Journal of Integer Sequences, Vol. 16 (2013),
Article 13.3.4

# Lambda Words: A Class of Rich Words Defined Over an Infinite Alphabet

Norman Carey
CUNY Graduate Center
365 Fifth Avenue
New York, NY 10016
USA
ncarey@gc.cuny.edu

**Abstract**

Lambda words are sequences obtained by encoding the differences between ordered elements of the form $i + j\theta$, where $i$ and $j$ are non-negative integers and $1 < \theta < 2$. Lambda words are right-infinite words defined over an infinite alphabet that have connections with Sturmian words, Christoffel words, and interspersion arrays. We show that Lambda words are infinite rich words. Furthermore, any Lambda word may be mapped onto a right-infinite word over a three-letter alphabet. Although the mapping preserves palindromes and non-palindromes of the Lambda word, the resulting Gamma word is not rich.

## 1 Introduction

Here we formalize the *Lambda word* [9], which is a sequence formed by encoding the differences between ordered elements of the form $i + j\theta$, where $i$ and $j$ are non-negative integers and $\theta$ is irrational, $1 < \theta < 2$. For example, when $\theta = \phi = \frac{\sqrt{5}+1}{2}$, the ordering begins

$$(0 + 0\phi) < (1 + 0\phi) < (0 + 1\phi) < (2 + 0\phi) < (1 + 1\phi) < \cdots .$$

Taking differences from $(0, 1, \phi, 2, 1 + \phi, \ldots)$ we get

$$1,\ \phi - 1,\ 2 - \phi,\ \phi - 1, \ldots .$$

There are four differences so far, but two of them are the same, and so the sequence of differences may be represented by the integer sequence, $(0, 1, 2, 1, \ldots)$. This is the beginning of $\Lambda_\phi$, the Lambda word generated by $\phi$ (A216763). Showing more terms,

$$\Lambda_\phi = (0, 1, 2, 1, 2, 3, 2, 2, 3, 2, 3, 4, 3, 2, 3, 4, 3, 3, 4, 3, 4, \ldots).$$

A Lambda word is defined over an infinite alphabet $\mathscr{A}$ of non-negative integers. Lambda words are related to Christoffel words [4] and Sturmian words [3, 12]. As we will show, they may be derived from interspersion arrays [23]. Lambda words developed from considerations of certain palindromic tonal spaces that contain important musical scales [10]. The main result of this paper is to show that Lambda words are rich words [19, 8].

Lambda words exhibit the following properties:

1. A Lambda word is a right-infinite word defined over an infinite alphabet [9, p. 46].
2. There are no recurrent letters (*a fortiori*, no recurrent factors) in a Lambda word.
3. The number of occurrences of any letter in a Lambda word may be calculated by means of convergents in the continued fraction expansion of $\theta$.
4. Palindromes in a Lambda word are on alphabets of no more than three letters.
5. A Lambda word is rich.
6. There exists a projection that maps a Lambda word onto a right-infinite word over a three-letter alphabet that preserves palindromes and non-palindromes.

In this paper, after definitions in §2, the structure of the Lambda word is outlined through examples in §3. In §4, properties of continued fractions are discussed. In §5, we demonstrate statements (2) and (3), and in §6, statement (4). The central result of the paper is to prove statement (5) in §7. Statement (6) is demonstrated in §8. We define Lambda words after some basic notation and definitions.

# 2 The Lambda word

Following Lothaire [27], we say that $\mathscr{A}$ is a set of symbols, finite or infinite. The symbols are referred to as *letters*. A finite or infinite sequence of letters from $\mathscr{A}$ is a *word*. A finite word over the alphabet $\mathscr{A}$ is an element of $\mathscr{A}^*$, the free monoid generated by $\mathscr{A}$. The monoid operation is concatenation where it is understood that the empty word, $\epsilon$, is a member of $\mathscr{A}^*$ and serves as the identity element. If $w$ is a word over $\mathscr{A}^*$ and $w = uv$, then $u$ is a *prefix* and $v$ a *suffix* respectively of $w$. Both $u$ and $v$ are *factors* of $w$, as is $x$ if $w = uxv$. The set of factors of a finite or infinite word $w$ is denoted by $\mathcal{F}(w)$. A *right-infinite word* over $\mathscr{A}$ is a map $h$ from the set of non-negative integers into $\mathscr{A}$ forming an infinite sequence, $h(0), h(1), \ldots, h(n), \ldots$, written with or without commas. If $w = x_1 x_2 \cdots x_k$ where $x_1, x_2, \ldots x_k \in \mathscr{A}$ then the *reversal* of $w$ is $\widetilde{w} = x_k \cdots x_2 x_1$. A *palindrome* is a word $p$ such that $p = \widetilde{p}$.

In what follows, $\mathscr{A}$ denotes an infinite alphabet of non-negative integers, elements of $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$. For letters $x, y \in \mathscr{A}$ we write $x \prec y$ to indicate lexicographic order, which, here, coincides with numerical order.

We define $\theta \in \mathbb{R} \setminus \mathbb{Q}$, $1 < \theta < 2$ and the set

$$S(\theta) = \{i + j\theta \mid i, j \in \mathbb{N}_0\}.$$

$S_\theta$ will denote the strictly increasing sequence obtained by sorting the elements of $S(\theta)$ in ascending order. The $n$-th term of $S_\theta$ is denoted by $s_n = i_n + j_n\theta$.

The first-order difference sequence of a given sequence $A$ is written as $\Delta A$. In particular, we will be interested in the difference sequence $\Delta S_\theta$ where the $n$-th term is denoted by $\delta_\theta(n) = s_{n+1} - s_n$. The differences $\delta_\theta(n)$ are generally decreasing as $n$ increases, but not strictly. Moreover, because $\Delta S_\theta$ contains repetitions, we may assign each element of $\Delta S_\theta$ to an integer (an element in $\mathscr{A}$) through a bijective mapping $\lambda$: Let $\lambda(\delta_\theta(0)) = 0$. If $n_1$ is the least value such that $\delta_\theta(n_1) \neq \delta_\theta(0)$, then $\lambda(\delta_\theta(n_1)) = 1$. Then $\lambda(\delta_\theta(n_2)) = 2$, where $n_2$ is the least value such that $\delta_\theta(n_2)$ is unequal to either $\delta_\theta(n_1)$ or $\delta_\theta(n_0)$, and so on. More than a simple coding of differences, the mapping $\lambda$ contains a deeper significance through a connection with the continued fraction expansion of $\theta$, to be discussed in §4.1.

Finally, $\Lambda_\theta = (\lambda(\delta_\theta(0)),\ \lambda(\delta_\theta(1)),\ \lambda(\delta_\theta(2)),\ \lambda(\delta_\theta(3)), \ldots)$. That is, $\Lambda_\theta$ is the word obtained from the sequence of differences in $\Delta S_\theta$, encoded by $\lambda$. We refer to $\Lambda_\theta$ as the *Lambda word generated by* $\theta$.

# 3 Examples of Lambda words

Let $\vartheta = \log_2 3$. The table below presents $\Lambda_\vartheta$ (A216448) as the encoded differences of the ascending sequence $S_\vartheta$.

| $n$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| $S_\vartheta$ | $0+0\vartheta$ | $1+0\vartheta$ | $0+1\vartheta$ | $2+0\vartheta$ | $1+1\vartheta$ | $3+0\vartheta$ | $0+2\vartheta$ | $2+1\vartheta$ |
| $\Delta S_\vartheta$ | $1-0\vartheta$ | $-1+1\vartheta$ | $2-1\vartheta$ | $-1+1\vartheta$ | $2-1\vartheta$ | $-3+2\vartheta$ | $2-1\vartheta$ | |
| $\Lambda_\vartheta$ | 0 | 1 | 2 | 1 | 2 | 3 | 2 | |

Because $1 - 0\vartheta$ is the "0-th" difference in $\Delta S_\vartheta$, it maps, under $\lambda$, to 0. When $n$ is 1 or 3, $\delta_\vartheta(n) = -1 + 1\vartheta$. Then $\lambda(\delta_\vartheta(1)) = \lambda(\delta_\vartheta(3)) = 1$. Similarly, $\lambda(\delta_\vartheta(2)) = \lambda(\delta_\vartheta(4)) = \lambda(\delta_\vartheta(6)) = 2$, and $\lambda(\delta_\vartheta(5)) = 3$, and so on.

According to Kimberling [24], the sequence of integers $i_n$ from $S_\theta$ is the *signature* of $\theta$, whereas $j_n$ gives the signature of $1/\theta$. Kimberling also introduces the *interspersion array.* Figure 1 shows the interspersion array associated with the signature sequence of $1/\vartheta^{-1}$. We modify Kimberling's definition of the array so as to include zero [23, p. 313]: An array $A = (a_{ij})$, $i \geq 0, j \geq 0$, of non-negative integers is an *interspersion* if

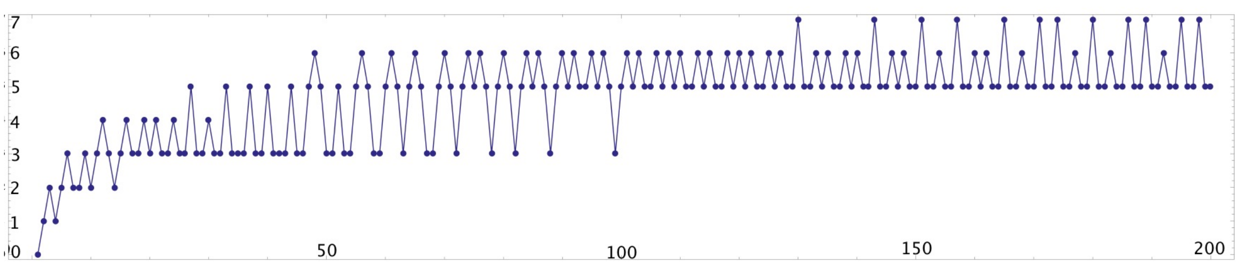
1. the rows of $A$ comprise a partition of the non-negative integers;
2. every row of $A$ is an increasing sequence;
3. every column of $A$ is an increasing (possibly finite) sequence;
4. if $(u_j)$ and $(v_j)$ are distinct rows of $A$ and if $p$ and $q$ are any indices for which $u_p < v_q < u_{p+1}$, then $u_{p+1} < v_{q+1} < u_{p+2}$.

According to Kimberling and Brown [25], an array such as the one shown in Figure 1 is *transposable*, because substituting $1/\theta$ for $\theta$ yields the transpose of the array. Adding one to each element in column one produces A022330, and the same addition on elements of the first row gives A022331. The relationship between transposable interspersions and Lambda words is this: Connecting the elements of the Figure 1 array in ascending order determines a sequence of vectors that begins $(1,0)$, $(-1,1)$, $(2,-1)$, $(-1,1)$, and these are also the coefficients for the successive elements of the difference sequence $\Delta S_\vartheta$. Labeling each distinct vector starting with 0 yields $\Lambda_\vartheta = (0,1,2,1,\ldots)$. (See A167267 for a transposable interspersion of the signature sequence of $\phi$.)

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 3 | 5 | 8 | 12 | 16 | 21 | 27 | 33 | 40 | 47 |
| 2 | 4 | 7 | 10 | 14 | 19 | 24 | 30 | 37 | 44 | 52 | |
| 6 | 9 | 13 | 17 | 22 | 28 | 34 | 41 | 49 | | | |
| 11 | 15 | 20 | 25 | 31 | 38 | 45 | 53 | | | | |
| 18 | 23 | 29 | 35 | 42 | 50 | | | | | | |
| 26 | 32 | 39 | 46 | 54 | | | | | | | |
| 36 | 43 | 51 | | | | | | | | | |
| 48 | | | | | | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $0 \to 1$ | $1 \to 2$ | $2 \to 3$ | $3 \to 4$ | $4 \to 5$ | $5 \to 6$ | $6 \to 7$ | $7 \to 8$ |
| $(1,0)$ | $(-1,1)$ | $(2,-1)$ | $(-1,1)$ | $(2,-1)$ | $(-3,2)$ | $(2,-1)$ | $(2,-1)$ |
| 0 | 1 | 2 | 1 | 2 | 3 | 2 | 2 |

Figure 1: Interspersion array for sequence $i(\vartheta^{-1})$ and its path of vectors.

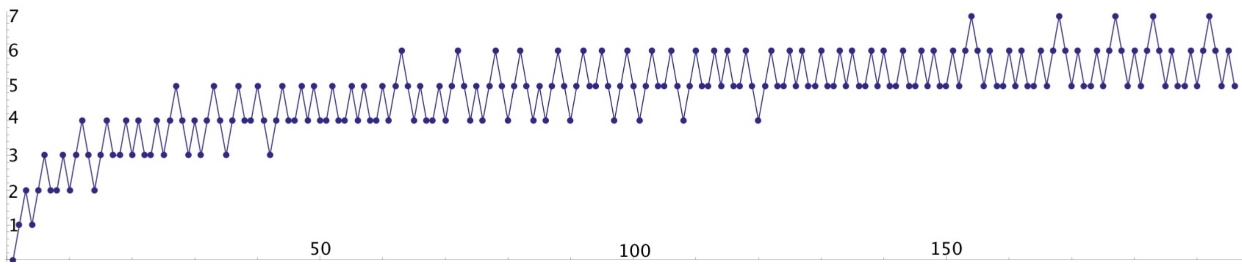



$\Lambda_\vartheta =$ 0, 1, 2, 1, 2, 3, 2, 2, 3, 2, 3, 4, 3, 2, 3, 4, 3, 3, 4, 3, 4, 3, 3, 4, 3, 3, 5, 3, 3, 4, 3, 3, 5, 3, 3, 3, 5, 3, 3, 5, 3, 3, 3, 5, 3, 3, 5, 6, 5, 3, 3, 5, 3, 3, 5, 6, 5, 3, 3, 5, 6, 5, 3, 5, 6, 5, 3, 3, 5, 6, 5, 3, 5, 6, 5, 6, 5, 3, 5, 6, 5, 3, 5, 6, 5, 6, 5, 3, 5, 6, 5, 6, 5, 5, 6, 5, 6, 5, 3, 5, 6, 5, 6, 5, 5, 6, 5, 6, 5, 6, 5, 5, 6, 5, 6, 5, 5, 6, 5, 6, 5, 6, 5, 5, 6, 5, 6, 5, 5, 7, 5, 5, 6, 5, 6, 5, 5, 6, 5, 6, 5, 5, 7, 5, 5, 6, 5, 6, 5, 5, 7, 5, 5, 6, 5, 5, 7, 5, 5, 6, 5, 6, 5, 5, 7, 5, 5, 6, 5, 5, 7, 5, 5, 7, 5, 5, 6, 5, 5, 7, 5, 5, 6, 5, 5, 7, 5, 5, 7, 5, 5, 6, 5, 5, 7, 5, 5, 7, 5, 5, ...

Figure 2: Graph of first 200 elements of $\Lambda_\vartheta$.

Figure 2 provides a graphical representation of the first 200 elements of $\Lambda_\vartheta$. It shows that $\Lambda_\vartheta$ gradually increases with $n$, but with repetitions and switchbacks, allowing for an abundance of palindromes. Figure 2 suggests other properties of Lambda words. For example, each integer appears a finite number of times: After smaller integers die out, larger ones take their place.

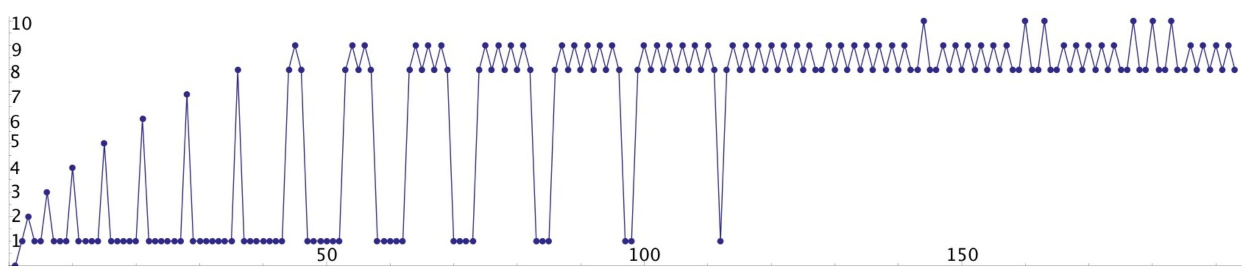
Of particular interest is A216763, the Fibonacci Lambda word, $\Lambda_\phi$, where $\phi = \frac{\sqrt{5}+1}{2}$. Figure 3 presents its first 196 elements. Because of the slow convergence of the continued fraction of $\phi$, successive members of $\Lambda_\phi$ differ by at most unity.




$\Lambda_\phi =$ 0, 1, 2, 1, 2, 3, 2, 2, 3, 2, 3, 4, 3, 2, 3, 4, 3, 3, 4, 3, 4, 3, 3, 4, 3, 4, 5, 4, 3, 4, 3, 4, 5, 4, 3, 4, 5, 4, 4, 5, 4, 3, 4, 5, 4, 4, 5, 4, 5, 4, 4, 5, 4, 4, 5, 4, 5, 4, 4, 5, 4, 5, 6, 5, 4, 5, 4, 4, 5, 4, 5, 6, 5, 4, 5, 4, 5, 6, 5, 4, 5, 6, 5, 4, 5, 4, 5, 6, 5, 4, 5, 6, 5, 5, 6, 5, 4, 5, 6, 5, 4, 5, 6, 5, 5, 6, 5, 4, 5, 6, 5, 5, 6, 5, 6, 5, 5, 6, 5, 4, 5, 6, 5, 5, 6, 5, 6, 5, 5, 6, 5, 5, 6, 5, 6, 5, 5, 6, 5, 6, 5, 5, 6, 5, 5, 6, 5, 6, 5, 5, 6, 5, 6, 7, 6, 5, 6, 5, 5, 6, 5, 6, 5, 5, 6, 5, 6, 7, 6, 5, 6, 5, 5, 6, 5, 6, 7, 6, 5, 6, 5, 6, 7, 6, 5, 6, 5, 5, 6, 5, 6, 7, 6, 5, 6, 5, . . .

Figure 3: Graph of first 196 elements of $\Lambda_\phi$.

The Lambda word generated by $\pi - 2$, A216764, presents a very different profile, as shown in Figure 4. The continued fraction expansion, $[1, 7, 15, 1, 292, \ldots]$, converges relatively rapidly. Note that the greatest difference between successive values so far is 7, which follows from the value of the second partial quotient of the expansion. Further on, the word will exhibit differences of 15, 292, and so on.




$\Lambda_{\pi-2} =$ 0, 1, 2, 1, 1, 3, 1, 1, 1, 4, 1, 1, 1, 1, 5, 1, 1, 1, 1, 1, 6, 1, 1, 1, 1, 1, 1, 7, 1, 1, 1, 1, 1, 1, 1, 8, 1, 1, 1, 1, 1, 1, 1, 8, 9, 8, 1, 1, 1, 1, 1, 1, 8, 9, 8, 9, 8, 1, 1, 1, 1, 1, 8, 9, 8, 9, 8, 9, 8, 1, 1, 1, 1, 8, 9, 8, 9, 8, 9, 8, 9, 8, 1, 1, 1, 8, 9, 8, 9, 8, 9, 8, 9, 8, 9, 8, 1, 1, 8, 9, 8, 9, 8, 9, 8, 9, 8, 9, 8, 9, 8, 1, 8, 9, 8, 9, 8, 9, 8, 9, 8, 9, 8, 9, 8, 9, 8, 8, 9, 8, 9, 8, 9, 8, 9, 8, 9, 8, 9, 8, 9, 8, 8, 10, 8, 8, 9, 8, 9, 8, 9, 8, 9, 8, 9, 8, 9, 8, 8, 10,8, 8, 10, 8, 8, 9, 8, 9, 8, 9, 8, 9, 8, 9, 8, 8, 10, 8, 8, 10, 8, 8, 10, 8, 8, 9, 8, 9, 8, 9, 8, 9, 8, . . .

Figure 4: Graph of first 193 elements of $\Lambda_{\pi-2}$.

# 4 Lambda words and continued fractions

The structure of a Lambda word is determined by the continued fraction expansion of $\theta$, its generating value. We demonstrate some of these relationships here, beginning with a theorem about differences in $\Delta S_\theta$. Standard theorems regarding continued fractions can be found in many texts [11, 22, 29, 26]. If $[t_0; t_1, \ldots, t_k, \ldots]$ is a continued fraction, then the $t_k$ are its *partial quotients* and $c_0 = [t_0]$, $c_1 = [t_0; t_1]$, $c_2 = [t_0; t_1, t_2]$, etc., are its *principal convergents*. There are two formal convergents for any continued fraction: $c_{-2} = 0/1$ and $c_{-1} = 1/0$. If $c_k = a_k/b_k$, then $\gcd(a_k, b_k) = 1$. Moreover, $c_{k+1} = (t_{k+1}a_k + a_{k-1})/(t_{k+1}b_k + b_{k-1})$. If $t_k > 1$ we define integer $t$ such that $1 \leq t < t_k$. Then $[t_0; t_1, \ldots, t_{k-1}, t]$ is an *intermediate convergent* of the continued fraction. In this paper, the term "convergent" means a principal convergent *or* an intermediate convergent. Khinchin [26] devises the term "best approximation of the second kind" for a rational $p/q$ of some real number $\theta$ if $|q\theta - p| < |q'\theta - p'|$ whenever $p \geq p'$. (A best approximation of the first kind involves the value $|\theta - p/q|$ and plays no role here. A "best approximation" will refer to one of the second kind.)

In addition to the best "two-sided" approximations described, Richards [30] further defines a best one-sided approximation: Take $p/q < \theta$. If $0 < q\theta - p < q'\theta - p'$ whenever $q \geq q'$, then $p/q$ is a best lower, or left, approximation. For any best left approximation it follows that $p = \lfloor q\theta \rfloor$. When $p/q$ is a best higher, or right approximation, $\theta < p/q$, and $q = \lfloor p/\theta \rfloor$. With $1 < \theta < 2$, any intermediate convergent is a best one-sided approximation and (with a single exception) any principal convergent is a best two-sided approximation. Certainly, any best two-sided approximation is the best on its side as well. (The exception is in the case of the formal convergent $0/1$. It is not a best left approximation because the convergent $1/1$ is also less than $\theta$ and has the same denominator.)

**Theorem 1.** *$|A - B\theta| \in \Delta S_\theta \iff A/B$ is a best approximation (of the second kind) of $\theta$.*

*Proof.* (See also [11, Theorem 1].)

$\Rightarrow$. We prove the case $\delta_\theta(n) = B\theta - A$, i.e., $A < B\theta$.

$$\delta_\theta(n) = s_{n+1} - s_n = (i_{n+1} + j_{n+1}\theta) - (i_n + j_n\theta) = B\theta - A.$$

Then $(j_{n+1} - j_n)\theta = B\theta$ and $i_n - i_{n+1} = A$, and so $s_{n+1} \geq B\theta$ and $s_n \geq A$. Define $g$ such that

$$g = (s_{n+1} - B\theta) = (s_n - A). \tag{1}$$

Then $g \in S_\theta$. Assume $B\theta - A > 0$ but $A/B$ is not a best left approximation. Then there exist successive best left approximations $p/q$ and $p'/q'$ with $q \leq B < q'$ such that $q\theta - p < B\theta - A$.

Then

$$0 < q\theta - p < B\theta - A.$$

Adding $A + g$ we obtain

$$A + g < A + g - p + q\theta < B\theta + g.$$

Clearly, $A + g - p + q\theta = s_{n'} \in S_\theta$, and so, by (1), $s_n < s_{n'} < s_{n+1}$. Contradiction. (For the case $\delta_\theta(n) = A - B\theta$, redefine $g$ as $\hat{g} = (s_{n+1} - A) = (s_n - B\theta)$.)

$\Leftarrow$. The second part of the proof is again by contradiction: Assume $A/B$ is a best left approximation. Now assume that $A = s_n$ and $B\theta = s_{n+k}$ with $k > 1$. Then we must have $s_{n+1} = i + j\theta$ such that

$$A < i + j\theta < B\theta \tag{2}$$

Then $j < B$. Because $\lfloor B\theta \rfloor = A$, we see that $i < A$.

Subtracting $A$ from (2) we obtain

$$0 < j\theta - (A - i) < B\theta - A. \tag{3}$$

Because $A/B$ is a best left approximation, there is no integer $k$ such that $j\theta - k < B\theta - A$. In particular, $j\theta - (A - i) \not< B\theta - A$. contradicting (3). Then if $A/B$ is a best left approximation, $\delta_\theta(n) = (B\theta - A)$, therefore $(B\theta - A) \in \Delta S_\theta$. □

## 4.1 The mapping $\lambda$ formalized

The continued fraction allows for an algorithmic form of the mapping $\lambda$: In the case of the formal convergent $1/0$, define $\lambda(1 - 0\theta) = \lambda(\delta_\theta(0)) = 0$. Otherwise, for each distinct difference $|A - B\theta|$, express $A/B$ as a finite continued fraction $[1; t_1, t_2, \ldots, t_k]$. Then $\lambda$ may be computed as

$$\lambda|A - B\theta| = \sum_{i=0}^{k} t_i.$$

| $\lvert A - B\vartheta\rvert$ | $A/B$ | $[1; t_1, \ldots, t_k]$ | $\sum_{i=0}^{k} t_i$ |
|---|---|---|---|
| 1 | 1/0 | | 0 |
| $\vartheta - 1$ | 1/1 | [1] | 1 |
| $2 - \vartheta$ | 2/1 | [1, 1] | 2 |
| $2\vartheta - 3$ | 3/2 | [1, 1, 1] | 3 |
| $5 - 3\vartheta$ | 5/3 | [1, 1, 1, 1] | 4 |
| $8 - 5\vartheta$ | 8/5 | [1, 1, 1, 2] | 5 |
| $7\vartheta - 11$ | 11/7 | [1, 1, 1, 2, 1] | 6 |
| $12\vartheta - 19$ | 19/12 | [1, 1, 1, 2, 2] | 7 |
| $27 - 17\vartheta$ | 27/17 | [1, 1, 1, 2, 2, 1] | 8 |
| $46 - 29\vartheta$ | 46/29 | [1, 1, 1, 2, 2, 2] | 9 |
| $65 - 41\vartheta$ | 65/41 | [1, 1, 1, 2, 2, 3] | 10 |

Figure 5: Computing $\lambda|A - B\vartheta|$ by sums of partial quotients.

Figure 5 shows the first eleven convergents of the continued fraction expansion of $\vartheta = \log_2 3$ under $\lambda$. Every non-negative integer is equal to exactly one of the summations, which is guaranteed for all $\theta$ by the second half of the proof of Theorem 1. As a result, the mapping $\lambda$ associates the difference $|A - B\theta|$ with the row of the Stern-Brocot tree on which $A/B$ is found [32, 20, pp. 116–117].

## 4.2 The Hurwitz chain

In what follows we will make use of an oft-noted connection between Farey series, continued fractions, and mediants [17, 30]. If $a/b < c/d$ are consecutive fractions in a Farey series then $bc - ad = 1 = \gcd(a, b) = \gcd(c, d)$. Then we say that $(a/b,\ c/d)$ form a *Farey pair*. Whenever $a/b < \theta < c/d$, the Farey pair is also a pair of best left-right approximations of $\theta$. The *Hurwitz chain* for $\theta$ contains all such pairs $(a/b, c/d)$ [21]. If $(a/b, c/d)$ belongs to the Hurwitz chain for $\theta$, then, if $a/b$ is an intermediate convergent, $c/d$ is the previous principal convergent and conversely. Otherwise, $a/b$ and $c/d$ are consecutive principal convergents. The mediant of these fractions, $(a + c)/(b + d)$ falls between them and is also a convergent. This implies that $(a + c)/(b + d)$ forms a member of the Hurwitz chain for $\theta$ with either $a/b$ or $c/d$.

Let $A/B$ and $p/q$ be a pair of best left-right approximations of $\theta$ such that $|A - B\theta| > |p-q\theta|$. Then either $(A/B, p/q)$ or $(p/q, A/B)$ belongs to the Hurwitz chain for $\theta$. If $A/B$ is a principal convergent, then $p/q$ is the principal convergent that immediately follows $A/B$, and if $A/B$ is an intermediate convergent, then $p/q$ is the principal convergent that immediately precedes $A/B$. Furthermore, $q > B$ when $A/B$ is a principal convergent, $q < B$ otherwise. Either way, $(A + p)/(B + q)$ is the first convergent that succeeds both.

# 5 Counting letters

A recurrent factor of an infinite word is a factor that appears infinitely often. Here we show that there are no recurrent letters, consequently, no recurrent factors in a Lambda word. We let $\lambda|A - B\theta| = x$ and let $|\Lambda_\theta|_x$ represent the number of times the letter $x$ occurs in $\Lambda_\theta$.

**Theorem 2.** $|\Lambda_\theta|_x = pq$.

*Proof.* We prove the theorem for $A/B < \theta$; the proof for $A/B > \theta$ is left as an exercise. First, let us define

$$\begin{array}{rcccc} X & = & B\theta & - & A \\ Y & = & p & - & q\theta \\ Z & = & (B+q)\theta & - & (A+p) \end{array}$$

Then $Z = X - Y$. It follows that $A/B < (A + p)/(B + q) < \theta < p/q$. Furthermore, $Y < X$ and $Z < X$. We define a subset $S(\theta)_x$ of $S(\theta)$ such that

$$S(\theta)_x = \{s_n \mid \delta_\theta(n) = X \text{ where } \lambda(X) = x\}.$$

That is, $s_n \in S(\theta)_x \iff s_n + X = s_{n+1}$. We will show that there are $pq$ elements in $S(\theta)_x$, and so $|\Lambda_\theta|_x = pq$. In order to do so, we prove that

$$S(\theta)_x = \{a + b\theta \mid A \le a < (A + p); 0 \le b < q\} \tag{4}$$

Because $s_n \in S(\theta)$ it immediately follows that $0 \le b$.

Further, $s_n + X \in S(\theta)$ implies $a + b\theta + (B\theta - A) \in S(\theta)$ and so $A \le a$.

If $s_n \in S(\theta)_x$ and $b - q \geq 0$, then it would follow that $s_n + Y = a + b\theta + (p - q\theta) \in S(\theta)$. But $Y < X$ and so $s_n < s_n + Y < s_n + X$, that is, $s_{n+1} \neq s_n + X$. Thus, $s_n \notin S(\theta)_x$, contrary to hypothesis. Therefore $b < q$.

If $a \geq A + p$, then $a + b\theta + ((B + q)\theta - (A + p)) = s_n + Z \in S(\theta)$. But $Z < X$ and so $s_n < s_n + Z < s_n + X$, contradicting $s_{n+1} = s_n + X$. Then $a < A + p$.

Therefore $A \leq a < (A + p)$ and $0 \leq b < q$ and so there are $pq$ elements in $S(\theta)_x$, i.e., $|\Lambda_\theta|_x = pq$. □

We note that, when $A/B > \theta$, we have

$$S(\theta)_x = \{a + b\theta \mid 0 \leq a < p; B \leq b < B + q\}. \tag{5}$$

# 6 Two-Letter factors of Lambda words

To determine the limit on the number of different letters in palindromes in a Lambda word, we explore the two-letter factors of the word, which we define as members of the set $\mathcal{F}_2(\Lambda_\theta)$. For $x, y \in \mathscr{A}$, we determine the conditions under which $xy \in \mathcal{F}_2(\Lambda_\theta)$.

## 6.1 Letter repetitions

To begin, we investigate the two-letter factors of the form $xx = x^2$. Let $t$ be the greatest integer such that $x^t$ is a factor of $\Lambda_\theta$. Let $t_k$ be the $k$-th partial quotient of the continued fraction expansion of $\theta = [1; t_1, \ldots, t_k, \ldots]$.

**Theorem 3.** *If $x = \lambda|A - B\theta|$ where $A/B$ is an intermediate convergent, $xx \notin \mathcal{F}_2(\Lambda_\theta)$. If $A/B$ is the principal convergent $a_k/b_k$, $xx \in \mathcal{F}_2(\Lambda_\theta)$. There are two special cases when $A/B$ is $a_{-1}/b_{-1}$ or $a_0/b_0$.*

*Proof.* Let $|A - B\theta| = X$. If a $xx$ is a factor of $\Lambda_\theta$, then $S(\theta)_x$ would have to include both $a + b\theta$ and $a + b\theta + X$. If $x^t$ is a factor, then $a + b\theta + (t - 1)X$ must be an element of $S(\theta)_x$. We can solve for $t$ by invoking the conditions on $a$ and $b$ in $S(\theta)_x$, as shown in (4) and (5).

$$\left.\begin{array}{c} A \leq a + (t-1)A < A + p \\ 0 \leq b + (t-1)B < q \end{array}\right\} \text{ by (4)}$$

$$\left.\begin{array}{c} 0 \leq a + (t-1)A < p \\ B \leq b + (t-1)B < B + q \end{array}\right\} \text{ by (5)}$$

We are interested in the maximal value for $t$, so we substitute minimal values for $a$ and $b$, which simplifies to two results

$$t < 1 + \frac{p}{A} \text{ and } t < 1 + \frac{q}{B}. \tag{6}$$

When $A/B$ is an intermediate convergent, $p/q$ is the previous principal convergent, and so $B > q$ (and $A > p$). Therefore $t < 1 + (q/B) < 2$, i.e., $t = 1$. Therefore, when $A/B$ is an intermediate convergent $xx \notin \mathcal{F}_2(\Lambda_\theta)$.

If $A/B$ is a principal convergent, then it is associated with the partial quotient $t_k$, and $p/q$ is the the next convergent, associated with $t_{k+1}$. Let $p'/q'$ represent the principal convergent that immediately precedes $A/B$. Then $q/B = t_{k+1} + (q'/B)$. Whenever $k > 0$, $q'/B \leq 1$, and so we can solve for $t$: $t < 1 + t_{k+1} + (q'/B)$. Then $t = 1 + t_{k+1}$. (Alternatively, $t < 1 + t_{k+1} + (p'/A)$ yields the same result for $k > 0$.) Because $t_{k+1} \geq 1$, then $t > 1$ and so whenever $A/B$ is a principal convergent, $xx \in \mathcal{F}_2(\Lambda_\theta)$.

There are two special cases that involve principal convergents when $k = -1$ or 0. In both cases, $t = t_{k+1}$: For $k = -1$, $A/B = 1/0$, $p/q = 1/1$. The first inequality of (6) resolves to $t < 2$, but the second entails division by zero and is rejected. Then $t = t_{k+1} = t_0 = 1$ and $xx \notin \mathcal{F}_2(\Lambda_\theta)$. For $k = 0$, $A/B = 1/1$, $p/q = (1 + t_1)/t_1$ The inequalities in (6) resolve to $t < 2 + t_1$ and $t < 1 + t_1$. Both must be satisfied but the second is more restrictive, so again $t = t_1 = t_{k+1}$ Then $xx \in \mathcal{F}_2(\Lambda_\theta)$ whenever $t_1 \geq 2$. □

## 6.2 Square-free two-letter factors

For $u, v, \in \mathscr{A}$ $u \neq v$, we examine conditions under which $uv \in \mathcal{F}_2(\Lambda_\theta)$. We let $r$ be a word over $\mathscr{A}$ such that $r \in \mathcal{F}(\Lambda_\theta)$. We begin by examining cases of pairs of letters $(u, v)$ with $u \prec v$ such that $vru \notin \mathcal{F}(\Lambda_\theta)$.

Recall that $X - Y = Z$. We define $W = X + Y$. We write $a/b \oplus c/d = (a+c)/(b+d)$ to indicate that the fraction on the right side is the mediant of the two on the left. Similarly, $a/b \ominus c/d = (a-c)/(b-d)$. We define a function $f$ on the difference $|A - B\theta|$ such that $f|A - B\theta| = A/B$. The pair $(f(X), f(Y))$ is a member of the Hurwitz chain for $\theta$. Then $f(X) \oplus f(Y) = f(Z)$ and so $(f(Y), f(Z))$ is also a member of the Hurwitz chain. It follows that $f(X) \ominus f(Y) = f(W)$ and so either $(f(W), f(X))$ or $(f(W), f(X))$ is a member of the Hurwitz chain as well. Certainly $f(W)$, $f(X)$, $f(Y)$, and $f(Z)$ are all convergents to $\theta$, and so by Theorem 1, $W$, $X$, $Y$, and $Z$, are differences in $\Delta S_\theta$. See Figure 6.

| | | | | |
|---|---|---|---|---|
| | | Differences | | |
| $W$ | $=$ | $X + Y$ | $=$ | $\mid (B-q)\theta - (A-p) \mid$ |
| $X$ | $=$ | | | $\mid B\theta - A \mid$ |
| $Y$ | $=$ | | | $\mid p - q\theta \mid$ |
| $Z$ | $=$ | $X - Y$ | $=$ | $\mid (B+q)\theta - (A+p) \mid$ |
| | | Ratios | | |
| $f(W)$ | $=$ | $f(X) \ominus f(Y)$ | $=$ | $\mid A - p \mid / \mid B - q \mid$ |
| $f(X)$ | $=$ | | | $A/B$ |
| $f(Y)$ | $=$ | | | $p/q$ |
| $f(Z)$ | $=$ | $f(X) \oplus f(Y)$ | $=$ | $(A+p)/(B+q)$ |
| | | Letters | | |
| $\lambda(W)$ | $=$ | $w$ | | |
| $\lambda(X)$ | $=$ | $x$ | | |
| $\lambda(Y)$ | $=$ | $y$ | | |
| $\lambda(Z)$ | $=$ | $z$ | | |

Figure 6: Differences, ratios, and letters that govern potential two-letter factors.

The following theorem provides groundwork for Theorem 5, which will show that $uv$ is a factor if and only if $f((U), f(V))$ belongs to the Hurwitz chain for $\theta$. As before, for difference $U$, we let $\lambda(U) = u$.

Let $r$ be a (possibly empty) factor of $\Lambda_\theta$.

**Theorem 4.** *If $(f(X), f(Y))$ belongs to the Hurwitz chain for $\theta$, $zrw \notin \mathcal{F}(\Lambda_\theta)$. Furthermore, $wz \notin \mathcal{F}_2(\Lambda_\theta)$.*

*Proof.* We prove the first part of the theorem by showing that the difference $W$ never comes after $Z$ in $\Delta S_\theta$. We prove the case $X = A - B\theta$:

$$\begin{array}{rcccc} W & = & X & + & Y \\ X & = & A & - & B\theta \\ Y & = & q\theta & - & p \\ Z & = & (A+p) & - & (B+q)\theta \end{array}$$

The smallest member of $S_\theta$ for which $\delta_\theta(n) = Z$ is $(B+q)\theta$. Therefore, if $W$ appears after $Z$, then there exists $\delta_\theta(m) = W$ such that $s_m \geq A + p > (B+q)\theta$. Consequently, $s_m > (B+q)\theta - Y = p + B\theta$. But $s_m < s_m + Y < s_m + X < s_m + W$ and at least one of these is true: $s_m + Y \in S_\theta$ and/or $s_m + X \in S_\theta$. Then $W \neq \delta_\theta(m)$ whenever $s_m \geq (B+q)\theta$. Therefore $zrw \notin \mathcal{F}(\Lambda_\theta)$.

It is also true that $wz$ is not a factor. The first part of the proof has shown that $\delta_\theta(n) \neq W$ when $s_n > (B+q)\theta$ and so we need only show that if $s_m = (B+q)\theta$, then $\delta_\theta(m-1) \neq W$. Clearly, $(B+q)\theta - Y \in S_\theta$. However, $(B+q)\theta - W < (B+q)\theta - Y < (B+q)\theta = s_m$ and so $W \neq \delta_\theta(m-1)$. □

Certainly if $z' \succeq z$ then $z'rw$ is not a factor of $\Lambda_\theta$.

Theorem 3 revealed the conditions under which $x^2 \in \mathcal{F}_2(\Lambda_\theta)$. The following theorem demonstrates the conditions under which $xy \in \mathcal{F}_2(\Lambda_\theta)$ is comprised of different letters. Let $u = \lambda(U)$, $v = \lambda(V)$, and $u \neq v$.

**Theorem 5.** *$uv \in \mathcal{F}_2(\Lambda_\theta) \iff (f(U), f(V))$ belongs to the Hurwitz chain for $\theta$.*

*Proof.* Theorem 4 has limited membership in $\mathcal{F}_2(\Lambda_\theta)$ for a fixed $\{x, y\}$ to a pair of letters in $\{w, x, y, z\}$. Then $U, V \in \{W, X, Y, Z\}$. Let $U = |A^* - B^*\theta|$, $V = |p^* - q^*\theta|$. By Theorem 1, $A^*/B^*$ and $p^*/q^*$ are convergents.

$\Rightarrow$ Assume $uv \in \mathcal{F}_2(\Lambda_\theta)$. If $(f(U), f(V))$ does not belong to the Hurwitz chain then at least one of these is false:

1. $|A^*q^* - B^*p^*| = 1$

2. $f(U)$ and $f(V)$ lie on opposite sides of $\theta$.

We test both conditions. Assume $|A^*q^* - B^*p^*| \neq 1$. Because $|Aq - Bp| = 1$, then $|A^*q^* - B^*p^*| = 1$ for all pairs of $\{f(W), f(X), f(Y), f(Z),\}$ except the pair $\{f(W), f(Z)\}$. But Theorem 4 has already shown that neither $zw$ nor $wz$ belong to $\mathcal{F}_2(\Lambda_\theta)$, therefore condition 1 is true, that is, $|A^*q^* - B^*p^*| = 1$.

Now assume that $f(U)$ and $f(V)$ are both greater than $\theta$. Let $\delta_\theta(n) = U$, $\delta_\theta(n+1) = V$, with $U > V$ (otherwise relabel $U \leftrightarrow V$). If $uv \in \mathcal{F}_2(\Lambda_\theta)$, then we have contiguous elements $s_n$, $s_n + U, s_n + U + V$. If $s_n = a + b\theta$, then $b \geq B^* + q^*$, so $s_n + V \in s_\theta$, but $s_n < s_n + V < s_n + U + V$, which implies that $U = V$, a contradiction. Similarly when $f(U)$ and $f(V)$ are both less than $\theta$. Therefore both conditions are true and so $uv \in \mathcal{F}_2(\Lambda_\theta)$ implies that $(f(U), f(V))$ is a member of the Hurwitz chain.

$\Leftarrow$ Now assume $(f(U), f(V))$ is a member of the Hurwitz chain for $\theta$. We show that $uv \in \mathcal{F}_2(\Lambda_\theta)$. We let $(f(U), f(V)) = (f(X), f(Y))$, remembering that $(f(X), f(Y))$ may represent *any* member of the Hurwitz chain for $\theta$. The limits on $S(\theta)x$ in Theorem 2 will be utilized to show that $\max(S(\theta)_x) + X \in S(\theta)_y$ and so $xy = uv \in \mathcal{F}_2(\Lambda_\theta)$. Let $P/Q$ be the principal convergent that immediately follows $p/q$. Again letting $X = B\theta - A$, $Y = p - q\theta$,

$$\max(S(\theta)_x) + X = (A + p - 1 + (q-1)\theta) + (B\theta - A) = p - 1 + (B + q - 1)\theta.$$

Then

$$p - 1 + (B + q - 1)\theta \in S(\theta))_y = \{a + b\theta \mid 0 \leq a < P; q \leq b < q + Q\}$$

because $0 \leq p - 1 < P$ and $q \leq (B + q - 1) < q + Q$. Then $uv \in \mathcal{F}_2(\Lambda_\theta$. We can also show that $\max(S(\theta)_x) - Y \in S(\theta)_y$ and so $vu$ also belongs to $\mathcal{F}_2(\Lambda_\theta)$. $\square$

For a fixed $(f(X), f(Y))$, members of the Hurwitz chain for $\theta$ always include both $(f(Y), f(Z))$ and $(f(X), f(Y))$. If $f(W)$ is on the same side of $\theta$ as $f(X)$, then $(f(W), f(Y))$ belongs to the Hurwitz chain. Otherwise, $(f(W), f(X))$ does. Then there are six two-letter square-free factors over $\{w, x, y, z\}$: $yz$, $xy$, $wx$ (or $wy$), and their reversals.

**Corollary 6.** *Palindromes in $\Lambda_\theta$ are over alphabets of at most three letters.*

*Proof.* By Theorem 4 it follows that if $z' \succeq z$ then $z'rw \notin \mathcal{F}(\Lambda_\theta)$ and so $z'rw\tilde{r}z' \notin \mathcal{F}(\Lambda_\theta)$ and $w\tilde{r}z'rw \notin \mathcal{F}(\Lambda_\theta)$. Then any palindrome in $\Lambda_\theta$ is over some alphabet $\mathscr{X} = \{x, y, z\} \subset \mathscr{A}$. $\square$

# 7 Rich words

The topic of palindromic complexity [1, 14, 15] concerns the question of the number of distinct palindromes contained in a word of length $k$. Words that achieve the maximal number of palindromes have been known alternately as "rich" words [1, 7, 19] or "full" words [5, 6]. We will use the term "rich" here, as it seems to be found more in current usage. It is known [13] that a word of length $k$ contains at most $k + 1$ distinct palindromes (including the empty word). A word that achieves this maximal number of distinct palindromes is a *rich* word [19]. Alternately, a rich word contains $k$ non-empty distinct palindromes. Thus, the word "large" is rich, as each of its five letters is a palindrome. The word "small" is also rich, because "s", "m", "a" "l" and "ll" constitute five non-empty palindromes. On the other hand, "edge" is not rich as it contains only three distinct non-empty palindromes. All words of up to length three are rich.

Droubay, Justin, and Pirillo [13] demonstrate an equivalent definition of rich words: A word is rich if each of its prefixes terminates with a uni-occurrent palindromic suffix (ups). For example, the relevant prefixes of the rich word "indeed" are "*i*", "i*n*", "in*d*", "ind*e*", "ind*ee*", "in*deed*", and the empty word. The italics capture the uni-occurrent palindromic suffixes. Because the reverse of a rich word is also rich, we can also say that a word is rich if each suffix contains a uni-occurrent palindromic prefix.

It follows that a word is rich if all of its factors are rich, which extends the concept to include infinite rich words. A right-infinite word is rich if each prefix of the word terminates in a ups. Recent studies of rich words have examined Sturmian [28], episturmian [18], and other words on finite alphabets. The results here extend the topic to the Lambda word, a word defined over an an infinite alphabet. By way of motivation, we again examine $\Lambda_\vartheta$ where $\vartheta = \log_2 3$. The beginning of this sequence was shown in Figure 2. As Figure 7 shows, all of its first twelve prefixes end with a ups.

$$
\begin{array}{r|l}
0 & \underline{0} \\
1 & 0,\ \underline{1} \\
2 & 0,\ 1,\ \underline{2} \\
3 & 0,\ \underline{1,\ 2,\ 1} \\
4 & 0,\ 1,\ \underline{2,\ 1,\ 2} \\
5 & 0,\ 1,\ 2,\ 1,\ 2,\ \underline{3} \\
6 & 0,\ 1,\ 2,\ 1,\ \underline{2,\ 3,\ 2} \\
7 & 0,\ 1,\ 2,\ 1,\ 2,\ 3,\ \underline{2,\ 2} \\
8 & 0,\ 1,\ 2,\ 1,\ 2,\ \underline{3,\ 2,\ 2,\ 3} \\
9 & 0,\ 1,\ 2,\ 1,\ \underline{2,\ 3,\ 2,\ 2,\ 3,\ 2} \\
10 & 0,\ 1,\ 2,\ 1,\ 2,\ 3,\ 2,\ 2,\ \underline{3,\ 2,\ 3} \\
11 & 0,\ 1,\ 2,\ 1,\ 2,\ 3,\ 2,\ 2,\ 3,\ 2,\ 3,\ \underline{4}
\end{array}
$$

Figure 7: Uni-occurrent palindromic suffixes in the first twelve prefixes of $\Lambda_\vartheta$.

The following proposition and four lemmas set up Theorem 12, which proves that $\Lambda_\theta$ is rich. We will show that every prefix of $\Lambda_\theta$ is terminated by a ups.

## 7.1 Additive complements

Two definitions are required: If $c + c' = h$, then $c'$ is the *complement* of $c$ with respect to $h$. A strictly increasing sequence $C = (c_1, \ldots, c_n)$ where complements are defined with respect to $c_1 + c_n$, is said to be a *sequence of additive complements* (or simply, "sequence of complements") if $c' \in C$ whenever $c \in C$. It follows that if $c_a < c_b$, then ${c_a}' > {c_b}'$ and also that ${c_k}' = c_{n+1-k}$ where $1 \le k \le n$.

**Proposition 7.** (Trivial). *$C$ is a sequence of complements iff $\Delta C$ is a palindrome.*

The proof is immediate from the definitions.

## 7.2 Beatty sequences

Setting $k \in \mathbb{N}$ allows for two values that approximate $\theta$:

$$\frac{\lfloor k\theta \rfloor}{k} < \theta < \frac{k}{\lfloor k(\theta^{-1}) \rfloor}.$$

These approximations include the “best approximations of the second kind” but form a larger set, comprised of the best left approximation for every denominator $k$, and the best right one for every numerator $k$. We also have

$$\begin{array}{rcl} \lfloor k\theta \rfloor & < & k\theta \\ \lfloor k(\theta^{-1}) \rfloor \theta & < & k. \end{array} \tag{7}$$

Summing the values in (7) gives rise to two integer sequences:

$$\begin{array}{rcl} K^{(-)} & = & \lfloor k\theta \rfloor + k \\ K^{(+)} & = & \lfloor k(\theta^{-1}) \rfloor + k \end{array}$$

where $k$ runs through the positive integers.

For the next lemma, we utilize the notion of complementary Beatty sequences, the word “complement” here used in a sense distinct from above. A Beatty sequence [2, 16] is an infinite sequence of integers $\lfloor k\alpha \rfloor$ where $\alpha$ is an irrational number. Complementary Beatty Sequences $\lfloor k\alpha \rfloor$ and $\lfloor k\beta \rfloor$ are formed whenever $1/\alpha + 1/\beta = 1$. Any positive integer belongs to exactly one of $\lfloor k\alpha \rfloor$ or to $\lfloor k\beta \rfloor$.

**Lemma 8.** *$K^{(-)}$ and $K^{(+)}$ form a complementary pair of Beatty sequences.*

*Proof.* If $\alpha = (\theta + 1)$, we get the Beatty Sequence $\lfloor k\alpha \rfloor = \lfloor k(\theta + 1) \rfloor = \lfloor k\theta + k \rfloor = \lfloor k\theta \rfloor + k = K^{(-)}$. Solving $1/(\theta + 1) + 1/\beta = 1$ for $\beta$ gives $\frac{\theta+1}{\theta}$, which yields Beatty Sequence $\lfloor k\beta \rfloor = \lfloor k(\frac{\theta+1}{\theta}) \rfloor = \lfloor \frac{k\theta}{\theta} + \frac{k}{\theta} \rfloor = k + \lfloor \frac{k}{\theta} \rfloor = K^{(+)}$. Consequently, every positive integer belongs either to $K^{(-)}$ or to $K^{(+)}$. $\square$

### 7.3 Nuclear sequences and palindromes

We utilize the limits shown in (7) to form two classes of contiguous subsequences of $S_\theta$, indexed by $K^{(-)}$ or $K^{(+)}$:

$$\begin{array}{l} N_{K^{(-)}} = (s_n \mid \lfloor k\theta \rfloor \leq s_n \leq k\theta) \\ N_{K^{(+)}} = (s_n \mid \lfloor k(\theta^{-1}) \rfloor \theta \leq s_n \leq k) \end{array}$$

Letting $K = K^{(-)} \cup K^{(+)}$, then $N_K$ refers to a member of either of these types of “nuclear” sequences. (It is not difficult to show that the nuclear sequences cover $S_\theta$, i.e., for all $n$, there is at least one $K$ such that $s_n \in N_K$. The result is not required in what follows.)

**Lemma 9.** *$N_K$ is a sequence of complements.*

*Proof.* We prove that both $N_{K^{(-)}}$ and $N_{K^{(+)}}$ are complementary sequences. We begin with $N_{K^{(-)}} = (s_n \mid \lfloor k\theta \rfloor \leq s_n \leq k\theta)$. Let $s'_n$ represent the complement of $s_n$. That is,

$$s_n + s'_n = \lfloor k\theta \rfloor + k\theta. \tag{8}$$

Clearly if $s_n \in N_{K(-)}$, $\lfloor k\theta \rfloor \le s'_n \le k\theta$. Now we show that $s'_n \in S_\theta$. Let $s_n = i + j\theta$, $s'_n = i' + j'\theta$. Then $\lfloor k\theta \rfloor \le i + j\theta \le k\theta$. This implies $\lfloor k\theta \rfloor = \lfloor i + j\theta \rfloor$, which gives $i \le \lfloor k\theta \rfloor$ and $j \le k$. Rewrite Equation (8) as

$$i + i' + (j + j')\theta = \lfloor k\theta \rfloor + k\theta$$

Because $\theta$ is irrational,

$$0 \le i + i' = \lfloor k\theta \rfloor$$
$$0 \le j + j' = k.$$

Then, $0 \le i' \le \lfloor k\theta \rfloor$ and $0 \le j' \le k$, which means $s'_n = i' + j'\theta \in S_\theta$ and therefore $s'_n \in N_{K(-)}$ whenever $s_n \in N_{K(-)}$. Then $N_{K(-)}$ is a sequence of complements.

Next we prove the same for $N_{K(+)} = (s_n \mid \lfloor k(\theta^{-1}) \rfloor \theta \le s_n \le k)$. Rewrite Equation (8) as

$$s_n + s'_n = \lfloor k(\theta^{-1}) \rfloor \theta + k. \tag{9}$$

and so if $s_n \in N_{K(+)}$, $\lfloor k(\theta^{-1}) \rfloor \theta \le s'_n \le k$.

If $s_n = i + j\theta \in N_{K(+)}$ we have $i + j\theta < k$, and so $i \le k$. Dividing the inequality $\lfloor k(\theta^{-1}) \rfloor \theta \le i + j\theta \le k$ through by $\theta$ we get $\lfloor \frac{k}{\theta} \rfloor \le \frac{i}{\theta} + j \le \frac{k}{\theta}$, and so $\lfloor \frac{k}{\theta} \rfloor = \lfloor \frac{i}{\theta} + j \rfloor$, or $j \le \lfloor \frac{k}{\theta} \rfloor$. Then

$$0 \le i + i' = k$$
$$0 \le j + j' = \lfloor k(\theta^{-1}) \rfloor.$$

Therefore both $i' \ge 0$ and $j' \ge 0$, and so $s'_n = i' + j'\theta \in S_\theta$ and $N_{K(+)}$ is also a sequence of complements. Thus for all $K$, $N_K$ is a sequence of complements of $S_\theta$. Then, by Proposition 7, $\Delta N_K$ is a palindrome. $\square$

We refer to $\Delta N_K$ as a *nuclear* palindrome.

## 7.4 Maximal sequences and palindromes

In general, it is possible to expand a nuclear palindrome, adding elements symmetrically to the left and the right as we will demonstrate below. First we simplify our notation and let

$$A = \begin{cases} \min(N_{K(-)}) \\ \max(N_{K(+)}) \end{cases} \quad B\theta = \begin{cases} \max(N_{K(-)}) \\ \min(N_{K(+)}) \end{cases}$$

It follows that $A + B\theta = \min(N_K) + \max(N_k)$ that is, complements in $N_K$ are defined with respect to $A + B\theta$. Furthermore, $A + B = K$. We can, in general, find longer complementary sequences of $A + B\theta$. Whenever $i + j\theta \in N_K$, $0 \le i \le A$ and $0 \le j \le B$. However, these limits hold on $i$ and $j$ whenever $s_n < \max(N_{K+1})$. Then let

$$T_K = (s_n \mid \max(N_K) \le s_n < \max(N_{K+1}))$$

$$T'_K = (s_n \mid s'_n \in T_K)$$

where, again, complements are defined with respect to $A + B\theta$. Clearly, the sequences $(T_K)$ and $(T'_K)$ have the same number of elements and $\min(T'_K) + \max(T_K) = A + B\theta$. Let $C_K = T'_K \cup N_k \cup T_K$. We call $C_K$ a "maximal" sequence.

**Lemma 10.** *$C_K$ is a sequence of complements.*

*Proof.* Like $N_K$, $C_K$ is a contiguous sequence of complements with respect to $A + B\theta$. If $s_n \in N_K$, then $s'_n \in N_K$. If $s_n \in T_K$ (resp. $T'_K$), then $s'_n \in T'_K$ (resp. $T_K$). $\square$

Consequently, by Proposition 7, $\Delta C_K$ is a (maximal) palindrome.

### 7.5 Medial subpalindromes

For palindrome $p$, if $p = xp'\widetilde{x}$ and $p' \neq \epsilon$, then $p'$ is a *medial subpalindrome* of $p$. (The term "central factor" is also used [13, p. 2]). It is easy to see that $p'$ itself is a palindrome. Clearly, $\Delta(N_K)$ is a medial subpalindrome of $\Delta(C_K)$ because

$$\Delta(C_K) = \Delta T'_K \cup \Delta(N_K) \cup \Delta T_K = \widetilde{\Delta T_K} \cup \Delta(N_K) \cup \Delta T_K.$$

(The reversal of $\Delta T_K$ is $\Delta T'_K$.)

Let $\Delta M$ be a medial subpalindrome of $\Delta(C_K)$ such that

$$\Delta N_K \subseteq \Delta M \subseteq \Delta C_K. \tag{10}$$

Let $S_M = (s_m \mid 0 \leq s_m \leq \max(M))$. Then $\Delta S_M$ is a non-empty prefix of $\Delta S_\theta$ and $\max(M) > 0$.

**Lemma 11.** *The palindrome $\Delta M$ is a ups of $\Delta S_M$.*

*Proof.* Clearly, $\Delta N_K$ is a medial subpalindrome in $\Delta M$ and $\Delta M$ is a palindromic suffix of $\Delta S_M$. Therefore if $\Delta N_K$ is uni-occurrent in $\Delta S_M$ then $\Delta M$ is also. The maximal values for $i$ and $j\theta$ in $S_M$ are $A$ and $B\theta$. Then $\Delta N_K$ is uni-occurrent in $\Delta S_M$ and $\Delta M$ is a ups in $\Delta S_M$. $\square$

**Theorem 12.** *$\Lambda_\theta$ is a rich word.*

*Proof.* From (10) it follows that $\max(N_K) \leq \max(M) \leq \max(C_K)$. Then $\max(M)$ is any member of $T_K$. Furthermore, for any $s_n > 0$, there is exactly one $T_K$ such that $s_n \in T_K$ because $\max(T_K) < \max(N_{K+1}) = \min(T_{K+1})$. Therefore every (non-empty) prefix of $\Delta S_M$ of $\Delta S_\theta$ ends with ups $\Delta M$. The bijection $\lambda$ guarantees that $\Lambda_\theta$ is a rich word. $\square$

## 8 A projection of $\Lambda_\theta$ onto $\Gamma_\theta$, a word over a three-letter alphabet

In Corollary 6, we found that any palindrome in $\Lambda_\theta$ is over a three-letter alphabet $\mathscr{X} = \{x, y, z\} \subset \mathscr{A}$. This suggests that we may map any Lambda word onto a word with only three symbols that preserves both palindromes and non-palindromes of the Lambda word. We construct a mapping $\gamma$ such that $\gamma(W) = \gamma(Z)$. We also require $\gamma(X) \neq \gamma(Y)$, where, as before $f(X)$ and $f(Y)$ form a member of the Hurwitz chain for $\theta$. Finally, both $\gamma(X)$ and $\gamma(Y)$ must be distinct from $\gamma(Z)$.

| $A/B$ | $\lambda\lvert A-B\vartheta\rvert$ | $\gamma\lvert A-B\vartheta\rvert$ |
|---|---|---|
| 1/0 | 0 | 0 |
| 1/1 | 1 | 1 |
| 2/1 | 2 | 2 |
| 3/2 | 3 | 0 |
| 5/3 | 4 | 1 |
| 8/5 | 5 | 2 |
| 11/7 | 6 | 1 |
| 19/12 | 7 | 0 |
| 27/17 | 8 | 1 |
| 46/29 | 9 | 2 |
| 65/41 | 10 | 1 |
| 84/53 | 11 | 2 |

Figure 8: The mappings $\lambda|A - B\vartheta|$ and $\gamma|A - B\vartheta|$.

Let $A/B$ be a convergent of $\theta$ and let $a_k/b_k$ be the principal convergent that immediately precedes $A/B$. Then the pair $(A/B, a_k/b_k)$ forms a member of the Hurwitz chain for $\theta$. We assign seed values to the first member of the chain $(1/1, 1/0)$:

$$\begin{array}{lclcl} \gamma|a_{-1} - b_{-1}\theta| & = & \gamma(1 - 0\theta) & = & 0 \\ \gamma|a_0 - b_0\theta| & = & \gamma(-1 + 1\theta) & = & 1 \end{array}$$

and then define recursively:

$$\gamma|A - B\theta| = 3 - (\gamma|a_k - b_k\theta| + \gamma|(A - a_k) - (B - b_k)\theta|). \tag{11}$$

Because $f(A/B) = f(a_k/b_k) \oplus f((A - a_k)/(B - b_k))$, we can interpret (11) as $\gamma(Z) = 3 - (\gamma(Y) + \gamma(X)$. We can rearrange (11) into

$$\gamma|(A - a_k) - (B - b_k)\theta| = 3 - (\gamma|A - B\theta| + \gamma|a_k - b_k\theta|) \tag{12}$$

which may be construed as $\gamma(W) = 3 - (\gamma(X) + \gamma(Y))$. The set $\{\gamma(X), \gamma(Y), \gamma(Z)\}$ maps bijectively onto $\{0, 1, 2\}$ and so does $\{\gamma(W), \gamma(X), \gamma(Y)\}$. Therefore,

$$\gamma(W) + \gamma(X) + \gamma(Y) = \gamma(X) + \gamma(Y) + \gamma(Z) = 3.$$

Moreover, for a fixed $(X, Y)$, $\gamma(W) = \gamma(Z)$ as required. Letting $\vartheta = \log_2 3$, Figure 8 shows the projection of the infinite alphabet $\mathscr{A}$ of $\Lambda_\vartheta$ onto $\{0, 1, 2\}$.

When $\theta = \phi$, the "golden" ratio, we can rewrite (11) and (12) as

$$\gamma|a_{k+1} - b_{k+1}\phi| = 3 - (\gamma|a_k - b_k\phi| + \gamma|a_{k-1} - b_{k-1}\phi|) = \gamma|a_{k-2} - b_{k-2}|\phi.$$

Because $\gamma|a_0 - b_0\theta| = 1$, it follows that $\gamma|a_k - b_k\phi| \equiv k + 1 \pmod 3$. That is, the successive values for $\gamma|A - B\phi|$ repeat the pattern (0, 1, 2). With any other value of $\theta$, intermediate convergents may come between $a_k/b_k$ and $a_{k+1}/b_{k+1}$. If $\gamma|a_k - b_k\theta| = n$, then, under $\gamma$, the images of the intermediate convergents that follow alternate between $n+1 \pmod 3$ and $n-1 \pmod 3$. When $\theta$ is a quadratic irrational, the successive values for $\gamma|A - B\theta|$ are eventually cyclic, whereas other values of $\theta$ exhibit no cyclic patterns. This, of course, reflects the patterning of partial quotients in the continued fraction expansions of these values.

The mapping $\gamma$ partitions the set of convergents of $\theta$ into three classes, in contrast to the well-observed partitioning into two classes that depends on whether $A/B$ is greater than or less than $\theta$. The properties of this tripartite classification could warrant further investigation.

The mapping defined in (11) projects $\Lambda_\theta$ onto the *Gamma word generated by* $\theta$, $\Gamma_\theta$, a word defined over $\{0, 1, 2\}$. The table below illustrates that projection when $\theta = \vartheta = \log_2 3$:

| $n$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\Lambda_\vartheta$ | 0 | 1 | 2 | 1 | 2 | 3 | 2 | 2 | 3 | 2 | 3 | 4 | 3 | 2 | 3 | 4 | 3 |
| $\Gamma_\vartheta$ | 0 | 1 | 2 | 1 | 2 | 0 | 2 | 2 | 0 | 2 | 0 | 1 | 0 | 2 | 0 | 1 | 0 |
| $n$ | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 |
| $\Lambda_\vartheta$ | 3 | 4 | 3 | 4 | 3 | 3 | 4 | 3 | 3 | 5 | 3 | 3 | 4 | 3 | 3 | 5 | 3 |
| $\Gamma_\vartheta$ | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 2 | 0 |

Although $\Gamma_\theta$ preserves palindromes and non-palindromes of $\Lambda_\theta$, $\Gamma_\theta$ is itself not rich: If any factor of a word is not rich, the word itself is not rich. The first six-letter factor of $\Gamma_\vartheta$, 012120, has no uni-occurrent palindromic suffix and is therefore not rich, a property that $\Gamma_\vartheta$ inherits. Certainly there are recurrent factors in the Gamma word but it is not known if all factors of all Gamma words are recurrent.

## 9 Acknowledgments

The author wishes to thank the editor and anonymous readers for their invaluable help. Thanks are also due to David Clampitt for many timely suggestions and corrections and to the other members of Micrologus, Emmanuel Amiot and Thomas Noll. Also, thanks to Srečko Brlek who provided insight into full (or rich) words at a conference at IRCAM.

## References

[1] J.-P. Allouche, M. Baake, J. Cassaigne, and D. Damanik, Palindrome complexity. *Theoret. Comput. Sci.* **292** (2003), 9–31.

[2] S. Beatty, A. Ostrowski, J. Hyslop, and A. C. Aitken, Solutions to Problem 3173. *Amer. Math. Monthly* **34** (1927), 159–160.

[3] J. Berstel and P. Séébold, Sturmian words, in *Algebraic Combinatorics on Words*, Encyc. Math. Appl., Vol. 90, Cambridge Univ. Press, 2002, pp. 45–110.

[4] J. Berstel, A. Lauve, C. Reutenauer, and F. V. Saliola, *Combinatorics on Words: Christoffel Words and Repetitions in Words*, CRM Monograph Series, Vol. 27, Amer. Math. Soc., 2008.

[5] A. Blondin Massé, S. Brlek, S. Labbé, and L. Vuillon, Palindromic complexity of codings of rotations, *Theoret. Comput. Sci.* **412** (2011), 6455–6463.

[6] S. Brlek, S. Hamel, M. Nivat, and C. Reutenauer: On the palindromic complexity of infinite words, *Internat. J. Found. Comput. Sci.* **15** (2004), 293–306.

[7] M. Bucci, A. de Luca, and A. De Luca, Rich and periodic-like words, in *DLT 2009, Proceedings of the 13th International Conference on Developments in Language Theory*, Lect. Notes in Comput. Sci., Vol. 5583, Springer, 2009, pp. 145–155.

[8] M. Bucci, A. De Luca, A. Glen, and L. Q. Zamboni, A new characteristic property of rich words, *Theoret. Comput. Sci.* **410** (2009), 2860–2863.

[9] N. Carey, On a class of locally symmetric sequences: The right infinite word $\Lambda_\theta$, in *Mathematics and Computation in Music*, Lect. Notes in Comp. Sci., Vol. 6726, Springer, 2011, pp. 42–55.

[10] N. Carey and D. Clampitt, Regions: A theory of tonal spaces in early medieval treatises, *J. Music Th.* **40** (1996), 113–147.

[11] N. Carey and D. Clampitt, Two theorems concerning rational approximations, *J. Math. Mus.* **6** (2012), 61–66.

[12] A. Carpi and A. de Luca, Central sturmian words: Recent developments, in *Developments in Language Theory*, Lect. Notes in Comp. Sci., Vol. 3572, Springer, 2005, pp. 36–56.

[13] X. Droubay, J. Justin, and G. Pirillo, Episturmian words and some constructions of de Luca and Rauzy, *Theoret. Comput. Sci.* **255** (2001), 539–553.

[14] X. Droubay, Palindromes in the Fibonacci word, *Inform. Process. Lett.* **55** (1995), 217–221.

[15] X. Droubay, and G. Pirillo, Palindromes and Sturmian words, *Theoret. Comput. Sci.* **223** (1999), 73–85.

[16] M. Gardner, *Penrose Tiles and Trapdoor Ciphers ... and the Return of Dr. Matrix*, W. H. Freeman, 1989.

[17] J. Goldman, Hurwitz sequences, the Farey process, and general continued fractions, *Adv. Math.* **72** (1988), 239–260.

[18] A. Glen, and J. Justin, Episturmian words: a survey, *RAIRO Theor. Inform. Appl.* **43** (2009), 402–433.

[19] A. Glen, J. Justin, S. Widmer, and L. Q. Zamboni, Palindromic richness, *European J. Combin.* **30** (2009), 510–531.

[20] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, 2nd ed., Addison-Wesley, 1994.

[21] A. Hurwitz, Über die angenäherte Darstellung der Zahlen durch rationale Brüche, *Math. Ann.* **44** (1894), 417–436.

[22] B. Jones, *The Theory of Numbers*, Holt, Rinehart, and Winston, 1955.

[23] C. Kimberling, Interspersions and dispersions, *Proc. Amer. Math. Soc.* **117** (1993), 313–321.

[24] C. Kimberling, Fractal sequences and interspersions, *Ars Combin.* **45** (1997), 157–168.

[25] C. Kimberling and J. Brown, Partial complements and transposable dispersions, *J. Integer. Seq.* **7** (2004), Article 04.1.6.

[26] A. Khinchin, *Continued Fractions*, trans. ed. by H. Eagle. Original Russian edition, 1935. Univ. Chicago Press, 1964.

[27] M. Lothaire, *Algebraic Combinatorics on Words*, Encyc. Math. Appl., Vol. 90, Cambridge Univ. Press, 2002.

[28] A. Luca, A. Glen, and L. Q. Zamboni, Rich, Sturmian, and trapezoidal words, *Theoret. Comput. Sci.* **407** (2008), 569–573.

[29] C. D. Olds, *Continued Fractions*, Random House, 1963.

[30] I. Richards, Continued fractions without tears, *Math. Mag.* **54** (1981) 163–171.

[31] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences, `http://oeis.org`.

[32] M. A. Stern, Über eine zahlentheoretische Funktion. *J. reine angew. Math.* **55** (1858) 193–220.

---

2013 *Mathematics Subject Classification* Primary 68R15; Secondary 06F05, 11J70, 11Y55.
*Keywords*: combinatorics on words, palindrome, rich word, continued fraction, Hurwitz chain.

---

(Concerned with sequences A022330, A022331, A167267, A216448, A216763, and A216764.)

---

Received September 15 2012; revised version received January 18 2013; February 17 2013. Published in *Journal of Integer Sequences*, March 2 2013.

---

Return to Journal of Integer Sequences home page.