A note on the Markoff condition and central words

Amy Glen, Aaron Lauve, Franco V. Saliola *

LaCIM, Université du Québec à Montréal, Case Postale 8888, succursale Centre-ville, Montréal (Québec) H3C 3P8, CANADA

Abstract

We define *Markoff words* as certain factors appearing in bi-infinite words satisfying the *Markoff condition*. We prove that these words coincide with *central words*, yielding a new characterization of *Christoffel words*.

Key words: combinatorial problems, Markoff condition, balanced words, central words, Christoffel words, palindromes.
1991 MSC: 68R15.

1. Introduction

In studying the minima of certain binary quadratic forms $AX^2+2BXY+CY^2$, Markoff [8,9] introduced a necessary condition that a bi-infinite word $\boldsymbol{s}$ must satisfy in order that it represent the continued fraction expansions of the two roots of $AX^2 + 2BX + C$. Over an alphabet $\{a, b\}$, his condition essentially states that each factor $x\widetilde{m}abmy$ occurring in $\boldsymbol{s}$, where $\widetilde{m}$ is the word m read in reverse and $\{x, y\} = \{a, b\}$, has the property that $x = b$ and $y = a$. We call such words m *Markoff words* in what follows. See Definition 2.

From [11] (see also [2, pg. 30]), it is known that the bi-infinite words satisfying the Markoff condition are precisely the *balanced words* of Morse and Hedlund [10]. After the work of A. de Luca [3,4], we know that palindromes now play a 'central' role in the study of such words. Here, we establish the following new characterization of a particular family of palindromes called *central words*.

Theorem 1 *A word is a Markoff word if and only if it is a central word.*

Central words hold a special place in the rich theory of *Sturmian words* (e.g., see [7, Chapter 2]). For instance, it follows from the work of de Luca and Mignosi [4,5] that central words coincide with the palindromic prefixes of standard Sturmian words.

As an immediate consequence of Theorem 1, we obtain a new characterization of *Christoffel words* in Corollary 7. Since the Markoff condition is relatively unknown, we discuss it and its relationship to Christoffel words at greater length in Section 5.

2. The Markoff condition

Fix an alphabet $\{a, b\}$. A finite sequence a_1, $a_2, \ldots, a_n$ of elements from $\{a, b\}$ is called a *word* of *length* n and is written $w = a_1a_2\cdots a_n$. The length of w is denoted by $|w|$ and we denote by $|w|_a$ (resp. $|w|_b$) the number of occurrences of the letter a (resp. b) in w.

A *right-infinite* (resp. *left-infinite*, *bi-infinite*) word over $\{a, b\}$ is a sequence indexed by $\mathbb{N}^+$ (resp. $\mathbb{Z}\backslash\mathbb{N}^+$, $\mathbb{Z}$) with values in $\{a, b\}$. For instance, a

* Corresponding author.
Email addresses: amy.glen@gmail.com (Amy Glen), saliola@gmail.com (Franco V. Saliola).
URLs: http://www.lacim.uqam.ca/∼glen (Amy Glen), http://www.lacim.uqam.ca/∼lauve (Aaron Lauve), http://www.lacim.uqam.ca/∼saliola (Franco V. Saliola).
[1] Thanks are due to C. Reutenauer and J. Berstel, who introduced the authors to Christoffel words at the Centre des Recherches Mathématiques (Montréal), March, 2007.

left-infinite word is represented by $\boldsymbol{u} = \cdots a_{-2}a_{-1}a_0$ and a right-infinite word by $\boldsymbol{v} = a_1a_2a_3\cdots$, and their *concatenation* gives the bi-infinite word $\boldsymbol{uv} = \cdots a_{-2}a_{-1}a_0a_1a_2a_3\cdots$. Infinite words are typically typed in boldface.

If $v = a_1a_2\cdots$ is a finite or a right-infinite word, then its *reversal* $\widetilde{v}$ is the word $\cdots a_2a_1$. Similarly, if $\boldsymbol{u}$ is a left-infinite word, then its reversal is the right-infinite word $\widetilde{\boldsymbol{u}}$. We define the reversal of a bi-infinite word $\boldsymbol{s} = \cdots a_{-2}a_{-1}a_0a_1a_2\cdots$ by $\widetilde{\boldsymbol{s}} = \cdots a_2a_1a_0a_{-1}a_{-2}\cdots$. A finite word w is a *palindrome* if $w = \widetilde{w}$.

A *factor* of a finite or infinite word w is a finite word v such that $w = uvu'$ for some words u, u'.

Definition 2 (1) *Suppose $\boldsymbol{s}$ is a bi-infinite word on the alphabet $\{a,b\}$. We say that $\boldsymbol{s}$ satisfies the* **Markoff condition** *if for each factorization $\boldsymbol{s} = \widetilde{\boldsymbol{u}}xy\boldsymbol{v}$ with $\{x,y\} = \{a,b\}$, one has either $\boldsymbol{u} = \boldsymbol{v}$ or $\boldsymbol{u} = my\boldsymbol{u}'$ and $\boldsymbol{v} = mx\boldsymbol{v}'$ for some finite word m (possibly empty) and right-infinite words $\boldsymbol{u}'$, $\boldsymbol{v}'$.*
(2) *A (finite) word m is a* **Markoff word** *if there exists a bi-infinite word $\boldsymbol{s}$ satisfying the Markoff condition with a factorization of the form $\boldsymbol{s} = \widetilde{\boldsymbol{u}}y\widetilde{m}xymx\boldsymbol{v}$, where $\{x,y\} = \{a,b\}$.*

Note that a bi-infinite word $\boldsymbol{s}$ satisfies the Markoff condition if and only if its reversal $\widetilde{\boldsymbol{s}}$ does, and $\boldsymbol{s}$ does not satisfy the Markoff condition if and only if $\boldsymbol{s}$ or $\widetilde{\boldsymbol{s}}$ contains a factor of the form $a\widetilde{m}abmb$ for some finite word m.

Words $\boldsymbol{s}$ satisfying the Markoff condition fall into four classes: the periodic class; two aperiodic classes; and an ultimately periodic class. See Section 5. An example of each type appears below.

$$\cdots(aabab)(aabab)(aabab)(aabab)\cdots \quad (1)$$
$$\cdots abaababababaababababaababababa\cdots \quad (2)$$
$$\cdots abaabababababaababababaabababa\cdots \quad (3)$$
$$\cdots(baaa)(baaa)baab(aaab)(aaab)\cdots \quad (4)$$

3. The balanced property

Observe that the above examples of bi-infinite words are "balanced" in the following sense.

Definition 3 *A finite or infinite word w over $\{a,b\}$ is said to be* **balanced** *if for any two factors u, v of w with $|u| = |v|$, we have $\big||u|_a - |v|_a\big| \leq 1$ (or equivalently, $\big||u|_b - |v|_b\big| \leq 1$), i.e., the number of a's (or b's) in each of u and v differs by at most 1.*

This notion dates back to the seminal work of Morse and Hedlund [10]. More recently, Reutenauer proved the equivalence between the Markoff condition and the above balanced property:

Proposition 4 [11, Theorem 3.1] *A bi-infinite word $\boldsymbol{s}$ satisfies the Markoff condition if and only if $\boldsymbol{s}$ is balanced.* □

In Section 5 we recount Reutenauer's Theorem 6.1 in [11], which gives a refinement of the balanced property and of the Markoff condition yielding the four classes illustrated by (1)–(4).

4. Central words

There exist several equivalent ways to define central words (see [7, Chapter 2]). Here we choose to use the following definition, as proved in [7, Proposition 2.2.34] using results from [4,5].

Definition 5 *A word w over $\{a,b\}$ is* **central** *if and only if awb and bwa are balanced.*

The following fact is especially pertinent.

Lemma 6 *Any central word is a palindrome.*

PROOF. If w is central, then awb and bwa are balanced by Definition 5. Arguing by contradiction, suppose w is not a palindrome. There exist words u, v, z and letters $\{x,y\} = \{a,b\}$ such that $w = uxv = zy\widetilde{u}$. But then

$$xwy = xuxvy = xzy\widetilde{u}y,$$

and the factors xux and $y\widetilde{u}y$ contradict the balanced property of xwy. □

Note. Lemma 6 also appears under a different guise in [5, Lemma 7]. Also see Corollary 2.2.9 in [7]. The above proof is easily adapted to show directly that Markoff words are palindromes.

We are now ready to prove Theorem 1: *a word is a Markoff word if and only if it is a central word.*

Proof of Theorem 1 Suppose m is a Markoff word. Let $\boldsymbol{s}$ be a bi-infinite word satisfying the Markoff condition for which $y\widetilde{m}xymx$ is a factor, where $\{x,y\} = \{a,b\}$. The reversal of this factor, namely $x\widetilde{m}yxmy$, is a factor of $\widetilde{\boldsymbol{s}}$, which also satisfies the Markoff condition. Therefore, the words amb and bma are factors of bi-infinite words satisfying the Markoff condition, and hence are balanced by Proposition 4. Thus m is central, by Definition 5.

Conversely, suppose m is a central word. Then m is a palindrome by Lemma 6, and moreover $amb = a\widetilde{m}b$ is balanced (Definition 5). Therefore the word $a\widetilde{m}bamb$ is also balanced and it can be viewed as a factor of some bi-infinite word satisfying the Markoff condition by Proposition 4—specifically, a bi-infinite word of the type represented in (1), with amb repeated bi-infinitely. Thus m is a Markoff word by Definition 2(2). □

An immediate corollary is a new characterization of Christoffel words (defined in the next section).

Corollary 7 *A word m is a Markoff word if and only if amb is a Christoffel word.*

PROOF. From [7, Chapter 2], a finite word amb is Christoffel word if and only if m is a central word, i.e., a Markoff word (by Theorem 1). □

5. Christoffel words

This section describes four classes of words satisfying the Markoff condition and how they naturally coincide with four classes of balanced words.

If a bi-infinite word $\boldsymbol{s}$ satisfies the Markoff condition, then it falls into exactly one of the following classes.

Let $\{x, y\} = \{a, b\}$.

(M_1) The lengths of the Markoff words m occurring in $\boldsymbol{s}$ are bounded and $\boldsymbol{s}$ cannot be written as $\widetilde{\boldsymbol{u}}xy\boldsymbol{u}$ for some word $\boldsymbol{u}$.

(M_2) The lengths of the Markoff words m occurring in $\boldsymbol{s}$ are unbounded and $\boldsymbol{s}$ cannot be written as $\widetilde{\boldsymbol{u}}xy\boldsymbol{u}$ for some word $\boldsymbol{u}$.

(M_3) There is exactly one $j \in \mathbb{Z}$ such that $\boldsymbol{s}$ has the factorization $\boldsymbol{s} = \widetilde{\boldsymbol{u}}s_js_{j+1}\boldsymbol{u}$ with $s_j \neq s_{j+1}$.

(M_4) $\boldsymbol{s}$ is not of type (M_1)–(M_3).

(Equivalently, $\boldsymbol{s}$ is in (M_4) iff there exist at least two $i \in \mathbb{Z}$ such that $\boldsymbol{s} = \widetilde{\boldsymbol{u}}s_is_{i+1}\boldsymbol{u}$ with $s_i \neq s_{i+1}$.)

The four examples (1)–(4) in Section 2 correspond, respectively, to the classes (M_1)–(M_4) above. We now turn to constructing words in each of the above classes. To achieve this, we present a geometric construction of Christoffel words, which allows for a description of balanced bi-infinite words.

Fix $p, q \in \mathbb{N}$, with p and q relatively prime. Let $\mathcal{P}$ denote the path in the integer lattice from $(0, 0)$ to (p, q) that satisfies: (i) $\mathcal{P}$ lies below the line segment $\mathcal{S}$ which begins at the origin and ends at (p, q); and (ii) the region in the plane enclosed by $\mathcal{P}$ and $\mathcal{S}$ contains no other points of $\mathbb{Z} \times \mathbb{Z}$ besides those of $\mathcal{P}$.

Each step in $\mathcal{P}$ moves from a point $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ to either $(x + 1, y)$ or $(x, y + 1)$, so we get a word $L(p, q)$ over the alphabet $\{a, b\}$ by encoding steps of the first type by the letter a and steps of the second type by the letter b. See Figure 1.

The word $L(p, q)$ is called the **(lower) Christoffel word** of slope $\frac{q}{p}$. The *upper Christoffel words* are defined analogously.

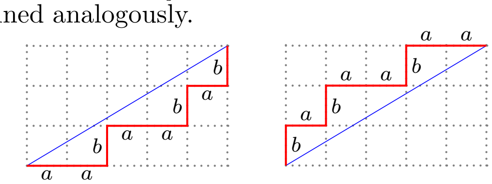

Fig. 1. The lower and upper Christoffel words of slope $\frac{3}{5}$ are $aabaabab$ and $babaabaa$, respectively.

For an introduction to the beautiful theory of Christoffel words, see [7, Chapter 2] or [1].

If the line segment $\mathcal{S}$ (as defined above) is replaced by a line ℓ, then the construction produces balanced bi-infinite words. Moreover, all balanced bi-infinite words can be obtained by modifying this construction; they fall naturally into the following four classes determined by ℓ.

(B_1) $\ell(x) = \frac{q}{p}x$ is a line of rational slope $\frac{q}{p}$ (these are the periodic balanced words, see (1)).

(B_2) ℓ is a line of irrational slope that does not meet any point of $\mathbb{Z} \times \mathbb{Z}$ (in (2), $\ell(x) = \frac{\pi}{4}x + e$).

(B_3) $\ell(x) = \alpha x$ is a line of irrational slope meeting exactly one point of $\mathbb{Z} \times \mathbb{Z}$ (in (3), $\ell(x) = \frac{\pi}{4}x$).

(B_4) The balanced words not of type (B_1)–(B_3).

Balanced bi-infinite words of type (B_4), represented in (4), are either of the form $\cdots xxyxx \cdots$ or $\cdots(ymx)(ymx)(ymy)(xmy)(xmy)\cdots$, where $\{x, y\} = \{a, b\}$ and m is a Markoff word. Hence, it is possible to adapt the geometric construction above to construct this class of balanced words also. See Figure 2.

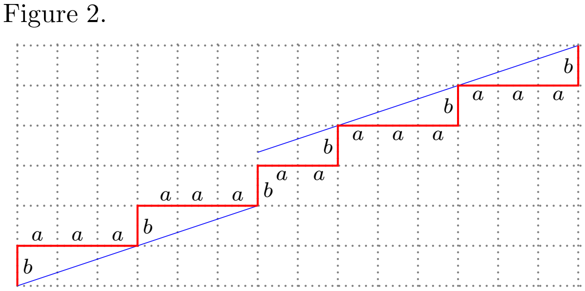

Fig. 2. Constructing example (4) $\cdots(baaa)baab(aaab)\cdots$.

As shown in [11], classes (B_1)–(B_4) are derived from Morse and Hedlund's description of balanced bi-infinite words [10] (also see Heinis [6]).

Proposition 8 [11, Theorem 6.1] *For* $1 \leq i \leq 4$*, one has the coincidences* $(M_i) = (B_i)$*.*

In closing, we mention that Markoff was interested in words over the alphabet $\{1, 2\}$ that satisfy the Markoff condition. For these words, he studied the continued fraction quantities

$$\lambda_i(\boldsymbol{s}) = s_i + [0, s_{i+1}, \cdots] + [0, s_{i-1}, s_{i-2}, \cdots]$$

and $\Lambda(\boldsymbol{s}) = \sup_i \lambda_i(\boldsymbol{s})$. Reutenauer [11, Theorem 7.2] showed that classes (M_1)–(M_4) correspond, respectively, to those $\boldsymbol{s}$ satisfying the Markoff condition with: $\Lambda(\boldsymbol{s}) < 3$; $\lambda_i(\boldsymbol{s}) < 3$ for all i but $\Lambda(\boldsymbol{s}) = 3$; $\Lambda(\boldsymbol{s}) = 3 = \lambda_i(\boldsymbol{s})$ for a unique $i \in \mathbb{Z}$; $\Lambda(\boldsymbol{s}) = 3 = \lambda_i(\boldsymbol{s})$ for at least two $i \in \mathbb{Z}$.

The set $\{\Lambda(\boldsymbol{s}) \mid \boldsymbol{s} \text{ is a bi-infinite word over } \mathbb{N}^+\}$, with none of the conditions on $\boldsymbol{s}$ originally imposed by Markoff, has become known as the **Markoff spectrum**. Results and open questions concerning the Markoff spectrum may be found in [2].

References

[1] J. Berstel, A. Lauve, C. Reutenauer, F. V. Saliola, Combinatorics on Words, CRM Proceedings & Lecture Notes, American Mathematical Society, Providence, RI, (to appear).

[2] T. W. Cusick, M. E. Flahive, The Markoff and Lagrange spectra, vol. 30 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 1989.

[3] A. de Luca, A combinatorial property of the Fibonacci words, Inform. Process. Lett. 12 (4) (1981) 193–195.

[4] A. de Luca, Sturmian words: structure, combinatorics, and their arithmetics, Theoret. Comput. Sci. 183 (1) (1997) 45–82.

[5] A. de Luca, F. Mignosi, Some combinatorial properties of Sturmian words, Theoret. Comput. Sci. 136 (2) (1994) 361–385.

[6] A. Heinis, Arithmetics and combinatorics of words of low complexity, Ph.D. Thesis, *University of Leiden*, 2001.

[7] M. Lothaire, Algebraic combinatorics on words, vol. 90 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 2002.

[8] A. Markoff, Sur les formes quadratiques binaires indéfinies, Math. Ann. 15 (3) (1879) 381–406.

[9] A. Markoff, Sur les formes quadratiques binaires indéfinies, Math. Ann. 17 (3) (1880) 379–399, (second memoire).

[10] M. Morse, G. A. Hedlund, Symbolic dynamics II. Sturmian trajectories, Amer. J. Math. 62 (1940) 1–42.

[11] C. Reutenauer, On Markoff's property and Sturmian words, Math. Ann. 336 (1) (2006) 1–12.