimed complexity up to the bound $n \leq 5$.

The proof of the main result, concerning 2-adic integers, is in Section 3. In Section 4 we prove that an associated real-valued function $\Phi_{\mathbb{R}}(1c_\alpha)$ is a *devil's staircase*. This function with the same series expansion explains the underlying idea when computing $\Phi(1c_\alpha)$.

2. Results

Theorem 1. *Let* $\alpha = [0; a_1, a_2, \ldots]$ *be the simple continued fraction expansion of the irrational number* α *with convergents* (p_k/q_k)*,* $1c_\alpha = \lceil (j+1)\alpha \rceil - \lceil j\alpha \rceil$ *for* $j = 0, 1, 2, \ldots$ *and* $1c_\alpha \in \mathbb{Z}_2$*. Then it holds in* $\mathbb{Z}_2$*:*

$$\Phi(1c_\alpha) = -\frac{1}{3} - \sum_{j=0}^{\infty} (-1)^{j+1} \frac{2^{q_{j+1}+q_j-1}}{3(3^{p_{j+1}} - 2^{q_{j+1}})(3^{p_j} - 2^{q_j})}.$$

Corollary 2. *Let α, (p_k/q_k) and $1c_\alpha$ be as in Theorem 1. Then it holds in $\mathbb{Z}_2$:*

$$-\frac{1}{\Phi(1c_\alpha)} = B_0 + \cfrac{A_1}{B_1 + \cfrac{A_2}{B_2 + \cfrac{A_3}{B_3 + \cdots}}},$$

where

$$\begin{aligned} B_0 &= 3, & & \\ B_1 &= -1, & A_1 &= 2^{q_1} = 2^{a_1}, \\ B_{k+1} &= 3^{(-1)^{k+1}} \cdot 3^{p_{k-1}} \cdot c_k, & A_{k+1} &= 2^{q_{k+1}-q_{k-1}}, \\ \text{and} \quad c_k &= \frac{3^{p_k a_{k+1}} - 2^{q_k a_{k+1}}}{3^{p_k} - 2^{q_k}} & & \text{for } k = 1, 2, 3, \ldots . \end{aligned}$$

Example 3. (see also Example 7). *Let $(F_k)_{k=0}^{\infty}$ be the Fibonacci Sequence defined by $F_0 := 0$, $F_1 := 1$ and $F_k = F_{k-1} + F_{k-2}$ for $k \geq 2$. Let γ denote the golden ratio: $\gamma = \frac{1+\sqrt{5}}{2}$. For the irrational $1/\gamma = 0.6180\cdots$, the following holds in $\mathbb{Z}_2$:*

$$\begin{aligned} -\frac{1}{\Phi(1c_{(1/\gamma)})} &= -\frac{1}{\sum_i \frac{2^{\lfloor i\gamma \rfloor}}{3^{1+i}}} \\ &= 3 + \cfrac{2^{F_1}}{-1 + \cfrac{2^{F_2}}{3^{F_0+1} + \cfrac{2^{F_3}}{3^{F_1-1} + \cfrac{2^{F_4}}{3^{F_2+1} + \cfrac{2^{F_5}}{3^{F_3-1} + \cfrac{2^{F_6}}{3^{F_4+1} + \cdots}}}}}}. \end{aligned}$$

This expansion is a new member of the family of remarkable sequences related to the golden ratio, but now in the 2-adic world. For instance, there is the famous expansion of the Rabbit Constant $\sum_{i=1}^{\infty} \frac{1}{2^{\lfloor i\gamma \rfloor}} = [0; 2^{F_0}, 2^{F_1}, 2^{F_2}, 2^{F_3}, \ldots] = 0.70980344\cdots$ (Davison [4]). Our expansion converges in $\mathbb{Z}_2$ but diverges in $\mathbb{R}$. However, the divergence is acceptable: it diverges by oscillation between two distinct irrational limit points ζ and $(\zeta - 1/6)$; the odd convergents approach $\zeta = 10.37012714\cdots$ and the even approach $(\zeta - 1/6)$. Defining a new map Φ^* which is dual to Φ, we get the following expansion, convergent in $\mathbb{R}$, which proves the irrationality of ζ (relation (23)).

$$\Phi^*_{\mathbb{R}}(1c_{(1/\gamma)}) = \zeta = \cfrac{-3^{F_0}}{2^{F_0+1} + \cfrac{-3^{F_1}}{2^{F_1-1} + \cfrac{3^{F_2}}{2^{F_2+1} + \cfrac{3^{F_3}}{2^{F_3-1} + \cfrac{3^{F_4}}{2^{F_4+1} + \cfrac{3^{F_5}}{2^{F_5-1} + \cfrac{3^{F_6}}{2^{F_6+1} + \cdots}}}}}}}.$$

An infinite word w has full complexity if there are 2^n different factors of length n for every $n > 0$. Let $D(n)$ denote the minimal number of digits such that the prefix of w with length $D(n)$ has 2^n different factors of length n. In the following examples we use prefixes of length $D(5)$, i.e., $D(5)$ digits are needed for finding all of the $2^5 = 32$ different factors of length 5 in $\Phi(1c_\alpha)$.

Example 4.

$$\alpha = \frac{\ln(3)}{27} = [0; 24, 1, 1, 2, 1, 3, 2, 1, \ldots] = 0.0406\cdots$$

$$(p_k/q_k)_{k=0}^{\infty} = (0, 1/24, 1/25, 2/49, 5/123, 7/172, 26/639, 59/1450, 85/2089, \ldots);$$

$$(A_k)_{k=1}^{\infty} = (16777216, 16777216, 33554432, \qquad 316912650057057350374175801344, ...),$$

$$(B_k)_{k=0}^{\infty} = (3, -1, 3, 1, 5066549580791889, 3, ...),$$

$1c_\alpha =$ 10000000000000000000000001000000000000000000000001
00000000000000000000000100000000000000000000000010
00000000000000000000001000000000000000000000000100
00000000000000000000001000000000000000000000001000
000000000000000,

$\Phi(1c_\alpha) =$ 10101010101010101010101000111000111000111000111001
101111101001000010110111001011101101011001111110010
001111110011100110010100100011000010100101010111100
001110100010001000001011111111101001100101000110011 0
110111001100000,

$D(5) =$ 215.

Example 5.

$$
\begin{aligned}
\alpha = &\quad \frac{\pi}{6} = [0; 1, 1, 10, 10, 1, 1, 1, \ldots] = 0.5235\cdots\ , \\
(p_k/q_k)_{k=0}^{\infty} = &\quad (0, 1, 1/2, 11/21, 111/212, 122/233, 233/445, 355/678, \ldots); \\
(A_k)_{k=1}^{\infty} = &\quad (2, 2, 1048576, 164550455732120604215496918255735050498 2 \\
&\qquad 735865633579863348609024, ...), \\
(B_k)_{k=0}^{\infty} = &\quad (3, -1, 3, 989527, 7713282525627257030267828369842165309 \\
&\qquad 7678658419723584295555, 59049, ...), \\
1c_\alpha = &\quad 11010101010101010101011010101010101010101011010101 \\
&\quad 01010101010101101010101010101010101101010101010101 \\
&\quad 01010110101, \\
\Phi(1c_\alpha) = &\quad 11010101010101010101011001100110101100100100011100 \\
&\quad 01011000011011111110100111001100110100101110111100 \\
&\quad 11100100000, \\
D(5) = &\quad 111.
\end{aligned}
$$

Example 6.

$$
\begin{aligned}
\alpha = &\quad \frac{1}{\ln(3)} = [0; 1, 10, 7, 9, 2, 2, \ldots] = 0.9102\cdots\ , \\
(p_k/q_k)_{k=0}^{\infty} = &\quad (0, 1, 10/11, 71/78, 649/713, 1369/1504, 3387/3721, \ldots); \\
(A_k)_{k=1}^{\infty} = &\quad (2, 1024, 1511157274518286468382 72, ...), \\
(B_k)_{k=0}^{\infty} = &\quad (3, -1, 174075, 43914238431643758422900358577, ...), \\
1c_\alpha = &\quad 11111111111011111111110111111111101111111111011111 \\
&\quad 11111011111111110111111111101111111111101111111111 \\
&\quad 0111111111101111111111011111111110, \\
\Phi(1c_\alpha) = &\quad 11111111111001100110001001111101001100101100111001 \\
&\quad 00100101011100100010010101001111000001110011111100 \\
&\quad 1011100001101001001101001010011011, \\
D(5) = &\quad 134.
\end{aligned}
$$

Example 7.

$$\alpha = \frac{2}{1+\sqrt{5}} = [0;1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,\ldots] = 0.6180\cdots,$$

$$(p_k/q_k)_{k=0}^{\infty} = (0,1,1/2,2/3,3/5,5/8,8/13,13/21,21/34,34/55,55/89,$$
$$89/144,144/233,233/377,377/610,610/987,\ldots);$$

$$(A_k)_{k=1}^{\infty} = (2,2,4,8,32,256,8192,2097152,17179869184,$$
$$36028797018963968,618970019642690137449562112,...),$$

$$(B_k)_{k=0}^{\infty} = (3,-1,3,1,9,3,81,81,19683,531441,31381059609,$$
$$5559060566555523,523347633027360537213511521,...),$$

$$1c_\alpha = 11011010110110101101011011010110110101101011011010$$
$$11010110110101101101011010110110101101101011010110$$
$$110101,$$

$$\Phi(1c_\alpha) = 11011110111001000110011110100010010111011101011000$$
$$10000011011111110000100100111000001001011010001011$$
$$101010,$$

$$D(5) = 106.$$

Example 8.

$$\alpha = \frac{\ln(2)}{\ln(3)} = [0;1,1,1,2,2,3,1,5,2,\ldots] = 0.6309\cdots,$$

$$(p_k/q_k)_{k=0}^{\infty} = (0,1,1/2,2/3,5/8,12/19,41/65,53/84,306/485,665/1054,\ldots);$$

$$(A_k)_{k=1}^{\infty} = (2,2,4,64,65536,144115188075855872,$$
$$36893488147419103232,...),$$

$$(B_k)_{k=0}^{\infty} = (3,-1,3,1,153,1497,609397039593657,177147,...),$$

$$1c_\alpha = 11011011010110110101101101101011011010110110110101$$
$$1011010110110110101101101011011010110110101101101,$$

$$\Phi(1c_\alpha) = 11011111110110100111110110010011110101010000001010$$
$$0000001011010101111111001011101100010100001000110,$$

$$D(5) = 99.$$

Example 9.

$$\begin{aligned}
\alpha = &\quad \ln(2) = [0; 1, 2, 3, 1, 6, 3, 1, 1, 2, \ldots] = 0.6931\cdots,\\
(p_k/q_k)_{k=0}^{\infty} = &\quad (0, 1, 2/3, 7/10, 9/13, 61/88, 192/277, 253/365,\\
&\qquad 445/642, 1143/1649, \ldots);\\
(A_k)_{k=1}^{\infty} = &\quad (2, 4, 512, 1024, 30223145490365729367654 4, ...),\\
(B_k)_{k=0}^{\infty} = &\quad (3, -1, 15, 217, 27, 36699009263416097247588 75, ...),\\
1c_\alpha = &\quad 11101101101110110110110111011011011011101101101101\\
&\quad 11011011011011101101101101110110110110111011011011\\
&\quad 10110110110111011011011011101101101101110110110110\\
&\quad 1110110110110111011011011011,\\
\Phi(1c_\alpha) = &\quad 11100101010111010000101010000111110000001100111001 1\\
&\quad 10010001110100100010010100101110101110011010011100\\
&\quad 10000010010011111111001101001000111000000100011001\\
&\quad 1101110101001110100000110110,\\
D(5) = &\quad 178.
\end{aligned}$$

Here are some additional values of the function $D(n)$:

n	$\frac{\ln(3)}{27}$ 0.0406...	$\frac{\pi}{6}$ 0.5235...	$\frac{1}{\ln(3)}$ 0.9102...	$\frac{2}{1+\sqrt{5}}$ 0.6180...	$\frac{\ln(2)}{\ln(3)}$ 0.6309...	$\ln(2)$ 0.6931...
1	2	3	12	3	3	4
2	28	25	14	13	17	6
3	30	48	32	17	43	19
4	65	66	86	55	47	42
5	215	111	134	106	99	178
6	252	335	263	211	304	448
7	715	629	896	909	614	553
8	1105	1615	1832	1644	1579	1806

3. Proofs of Theorem 1 and Corollary 2

In this section, if not otherwise stated, p/q denotes any rational number in reduced fraction form with $0 < p/q \leq 1$ (the denominator can be even), α denotes any irrational number with $0 < \alpha < 1$, and p_k/q_k are its convergents. We denote by $\mathbb{A}^*$ the set of finite words over $\mathbb{A}$, by ε the empty word, by $\ell(w)$ the length of the word w, and by $h(w)$ its height, i.e., the number of 1's in w. A word (finite or infinite) is called *balanced* if the height of any two factors of the same length differ by at most 1. Sturmian words are aperiodic, balanced, infinite words (see [8]).

Definition 10. Let M_2 denote the set of the 2-adic integers having a rational or irrational upper mechanical word as digits structure:

$$M_2 := \{m_x \in \mathbb{Z}_2 \,:\, m_x = \lceil (j+1)x \rceil - \lceil jx \rceil \textit{ for } j = 0,1,2,\ldots \textit{ and } 0 < x \leq 1\}.$$

If $x = \alpha$, then $m_\alpha = 1c_\alpha$. We have to compute $\Phi(m_\alpha)$ in terms of convergents p_k/q_k. Note that $x \neq 0$ because $\Phi(0) := 0$ (an infinite word of 0's).

If $x = p/q$, then $m_{p/q}$ is a purely periodic balanced infinite word. The period $\overline{m}_{p/q}$, a finite word, has length q and height p. This word is known as the *Christoffel word* ([8]). It always starts with 1 and ends with 0, and thus $\overline{m}_{p/q} = 1z0$. The central word z is a palindromic word with $\ell(z) = q-2$ and $h(z) = p-1$. *For example,* $m_{5/7} = 11101101110110111011011101101110\cdots$ and $\overline{m}_{5/7} = 1110110$.

Lemma 11. *For all $x \in M_2$,*

$$d_i = \left\lfloor \frac{i}{x} \right\rfloor \qquad (i = 0,1,2,\ldots).$$

Proof. The d_i are those $j \in \mathbb{N}_0$ for which $\lceil (j+1)x \rceil - \lceil jx \rceil = 1$.

a) Let $0 < x < 1$. Fix any $i \in \mathbb{N}_0$. There exists a j such that $jx \leq i < (j+1)x$; thus $j \leq i/x < j+1$. So $j = \lfloor \frac{i}{x} \rfloor$.

b) If $x = 1$, then $j = d_i = i$. □

Lemma 12. *For all $m_{p/q} \in M_2$,*

$$\Phi(m_{p/q}) = \frac{3^p}{2^q - 3^p} \sum_{i=0}^{p-1} \frac{1}{3^{1+i}} \cdot 2^{\lfloor i \cdot \frac{q}{p} \rfloor}.$$

Proof. Let $i = r + np$ and $0 \leq r < p$. Then by (2) and Lemma 11,

$$\begin{aligned}
\Phi(m_{p/q}) \qquad &= -\sum_{i=0}^{\infty} \frac{1}{3^{1+i}} \cdot 2^{\lfloor i \cdot \frac{q}{p} \rfloor} = -\sum_{n=0}^{\infty} \frac{2^{nq}}{3^{np}} \sum_{r=0}^{p-1} \frac{2^{\lfloor r \cdot \frac{q}{p} \rfloor}}{3^{1+r}} \\
&= \frac{3^p}{(2^q - 3^p)} \sum_{r=0}^{p-1} \frac{1}{3^{1+r}} \cdot 2^{\lfloor r \cdot \frac{q}{p} \rfloor}
\end{aligned}$$

since $\left| \frac{2^q}{3^p} \right|_2 < 1$. □

Definition 13. (Halbeisen, Hungerbühler [5]). The function $\varphi : \mathbb{A}^* \to \mathbb{N}_0$ is defined recursively by

$$\begin{aligned} \varphi(\varepsilon) &= 0; \\ \varphi(w0) &= \varphi(w); \\ \varphi(w1) &= 3\varphi(w) + 2^{\ell(w)}. \end{aligned}$$

Using the pointer notation d_i, we get ([5])

$$\varphi(w) = \sum_{i=0}^{h(w)-1} 3^{h(w)-1-i} 2^{d_i}. \tag{7}$$

Further ([5]), for all $u, v \in \mathbb{A}^*$,

$$\varphi(uv) = 3^{h(v)}\varphi(u) + 2^{\ell(u)}\varphi(v). \tag{8}$$

Lemma 14. *For all $m_{p/q} \in M_2$,*

$$\Phi(m_{p/q}) = \frac{\varphi(\overline{m}_{p/q})}{2^q - 3^p}\ .$$ [3]

Proof. Apply Lemma 12 and (7). □

Clearly $\Phi(m_{p/q}) \in \mathbb{Q}_{\text{odd}}$.

It is a main fact in the theory of words that the Christoffel words $\overline{m}_{p_k/q_k}$ (p_k/q_k are the convergents of α) converge to the word $1c_\alpha$:

Lemma 15. *Let $v_k := \overline{m}_{p_k/q_k}$ for $k \geq 2$ and $v_0 := 0$, $v_1 := 1(0)^{a_1-1}$.* [4]
Then $m_\alpha = 1c_\alpha = \lim_{k\to\infty} v_k$. In addition, for $k \geq 1$,

$$v_{k+1} = \begin{cases} v_k^{a_{k+1}} v_{k-1} & \text{if } k \text{ odd;} \\ v_{k-1} v_k^{a_{k+1}} & \text{if } k \text{ even.} \end{cases}$$

Proof. The statement is part of Exercise 2.2.10 in Lothaire [8]. □

If $v_k \to m_\alpha$ then we have $\Phi(v_k) \to \Phi(m_\alpha)$. We now construct a new sequence $(-P_k/Q_k)_{k=0}^{\infty}$, slightly different from $\Phi(v_k)$, but with the same property: $(-P_k/Q_k) \to \Phi(m_\alpha)$.

The following function g has its origin in a devil's staircase (see Section 4).

---

[3] It is easy to prove that $\varphi(\overline{m}_{p/q})$ is the same quantity as $M_{\ell,n}$ in [5], Corollary 1 (with $\ell = q, n = p$).

[4] $(0)^{a_1-1}$ means $(a_1 - 1)$ times 0.

Definition 16. (The function "right-gap"). *Let* $g := \mathbb{Q} \cap [0,1] \to \mathbb{Q}_{odd}$ *be defined by*

$$g\left(\frac{p}{q}\right) := \frac{1}{3} \cdot \frac{2^{q-1}}{3^p - 2^q}.$$

If $p_k/q_k \to \alpha$, then $q_k \to \infty$. Thus $g(p_k/q_k)$ converges to 0 since $|g(p_k/q_k)|_2 = 2^{1-q_k}$. Consequently, the sequence

$$\begin{aligned} &\Phi(v_0) + g\left(\tfrac{p_0}{q_0}\right), \quad \Phi(v_1), \quad \Phi(v_2) + g\left(\tfrac{p_2}{q_2}\right), \quad \Phi(v_3), \\ &\qquad\qquad\qquad\qquad \Phi(v_4) + g\left(\tfrac{p_4}{q_4}\right), \quad \Phi(v_5), \cdots \end{aligned} \tag{9}$$

converges to $\Phi(m_\alpha)$. The terms are

$$\begin{aligned} &-\frac{3\varphi(v_0) - 2^{q_0-1}}{3(3^{p_0} - 2^{q_0})}, \quad -\frac{\varphi(v_1)}{3^{p_1} - 2^{q_1}}, -\frac{3\varphi(v_2) - 2^{q_2-1}}{3(3^{p_2} - 2^{q_2})}, \\ &\qquad\qquad -\frac{\varphi(v_3)}{3^{p_3} - 2^{q_3}}, -\frac{3\varphi(v_4) - 2^{q_4-1}}{3(3^{p_4} - 2^{q_4})}, \ldots . \end{aligned} \tag{10}$$

We write P_k for the numerators and Q_k for the denominators; the "-" sign remains:

$$-\frac{P_0}{Q_0} = -\frac{-1}{-3}, \quad -\frac{P_1}{Q_1} = -\frac{1}{3 - 2^{q_1}}, \quad -\frac{P_2}{Q_2}, -\frac{P_3}{Q_3}, -\frac{P_4}{Q_4}, \ldots .$$

In conclusion, we have the following lemma.

Lemma 17.

$$\lim_{k\to\infty} \Phi(v_k) = -\lim_{k\to\infty} \frac{P_k}{Q_k} = \Phi(m_\alpha).$$

Proof. The statement follows from (9). □

Lemma 18. *For* $k \geq 1$*,*

$$\begin{aligned} P_{k+1} &= 3^{(-1)^{k+1}} 3^{p_{k-1}} c_k P_k + 2^{q_{k+1}-q_{k-1}} P_{k-1}, \\ Q_{k+1} &= 3^{(-1)^{k+1}} 3^{p_{k-1}} c_k Q_k + 2^{q_{k+1}-q_{k-1}} Q_{k-1}, \\ \textit{where} \quad c_k &:= \frac{3^{p_k a_{k+1}} - 2^{q_k a_{k+1}}}{3^{p_k} - 2^{q_k}}. \end{aligned}$$

Proof. We divide the proof in four parts.

(a) The relation for Q_{k+1} follows from the identity

$$3^{p_{k+1}} - 2^{q_{k+1}} = 3^{p_{k-1}}(3^{p_k a_{k+1}} - 2^{q_k a_{k+1}}) + 2^{q_k a_{k+1}}(3^{p_{k-1}} - 2^{q_{k-1}}).$$

Recall that $p_{k+1} = a_{k+1}p_k + p_{k-1}$ and $q_{k+1} = a_{k+1}q_k + q_{k-1}$.

(b) By induction from (8),

$$\varphi(v_k^{a_{k+1}}) = c_k\varphi(v_k).$$

(c) Let k be odd.

$$\varphi(v_{k+1}) \quad = \quad \varphi(v_k^{a_{k+1}}v_{k-1}) = 3^{p_{k-1}}c_k\varphi(v_k) + 2^{q_k a_{k+1}}\varphi(v_{k-1}).$$

Then

$$\begin{aligned} P_{k+1} = &\quad 3\varphi(v_{k+1}) - 2^{q_{k+1}-1} \\ = &\quad 3\cdot 3^{p_{k-1}}c_k\varphi(v_k) + 2^{q_k a_{k+1}}(3\varphi(v_{k-1}) - 2^{q_{k-1}-1}). \end{aligned}$$

(d) Let k be even. The word z_k in $v_k = 1z_k0$ is *central* and, by (Proposition 2.2.15, Lothaire [8]), the words $s_k := z_k10$ and $s'_k = z_k01$ are *standard words* . The standard sequence is defined by $s_{-1} := 1$, $s_0 := 0$ and $s_n = s_{n-1}^{t_n}s_{n-2}$ for $n \geq 1$ (Lothaire [8]), where $\alpha = [0; 1 + t_1, t_2, \ldots]$ is the continued fraction expansion.[5] For k even, there are the bijections

$$\begin{aligned} v_k = 1z_k0 &\quad\longleftrightarrow\quad s_k = z_k10, \\ v_{k+1} = 1z_{k+1}0 &\quad\longleftrightarrow\quad s_{k+1} = z_{k+1}01. \end{aligned}$$

So we get

$$\begin{array}{rcccl} v_k = 1z_k0 & \longrightarrow & s_k & = & z_k10 \\ & & & \Downarrow & \\ v_{k+1} = 1(z_k10)^{a_{k+1}}z_{k-1}0 & \longleftarrow & s_{k+1} & = & (z_k10)^{a_{k+1}}z_{k-1}01, \end{array}$$

and

$$\begin{aligned} P_{k+1} &= \varphi(v_{k+1}) = \varphi(1(z_k10)^{a_{k+1}}z_{k-1}0) = 3^{p_{k+1}-1} + 2\varphi((z_k10)^{a_{k+1}}z_{k-1}) \\ &= 2\cdot 3^{p_{k-1}-1}c_k(3\varphi(z_k) + 2^{q_k-2}) + 2^{q_k a_{k+1}+1}\varphi(z_{k-1}) + 3^{p_{k+1}-1}. \end{aligned}$$

But $\varphi(v_k) = \varphi(1z_k0) = 3^{p_k-1} + 2\varphi(z_k)$, thus

$$\begin{aligned} 3\varphi(z_k) &= 2^{-1}\cdot 3\varphi(v_k) - 2^{-1}\cdot 3^{p_k}, \\ 2\varphi(z_{k-1}) &= \varphi(v_{k-1}) - 3^{p_{k-1}-1}. \end{aligned}$$

Hence,

$$\begin{aligned} P_{k+1} = 3^{p_{k-1}-1}c_k(3\varphi(v_k) + 2^{q_k-1}) + 2^{q_k a_{k+1}}\varphi(v_{k-1}) - 3^{p_{k-1}-1}c_k3^{p_k} \\ -2^{q_k a_{k+1}}3^{p_{k-1}-1} + 3^{p_{k+1}-1}. \end{aligned}$$

Using the obvious identity $3^{p_{k-1}-1}c_k2^{q_k-1} = 3^{p_{k-1}-1}c_k2^{q_k} - 3^{p_{k-1}-1}c_k2^{q_k-1}$, we get $P_{k+1} = 3^{p_{k-1}-1}c_k(3\varphi(v_k) - 2^{q_k-1}) + 2^{q_{k+1}-q_{k-1}}\varphi(v_{k-1}) = 3^{(-1)^{k+1}}3^{p_{k-1}}c_kP_k$ $+ 2^{q_{k+1}-q_{k-1}}P_{k-1}$. □

[5] We use t_n instead of d_n

For $k \geq 1$, let

$$B_{k+1} := 3^{(-1)^{k+1}} 3^{p_{k-1}} c_k \qquad \text{and} \qquad A_{k+1} := 2^{q_{k+1}-q_{k-1}}, \tag{11}$$

so Lemma 18 can be written as

$$\begin{aligned} P_{k+1} &= B_{k+1} P_k + A_{k+1} P_{k-1}, \\ Q_{k+1} &= B_{k+1} Q_k + A_{k+1} Q_{k-1}. \end{aligned} \tag{12}$$

Lemma 19.

$$P_{k+1} Q_k - P_k Q_{k+1} = (-1)^{k+1} 2^{q_{k+1}+q_k-1} \qquad (k \geq 0).$$

Proof. (a) For $k = 0$: $\quad P_1 Q_0 - P_0 Q_1 = -2^{q_1} = (-1)^{0+1} 2^{q_1+q_0-1}$.

(b) For $k \geq 1$: $\quad P_{k+1} Q_k - P_k Q_{k+1} = -A_{k+1}(P_k Q_{k-1} - P_{k-1} Q_k)$ by (12).

Therefore,
$$\begin{aligned} P_2 Q_1 - P_1 Q_2 &= -A_2(P_1 Q_0 - P_0 Q_1) = -2^{q_2-q_0}(-2^{q_1}) = 2^{q_2+q_1-1}, \\ P_3 Q_2 - P_2 Q_3 &= -A_3(P_2 Q_1 - P_1 Q_2) = -2^{q_3-q_1} 2^{q_2+q_1-1} = -2^{q_3+q_2-1}, \\ &\ \vdots \end{aligned}$$

We omit the induction. □

Lemma 20.

$$\frac{P_{k+1}}{Q_{k+1}} = \frac{P_0}{Q_0} + \sum_{j=0}^{k} (-1)^{j+1} \frac{2^{q_{j+1}+q_j-1}}{3(3^{p_{j+1}} - 2^{q_{j+1}})(3^{p_j} - 2^{q_j})} \qquad (k \geq 0).$$

Proof. By Lemma 19, the difference between consecutive terms is

$$\frac{P_{k+1}}{Q_{k+1}} - \frac{P_k}{Q_k} = (-1)^{k+1} \frac{2^{q_{k+1}+q_k-1}}{Q_{k+1} Q_k} \qquad (k \geq 0).$$ □

We now complete the proof of Theorem 1.

Proof of Theorem 1. For $k \to \infty$ the sum in Lemma 20 converges, since the terms added have 2-adic norm $2^{1-q_{j+1}-q_j}$ which converges to 0 for increasing j. This fact is sufficient to guarantee the convergence of a series in $\mathbb{Z}_2$. The statement of Theorem 1 follows immediately from Lemma 17. □

Lemma 21. *For $k \geq 0$, there holds*

$$\Phi(m_\alpha) = -\frac{P_k}{Q_k} - (-1)^k \sum_{j=0}^{\infty} (-1)^{j+1} \frac{2^{q_{k+j+1}+q_{k+j}-1}}{3(3^{p_{k+j+1}} - 2^{q_{k+j+1}})(3^{p_{k+j}} - 2^{q_{k+j}})}.$$

Proof. The statement follows from Lemma 20 and Lemma 17. □

Proof of Corollary 2. We show that $\frac{Q_0}{P_0}, \frac{Q_1}{P_1}, \frac{Q_2}{P_2}, \frac{Q_3}{P_3}, \ldots$ are the convergents of a generalized continued fraction expansion for $\frac{-1}{\Phi(m_\alpha)}$. Indeed, Lemma 19 is the determinant formula for this expansion. A_k, B_k are defined for $k \geq 2$ in (11). We define $\frac{Q_0}{P_0} := B_0$, so $B_0 = 3$. From Lemma 19 we get $P_1Q_0 - P_0Q_1 = -2^{q_1} = -A_1$, so $A_1 = 2^{q_1}$. Finally, $\frac{Q_1}{P_1} = \frac{B_1B_0+A_1}{B_1}$ and $\frac{Q_1}{P_1} = 3 - 2^{q_1}$ yield $B_1 = -1$. □

4. A Devil's Staircase

In this section we leave the 2-adic world and consider Φ as a real-valued function, now called $\Phi_{\mathbb{R}}$. It is in this context where the right-gap function actually appears (Definition 16).

Using the absolute value as the norm, the proof of Lemma 12 fails. The series $\sum_{n=0}^{\infty} \frac{2^{nq}}{3^{np}}$ converges if and only if $(2^q/3^p) < 1$ or equivalently, if and only if $\frac{ln(2)}{ln(3)} < p/q \leq 1$.

Definition 22. Let $f := \mathbb{Q} \cap \left(\frac{\ln(2)}{\ln(3)}, 1\right] \to \mathbb{R}$ be defined by

$$f\left(\frac{p}{q}\right) := \Phi_{\mathbb{R}}(m_{p/q}) = \frac{\varphi(\overline{m}_{p/q})}{2^q - 3^p} \qquad (p, q \text{ coprime}).$$

Note that $\Phi_{\mathbb{R}}(m_{p/q}) = -\sum_i \frac{1}{3^{i+1}} 2^{d_i}$ is now a negative rational number when calculated over the infinite word $m_{p/q}$.

A plot of the function f reveals the structure of a devil's staircase. There is a gap associated with any rational of the domain.

Lemma 23. *For* $\frac{p'}{q'}, \frac{p}{q} \in \mathbb{Q} \cap (\frac{ln(2)}{ln(3)}, 1]$ *and* g *as in Definition 16,*

$$\textit{if} \quad \frac{p'}{q'} > \frac{p}{q}, \quad \textit{then} \quad f(\frac{p'}{q'}) > f(\frac{p}{q}) + g(\frac{p}{q}).$$

Proof. Fix $\frac{p}{q}$. We choose a Farey sequence of any order $N \geq q$ and suppose that $\frac{p}{q}$ and $\frac{p'}{q'}$ are a Farey pair: $\frac{p'}{q'}$ is the right neighbor of $\frac{p}{q}$. Hence, $pq' - p'q = -1$. By (2) and Lemma 11, we have

$$\begin{aligned} f(\frac{p}{q}) &= -\sum_{i=0}^{\infty} \frac{1}{3^{1+i}} 2^{\lfloor i\frac{q}{p} \rfloor} = -\sum_{j=1}^{\infty} \frac{1}{3^{1+jp}} 2^{jq} - \sum_{i \neq jp} \frac{1}{3^{1+i}} 2^{\lfloor i\frac{q}{p} \rfloor}. \\ f(\frac{p'}{q'}) &= -\sum_{j=1}^{\infty} \frac{1}{3^{1+jp}} 2^{\lfloor jp\cdot\frac{q'}{p'} \rfloor} - \sum_{i \neq jp} \frac{1}{3^{1+i}} 2^{\lfloor i\frac{q'}{p'} \rfloor}. \end{aligned}$$

As $\lfloor i\frac{q}{p}\rfloor \geq \lfloor i\frac{q'}{p'}\rfloor$,

$$f(\frac{p'}{q'}) - f(\frac{p}{q}) \geq \sum_{j=1}^{\infty} \frac{1}{3^{1+jp}}(2^{jq} - 2^{\lfloor jp\cdot\frac{q'}{p'}\rfloor}).$$

Since $pq' - p'q = -1, \quad jp \cdot \frac{q'}{p'} = jq - \frac{j}{p'}$. Hence,

$$\begin{array}{lcll} \lfloor jp \cdot \dfrac{q'}{p'}\rfloor & \leq & jq - 1 & \text{if} \quad j \leq p'; \\ \lfloor jp \cdot \dfrac{q'}{p'}\rfloor & < & jq - 1 & \text{if} \quad j > p'. \end{array}$$

Consequently,

$$f(\frac{p'}{q'}) - f(\frac{p}{q}) > \frac{1}{2}\sum_{j=1}^{\infty} \frac{1}{3^{1+jp}} 2^{jq} = \frac{1}{3} \cdot \frac{2^{q-1}}{3^p - 2^q} = g(\frac{p}{q}). \qquad \square$$

Lemma 23 proves that f is strictly increasing over the rationals.

Lemma 24. *Let $(x_i)_{i=0}^{\infty}$ be a sequence of rationals converging to α, and let g be as in Definition 16. Then*

$$\lim_{i\to\infty} g(x_i) = 0.$$

Proof. We have $\alpha > \frac{ln(2)}{ln(3)}$. Let $c \in (0, \alpha - \frac{ln(2)}{ln(3)})$. There exists an index i_0 such that $x_i > \frac{ln(2)}{ln(3)} + c$ for all $i \geq i_0$. We assume $i \geq i_0$ and $x_i := a_i/b_i$, written in reduced fraction form. The convergents p_k/q_k are the best approximation of α:

$$\text{if} \quad \left|\alpha - \frac{a_i}{b_i}\right| < \left|\alpha - \frac{p_k}{q_k}\right| \quad \text{for some } k\text{, then} \quad b_i \geq q_k.$$

For increasing k, $q_k \to \infty$. So $b_i \to \infty$. Note that $a_i/b_i > \frac{ln(2)}{ln(3)} + c$. Hence $a_i > b_i\frac{ln(2)}{ln(3)} + b_i c$. Then

$$\frac{2^{b_i}}{3^{a_i}} < \frac{2^{b_i}}{3^{b_i\cdot\frac{ln(2)}{ln(3)}}} \cdot \frac{1}{3^{b_i c}} = \left(\frac{1}{3^c}\right)^{b_i}.$$

Since $b_i \to \infty$ and $1/3^c < 1$, $\lim_{b_i\to\infty} \frac{2^{b_i}}{3^{a_i}} = 0$. Now,

$$g(\frac{a_i}{b_i}) = \frac{1}{6} \cdot \frac{\frac{2^{b_i}}{3^{a_i}}}{1 - \frac{2^{b_i}}{3^{a_i}}} \quad \text{and} \quad \lim_{i\to\infty} g(\frac{a_i}{b_i}) = 0 \quad \text{as claimed.} \qquad \square$$

We show that the series expansion of Theorem 1 converges also in $\mathbb{R}$. We see that the series of Theorem 1 can be written formally as

$$\sum_{j=0}^{\infty}(-1)^{j+1}\frac{2^{q_{j+1}+q_j-1}}{3(3^{p_{j+1}}-2^{q_{j+1}})(3^{p_j}-2^{q_j})} = \sum_{j=0}^{\infty}(-1)^{j+1}\cdot 6\cdot g(\frac{p_j}{q_j})g(\frac{p_{j+1}}{q_{j+1}}). \qquad (13)$$

Lemma 25. *The following limit exists:*

$$\lim_{k\to\infty}\sum_{j=0}^{k}(-1)^{j+1}\cdot 6\cdot g(\frac{p_j}{q_j})g(\frac{p_{j+1}}{q_{j+1}}).$$

Proof. The $\frac{p_j}{q_j}$ are the convergents of $\alpha > \frac{ln(2)}{ln(3)}$. There exists an index j_0 such that $\frac{p_j}{q_j} > \frac{ln(2)}{ln(3)}$ and consequently, $3^{p_j} > 2^{q_j}$ for all $j \geq j_0$.
We show that $\lim_{k\to\infty}\sum_{j=j_0}^{k}(-1)^{j+1}g(\frac{p_j}{q_j})g(\frac{p_{j+1}}{q_{j+1}})$ exists.
By the criterion of Leibniz for alternating series, it is sufficient that the absolute terms $\left|(-1)^{j+1}g(\frac{p_j}{q_j})g(\frac{p_{j+1}}{q_{j+1}})\right| = g(\frac{p_j}{q_j})g(\frac{p_{j+1}}{q_{j+1}})$ decrease strictly monotone to 0. In fact, it is an easy check that for $j \geq j_0$:

$$g(\frac{p_j}{q_j}) > g(\frac{p_{j+2}}{q_{j+2}}) \iff 2^{q_j}3^{p_{j+2}} > 2^{q_{j+2}}3^{p_j} \iff \frac{p_{j+1}}{q_{j+1}} > \frac{\ln(2)}{\ln(3)}. \qquad (14)$$

The last term is equivalent to $3^{p_{j+1}} > 2^{q_{j+1}}$. □

By Lemma 24, the g_i's approach 0. The real-valued sequence (9) and $(\Phi_{\mathbb{R}}(v_k))_{k=0}^{\infty}$ converge to the same limit $\Phi_{\mathbb{R}}(m_\alpha)$. It follows that Lemma 17 also holds for real numbers. The number (-1/$\Phi_{\mathbb{R}}(m_\alpha)$) can be calculated with the real-valued continued fraction of Corollary 2. So $\Phi_{\mathbb{R}}(m_\alpha)$ is irrational. We extend f to a function F over the whole interval $(\frac{\ln(2)}{\ln(3)}, 1]$.

Definition 26. Let $F := (\frac{ln(2)}{ln(3)}, 1] \to \mathbb{R}$ be defined by

$$\begin{aligned} F(x) &:= \Phi_{\mathbb{R}}(m_x) = -\lim_{k\to\infty}\frac{P_k}{Q_k} && \text{if } x \text{ is irrational;} \\ F(p/q) &:= \Phi_{\mathbb{R}}(m_{p/q}) = \frac{\varphi(\overline{m}_{p/q})}{2^q - 3^p} && (p, q \text{ coprime}). \end{aligned}$$

Lemma 27. *$F(x)$ is a strictly monotone increasing function. Furthermore, $F(x)$ is continuous at $x = \alpha$.*

Proof. By Lemma 25 and (13), Lemma 21 holds in $\mathbb{R}$:

$$\Phi_{\mathbb{R}}(m_\alpha) - (-\frac{P_k}{Q_k}) = -(-1)^k \sum_{j=0}^{\infty} (-1)^{j+1} \frac{2^{q_{k+j+1}+q_{k+j}-1}}{3(3^{p_{k+j+1}} - 2^{q_{k+j+1}})(3^{p_{k+j}} - 2^{q_{k+j}})} \qquad (k \geq 0).$$

By (14), there exists a sufficiently large k_0 such that $(p_k/q_k) > \frac{ln(2)}{ln(3)}$ for all $k > k_0$. For $k > k_0$, the absolute values of the terms added converge strictly monotone to 0, and $(-P_k/Q_k)$ approaches $\Phi_{\mathbb{R}}(m_\alpha) = F(\alpha)$. Therefore, $-P_{2k}/Q_{2k} < F(\alpha) < -P_{2k+1}/Q_{2k+1}$ for all $2k > k_0$. This inequality and Lemma 23 prove that $F(x)$ is strictly monotone everywhere.

Recall that $-P_{2k}/Q_{2k} = F(p_{2k}/q_{2k}) + g(p_{2k}/q_{2k})$ and $-P_{2k+1}/Q_{2k+1} = F(p_{2k+1}/q_{2k+1})$. Choose p/q such that $p_{2k}/q_{2k} < p/q < p_{2k+2}/q_{2k+2}$. Then $-P_{2k}/Q_{2k} < F(p/q) < F(\alpha)$. Consequently, we have $F([p/q, p_{2k+1}/q_{2k+1}]) \subset [-P_{2k}/Q_{2k}, -P_{2k+1}/Q_{2k+1}]$. For any given $\epsilon > 0$, there is a sufficiently large k such that $[-P_{2k}/Q_{2k}, -P_{2k+1}/Q_{2k+1}]$ lies entirely inside an ϵ-neighborhood of $F(\alpha)$. This proves the continuity at $x = \alpha$. □

The previous lemmas prove that the function $F := \left(\frac{\ln(2)}{\ln(3)}, 1\right] \to \mathbb{R}$

- has range $F\left((\frac{\ln(2)}{\ln(3)}, 1]\right) \subset (-\infty, -1]$; [6]
- is strictly monotone increasing
- maps rationals to rationals;
- maps irrationals to irrationals;
- is discontinuous at every rational;
- is continuous at every irrational.

The function is similar to other devil's staircases. Perhaps the first one of this type was given by Bőhmer [3], proving the transcendence of certain dyadic fractions. It seems that F additionally maps irrationals to transcendental numbers. We have no proof.

What happens when $0 < \alpha < \frac{\ln(2)}{\ln(3)}$?

First of all, Lemma 12, interpreted in $\mathbb{R}$, is no longer true. But all is not lost. Let (p_j/q_j) be the convergents of α. Then $\frac{p_j}{q_j} < \frac{\ln(2)}{\ln(3)}$ for all sufficiently large j, so that the relation (14) simply can be inverted, substituting $>$ by $<$. So Lemma 25 is still valid because $3^{p_{j+1}} < 2^{q_{j+1}}$ implies that $g(\frac{p_j}{q_j})$ and $g(\frac{p_{j+1}}{q_{j+1}})$ are both negative. If

[6] The range is an uncountable, nowhere dense null set.

$j \to \infty$, then $g(\frac{p_j}{q_j}) \to (-1/6)$, so Lemma 24 is no longer valid. The real-valued sequence (9) no longer converges to $\Phi_{\mathbb{R}}(m_\alpha)$: the terms with odd index still converge to $\Phi_{\mathbb{R}}(m_\alpha)$, those with even index converge to $\Phi_{\mathbb{R}}(m_\alpha) - \frac{1}{6}$. The limit (in $\mathbb{R}$) of Lemma 17 does not exist. In fact, the real-valued sequence $(-P_k/Q_k)$ has exactly two limit points.

It is possible to extend F artificially to the left side of $\frac{\ln(2)}{\ln(3)}$. Since the limit in Definition 26 no longer exists, we define $F(\alpha)$ as the upper limit of $(-P_k/Q_k)$. The real-valued expansions of Theorem 1 and Corollary 2 remain still useful provided we use approximations that stop at an odd index. Note that $F(x) > 0$ is at the left and $F(x) < 0$ is at the right side of $\frac{\ln(2)}{\ln(3)}$. Furthermore, F diverges at $x = \frac{\ln(2)}{\ln(3)}$, the odd approximations in Theorem 1 approach $-\infty$ and the even $+\infty$, while in Corollary 2 both approximations approach 0.

A plot of the artificially extended F shows a positive, strictly monotone increasing devil's staircase with gaps at the left side of the rationals, a very different behavior from the original F. So we abandon F and construct a new function F^*, specially for $0 < x < \frac{\ln(2)}{\ln(3)}$, which will have a convergent series expansion.

First we define

$$F^*(p_k/q_k) := \Phi^*_{\mathbb{R}}(m_{p_k/q_k}) := \Phi^*_{\mathbb{R}}(v_k) := \frac{\varphi(\overline{m}_{p_k/q_k})}{2^{q_k} - 3^{p_k}}.$$

The last term is the same number as in Lemma 14, but this is no longer the same as $\Phi_{\mathbb{R}}(m_{p_k/q_k})$ since Lemma 12 and Lemma 14 are false for $0 < p_k/q_k < \frac{\ln(2)}{\ln(3)}$.

The sequence (9) is no longer appropriate. This time we get the best approximation of $F^*(\alpha)$ by

$$\Phi^*_{\mathbb{R}}(v_0), \quad \Phi^*_{\mathbb{R}}(v_1) + g'(\frac{p_1}{q_1}), \quad \Phi^*_{\mathbb{R}}(v_2), \quad \Phi^*_{\mathbb{R}}(v_3) + g'(\frac{p_3}{q_3}), \quad \Phi^*_{\mathbb{R}}(v_4), \ \ldots \tag{15}$$

with the new left-gap $g'(\frac{p_k}{q_k}) := g(\frac{p_k}{q_k}) + \frac{1}{6} = \frac{1}{2} \cdot \frac{3^{p_k-1}}{3^{p_k}-2^{q_k}}$, which now approaches 0 when $k \to \infty$.

The sequences (15) and $(\Phi^*_{\mathbb{R}}(v_k))_{k=0}^{\infty}$ converge to the same limit, if such a limit exists. The new terms are

$$\frac{\varphi(v_0)}{2^{q_0} - 3^{p_0}}, \ \frac{2\varphi(v_1) - 3^{p_1-1}}{2(2^{q_1} - 3^{p_1})}, \ \frac{\varphi(v_2)}{2^{q_2} - 3^{p_2}}, \ \frac{2\varphi(v_3) - 3^{p_3-1}}{2(2^{q_3} - 3^{p_3})}, \ \frac{\varphi(v_4)}{2^{q_4} - 3^{p_4}}, \ \ldots. \tag{16}$$

We write P'_k for the numerators and Q'_k for the denominators:

$$\frac{P'_0}{Q'_0} = \frac{0}{1}, \quad \frac{P'_1}{Q'_1} = \frac{1}{2(2^{q_1} - 3)}, \quad \frac{P'_2}{Q'_2}, \ \frac{P'_3}{Q'_3}, \ \frac{P'_4}{Q'_4}, \ \ldots.$$

Compare the sequence (16) with (10). There is a duality: the substitutions

$$2 \longleftrightarrow 3, \quad p_k \longleftrightarrow q_k \tag{17}$$

and $k \longrightarrow k+1$ map (16) to (10) for $k \geq 0$; only the first term $-\frac{-1}{-3}$ in (10) is left out. Hence we can expect that our new series expansion is dual to the one given in Theorem 1 with the same substitutions (17).

In fact, the Lemmas 18 and 19 interpreted in $\mathbb{R}$ now have a dual version with the same substitutions. Lemma 18' will be as follows:

For $k \geq 1$,

$$\begin{aligned} P'_{k+1} &= 2^{(-1)^k} 2^{q_{k-1}} c_k P'_k + 3^{p_{k+1}-p_{k-1}} P'_{k-1}, \\ Q'_{k+1} &= 2^{(-1)^k} 2^{q_{k-1}} c_k Q'_k + 3^{p_{k+1}-p_{k-1}} Q'_{k-1}, \\ \text{where} \quad c_k &:= \frac{2^{q_k a_{k+1}} - 3^{p_k a_{k+1}}}{2^{q_k} - 3^{p_k}}. \end{aligned}$$

Note that $(-1)^k$ instead of $(-1)^{k+1}$. The proof has four parts as in Lemma 18: sections (a) and (b) do not change; the easy section (c) now will be for k even; the harder section (d) will be for k odd, using

$$v_{k+1} = 1 z_{k-1} (10 z_k)^{a_{k+1}} 0 \quad \text{instead of} \quad v_{k+1} = 1 (z_k 01)^{a_{k+1}} z_{k-1} 0.$$ [7]

Instead of (11), we define

$$B'_{k+1} := 2^{(-1)^k} 2^{q_{k-1}} c_k \qquad \text{and} \qquad A'_{k+1} := 3^{p_{k+1}-p_{k-1}}. \tag{18}$$

Then

$$\begin{aligned} P'_{k+1} &= B'_{k+1} P'_k + A'_{k+1} P'_{k-1}, \\ Q'_{k+1} &= B'_{k+1} Q'_k + A'_{k+1} Q'_{k-1}. \end{aligned} \tag{19}$$

Further, Lemma 19' will say

$$P'_{k+1} Q'_k - P'_k Q'_{k+1} = (-1)^k 3^{p_{k+1}+p_k-1} \qquad (k \geq 0). \tag{20}$$

Finally, there holds Lemma 20':

$$\frac{P'_{k+1}}{Q'_{k+1}} = \sum_{j=0}^{k} (-1)^j \frac{3^{p_{j+1}+p_j-1}}{2(2^{q_{j+1}} - 3^{p_{j+1}})(2^{q_j} - 3^{p_j})} \qquad (k \geq 0). \tag{21}$$

Note that $P'_0/Q'_0 = 0$ by (16). For $k \to \infty$ the sum (21) converges, since

$$\begin{aligned} &\sum_{j=0}^{k} (-1)^j \frac{3^{p_{j+1}+p_j-1}}{2(2^{q_{j+1}} - 3^{p_{j+1}})(2^{q_j} - 3^{p_j})} \\ &\qquad = \sum_{j=0}^{k} (-1)^j \cdot 6 \cdot g'\left(\frac{p_j}{q_j}\right) g'\left(\frac{p_{j+1}}{q_{j+1}}\right) \qquad (k \geq 0). \end{aligned}$$

[7] See Exercise 2.2.10 in [8]

Consequently, if $0 < \alpha < \frac{\ln(2)}{\ln(3)}$, then

$$\lim_{k\to\infty} \frac{P'_k}{Q'_k} = \sum_{j=0}^{\infty} (-1)^j \frac{3^{p_{j+1}+p_j-1}}{2(2^{q_{j+1}} - 3^{p_{j+1}})(2^{q_j} - 3^{p_j})}. \tag{22}$$

This series can be written as a generalized continued fraction with convergents P'_k/Q'_k. Define $\frac{P'_0}{Q'_0} := B'_0 = 0$. By (20), we get $P'_1Q'_0 - P'_0Q'_1 = 3^{p_1+p_0-1} = 1 = -A'_1$, so $A'_1 = -1$. Finally, $\frac{P'_1}{Q'_1} = \frac{B'_1B'_0+A'_1}{B'_1}$ and $\frac{P'_1}{Q'_1} = \frac{1}{2(2^{q_1}-3)}$ yield $B'_1 = -2(2^{q_1} - 3)$. These start values, together with (18) and (19), give the expansion

$$\lim_{k\to\infty} \frac{P'_k}{Q'_k} = \cfrac{-1}{-2(2^{q_1} - 3) + \cfrac{-3^{p_2}}{B'_2 + \cfrac{A'_3}{B'_3 + \cfrac{A'_4}{B'_4 + \cdots}}}}. \tag{23}$$

Here is the new function we were looking for:

Definition 28. *Let* $F^* := \left[0, \frac{\ln(2)}{\ln(3)}\right) \to \mathbb{R}$ *be defined by*

$$\begin{aligned} F^*(x) &:= \Phi^*_{\mathbb{R}}(m_x) = \lim_{k\to\infty} \frac{P'_k}{Q'_k} && \text{if } x \text{ is irrational;} \\ F^*(p/q) &:= \Phi^*_{\mathbb{R}}(m_{p/q}) = \frac{\varphi(\overline{m}_{p/q})}{2^q - 3^p} && (p, q \text{ coprime}). \end{aligned}$$

The function $F^* := \left[0, \frac{\ln(2)}{\ln(3)}\right) \to \mathbb{R}$ (a devil's staircase)

- has range $F^*\left(\left[0, \frac{\ln(2)}{\ln(3)}\right)\right) \subset [0, +\infty)$; [8]
- is strictly monotone increasing;
- maps rationals to rationals;
- maps irrationals to irrationals;
- is discontinuous at every rational;
- is continuous at every irrational.

The proof of monotony and continuity is similar to the one of Lemma 27. So we omit the details.

[8] The range is an uncountable, nowhere dense null set.

References

[1] D. J. Bernstein, *A non-iterative 2-adic statement of the* $3x+1$ *conjecture*, Proc. Amer. Math. Soc. **121** (1994), 405–408. Also available at http://cr.yp.to/mathmisc.html.

[2] D. J. Bernstein and J. C. Lagarias, *The* $3x+1$ *conjugacy map*, Canadian Journal of Mathematics **48** (1996), 1154–1169. Also available at http://www.math.lsa.umich.edu/∼lagarias.

[3] P. E. Bőhmer, *Űber die Transzendenz gewisser dyadischer Brűche*, Mathematische Annalen **96** (1927), 367–377, 735. Also available at http://www.digizeitschriften.de.

[4] J. L. Davison, *A series and its associated continued fraction*, Proc. Amer. Math. Soc. **63** (1977), 29–32.

[5] L. Halbeisen and N. Hungerbűhler, *Optimal bounds for the length of rational Collatz cycles*, Acta Arithmetica **78** (1997), 227–239. Also available at http://www.iam.unibe.ch/∼halbeis.

[6] J. C. Lagarias, *The* $3x+1$ *problem and its generalizations*, Amer. Math. Monthly **92** (1985), 3–23. Also available at http://www.math.lsa.umich.edu/∼lagarias.

[7] J. C. Lagarias, $3x + 1$ *problem and related problems*, http://www.math.lsa.umich.edu/∼lagarias.

[8] M. Lothaire. *Algebraic Combinatorics on Words*, volume **90** of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge, (2002). Also available at http://en.wikipedia.org/wiki/Sturmian_word.

[9] K. G. Monks and J. Yazinski, *The autoconjugacy of the* $3x+1$ *function*, Discrete Mathematics **275** (2004), 219–236. Also available at http://www.scranton.edu/∼monks.