THE 2-ADIC, BINARY AND DECIMAL PERIODS OF $1/3^k$ APPROACH FULL COMPLEXITY FOR INCREASING k

Josefina López
Villa Abecia, Sud Cinti, Chuquisaca, Bolivia
josefinapedro@hotmail.com

Peter Stoll
Villa Abecia, Sud Cinti, Chuquisaca, Bolivia
josefinapedro@hotmail.com

Received: 2/28/11, Revised: 12/10/11, Accepted: 3/23/12, Published: 4/18/12

Abstract

An infinite word x over an alphabet with b letters has full complexity if for each $m \in \mathbb{N}$ all the b^m words of length m are factors of x. We prove that the periods of $\pm 1/3^k$ in the 2-adic expansion approach full complexity for increasing k: For any $m \in \mathbb{N}$, the periods for $k > \lceil (m+1)\ln(2)/\ln(3) \rceil$ have complexity 2^m. Amazingly, these 2^m words occur in the period almost the same number of times. On the way, first we prove the same for the binary period. We get a similar result for the decimal period of $1/3^k$.

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 $x_0 x_1 x_2 \cdots$ of 1's and 0's. The x_k are the digits of x, written from left to right. For instance, $-1 = 1111\cdots$ and $1 = 1000\cdots$, since $-1+1 = 0$. The 2-adic norm $|\cdot|_2$ in $\mathbb{Z}_2$ is given by $|x|_2 := 2^{-n}$ if $x \neq 0$, where x_n is the first nonzero digit of x, and $|x|_2 := 0$ if $x = 0$. Let $0 \le d_0 < d_1 < d_2 < \cdots$ be a finite or infinite sequence of nonnegative integers defined by $d_i := k$ whenever $x_k = 1$ for a 2-adic integer $x = x_0 x_1 \cdots x_k \cdots$. 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. Let $T^k(x)$ denote the k-th iterate of T and $T^0(x) := x$. The sequence

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

is called the *parity vector* of $x \in \mathbb{Z}_2$ (Lagarias [3]). The parity vector can be regarded as the 2-adic integer

$$v_x = \sum_{k=0}^{\infty} (T^k(x) \mod 2) \cdot 2^k.$$

Conversely, each $x \in \mathbb{Z}_2$ is uniquely determined by its parity vector v_x. Indeed, there is an explicit formula (Bernstein [1]):

$$x = \Phi(v_x), \text{ where } \quad \Phi(2^{d_0} + 2^{d_1} + 2^{d_2} + \cdots) := -\sum_{i \geq 0} \frac{1}{3^{i+1}} 2^{d_i}. \tag{1}$$

The periodicity conjecture of the famous $3x+1$ problem states that the parity vector v_x is eventually periodic for all $x \in \mathbb{Z}_2$ (Bernstein, Lagarias [2]). This conjecture is still open. If the parity vector of some $x \in \mathbb{Z}_2$ is eventually periodic, then $x = \Phi(v_x)$ is also eventually periodic. This follows from (1) by computing the corresponding geometric series using the 2-adic norm. For instance, $-1/3 = 101010\cdots$ since $2^0+2^2+\cdots = 1/(1-4)$. If the periodicity conjecture is true, Φ maps each aperiodic v_x onto an aperiodic 2-adic integer $x \in \mathbb{Z}_2$. In a former paper ([5]), we calculated $\Phi(v_x)$ over a Sturmian word (the digits of v_x are Sturmian). These aperiodic words are very similar to the periodic words by their "almost-periodicity". We accumulated evidence that $\Phi(v_x)$ is an infinite word of full complexity. That was reason enough for searching more in this direction.

George Pólya, in his famous book "How to solve it", suggested that "When you can't solve a problem" —and we cannot prove the periodicity conjecture— "then there is an easier problem you can solve: find it". That is what we did. We prove that the terms $-2^{d_i}/3^{i+1}$ added in (1) approach full complexity for increasing i. The complexity function of a word x counts for each nonnegative integer $m \geq 0$ the number $P(x,m)$ of different factors of length m in x.[1] The number $P(x,1)$ counts the different letters appearing in x. In our case, the alphabet is $\{0,1\}$, so $P(x,1) = 2$. If x is purely periodic with a period of length m, we have $P(x,m) = m$. If x is Sturmian, we have $P(x,m) = m+1$ for every $m \geq 0$ (Lothaire [6]). An infinite word x over $\{0,1\}$ has full complexity if for every nonnegative integer m all the 2^m words of length m are factors of x. We prove that for every $m \in \mathbb{N}$ the period p of $-1/3^{i+1}$ has complexity $P(x,m) = 2^m$ for all $(i+1) > \lceil (m+1)\ln(2)/\ln(3) \rceil$. The d_i's in (1) introduce d_i leading 0's for each $-1/3^{i+1}$. In this sense, our best result is Corollary 4 which gives the necessary and sufficient condition for complexity 2^m.

Luckily, there is a helpful relation between the 2-adic and the binary expansion for reduced fractions with odd denominator of the unit interval. This relation is given by Lemma 19 in Section 4: The 2-adic period of $-1/3^{i+1}$ is the reversal of the binary period of $+1/3^{i+1}$. Therefore, the main part of this paper is in Section 3 where we study the binary expansion of $1/3^k$. The transcription from binary to 2-adic periods is in Section 4, and our main result appears there as Corollary 4 of

[1] Lothaire [6] A word f is called a factor of a word x if there exist words u, v such that $x = ufv$.

Theorem 1. Most things we deal with in Section 3 are well known facts of elementary arithmetic, and the recursive method used there can be generalized to other bases and also to other fractions $1/p$ in base 10 for primes p different from 2 and 5. As such an application should be seen Section 5 where we study the decimal expansion of $1/3^k$. The main tool for establishing conditions for complexity is Proposition 23 in Section 6.

2. Results

Theorem 1. *Let $k \in \mathbb{N}$ and $m \in \mathbb{N} \setminus \{1\}$. The set $X_m(k)$ of words with length m in the binary expansion of $1/3^k$ has 2^m elements if and only if $k \geq k_0$, where*

$$k_0 = \left\lceil (m+1)\frac{\ln(2)}{\ln(3)} \right\rceil.$$

All these 2^m words can be found in the enlarged period $\overline{p}_k := p_k(0)^{m-1}$, which is the period p_k of $1/3^k$ concatenated with the first $m-1$ zeros of p_k.

The words $n \in X_m(k)$, $k \geq k_0$, occur in $\overline{p}_k$ with almost the same frequency $f(n)$:[2]

$$|f(n) - f(n')| \leq 2 \quad \textit{for all} \quad n, n' \in X_m(k).$$

Theorem 2. *Let k, k_0 and m be as in Theorem 1. If $k > k_0$, then the binary period p_k of $1/3^k$ has complexity 2^m.*

Theorem 3. *The binary period of $1/3^k$*

$$p_k = x_1 x_2 \cdots x_i \cdots x_{\ell_k} \qquad (\ell_k = 2 \cdot 3^{k-1})$$

has the following properties:

1. *The digits are given by $x_i = (2^i \bmod 3^k) \bmod 2$ or equivalently, the digit x_i is 1 if and only if the remainder $2^i \bmod 3^k$ is odd.*

2. *The number n_0 of 1's at an even position exceeds the number n_1 of 1's at an odd position by 1: $n_0 - n_1 = 1$.*

3. *The word p_k has the same number of 1's and 0's.*

4. *The right half of p_k is complementary to the left half of p_k: $x_{i+\ell_k/2} + x_i = 1$ $(0 \leq i \leq \ell_k/2)$.*

5. *In the left half of p_k, the number of 1's at an even position is equal to the number of 1's at an odd position.*

[2]The frequency $f(n)$ is the number of times a given word n occurs as factor of the enlarged period.

6. *There are exactly* $\lfloor k \cdot \ln(3)/\ln(2) \rfloor$ *leading* 0*'s in* p_k.

Corollary 4. *Let* $k \in \mathbb{N}$ *and* $m \in \mathbb{N} \setminus \{1\}$. *The set* $X_m(k)$ *of words with length* m *in the 2-adic expansion of* $-1/3^k$ *has* 2^m *elements if and only if* $k \geq k_0$, *where*

$$k_0 = \left\lceil (m+1)\frac{\ln(2)}{\ln(3)} \right\rceil.$$

All these 2^m *words can be found in* $\overline{\tilde{p}}_k := (0)^{m-1}\tilde{p}_k$, *which is the period* $\tilde{p}_k$ *of* $-1/3^k$ *extended to the left by* $m-1$ *zeros.*

The words $n \in X_m(k)$, $k \geq k_0$, *occur in* $\overline{\tilde{p}}_k$ *with almost the same frequency* $f(n)$:

$$|f(n) - f(n')| \leq 2 \quad \textit{for all} \quad n, n' \in X_m(k).$$

Corollary 5. *Let* k, k_0 *and* m *be as in Corollary 4. If* $k > k_0$, *then the 2-adic period* $\tilde{p}_k$ *of* $-1/3^k$ *has complexity* 2^m.

Corollary 6. *Let* k, k_0 *and* m *be as in Corollary 4. Let* p_k^+ *be the period of the (not purely periodic) 2-adic expansion* $1/3^k = 1p_k^+ p_k^+ \cdots$. *If* $k > k_0$, *then the 2-adic period* p_k^+ *has complexity* 2^m.

Theorem 7. *Let* $k \in \mathbb{N}$ *and* $m \in \mathbb{N}$. *The set* $X_m(k)$ *of words with length* m *in the decimal expansion of* $1/3^k$ *has* 10^m *elements if and only if* $k \geq k_0$, *where*

$$k_0 = \left\lceil m\frac{\ln(10)}{\ln(3)} \right\rceil + 2.$$

All these 10^m *blocks of* m *digits can be found in the enlarged period* $\overline{u}_k := u_k(0)^{m-1}$, *which is the period* u_k *of* $1/3^k$ *concatenated with the first* $m-1$ *zeros of* u_k.[3]

The words $n \in X_m(k)$, $k \geq k_0$, *occur in* $\overline{u}_k$ *with almost the same frequency* $f(n)$:

$$|f(n) - f(n')| \leq 1 \quad \textit{for all} \quad n, n' \in X_m(k).$$

Corollary 8. *Let* k, k_0 *and* m *be as in Theorem 7. If* $k > k_0$, *then the decimal period* u_k *of* $1/3^k$ *has complexity* 10^m.

3. The Binary Expansion

In this section, we use the standard notation for numbers in base 2 with the most significant digit at the left. For instance, $1101 = 2^3 + 2^2 + 2^0 = 13$ and $0.1101 = 2^{-1} + 2^{-2} + 2^{-4} = 13/16$.

The binary expansion of $1/3^k$ is a purely periodic fraction. The period is a word

$$p_k = x_1 x_2 \cdots x_i \cdots x_{\ell_k} \qquad (\ell_k = 2 \cdot 3^{k-1})$$

[3] If $m = 1$, then $(0)^{m-1} := \epsilon$ (the empty word).

of digits 0 and 1 with length $\ell_k = 2 \cdot 3^{k-1}$:

$$\frac{1}{3^k} = \sum_{i=1}^{\infty} \frac{x_i}{2^i} = \frac{2^{\ell_k}}{2^{\ell_k}-1} \sum_{i=1}^{\ell_k} x_i \cdot \frac{1}{2^i}. \tag{2}$$

For example,
$p_1 = 01$;
$p_2 = 000111$;
$p_3 = 000010010111101101$;
$p_4 = 000000110010100100010110000111111001101011011101001111$.

We get the period p_{k+1} dividing $1/3^k$ by 3, executing a long division by the binary integer 11. As $\ell_{k+1} = 3\ell_k$, we write the word p_k three times, building the word $p_k p'_k p''_k$ as dividend, to obtain the full period p_{k+1}. The long division in short form looks like this:

$$\begin{array}{r|l} & y_1 \cdots y_i \cdots y_{\ell_k}\; y'_1 \cdots y'_i \cdots y'_{\ell_k}\; y''_1 \cdots y''_i \cdots y''_{\ell_k} \\ \hline 11) & x_1 \cdots x_i \cdots x_{\ell_k}\, x'_1 \cdots x'_i \cdots x'_{\ell_k}\, x''_1 \cdots x''_i \cdots x''_{\ell_k} \\ & r_1 \cdots r_i \cdots r_{\ell_k}\; r'_1 \cdots r'_i \cdots r'_{\ell_k}\; r''_1 \cdots r''_i \cdots r''_{\ell_k} \end{array}$$

The divisor 11 and the dividend $p_k p_k p_k$ are in the second line. The quotient is in the first line, where the "0." has been omitted. At each step, after bringing down the next x_i, we have a remainder r_i. These remainders are in the third line. We define $r_0 := 0$ at the start of the division. Note that $x_i = x'_i = x''_i$ for each $i \in L_k$, where $L_k := \{1, 2, \ldots, \ell_k\}$.

There are only three possible remainders less than 11: $\{0, 1, 10\}$. Therefore, $\{r_i, r'_i, r''_i\} \subseteq \{0, 1, 10\}$. Obviously $r''_{\ell_k} = 0$. Since $\ell_{k+1} = 3\ell_k$, $r_{\ell_k} \neq 0$ and $r'_{\ell_k} \neq 0$. It is a simple, but important fact that the remainders r_i, r'_i and r''_i are a permutation of $(0, 1, 10)$.

Lemma 9. *Let $L_k := \{1, 2, \ldots, \ell_k\}$ and $i \in L_k$. The sequence (r_i, r'_i, r''_i) is a permutation of $(0, 1, 10)$.*

Proof. (a) The Lemma holds for $i = \ell_k$.
When bringing down the first digit x'_1 of p'_k, we take care of the previous remainder $r_{\ell_k} \neq 0$. If $r_{\ell_k} = 1$, the division continues by dividing the binary integer $1x'_1 = 1 \cdot 2^1 + x'_1 \cdot 2^0$ by 11. If $r_{\ell_k} = 10$, then we continue with $10x'_1 = 1 \cdot 2^2 + 0 \cdot 2^1 + x'_1 \cdot 2^0$. The division can be done by bringing down a single digit x'_1 but also by bringing down two or more digits in one step. When we bring down all the digits of p'_k in one step, the division by 11 continues with the binary integers

$$\begin{aligned} 1x'_1x'_2 \cdots x'_{\ell_k} &= 1 \cdot 2^{\ell_k} + x_1 x_2 \cdots x_{\ell_k} \quad \text{for } r_{\ell_k} = 1, \text{ or} \\ 10x'_1x'_2 \cdots x'_{\ell_k} &= 1 \cdot 2^{\ell_k+1} + 0 \cdot 2^{\ell_k} + x_1 x_2 \cdots x_{\ell_k} \quad \text{for } r_{\ell_k} = 10, \end{aligned}$$

since $p_k = p'_k$. The new remainder is r'_{ℓ_k} in both cases. As ℓ_k is even, we have $2^{\ell_k} \equiv 1 \pmod 3$, so that $r'_{\ell_k} \equiv 1 + r_{\ell_k} \pmod 3$ for $r_{\ell_k} = 1$, and $r'_{\ell_k} \equiv 2 + r_{\ell_k} \pmod 3$ for $r_{\ell_k} = 10$. Hence $(r_{\ell_k}, r'_{\ell_k}, r''_{\ell_k})$ is either $(1, 10, 0)$ or $(10, 1, 0)$.

(b) The Lemma holds for $i < \ell_k$.

We have $x_1 x_2 \cdots x'_i = x_1 \cdots x_{\ell_k} x'_1 \cdots x'_i$. The remainder of the left side is r'_i. The remainder of the right side is equal to the remainder of $r_{\ell_k} x_1 \cdots x_i$, where r_{ℓ_k} is either 1 or 10. Dividing the binary integer $r_{\ell_k} x_1 \cdots x_i$ by 3, we get the remainder $r_{\ell_k} \cdot 2^i + r_i$ modulo 3. Hence $r'_i \equiv r_{\ell_k} \cdot 2^i + r_i \pmod 3$. In the same way, $x_1 \cdots x''_i = x_1 \cdots x'_{\ell_k} x''_1 \cdots x''_i$ yields $r''_i \equiv r'_{\ell_k} \cdot 2^i + r_i \pmod 3$. There are 12 possibilities:

	r_{ℓ_k}	r'_{ℓ_k}	r_i	r'_i	r''_i
i even	1	10	0	1	10
	1	10	1	10	0
	1	10	10	0	1
	10	1	0	10	1
	10	1	1	0	10
	10	1	10	1	0
i odd	1	10	0	10	1
	1	10	1	0	10
	1	10	10	1	0
	10	1	0	1	10
	10	1	1	10	0
	10	1	10	0	1

In all cases, (r_i, r'_i, r''_i) is a permutation of $(0, 1, 10)$. □

A word x of length m entirely contained in p_k has two exact copies (x' and x'') in the dividend $p_k p'_k p''_k$: x' is factor of p'_k, and x'' is factor of p''_k. The input remainders r_{i-1}, r'_{i-1} and r''_{i-1}, and also the output remainders r_{i+m-1}, r'_{i+m-1} and r''_{i+m-1}, are all different by Lemma 9, so that the corresponding words y, y' and y'' in the quotient are all different too.

It is convenient to use base 10 numbers. Each word x of length m receives a list number n from 0 to $2^m - 1$ which is its value when read as a base 2 integer. For instance, if $m = 3$, the list number 6 is the word $x = 110$. When translating a list number back into a word, we take care of eventually leading 0's. For instance, if $m = 4$, the list number 5 is the word $x = 0101$. Dividing n by 3, we get the quotient q which is the list number of the word y. The possible remainders are 0,1 and 2.

We identify the words x or y with their respective list numbers n or q for a given $m \in \mathbb{N} \setminus \{1\}$. We denote the set of list numbers by

$$M_{binary} := \{0, 1, \ldots, 2^m - 1\}. \tag{3}$$

Table 1 shows the division by 3 for any $m \in \mathbb{N} \setminus \{1\}$. The construction of Table 1 is quite easy. Write the list numbers from 0 to $2^m - 1$ in the first column. Write each of them three times consecutively in column 2 starting with (0,0,0). When column 2 is filled up, continue with the remaining numbers in column 4 and finally in column 6. Write the remainders (0,1,2) in this order, filling up columns 3, 5 and 7.

	$r_{i-1}=0$		$r_{i-1}=1$		$r_{i-1}=2$	
n	q	r_{i+m-1}	q	r_{i+m-1}	q	r_{i+m-1}
0	0	0	$\lfloor \frac{2^m}{3} \rfloor$	$m \bmod 2+1$	$\lfloor \frac{2\cdot 2^m}{3} \rfloor$	$(m+1) \bmod 2+1$
1	0	1	$\lfloor \frac{1+2^m}{3} \rfloor$		$\lfloor \frac{1+2\cdot 2^m}{3} \rfloor$	
2	0	2				
3	1	0				
4	1	1				
5	1	2				
6	2	0				
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	$\lfloor \frac{n}{3} \rfloor$	$n \bmod 3$	$\lfloor \frac{n+2^m}{3} \rfloor$	$(n+2^m) \bmod 3$	$\lfloor \frac{n+2\cdot 2^m}{3} \rfloor$	$(n+2\cdot 2^m) \bmod 3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
2^m-1	$\lfloor \frac{2^m-1}{3} \rfloor$	$m \bmod 2$	$\lfloor \frac{2\cdot 2^m-1}{3} \rfloor$	$(m+1) \bmod 2$	$\lfloor \frac{3\cdot 2^m-1}{3} \rfloor$	2

Table 1: M_{binary}, quotients and remainders.

Definition 10. Let $m \in \mathbb{N} \setminus \{1\}$. The set of triads for m is defined by

$$S_{binary} := \left\{ \left(\left\lfloor \frac{n}{3} \right\rfloor, \left\lfloor \frac{n+2^m}{3} \right\rfloor, \left\lfloor \frac{n+2\cdot 2^m}{3} \right\rfloor \right) \middle| \, n \in M_{binary} \right\}.$$

We exclude $m=1$, because $q \in \{0,1\}$ would generate two times the same number in the triad. For $m>1$, the numbers of the triad are all different.

Definition 11. Let $m \in \mathbb{N} \setminus \{1\}$. Let $p = p_k p_k \cdots = x_1 x_2 \cdots$ be the infinite word of the binary expansion of $1/3^k$. The enlarged period of $1/3^k$ is the word $\overline{p}_k = p_k x_1 x_2 \cdots x_{m-1}$. We denote by $X_m(k)$ the set of (different) words of length m appearing in p. In the context it will be clear if its elements are words or list numbers identifying them.

Lemma 12. *For each $x \in X_m(k)$, the word x is factor of $\overline{p}_k$.*

Proof. Let $x = x_i \cdots x_{i+m-1} \in X_m(k)$. Take $j := i \bmod \ell_k$. If $j \neq 0$, then $x_j \cdots x_{j+m-1}$ is factor of $\overline{p}_k$. If $j=0$, then i is the last digit in some p_k, and the word $x_{\ell_k} x_1 x_2 \cdots x_{m-1}$ is factor of $\overline{p}_k$. □

Let $\overline{p}_k$ be the enlarged period of $1/3^k$ for some m. We determine the enlarged period $\overline{p}_{k+1}$ by executing the long division "$11)p_k p_k \overline{p}_k = \overline{p}_{k+1}$". We start with an example. In Example 13, we divide the enlarged period $\overline{p}_2$ of $1/9$ by 3 to get the enlarged period $\overline{p}_3$ of $1/27$. We do this for m equal to 2, 3 and 4. In the three cases, the dividend below the division line is $p_2 p_2 \overline{p}_2$, and the quotient above the division line is $\overline{p}_3$. The smaller numbers are list numbers which can be read in dividend and quotient by grouping the digits in blocks of 2, 3 or 4 digits, corresponding to m.

Example 13. Consider

$$\begin{array}{ccccccccccccccccccc}
\scriptstyle 00 & \scriptstyle 0 & \scriptstyle 1 & \scriptstyle 2 & \scriptstyle 0 & \scriptstyle 1 & \scriptstyle 2 & \scriptstyle 1 & \scriptstyle 3 & \scriptstyle 3 & \scriptstyle 3 & \scriptstyle 2 & \scriptstyle 1 & \scriptstyle 3 & \scriptstyle 2 & \scriptstyle 1 & \scriptstyle 2 & & \\
000 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & & \\
\hline
000 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & & \\
\scriptstyle 00 & \scriptstyle 1 & \scriptstyle 3 & \scriptstyle 3 & \scriptstyle 2 & \scriptstyle 0 & \scriptstyle 0 & \scriptstyle 1 & \scriptstyle 3 & \scriptstyle 3 & \scriptstyle 2 & \scriptstyle 0 & \scriptstyle 0 & \scriptstyle 1 & \scriptstyle 3 & \scriptstyle 3 & \scriptstyle 2 & & \\
\\
\scriptstyle 0 & \scriptstyle 0 & \scriptstyle 1 & \scriptstyle 2 & \scriptstyle 4 & \scriptstyle 1 & \scriptstyle 2 & \scriptstyle 5 & \scriptstyle 3 & \scriptstyle 7 & \scriptstyle 7 & \scriptstyle 6 & \scriptstyle 5 & \scriptstyle 3 & \scriptstyle 6 & \scriptstyle 5 & \scriptstyle 2 & \scriptstyle 4 & \\
000 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & \\
\hline
000 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & \\
\scriptstyle 0 & \scriptstyle 1 & \scriptstyle 3 & \scriptstyle 7 & \scriptstyle 6 & \scriptstyle 4 & \scriptstyle 0 & \scriptstyle 1 & \scriptstyle 3 & \scriptstyle 7 & \scriptstyle 6 & \scriptstyle 4 & \scriptstyle 0 & \scriptstyle 1 & \scriptstyle 3 & \scriptstyle 7 & \scriptstyle 6 & \scriptstyle 4 & \\
\\
 & \scriptstyle 0 & \scriptstyle 1 & \scriptstyle 2 & \scriptstyle 4 & \scriptstyle 9 & \scriptstyle 2 & \scriptstyle 5 & \scriptstyle 11 & \scriptstyle 7 & \scriptstyle 15 & \scriptstyle 14 & \scriptstyle 13 & \scriptstyle 11 & \scriptstyle 6 & \scriptstyle 13 & \scriptstyle 10 & \scriptstyle 4 & \scriptstyle 8 \\
000 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
\hline
000 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
 & \scriptstyle 1 & \scriptstyle 3 & \scriptstyle 7 & \scriptstyle 14 & \scriptstyle 12 & \scriptstyle 8 & \scriptstyle 1 & \scriptstyle 3 & \scriptstyle 7 & \scriptstyle 14 & \scriptstyle 12 & \scriptstyle 8 & \scriptstyle 1 & \scriptstyle 3 & \scriptstyle 7 & \scriptstyle 14 & \scriptstyle 12 & \scriptstyle 8
\end{array}.$$

For instance, if $m = 3$, we can read the following words in the dividend from left to right: 000, 001, 011, 111, 110, 100, 000, ..., 110, 100. The corresponding list numbers are 0, 1, 3, 7, 6, 4, 0, ..., 6, 4. In the quotient, we can read the list numbers 0, 0, 1, 2, 4, 1, 2, 5, ..., 2, 4. For $m = 3$, the 7 in the dividend generates 2, 7 and 5 in the quotient. Indeed, the list number 7 generates the triad (2,5,7) in Table 1. Note that these numbers appear permuted. In Example 13, we have $X_2(3) = \{0, 1, 2, 3\} = X_2(2)$. Further, $X_3(3) = \{0, 1, 2, 3, 4, 5, 6, 7\}$ and $X_3(2) = \{0, 1, 3, 4, 6, 7\}$. Finally, $X_4(3) = \{0, 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15\}$ and $X_4(2) = \{1, 3, 7, 8, 12, 14\}$. If $m > 6$, the words in the dividend (and also in the quotient) will be overlapping because the length of p_2 is 6. The example shows that $X_m(k)$ determines $X_m(k + 1)$ by means of table 1.

Proposition 14. *We have*

$$X_m(1) = \left\{ \left\lfloor \frac{2^m}{3} \right\rfloor, \left\lfloor \frac{2 \cdot 2^m}{3} \right\rfloor \right\}, \text{ where } m \in \mathbb{N} \setminus \{1\}.$$

Proof. We have $p_1 = 01$. Let $p := p_1 p_1 \cdots = x_1 x_2 \cdots$. Thus $\overline{p}_1 = p_1 x_1 x_2 \cdots x_{m-1}$.

There are only two words of length m: $n^{(1)} := 01\cdots$ and $n^{(2)} := 10\cdots$, both of length m.

Word $n^{(1)}$. If m is odd, then the last digit of $n^{(1)}$ is 0. Hence $n^{(1)} = 2^1 + 2^3 + 2^5 + \cdots + 2^{m-2}$. We get $n^{(1)} = (2^m - 2)/3$. So $n^{(1)} = \lfloor 2^m/3 \rfloor$ since $2^m \equiv 2 \pmod 3$. If m is even, then the last digit of $n^{(1)}$ is 1. Hence $n^{(1)} = 2^0 + 2^2 + 2^4 + \cdots + 2^{m-2}$. We get $n^{(1)} = (2^m - 1)/3 = \lfloor 2^m/3 \rfloor$ since $2^m \equiv 1 \pmod 3$.

Word $n^{(2)}$. For odd m, $n^{(2)} = 2^0 + 2^2 + \cdots + 2^{m-1} = (2^{m+1} - 1)/3 = \lfloor 2^{m+1}/3 \rfloor$. For even m, $n^{(2)} = 2^1 + 2^3 + \cdots + 2^{m-1} = (2^{m+1} - 2)/3 = \lfloor 2^{m+1}/3 \rfloor$. □

Proposition 15. *We have*

$$X_m(k+1) = \bigcup_{n \in X_m(k)} \left\{ \left\lfloor \frac{n}{3} \right\rfloor, \left\lfloor \frac{n+2^m}{3} \right\rfloor, \left\lfloor \frac{n+2\cdot 2^m}{3} \right\rfloor \right\}, \text{ where } k \in \mathbb{N},\ m \in \mathbb{N} \setminus \{1\}.$$

Proof. We execute the long division "$11)p_k p_k \overline{p}_k = \overline{p}_{k+1}$". The words are identified by their list numbers in M_{binary}. We use the notation $n^{(i)}$ and $q^{(i)}$ for words of length m having their first binary digit at position i.

The set of all words of length m in $p_k p_k \overline{p}_k$ is given by $X_m(k) = \{n^{(i)} \mid i \in L_k\}$. We have $n^{(i)} = n^{(i+\ell_k)} = n^{(i+2\ell_k)}$. Their input remainders r_{i-1}, $r_{i+\ell_k-1}$ and $r_{i+2\ell_k-1}$, and also their output remainders r_{i+m-1}, $r_{i+\ell_k+m-1}$ and $r_{i+2\ell_k+m-1}$, are all different by Lemma 9. It follows that the corresponding quotients $q^{(i)}$, $q^{(i+\ell_k)}$ and $q^{(i+2\ell_k)}$ are all different too.

Conversely, if $s^{(j)}$ is a factor of length m in $\overline{p}_{k+1}$ for some $j \in \{1, 2, \ldots, 3\ell_k\}$, then $s^{(j)} \in \{s^{(j \bmod \ell_k)}, s^{(j \bmod \ell_k + \ell_k)}, s^{(j \bmod \ell_k + 2\ell_k)}\}$. These factors are generated by $n^{(j \bmod \ell_k)} = n^{(j \bmod \ell_k + \ell_k)} = n^{(j \bmod \ell_k + 2\ell_k)}$. All these n's are equal to $n^{(j)}$. So,

$$X_m(k+1) = \bigcup_{i \in L_k} \{q^{(i)}, q^{(i+\ell_k)}, q^{(i+2\ell_k)}\}.$$

The q's are a permutation of $(\lfloor n^{(i)}/3 \rfloor, \lfloor (n^{(i)} + 2^m)/3 \rfloor, \lfloor (n^{(i)} + 2 \cdot 2^m)/3 \rfloor)$ (Table 1). □

Proposition 16. *Let $k \in \mathbb{N}$ and $m \in \mathbb{N} \setminus \{1\}$.*

$$X_m(k) = \left\{ \left\lfloor t \cdot \frac{2^m}{3^k} \right\rfloor \;\middle|\; t \not\equiv 0 \pmod 3 \quad \text{and} \quad 0 < t < 3^k \right\}.$$

Proof. Proposition 14 and Proposition 15 define a tree of triads. Note that the two numbers of $X_m(1)$ are elements of the triad generated by 0 in S_{binary}. Any $n \in M_{binary}$ generates the following tree of triads (we show only the levels $k = 1$ and $k = 2$):

$$\left(\left\lfloor \frac{n}{3} \right\rfloor, \left\lfloor \frac{n+2^m}{3} \right\rfloor, \left\lfloor \frac{n+2\cdot 2^m}{3} \right\rfloor\right)$$

$$\swarrow \qquad \downarrow \qquad \searrow$$

$$\overbrace{\left\lfloor \frac{n}{9} \right\rfloor, \left\lfloor \frac{n+3\cdot 2^m}{9} \right\rfloor, \left\lfloor \frac{n+6\cdot 2^m}{9} \right\rfloor} \overbrace{\left\lfloor \frac{n+1\cdot 2^m}{9} \right\rfloor, \left\lfloor \frac{n+4\cdot 2^m}{9} \right\rfloor, \left\lfloor \frac{n+7\cdot 2^m}{9} \right\rfloor} \overbrace{\left\lfloor \frac{n+2\cdot 2^m}{9} \right\rfloor, \left\lfloor \frac{n+5\cdot 2^m}{9} \right\rfloor, \left\lfloor \frac{n+8\cdot 2^m}{9} \right\rfloor}$$

The 3^k numbers written at level k have the form $\lfloor (n + t \cdot 2^m)/3^k \rfloor$, where $t \in \{0, 1, 2, \cdots, 3^k - 1\}$; eventually not all of them are different. The left branch of the tree corresponds to $t \equiv 0 \pmod 3$, the central branch to $t \equiv 1 \pmod 3$ and the right branch to $t \equiv 2 \pmod 3$. We substitute $n = 0$ and cut off the left branch. All the remaining numbers at level k are elements of $X_m(k)$ by Proposition 14 and Proposition 15. □

Note that $\mathrm{card}(X_m(k)) \leq \frac{2}{3} \cdot 3^k$.

Lemma 17. *There is a unique* $t_0 \equiv 2 \pmod 3$ *such that*

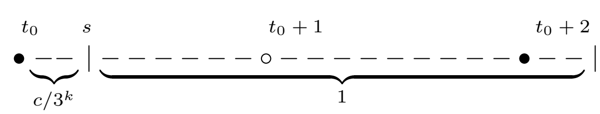
$$\left\lceil t_0 \cdot \frac{2^m}{3^k} \right\rceil - t_0 \cdot \frac{2^m}{3^k} \in \left\{ \frac{1}{3^k}, \frac{2}{3^k} \right\}.$$

Proof. The congruence $t \cdot 2^m + 1 \equiv 0 \pmod{3^k}$ has the unique solution $t' \equiv -1/2^m \pmod{3^k}$, and $t \cdot 2^m + 2 \equiv 0 \pmod{3^k}$ has the unique solution $t'' \equiv -2/2^m \pmod{3^k}$. Hence $t' 2^m/3^k + 1/3^k = \lceil t' 2^m/3^k \rceil$ and $t'' 2^m/3^k + 2/3^k = \lceil t'' 2^m/3^k \rceil$, where t' and t'' are taken as positive integers less than 3^k. Both are positive integers relatively prime to 3^k. Either $t' \equiv 1 \pmod 3$ and $t'' \equiv 2 \pmod 3$, or $t' \equiv 2 \pmod 3$ and $t'' \equiv 1 \pmod 3$. The claimed t_0 is either t'' or t'. □

Proposition 18. *Let* $k \in \mathbb{N}$ *and* $m \in \mathbb{N} \setminus \{1\}$*.*

$$X_m(k) = M_{binary} \quad \textit{if and only if} \quad 0 < \frac{2^m}{3^k} < \frac{1}{2}.$$

Proof. Let $t_0 \equiv 2 \pmod 3$ be as in Lemma 17. Then $t_0 + 1 \equiv 0 \pmod 3$, and $t_0 + 1$ is left out in $X_m(k)$ by Proposition 16. By Proposition 23, paragraph 1, $\{\lfloor t 2^m/3^k \rfloor \mid t \in \{0, 1, \cdots, 3^k - 1\}\} = M_{binary}$ if and only if $0 < 2^m/3^k < 1$. As shown in the following figure, we can avoid that the integer $s := \lfloor (t_0 + 1) 2^m/3^k \rfloor$ is left out in $X_m(k)$ by the condition $2 \cdot 2^m/3^k < 1 + c/3^k$ for the corresponding $c \in \{1, 2\}$ of Lemma 17, having now $s = \lfloor (t_0 + 2) 2^m/3^k \rfloor$.



By Lemma 22, there is no other $t \equiv 2 \pmod 3$ such that $t 2^m/3^k$ is closer to $\lceil (t 2^m/3^k \rceil$ than for t_0. It follows that the condition $2^{m+1} < 3^k + c$ is necessary and sufficient to guarantee $X_m(k) = M_{binary}$.

Recall that $m \geq 2$ and $k \geq 1$. Obviously, $2^{m+1} = 3^k$ is impossible. We show that $2^{m+1} = 3^k + 1$ is impossible too. Note that in this case $m+1$ is even since $2^{m+1} \equiv 1 \pmod 3$. We get $(2^{(m+1)/2} - 1)(2^{(m+1)/2} + 1) = 3^k$. Only one of the consecutive integers $(2^{(m+1)/2} - 1)$, $2^{(m+1)/2}$ and $(2^{(m+1)/2} + 1)$ is divisible by 3, so either $2^{(m+1)/2} - 1 = 1$ and $2^{(m+1)/2} + 1 = 3^k$, or $2^{(m+1)/2} - 1 = 3^k$ and $2^{(m+1)/2} + 1 = 1$. We get the excluded value $m = 1$ (and $k = 1$).

Hence $2^{m+1} < 3^k$. □

Proof of Theorem 1. By Proposition 18, $X_m(k) = M_{binary}$ if and only if $0 < 2^m/3^k < 1/2$. Thus $k > (m+1)\cdot\ln(2)/\ln(3)$. Hence $k_0 = \lceil (m+1)\cdot\ln(2)/\ln(3)\rceil$.

The same condition yields $m < k\cdot\ln(3)/\ln(2) - 1$, so $m \le \lfloor k\cdot\ln(3)/\ln(2)\rfloor - 1$. The number z of leading zeros in (2) is given by $z = \max\{i \mid 1/2^i > 1/3^k\}$, so $z = \lfloor k\cdot\ln(3)/\ln(2)\rfloor$. Thus $m-1 < z$ and $\overline{p}_k = p_k(0)^{m-1}$. All the 2^m words can be found in $\overline{p}_k$ by Lemma 12.

It remains to prove the last part of Theorem 1. We have $k \ge k_0$, so $X_m(k) = M_{binary}$. There is a $q \in \mathbb{N}$ such that $3^k = q\cdot 2^m + r$, where $0 < r < 2^m$. Then $q = \lfloor 3^k/2^m\rfloor$. By Proposition 23, paragraph 3, there are r different numbers in M_{binary} each of them occurring exactly $q+1$ times in $S_0 = (\lfloor t2^m/3^k\rfloor)_{t=0}^{3^k-1}$; each of the remaining $2^m - r$ numbers of M_{binary} occurs exactly q times in S_0. Clearly,

$$0 \le |f(n) - f(n')| \le 1 \quad \text{for all} \quad n, n' \in M_{binary}.$$

By Propositions 16 and 18, when leaving out each third term of S_0, we get a sequence S_0^* which also contains all the elements of $X_m(k)$.

In any sequence with s elements, by leaving out each third term, we eliminate $\lfloor s/3\rfloor$ or $\lceil s/3\rceil$ elements as suggested in the following scheme (the crosses signify "eliminated"):

s:	1	2	3	4	5	6	7	8	9	10	11
	x			x			x			x	
		x			x			x			x
			x			x			x		

Each number $n \in M_{binary}$ repeats in S_0 either q or $q+1$ times. In S_0^*, we have eliminated $\lfloor q/3\rfloor$ or $\lceil q/3\rceil$ elements for some $n \in M_{binary}$ occurring q times, and $\lfloor (q+1)/3\rfloor$ or $\lceil (q+1)/3\rceil$ elements for another $n' \in M_{binary}$ occurring $q+1$ times. We get

$$\begin{aligned} f(n) &\in \left\{q - \left\lfloor \frac{q}{3}\right\rfloor, q - 1 - \left\lfloor \frac{q}{3}\right\rfloor\right\}, \quad \text{and} \\ f(n') &\in \left\{q + 1 - \left\lfloor \frac{q+1}{3}\right\rfloor, q - \left\lfloor \frac{q+1}{3}\right\rfloor\right\}. \end{aligned}$$

It is an easy check that in the four possible cases holds $|f(n) - f(n')| \le 2$. □

Proof of Theorem 2. Let $p_{k_0}p_{k_0}\overline{p}_{k_0} = p_{k_0}p_{k_0}(p_{k_0}t)$ be the dividend and $p_{k_0+1}t$ the quotient, where t is the tail of $m-1$ zeros. We have $X_m(k_0) = M_{binary}$. By Proposition 15, $X_m(k_0+1) = \bigcup_{n\in X_m(k_0)}\{\lfloor \frac{n}{3}\rfloor, \lfloor \frac{n+2^m}{3}\rfloor, \lfloor \frac{n+2\cdot 2^m}{3}\rfloor\}$.

Every $n \in M_{binary}$ generates a triad. Conversely, the number a (respectively, b or c) of a triad (a, b, c) is generated by three consecutive list numbers as suggested in Table 2.

We define $T := \{n \in M_{binary} \mid$ the last binary digit of n is in $t\}$, and $T' := \{q \in M_{binary} \mid$ the last binary digit of q is in $t\}$ for the corresponding quotient. Every $q \in T'$ is generated by three consecutive list numbers. Since at least one of three

n-2	a						c			b	
n-1	a	b			a		c			b	c
n	a	b	c	<u>or</u>	a	b	c	<u>or</u>	a	b	c
n+1		b	c		a	b			a		c
n+2			c			b			a		

Table 2: Three consecutive list numbers (m odd).

consecutive list numbers —say z— must have its last binary digit equal to 1, the number q is generated by $z \notin T$ so that we do not need T' anymore. $\square$

Proof of Theorem 3. 1. Let $2^i = q_i \cdot 3^k + r_i$, where $0 < r_i < 3^k$ and $i \in \mathbb{N}$. Hence $q_i \equiv r_i \pmod 2$. Note that $r_i = 2^i \bmod 3^k$. We prove by induction the following statement:

$$r_i \equiv x_i \pmod 2 \text{ and } q_i = x_1 x_2 \cdots x_i \text{ (written as a binary integer).}$$

The statement holds for $i = 1$. Recall that $k \geq 1$. We have $2 = q_1 \cdot 3^k + r_1$, where $0 < r_1 < 3^k$. So $r_1 = 2$. Further $x_1 = 0$ since $1/3^k < 1/2$. Hence $r_1 \equiv x_1 \pmod 2$, and $q_1 = x_1$ since $q_1 = 0$.

Assume the statement holds for some i. Let $2^{i+1} = q_{i+1} \cdot 3^k + r_{i+1}$, where $0 < r_{i+1} < 3^k$. So $2^{i+1} = 2q_i \cdot 3^k + 2r_i$.

If $2r_i < 3^k$ then $r_{i+1} = 2r_i$, so $r_{i+1} \equiv 0 \pmod 2$. Further $q_{i+1} = 2q_i = x_1 \cdots x_i 0$, where $x_{i+1} = 0$. If $2r_i > 3^k$ then $r_{i+1} = 2r_i - 3^k$, so $r_{i+1} \equiv 1 \pmod 2$. Further $q_{i+1} = 2q_i + 1 = x_1 \cdots x_i 1$, where $x_{i+1} = 1$.

In both cases, $r_{i+1} \equiv x_{i+1} \pmod 2$ and $q_{i+1} = x_1 \cdots x_{i+1}$.

2. Write the remainders $2^i \bmod 3^k$ as a sequence R of increasing natural numbers:

$R :=$	(1,	2,	4,	5,	7,	8,	…	$1+3s$,	$2+3s$,	…	3^k-2,	3^k-1)
1's:	1			1	1			$\spadesuit$	$\clubsuit$		1	
								s even $\nwarrow$	$\nearrow$ s odd			
								1				

In the second line are written the 1's corresponding to the odd remainders by property 1. If $r = 1 + 3s$ is odd, then i is odd too since $2^i \equiv 1 + 3s \pmod{3^k}$ implies $2^i \equiv 1 \pmod 3$. If $r = 2 + 3s$ is odd, then i is even since $2^i \equiv 2 + 3s \pmod{3^k}$ implies $2^i \equiv 2 \pmod 3$. Therefore, the odd remainders $(1, 5, 7, \ldots, 3^k - 2)$ have the alternating positions (even, odd, even, odd,…,even). The last remainder in this sequence corresponds to an even position i, since there are $\ell_k/2 = 3^{k-1}$ odd remainders and $\lfloor 3^{k-1}/2 \rfloor$ pairs (even, odd). Hence $n_0 - n_1 = 1$.

3. There are $\ell_k/2 = 3^{k-1}$ odd remainders, each of them producing a 1. The remaining 3^{k-1} remainders produce a 0.

4. We have $2^{\ell_k} \equiv 1 \pmod{3^k}$. So $2^{\ell_k/2} \equiv -1 \pmod{3^k}$ since ℓ_k is even. Thus $2^{\ell_k/2+1} \equiv 1 \pmod{3^k}$. Hence,

$$\begin{array}{rcllrcl} 2^1 & \equiv & -1 \pmod{3^k} & \text{and} & 2^{\ell_k/2+1} & \equiv & 1 \pmod{3^k}; \\ 2^2 & \equiv & 1 \pmod{3^k} & \text{and} & 2^{\ell_k/2+2} & \equiv & -1 \pmod{3^k}; \\ & \vdots & & & & \vdots & \\ 2^{\ell_k/2} & \equiv & -1 \pmod{3^k} & \text{and} & 2^{\ell_k} & \equiv & 1 \pmod{3^k}. \end{array}$$

Then $2^i + 2^{\ell_k/2+i} \equiv 0 \pmod{3^k}$ for all $i \in (1, 2, \ldots, \ell_k/2)$. It follows $r_i + r_{\ell_k/2+i} = 3^k$. As 3^k is odd, one of the two remainders is odd and the other one is even. By property 1, $x_i + x_{\ell_k/2+i} = 1$.

5. The word $w := (0)^{\ell_k/2}(1)^{\ell_k/2}$ has the properties 2, 3 and 4 of Theorem 3. We start with w and try to construct p_k moving digits from one side to the other without violating the properties 2, 3 and 4. The positions i in the left half and the corresponding positions $\ell_k/2+i$ in the right half of w have different parity. We cannot exchange some 1 with any 0 because the property 4 remains valid if and only if 1 and 0 are placed at corresponding positions $\ell_k/2+i$ and i. When we exchange a 1 with the corresponding 0 at the left side, either n_0 or n_1 is decreasing 1 —initially, we had $n_0 - n_1 = 1$— so either $n_0 - 1 - (n_1 + 1) = -1$ or $n_0 + 1 - (n_1 - 1) = 3$ by property 3.

Therefore, we have to exchange the digits by pairs, so that the number of 1's in the left half must be even. When we move a pair of 1's, both having an even position, $n_0 - n_1$ will be altered: n_0 increases 2, and n_1 decreases 2, so $n_0 - n_1 = 5$. If both have initially an odd position, n_0 decreases 2, and n_1 increases 2, so $n_0 - n_1 = -3$. When we move two 1's to the left side, their starting positions must have different parity so that $n_0 - n_1 = 1$ is preserved.

6. By property 1, $\max\{i \mid 2^i < 3^k\} = \lfloor k \cdot \ln(3)/\ln(2) \rfloor$. □

4. The 2-Adic Expansion

Lemma 19. *Let $p_k = x_1 x_2 \cdots x_{\ell_k}$ be the period of the binary expansion of $1/3^k$ ($k \in \mathbb{N}$). The 2-adic expansion of $-1/3^k$ is purely periodic too, and its period $\tilde{p}_k$ is the reversal of p_k: $\tilde{p}_k = x_{\ell_k} \cdots x_2 x_1$.*

Proof. The ring $\mathbb{Z}_2$ of 2-adic integers has no zero divisors, so there is the field of fractions $\mathbb{Q}_2$ of 2-adic numbers. Every non-integer 2-adic number can be written as $\sum_{i=-s}^{\infty} x_i \cdot 2^i$, where $s > 0$ is some natural number. The binary expansion (2) can be written as

$$-\frac{1}{3^k} = \frac{-1}{2^{\ell_k} - 1} \cdot 2^{\ell_k} \sum_{i=1}^{\ell_k} x_i \frac{1}{2^i}.$$

We evaluate this expression term by term in the field $\mathbb{Q}_2$ of 2-adic numbers.

We get the 2-adic integer

$$\left.\frac{-1}{2^{\ell_k} - 1}\right|_{2-adic} = 1\ \underbrace{00\cdots0}_{\ell_k - 1\text{ zeros}}\ 1\underbrace{00\cdots0}\cdots,$$

since $-1 = 111\cdots$, $2^\ell_k - 1 = \underbrace{11\cdots 1}_{\ell_k \text{ ones}}$ and

$$\underbrace{11\cdots 1}_{\ell_k \text{ ones}} \times 1\ \underbrace{00\cdots 0}_{\ell_k - 1\,\text{zeros}}\ 1\underbrace{00\cdots 0}\cdots\ =\ \underbrace{111\cdots}_{-1}.$$

Further, we get the 2-adic non-integer

$$\sum_{i=1}^{\ell_k} x_i \frac{1}{2^i}\bigg|_{2-adic} = \sum_{i=-\ell_k}^{\infty} x_i 2^i = \sum_{i=-\ell_k}^{-1} x_i 2^i\ + 0,$$

where $x_{-i} := x_i$ for $-\ell_k \le i \le -1$, and $x_i = 0$ for $i \ge 0$. Note that the word $x_{-\ell_k}\cdots x_{-2}x_{-1}$ is the reversal of p_k.

We multiply the last sum by 2^{ℓ_k} to get the 2-adic integer $\tilde{p}_k 00\cdots$, where $\tilde{p}_k = x_{\ell_k}\cdots x_2 x_1$. Finally,

$$1\ \underbrace{00\cdots 0}_{\ell_k - 1\,\text{zeros}}\ 1\underbrace{00\cdots 0}\cdots\ \times\ \tilde{p}_k 00\cdots\ =\ \tilde{p}_k\tilde{p}_k\cdots\ .$$

□

Proof of Corollary 4. By Lemma 19, the word $(0)^{m-1}\tilde{p}_k$ is the reversal of the enlarged period $p_k(0)^{m-1}$. If one of the two words has complexity 2^m, then also the other one so that the necessary and sufficient condition of Theorem 1 holds here too. The frequency of non-palindromic words will possibly change but the difference of frequencies will not. □

Proof of Corollary 5. Apply Lemma 19. □

Proof of Corollary 6. We define the rotate left operator ρ by $\rho(x_1x_2x_3\cdots x_{\ell_k}) := x_2x_3\cdots x_{\ell_k}x_1$. Then

$$-\frac{1}{3^k}\bigg|_{2-adic} = \tilde{p}_k\tilde{p}_k\cdots = 1\rho(\tilde{p}_k)\rho(\tilde{p}_k)\cdots\ .$$

We need the additive inverse $+1/3^k\big|_{2-adic}$. Such a "change of sign" can be done in $\mathbb{Z}_2$ by exchanging 1's and 0's with exception of the first 1. For instance, $-1/3 = 101010\cdots$ and $1/3 = 1101010\cdots$. So

$$\frac{1}{3^k}\bigg|_{2-adic} = 1p_k^+ p_k^+\cdots\ ,$$

where p_k^+ denotes $\rho(\tilde{p}_k)$ with its 1's and 0's exchanged: $p_k^+ + \rho(\tilde{p}_k) = \underbrace{11\cdots 1}_{\ell_k \text{ ones}}$.

The rotate left operator ρ preserves the set of words of length m by moving the first word of $\tilde{p}_k$ to the end of $\rho(\tilde{p}_k)$; it also preserves the complementarity of left and right halves. Hence, if $\tilde{p}_k$ has complexity 2^m, then $\rho(\tilde{p}_k)$ has complexity 2^m too. Since all 2^m words of length m are factors of $\rho(\tilde{p}_k)$, p_k^+ contains the same words. We can even say where they can be found: we get $p_k^+ = \rho^{\ell_k/2}(\tilde{p}_k)$ by iterating ρ on $\tilde{p}_k$ exactly $\ell_k/2$ times. □

5. The Decimal Expansion

The decimal expansion of $1/3^k$ is purely periodic. Let ℓ_k denote the length of the period u_k. Then $\ell_1 = 1$ and $\ell_k = 3^{k-2}$ for $k \geq 2$, so ℓ_k is now odd. The words are over the alphabet $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Furthermore, $10^i \equiv 1 \pmod 3$ for all positive integers i. Since the long division in base 10 is essentially the same algorithm as in base 2, we have the following Lemma similar to Lemma 9.

Lemma 20. *Let $k \in \mathbb{N} \setminus \{1\}$ and $i \in \{1, 2, \ldots, 3^{k-2}\}$. The sequence (r_i, r_i', r_i'') is a permutation of* $(0, 1, 2)$.

Since $10^i \equiv 1 \pmod 3$, the proof is simpler and leads to only six possibilities (corresponding to the six permutations of $(0, 1, 2)$), equivalent to the case "i even" in base 2.

We define for $m \in \mathbb{N}$:

$$M_{decimal} := \{0, 1, \ldots, 10^m - 1\}; \tag{4}$$

$$S_{decimal} := \left\{ \left(\left\lfloor \frac{n}{3} \right\rfloor, \left\lfloor \frac{n + 10^m}{3} \right\rfloor, \left\lfloor \frac{n + 2 \cdot 10^m}{3} \right\rfloor \right) \middle| \, n \in M_{decimal} \right\}. \tag{5}$$

There is no special case $m = 1$ in base 10. The numbers of any triad of $S_{decimal}$ are all different, even in the case $m = 1$.

Definition 11 is still valid if we substitute "$m \in \mathbb{N} \setminus \{1\}$" by "$m \in \mathbb{N}$", "$p_k$" by "$u_k$" and "binary" by "decimal". If $m = 1$, then $(0)^{m-1}$ is defined as the empty word. Lemma 12 holds when we substitute $\overline{p}_k$ by $\overline{u}_k$.

For any $m \in \mathbb{N}$, the list number $0 \in M_{decimal}$ generates a tree. When we choose $m = 2$, the tree looks like this:

(0,33,66)

(0,33,66) (11,44,77) (22,55,88)

(0,33,66) (11,44,77) (22,55,88) (3,37,70) (14,48,81) (25,59,92) (7,40,74) (18,51,85) (29,62,96)

(1,34,67) (12,45,79) (23,56,90)

The length of the period u_k is regular for $k \geq 2$ (i. e., $\ell_{k+1} = 3\ell_k$). For p prime, the length of the period of $1/p^k$ is usually $(p-1)p^{k-1}$. There are three known exceptions, and $p = 3$ is one of them.[4] We have $1/9 = 0.111\cdots$ and $u_2 = 1$. At level $k = 2$, only the word 11 really appears in 1/9. Therefore, it is convenient to use the subtree generated by 11 because 11 generates the first triad $(3, 37, 70)$. We have $1/27 = 0.037037\cdots$, so $\overline{u}_3 = 0370$ for $m = 2$. In $\overline{u}_3$, there are the three words 03, 37 and 70 corresponding to the triad $(3, 37, 70)$ at level $k = 3$.

Further, $1/81 = 0.012345679012345679\cdots$. Then $\overline{u}_4 = 0123456790$ for $m = 2$. There are nine words in $\overline{u}_4$ corresponding to the triads $(1, 34, 67)$, $(12, 45, 79)$ and $(23, 56, 90)$ at level $k = 4$.

[4] Wikipedia, electronic: *http://en.wikipedia.org/wiki/Repeating_decimal.*

Proof of Theorem 7. For any $m \in \mathbb{N}$, define $1_m := \sum_{j=0}^{m-1} 10^j$. Instead of Proposition 14 and Proposition 15, we have now

$$X_m(3) = \left\{ \left\lfloor \frac{1_m}{3} \right\rfloor, \left\lfloor \frac{1_m + 10^m}{3} \right\rfloor \left\lfloor \frac{1_m + 2 \cdot 10^m}{3} \right\rfloor \right\}, \text{ and} \tag{6}$$

$$X_m(k+1) = \bigcup_{n \in X_m(k)} \left\{ \left\lfloor \frac{n}{3} \right\rfloor, \left\lfloor \frac{n + 10^m}{3} \right\rfloor, \left\lfloor \frac{n + 2 \cdot 10^m}{3} \right\rfloor \right\} \quad (k \geq 3). \tag{7}$$

It follows that the list numbers of $X_m(k)$ $(k \geq 3)$ have the form

$$X_m(k) = \left\{ \left\lfloor \frac{1_m + t \cdot 10^m}{3^{k-2}} \right\rfloor \,\middle|\, t \in \{0, 1, 2, \cdots, 3^{k-2} - 1\} \right\}.$$

We apply Proposition 23, paragraphs 1 and 2: $X_m(k) = M_{decimal}$ if and only if $0 < 10^m/3^{k-2} < 1$. This condition yields $k_0 = \lceil m \cdot \ln(10)/ln(3) \rceil + 2$. Additionally, $m \leq \lfloor (k-2)\ln(3)/\ln(10) \rfloor$. The number of leading zeros in $1/3^k$ is $z = \lfloor k \cdot \ln(3)/\ln(10) \rfloor$, so that $m - 1 < z$. Finally, $|f(n) - f(n')| \leq 1$ follows directly from Proposition 23, paragraph 3. $\square$

The proof of Corollary 8 is almost identical to that of Theorem 2. Instead of "Since at least one of three consecutive list numbers —say z— must have its last binary digit equal to 1" in the last part of the proof of Theorem 2, we have now "Since at least two of three consecutive list numbers —say z— must have its last digit different from 0". Obviously, instead of M_{binary}, Proposition 14 and Proposition 15, we have now $M_{decimal}$, expression (6) and expression (7).

6. A Helpful Sequence

Lemma 21. *Let $a/b \in \mathbb{Q}$ be a positive fraction in lowest terms: $a \in \mathbb{N}$, $b \in \mathbb{N} \setminus \{1\}$, and $\gcd(a, b) = 1$. Let $(\lfloor t \cdot \frac{a}{b} \rfloor)_{t=0}^{b-1}$ be a finite sequence of non-negative integers, where $t \in B := \{0, 1, \ldots, b-1\}$. Then the following hold:*

If $0 < \frac{a}{b} < 1$, then $\lfloor (t+1)\frac{a}{b} \rfloor - \lfloor t\frac{a}{b} \rfloor \leq 1$ for all $t, t+1 \in B$;

If $1 < \frac{a}{b}$, then there is some $t \in B$ such that $\lfloor (t+1)\frac{a}{b} \rfloor - \lfloor t\frac{a}{b} \rfloor \geq 2$.

Proof. Note that $b \geq 2$.
1. Let $0 < a/b < 1$. By the inequality $\lfloor x + y \rfloor \leq \lfloor x \rfloor + \lfloor y \rfloor + 1$ for $x, y \in \mathbb{R}$, we have $\lfloor t\frac{a}{b} + \frac{a}{b} \rfloor \leq \lfloor t\frac{a}{b} \rfloor + 1$ for all $t, t+1 \in B$.

2. Let $2 < a/b$. Then $t_0 = 0$ is the claimed t because $\lfloor a/b \rfloor - 0 \geq 2$.

3. Let $1 < a/b < 2$. Then $\frac{1}{b} \leq \frac{a-b}{b} < 1$ since $a - b \geq 1$. We write $\frac{a}{b} = 1 + \frac{a-b}{b}$. So

$$\left\lfloor t\frac{a}{b} \right\rfloor = t + \left\lfloor t\frac{a-b}{b} \right\rfloor. \tag{8}$$

There is a unique $t_0 \in B$ such that

$$\frac{t_0(a-b)}{b} + \frac{1}{b} = s$$

is an integer because $ta + 1 \equiv 0 \pmod{b}$ has the solution $t_0 \equiv -\frac{1}{a} \pmod{b}$. We prove that $t_0 \in B$ is the claimed t.

We have

$$\Big\lfloor \frac{t_0(a-b)}{b} + \frac{1}{b} \Big\rfloor = \frac{t_0(a-b)}{b} + \frac{1}{b} = s; \text{ hence } \Big\lfloor \frac{t_0(a-b)}{b} \Big\rfloor = s - 1.$$

Furthermore,

$$\Big\lfloor \frac{t_0(a-b)}{b} + \frac{a-b}{b} \Big\rfloor = \Big\lfloor \frac{t_0(a-b)}{b} + \frac{1}{b} + \frac{a-b-1}{b} \Big\rfloor = s + \Big\lfloor \frac{a-b-1}{b} \Big\rfloor = s.$$

We apply relation (8):

$$\begin{aligned} \Big\lfloor (t_0+1)\frac{a}{b} \Big\rfloor = t_0 + 1 + \Big\lfloor t_0 \frac{a-b}{b} + \frac{a-b}{b} \Big\rfloor &= t_0 + 1 + s, \\ \Big\lfloor t_0 \frac{a}{b} \Big\rfloor = t_0 + \Big\lfloor t_0 \frac{a-b}{b} \Big\rfloor &= t_0 + s - 1. \end{aligned}$$

Finally,

$$\Big\lfloor (t_0+1)\frac{a}{b} \Big\rfloor - \Big\lfloor t_0 \frac{a}{b} \Big\rfloor = 2.$$

□

Lemma 22. *Let $a/b \in \mathbb{Q}$ be a positive fraction in lowest terms: $a \in \mathbb{N}$, $b \in \mathbb{N} \setminus \{1\}$, and $\gcd(a,b) = 1$. Let $A := \{0,1,2,\ldots,a-1\}$, $B := \{0,1,2,\ldots,b-1\}$ and $n \in A$. Then*

$$D_{floor} := \Big\{ \frac{ta+n}{b} - \Big\lfloor \frac{ta+n}{b} \Big\rfloor \Bigm| t \in B \Big\} = \Big\{ \frac{0}{b}, \frac{1}{b}, \frac{2}{b} \ldots, \frac{b-1}{b} \Big\}.$$

Proof. As a is relatively prime to b, the positive integer a generates the additive cyclic group of order b in the finite ring $\mathbb{Z}/b\mathbb{Z}$. So $\{(ta) \bmod b \,|\, t \in B\} = B$. For a fix $n \in A$, $((ta+n) \bmod b)_{t=0}^{b-1}$ permutes the elements of B. Then also $\{(ta+n) \bmod b \,|\, t \in B\} = B$. Finally,

$$D_{floor} = \Big\{ \frac{(ta+n) \bmod b}{b} \Bigm| t \in B \Big\} = \Big\{ \frac{0}{b}, \frac{1}{b}, \frac{2}{b} \ldots, \frac{b-1}{b} \Big\}.$$

□

Proposition 23. *Let $a/b \in \mathbb{Q}$ be a positive fraction in lowest terms: $a \in \mathbb{N}$, $b \in \mathbb{N} \setminus \{1\}$, and $\gcd(a,b) = 1$. Let $A := \{0,1,2,\ldots,a-1\}$, $B := \{0,1,2,\ldots,b-1\}$ and $n \in A$. Then the following hold:*

1. $\Big\{ \Big\lfloor t \cdot \frac{a}{b} \Big\rfloor \Bigm| t \in B \Big\} = A \quad \Longleftrightarrow \quad 0 < \frac{a}{b} < 1.$

2. $\Big\{\Big\lfloor \dfrac{ta+n}{b}\Big\rfloor \,\Big|\, t\in B\Big\} = A \quad\Longleftrightarrow\quad \Big\{\Big\lfloor t\cdot \dfrac{a}{b}\Big\rfloor \,\Big|\, t\in B\Big\} = A.$

3. *Let* $0<\frac{a}{b}<1$ *and* $r := b \bmod a$. *In the sequence*

$$S_n := \Big(\Big\lfloor \frac{ta+n}{b}\Big\rfloor\Big)_{t=0}^{b-1},$$

each element of A *occurs either* $\lceil b/a\rceil$ *or* $\lfloor b/a\rfloor$ *times. There are exactly* r *elements of* A *occurring* $\lceil b/a\rceil$ *times in* S_n*; each of the remaining* $a-r$ *elements occurs* $\lfloor b/a\rfloor$ *times.*

Proof. 1. By Lemma 21.

2. We define

$$S_n' := \Big(\frac{ta+n}{b}\Big)_{t=0}^{b-1}.$$

By Lemma 22, the sequence S_n' contains exactly one integer. We calculate the corresponding t solving $ta+n\equiv 0 \pmod b$. So $t\equiv -\frac{1}{a}n \pmod b$. This linear function yields a unique $t\in B$ for each $n\in A$ and different t's for different n's. We denote this special t by t_n.

By Lemma 22, the sequence S_n' contains exactly one term $(ta+n)/b$ such that $(ta+n)/b+1/b$ is integer. We calculate the corresponding t solving $ta+n+1\equiv 0 \pmod b$. So $t\equiv -\frac{1}{a}n-\frac{1}{a} \pmod b$. We denote this special $t\in B$ by t_n'. Note that $t_n'\equiv t_n-\frac{1}{a} \pmod b$.

We get S_{n+1}' $(n\neq a-1)$ by adding $1/b$ to all terms of S_n' so that an integer is generated in S_{n+1}' at the position t_n' of S_n' and nowhere else by Lemma 22. Hence $t_{n+1}=t_n'$ $(n\neq a-1)$. It follows that the sequences S_n and S_{n+1} have the same terms, excepted at the position $t_{n+1}=t_n'$:

$$\Big\lfloor \frac{t_{n+1}a+n+1}{b}\Big\rfloor = \frac{t_{n+1}a+n+1}{b} = \frac{t_n'a+n}{b}+\frac{1}{b} = \Big\lfloor \frac{t_n'a+n}{b}\Big\rfloor + 1.$$

Let $\{\lfloor ta/b\rfloor \mid t\in B\} =$ A. Note that $t_0=0$ and $t_0'\neq t_0$. Further, $t_n'\neq 0$ for all $n\in A\setminus\{0\}$ because $-\frac{1}{a}n-\frac{1}{a}\equiv 0 \pmod b$ yields $n=b-1$, but $b-1\notin A$. The sequence S_0 has two or more leading zeros. When going up from n to $n+1$, no integer gets lost. Indeed, let

$$u_n' := \frac{(t_n'-1)a+n}{b} \quad\text{and}\quad v_n' := \frac{t_n'a+n}{b}$$

be two consecutive terms in S_n'. Then $v_n'+1/b := s+1$ is an integer —the only one— in S_{n+1}'. The integer $\lfloor v_n'\rfloor = s$ does not get lost in S_{n+1} since $\lfloor u_n'\rfloor = s$:

$$\Big\lfloor \frac{(t_n'-1)a+n}{b}\Big\rfloor = \Big\lfloor \frac{t_n'a+n}{b}+\frac{1}{b}-\frac{1+a}{b}\Big\rfloor = \Big\lfloor s+1-\frac{1+a}{b}\Big\rfloor = s.$$

Conversely, if $\{\lfloor (ta+n)/b\rfloor \mid t\in B\} = A$ for some $n\in A$, then there is at least one integer $s\in S_n$ occurring twice (or more times): $\lfloor u_n'\rfloor = \lfloor v_n'\rfloor = s$ at the positions

$t'_n - 1$ and t'_n. Since $t_n = t'_{n-1}$, we can go down from n to $n-1$ without loosing any of the integers of S_n. If $n = a-1$, then $t_{a-1} \equiv -1 + 1/a \pmod{b}$. It follows $t'_{a-1} \equiv -1 \pmod{b}$, so the integers at the positions $b-2$ and $b-1$ in S_{a-1} are always equal.

3. There is a $q \in \mathbb{N}$ such that $b = qa + r$ and $0 < r < a$. Clearly $q = \lfloor b/a \rfloor$. For an interval $I = [i, i+1)$, $i \in A$, we have

$$1 = q \cdot \frac{a}{b} + \frac{r}{b} \quad \text{and} \quad 0 < r < a.$$

Let D_{floor} be as in Lemma 22. Then $\{0/b, 1/b, 2/b, \dots, (r-1)/b\} \subset D_{floor}$. Note that the interval I is closed at the left and open at the right. For every $r' \in \{0/b, 1/b, 2/b, \dots, (r-1)/b\}$, the corresponding interval I —corresponding to a t such that $r' = (ta+n)/b - \lfloor (ta+n)/b \rfloor$— contains $\lfloor b/a \rfloor$ times the closed interval of length a/b and for any other $r'' \in D_{floor}$ only $(\lfloor b/a \rfloor - 1)$ times. Note that q intervals have $q+1$ endpoints. Consequently, there are r different integers in A each of them occurring exactly $(\lfloor b/a \rfloor + 1)$ times in S_n; each of the remaining $a - r$ integers occurs exactly $\lfloor b/a \rfloor$ times. □

References

[1] D. J. Bernstein, *A non-iterative 2-adic statement of the* $3x+1$ *conjecture*, Proc. Amer. Math. Soc. **121** (1994), 405–408. 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. Available at http://www.math.lsa.umich.edu/∼lagarias.

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

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

[5] Josefina Lopez and Peter Stoll. *The 3x+1 conjugacy map over a Sturmian word*, Integers **9** (2009), A13, 141–162. Available at http://www.integers-ejcnt.org/vol9.html.

[6] M. Lothaire. *Algebraic Combinatorics on Words*, volume **90** of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge, (2002).