Math. Ann. 336, 1–12 (2006) **Mathematische Annalen**

DOI: 10.1007/s00208-005-0677-8

On Markoff's property and Sturmian words

Christophe Reutenauer

Received: 11 November 2004 / Revised version: 13 April 2005 /
Published online: 19 June 2006 –

1. Introduction

The aim of the present article is to prove the equivalence between two classifications of a family of two-sided sequences, called Sturmian sequences. The first classification is due to Morse and Hedlund, in their work in Symbolic Dynamics [MH]. We obtain the second by refining a property due to Markoff [M1], [M3] in his study of minima of quadratic forms and inequalities relating real numbers and convergents of continued fractions.

The fact that Markoff's original property and the balance property of Morse and Hedlund define the same family of two-sided sequences is stated by Cusick and Flahive in their very useful book [CF], but not proved; we offer a proof here, based on the theory of Sturmian morphisms (see the theory of Berstel and Séébold [BS]).

Surprisingly, Markoff's property is not mentioned by specialists of Sturmian sequences (see e.g. [BS], [AS], [A]), although his third article [M3] is. In that article, he answers a question of Bernoulli [B], whose article is generally considered as the first on Sturmian sequences, followed in the 19th century by Christoffel [C1], [C2], Smith [S], by Morse-Hedlund [MH] in 1940, and many articles at the end of the 20th century and after.

We give Markoff's property on two-sided infinite words in Section 3, together with the balance property of Morse and Hedlund. Th. 3.1 states their equivalence. In section 4, we refine Markoff's property by introducing a parameter; this leads to a classification of the two-sided infinite words satisfying Markoff's property into 4 classes. In Section 5, the classification of Morse and Hedlund (slightly modified) into 4 classes is recalled, and in Section 6 we prove the equivalence of the two classifications (Th. 6.1). In Section 7, we associate to each two-sided infinite word over the positive integers some real number, obtained as the supremum of all the "two-sided continued fractions" associated to the word. We give

C. Reutenauer
Université du Québec à Montréal, Département de mathématiques, Case postale 8888, succursale Centre-Ville, Montréal (Québec) Canada, H3C 3P8
(e-mail: christo@math.uqam.ca)

Supported by NSERC.

a proof of Markoff's result stating that this real number is ≤ 3 if and only if w satisfies Markoff's property (Th.7.1), hence equivalently, is balanced; the crucial fact states that if the real number associated to A is ≤ 3, then $A \in \{1, 2\}^{\mathbb{Z}}$, and the maximal blocks of 1 and 2's in A are of even length (then w is defined by taking half of all these blocks). As mentioned, these results are not new, due to Markoff; however, the statements given here are slightly different from (and we believe simpler than) those in the literature, as [D], [CF]; moreover, we believe that they are of great interest for combinatorialists of words, especially specialists of Sturmian sequences; that's why, we have included the proofs of these results in the present article, obtained by working out the existing proofs [M1],[D], [CF].

Finally we give an analytical formulation to the classification into 4 classes of balanced words (Th. 7.2).

We note that Markoff's theory of minima of binary quadratic forms and continued fractions is enriched by the theory of Sturmian sequences. For Markoff's theory, see also [D] and [CF]. See also [Y] for a result related to Th. 7.1. In [CF] are defined the Markoff spectrum and the Lagrange spectrum, which coincide for real numbers ≤ 3 and lead to Sturmian sequences. Note that another spectrum, the Cassaigne spectrum, is also related to Sturmian sequences [C].

The author is grateful to François Bergeron and André Joyal for many stimulating discussions related to the present work. Thanks also to Valérie Berthé and to the referee for his suggestions leading to improvements of the presentation.

2. Some definitions

We consider words, left infinite words, right infinite words and two-sided infinite words on the finite set $\{a, b\}$, called the *alphabet*. Let us recall their definitions: a *word* is a finite sequence on $\{a, b\}$; a *right infinite word* is a mapping $\mathbb{N} \to \{a, b\}$, a *left infinite* word is a mapping $\mathbb{Z}_- \to \{a, b\}$ (where $\mathbb{Z}_- = \{n \in \mathbb{Z} \mid n \leq 0\}$) and a *two-sided infinite* word is a mapping $\mathbb{Z} \to \{a, b\}$ modulo the shift on $\{a, b\}^{\mathbb{Z}}$. Notations are respectively $a_1 \dots a_n$, $a_0 a_1 a_2 \dots$, $\dots a_{-2} a_{-1} a_0$, and $\dots a_{-1} a_0 a_1 a_2 \dots$ *Concatenation* is partially defined for these words. For example, one may concatenate two finite words, a left infinite word and a right infinite word, and so on. An equality as $w = upqv$, where w is a two-sided infinite word, means implicitly that u (resp. v) is a left (resp. right) infinite word, and p, q are finite words. This is in general a self-evident notation, which we use without further comment.

We also use the notation p^∞ (resp. ${}^\infty p$) for the right (resp. left) infinite word obtained by repeating iteratively the finite word p. Likewise, ${}^\infty p^\infty$ denotes the analogous two-sided infinite word. This exponent notation will also be used for finite exponents; actually there is no difference between p^n and ${}^n p$, but it makes things more symmetrical.

We denote by $\tilde{w}$ or $w\tilde{}$ the *mirror image* or *reversal* of w; it is a finite word, if w is; it is a left infinite word if w is a right infinite word. The words *prefix* and *suffix* are used in their self-evident meaning. Thus b^∞ is a suffix of ${}^\infty ab^\infty$ and aa is a prefix of a^∞.

A word m is a *factor* of a word w if for some words u, v (which may be empty), one has $w = umv$. In other words, a factor is a prefix of a suffix (or symmetrically, a suffix of a prefix).

3. The Markoff property

Let w be a two-sided infinite word on $\{a, b\}$. We say that w has *property* (M) if for any factorization $w = uxyv$, where $\{x, y\} = \{a, b\}$, one has

- either $u = \tilde{v}$
- or there is a factorization $u = u'ym$, $v = \tilde{m}xv'$.

Note that, by our conventions in Section 2, u, u' (resp. v, v') are left (resp. right) infinite words and m is a finite word.

Examples. 1. $w = {}^\infty aba^\infty$ has property (M). Indeed, the factorizations to consider are ${}^\infty aaba^\infty$ and ${}^\infty abaa^\infty$, which both satisfy the property (with $u = \tilde{v}$).
2. $w = {}^\infty(aabab)^\infty$. Among several cases, one factorization is ${}^\infty(aabab)a(aba)ba(aba)b(aabab)^\infty$, which satisfies property (M) with $m = aba$, $u' = {}^\infty(aabab)$, $v' = (aabab)^\infty$.
3. $w = {}^\infty ab^\infty$ has not property (M), since w contains the factor $aabb$.
aabb

Note that in property (M), if m exists, then it is unique. This leads to the following variant of property (M): a two-sided infinite word w on $\{a, b\}$ has this property, if and only if for any factorization $w = \dots a_{-3}a_{-2}a_{-1}xya_1a_2a_3\dots$, with $a_i \in \{a, b\}$ and $\{x, y\} = \{a, b\}$, if there exists i with $a_{-i} \neq a_i$, then $a_{-i} = y$ and $a_i = x$ for the smallest of these i. Note that the m is the previous definition is then defined by $\tilde{m} = a_1 \dots a_{i-1}$, $m = a_{1-i} \dots a_{-1}$. This variant of property (M) is essentially that appearing in Markoff's article [M1] p. 397 and [M3] p.28; see also [D] p. 85–87, [CF] p.6–7.

Following [MH], we say that a word on $\{a, b\}$ is *balanced* if for any finite factors u, v of equal length, the number of $a's$ in u and in v differ at most by 1 (equivalently, the number of $b's$, since u, v must have equal length). The following result is given without proof in [CF] p.30 (balanced $\Leftrightarrow$ Markoff balanced, in their terminology, see p.28–29); the authors claim that "rather tedious combinatorial arguments" are necessary to give a direct proof, and that this result simplifies their approach to the theory of Markoff. We use below some results on Sturmian morphisms to give a direct proof.

Theorem 3.1. *A two-sided infinite word w on the alphabet $\{a, b\}$ has property (M) if and only if it is balanced.*

We use the following lemma.

Lemma 3.1. *If w has property (M), then aa, bb are not simultaneously factors of w.*

A proof of this lemma will be given at the end of this Section.

A *Sturmian morphism* is an endomorphism of the free monoid $\{a, b\}^*$ generated by $\{a, b\}$, such that a, b are mapped onto nonempty finite words, and which preserves right infinite Sturmian words when acting by substitution; for our purposes, it is enough to know that right infinite Sturmian words coincide with balanced right infinite words. For this and topics on Sturmian words and morphisms, see [BS].

Following these authors, denote by G and D the morphisms defined by: $D(a) = ba$, $D(b) = b$, $G(a) = a$, $G(b) = ab$. They are Sturmian ([BS] Cor. 2.2.19). Since finite balanced words coincide with finite factors of right infinite Sturmian words ([BS] Prop. 2.1.17), a Sturmian morphism sends balanced finite words onto balanced finite words. From this we deduce that, likewise, it sends balanced infinite words (left, right or two-sided) onto balanced words. We need this remark in the proof of Theorem 3.1.

Proof of Theorem. 1. If w is balanced, for any factorization $w = u'y'mxy\tilde{m}x'v'$ with $\{x, y\} = \{a, b\}$, $x', y' \in \{a, b\}$, we cannot have $y' = x$ and $x' = y$; indeed, otherwise, w has the two factors xmx and $y\tilde{m}y$ and w is not balanced. Hence either $y' = x'$, and we find such a factorization of w with a longer m, or $y' = y$, $x' = x$. Hence w has property (M).

2a. Let w have property (M). By the Lemma, w has not simultaneously the factors aa and bb. We prove that $w = Dw'$ or Gw' for some two-sided infinite word w' having property (M). We do that by assuming that w has not the factor aa, the other case being similar. Then each a in w follows some b; we may therefore write w as an infinite product of ba and b, and this factorization is unique; thus $w = Dw'$ for some two-sided infinite word w' (recall that $D(a) = ba$, $D(b) = b$). Note that w' is unique. For use below, we make the following remark: each b in w' corresponds to a unique b in w, and the latter must be followed by another b in w, since $D(a)$ and $D(b)$ both begin by b.

By contradiction, suppose that w' has not property (M). Then either

(1) $w' = u'amab\tilde{m}bv'$

or

(2) $w' = u'bmba\tilde{m}av'$.

These two cases are not symmetrical, since D is not. We may write

$$m = b^{i_n}ab^{i_{n-1}} \dots b^{i_1}ab^{i_0}.$$

In case (1), w' contains the factor

$$amab\tilde{m}b = ab^{i_n}ab^{i_{n-1}} \dots b^{i_1}ab^{i_0}abb^{i_0}ab^{i_1} \dots b^{i_{n-1}}ab^{i_n}b$$

and w contains its image under D, that is

$$bab^{i_n}bab^{i_{n-1}}\dots b^{i_1}bab^{i_0}babb^{i_0}bab^{i_1}\dots b^{i_{n-1}}bab^{i_n}b.$$

By the above remark, this last b in w is followed by some b. Hence we reach a contradiction to property (M), since w contains

$$ab^{i_n+1}a\dots b^{i_1+1}ab^{i_0+1}abb^{i_0+1}ab^{i_1+1}\dots ab^{i_n+1}b = apab\tilde{p}b.$$

Consider case (2): then w' contains the factor $bb^{i_n}a\dots b^{i_1}ab^{i_0}bab^{i_0}ab^{i_1}\dots ab^{i_n}a$ and we see directly a contradiction since w contains $bb^{i_n}ba\dots b^{i_1}bab^{i_0}bba\, b^{i_0}bab^{i_1}\dots\ bab^{i_n}ba = bpba\tilde{p}a$.

b. Now, we prove by induction on n that for each factor of length n of a word w having property (M), this factor is balanced. This will prove the theorem. So let u be a factor of length n of some two-sided infinite word w having property (M). We know by 2a. that $w = Dw'$ or $w = Gw'$ for some word having property (M).

We consider only the case $w = Dw'$, the other being similar. If u does not begin by letter a, then $u = Du'$, where u' is a uniquely defined factor of w; then either $u = b^n$ is balanced, or u has some a and u' is shorter than u, which implies by induction that u' is balanced, and so is $u = Du'$. If u begins by letter a, then either $u = ab^{n-1}$ and u is balanced by inspection; or u has a second letter a, so that $bu = Du'$ for some u', which is shorter than u; u' is a factor of w', since bu is a factor of w, because each letter a in w follows some letter b, since $w \in Im(D)$; hence, by induction, u' is balanced and so is bu, hence u. □

Proof of Lemma. Suppose that w has the two factors aa and bb. Choose such factors closest possible. Then $w = uaa(ba)^k bbv$ (for example). If $k = 0$, w contains $aabb$ and has not property (M). Hence $k \geq 1$. Thus $w = uaab(ab)^k bv$. Write $ua = u'\,{}^i(ba)$ with maximal i, which may be infinite.

Suppose that $i = 0$. Then u ends with a and w contains $aaabab$, which contradicts property (M). Suppose that $1 \leq i < k$. Then w contains ${}^i(ba)ab(ab)^i ab$, hence w contains either $b\,{}^i(ba)ab(ab)^i ab$ or $a\,{}^i(ba)ab(ab)^i ab$. In the first case, w contains $bb(ab)^{i-1}aa$, which contradicts the minimality of k; in the second case, w contains by property (M) also the word $ba\,{}^i(ba)ab(ab)^i\, ab$, which contradicts the maximality of i. Suppose that $i \geq k$. Then w contains ${}^k(ba)ab(ab)^k b$, hence by (M) also $b\,{}^k(ba)ab(ab)^k b$; thus w contains $bb(ab)^{k-1}aa$, which contradicts the minimality of k. □

4. Markoff's property refined

We define 4 properties (M_i), $i = 1, 2, 3, 4$, which refine Markoff's property (M), which are mutually exclusive, and whose union is (M).

We say that w (a two-sided infinite word on $\{a, b\}$) has *property* (M_1) if there exists an integer N such that for any factorization $w = uxyv$, with $\{x, y\} = \{a, b\}$, there exists a finite word m of length $\leq N$ such that $u = u'ym$, $v = \tilde{m}xv'$ for some words u', v'. For given w having property (M_1), we denote by $N(w)$ the smallest possible N in the definition above (if no factorization $w = uxyv$ exists, that is, if $w = {}^{\infty}a^{\infty}$ or ${}^{\infty}b^{\infty}$, $N(w) = -1$). As examples, ${}^{\infty}(ab)^{\infty}$ and ${}^{\infty}(aab)^{\infty}$ have property (M_1); for instance, $w = {}^{\infty}(aab)^{\infty}$ has the factorization ${}^{\infty}(aab)\underline{a}a\underline{b}\,\underline{a}a\underline{b}(aab)^{\infty}$ and $N(w) = 1$, as the reader may verify.

Now, we say that w has *property* (M_2) if w has not property (M_1) and if for any factorization $w = uxyv$ as above, there is a finite word m such that $u = u'ym$, $v = \tilde{m}xv'$ (contrarily to (M_1), we do not require that $|m|$ be bounded, and in fact, arbitrarily long m must exist for the given w).

We say that w has property (M_3) if w has property (M) and has a unique factorization $w = uxy\tilde{u}$ with $\{x, y\} = \{a, b\}$. Note that "unique" here means the following: if w is represented by $\ldots a_{-2}a_{-1}a_0a_1a_2\ldots$ there is a unique i in $\mathbb{Z}$ such that $a_i \neq a_{i+1}$ and that $\ldots a_{i-2}a_{i-1}$ is the reversal of $a_{i+2}a_{i+3}\ldots$. Examples for (M_2) and (M_3) may not be immediately produced, and will be given below.

Finally, we say that w has property (M_4) if w has property (M), but none of the properties (M_1), (M_2), (M_3). In other words, w has property (M), and has at least two distinct factorizations $w = uxy\tilde{u}$, $\{x, y\} = \{a, b\}$. For example, $w = {}^{\infty}ab\,a^{\infty}$ has property (M_4), since it has the two factorizations ${}^{\infty}a\ ab\ a^{\infty}$ and ${}^{\infty}a\ ba\ a^{\infty}$ which are of the form $u\,xy\,\tilde{u}$.

5. The Morse-Hedlund classification

We know by Th.3.1 that property (M) is equivalent to the balance property of [MH]. In this article, Morse and Hedlund give a classification into 3 classes of balanced two-sided infinite words, which they call "periodic", "Sturmian" and "skew". According to their work, the Sturmian case is naturally divided into 2 subcases. We then obtain 4 classes, which will be shown to coincide with the classes defined by the refined Markoff's properties.

Let us give the definitions of Morse and Hedlund. Let w be a balanced two-sided infinite word on $\{a, b\}$. Then w has *property* (MH_1) if it is periodic. It has *property* (MH_4) if it is ultimately periodic at the left and at the right, without being periodic (this case is called the *skew case* by Morse and Hedlund). In order to define (MH_2) and (MH_3), we need some definitions.

We consider lattice paths in the plane $\mathbb{R}^2$, whose steps are segments $[(i, j), (i+1, j)]$ or $[(i, j), (i, j+1)]$, and whose coordinates never decrease. Such a path is encoded by a word on $\{a, b\}$, where a, b correspond respectively to the above steps. Now, to each line in $\mathbb{R}^2$ with irrational slope, associate the two-sided infinite word on $\{a, b\}$ which encodes the path which is either above the line, or below it (it may touch the line), and such that no lattice point lies in the interior of the figure delimited by the path and the line: the path above defines the *upper Christoffel* *word* associated to the line, and the path below defines the *lower Christoffel word*. See Figure 1.

It follows from the work of Morse and Hedlund [MH] that Christoffel words are balanced. We say that w has property (MH_2) if it is an upper or lower Christoffel word corresponding to a line which contains no lattice point; and that w has property (MH_3) if it is an upper or lower Christoffel word corresponding to a line which contains a lattice point (necessarily unique, since the line has irrational slope).

Remarks. 1. Christoffel words (associated to lines with irrational slope) correspond to the case called Sturmian by Morse and Hedlund.
2. The same construction, but with lines having rational slopes, corresponds to the periodic case, as it follows from the work of Morse and Hedlund.

For more on balanced two-sided infinite words, the reader may also read the nice thesis [H].

6. Equivalence of the two classifications

Theorem 6.1. $(M_i) \Leftrightarrow (MH_i)$

We prove the theorem by showing in the 5 lemmas below that (MH_i) implies (M_i). This will suffice since the properties (M_i) are mutually disjoint, and since we already have Th.3.1.

Lemma 6.1. *If w is a two-sided periodic infinite balanced word, then it has property (M_1).*

Proof. Let $\{x, y\} = \{a, b\}$ and suppose that $w = u\,xy\,v$. Since w is periodic, we may consider the periodic pattern which ends at x and begins at y; thus ux, yv may be written $ux = {}^{\infty}(ypx)ypx$, $ypx\,(ypx)^{\infty} = yv$, where p is a finite word. Now w has property (M) by Th.3.1. Hence necessarily $u = u'ym$, $v = \tilde{m}xv'$ with $|m| \le |p| \le$ period of w minus 2. □

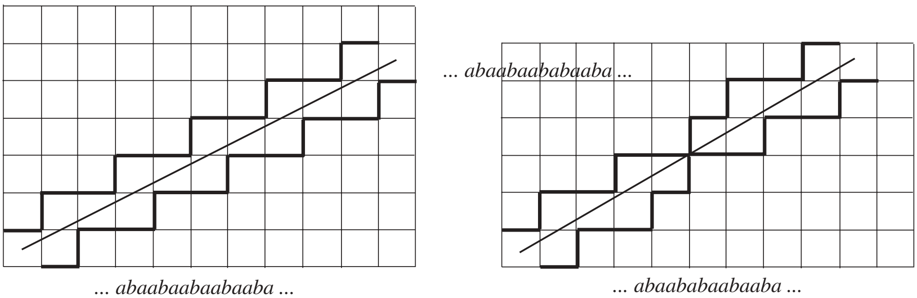

Fig. 1. Upper and lower Christoffel words

Lemma 6.2. *If w has property (M_1), then w is periodic.*

Proof. We claim that if $w = Dw'$ or $w = Gw'$, where G, D have been defined in Section 3, and if w has property (M_1), then w' has property (M_1), and if moreover $w \neq {}^{\infty}a^{\infty}$ and $w \neq {}^{\infty}b^{\infty}$, then $N(w') < N(w)$. Taking this claim for granted, we prove the lemma by induction on $N(w)$. If $N(w) = -1$, then $w = {}^{\infty}x^{\infty}$, $x \in \{a, b\}$, and we are done. If $N(w) \geq 0$, then we may by Lemma 3.1 find w' such that $w = Gw'$ or Dw'. Then $N(w') < N(w)$ by the claim, and we conclude, since w' is periodic by induction, hence w too.

Let us prove the claim. We may suppose that $w = Dw'$, the case $w = Gw'$ being symmetric. If w' has no factor ab or ba, then $w' = {}^{\infty}a^{\infty}$ or ${}^{\infty}b^{\infty}$, hence is periodic; if moreover w is not of this form, we have $-1 = N(w') < N(w)$. Let now w' have some factorization $w' = u'bm'ab\tilde{m}'av'$, with a finite m'; then (recalling that $D(a) = ba$, $D(b) = b$) $w = Dw' = D(u')bD(m')babD(\tilde{m}')baD(v')$; since $(D(m')b)^{\sim} = D(\tilde{m}')b$ (as the reader may easily verify), we have $w = \ldots bmab\tilde{m}a\ldots$, with $m = D(m')b$ longer than m'. If on the other hand, w' has a factorization $w' = uam'ba\tilde{m}'bv$, then $w = \ldots baD(m')bbaD(\tilde{m}')bb\ldots$ (the last b since $D(v)$ begins by b), hence $w = \ldots amba\tilde{m}b\ldots$, with $m = D(m')b$ and thus $\tilde{m} = D(\tilde{m}')b$; again m' is shorter than m. It remains to show that w' has no factorization $w' = uba\tilde{u}$ or $uab\tilde{u}$; but in the first case $w = D(u)bbaD(\tilde{u})$, and in the second $w = D(u)babD(\tilde{u})$, and we reach a contradiction since $(D(u)b)^{\sim} = D(\tilde{u})$ for any left infinite word u. Thus w' satisfies (M_1) and $N(w') < N(w)$. □

Lemma 6.3. *If w satisfies (MH_2), then w has property (M_2).*

Proof. Given two parallel lines in the plane, we call *strip* the closed set lying between them. Its *width* is the length of the segment obtained by intersection of the two lines with the line $x + y = 0$. Consider a strip of width $\sqrt{2}$. Suppose it has a central symmetry around a point of the form $(x, y) + \left(\frac{1}{2}, \frac{1}{2}\right)$, where $x, y \in \mathbb{Z}$. Then the two lines contain the points $(x + 1, y)$ and $(x, y + 1)$.

Now consider a path in the plane corresponding to a Christoffel word w. It is well-known that w is contained in a unique strip of width $\sqrt{2}$, which is the convex hull of the path. One of the two lines bordering the strip is the line defining the Christoffel word w. Moreover, by density, the strip is completely determined by any right (or left) infinite part of the path.

Suppose now that w is a Christoffel word and has a factorization $w = uxy\tilde{u}$. Then we conclude that the strip has a central symmetry as above. Hence the line defining the Christoffel word w has an integer point and w satisfies (MH_3), and not (MH_2). This shows that if w has property (MH_2), then w has not property (M_3) or (M_4), hence has property (M_1) or (M_2). Suppose by contradiction that w has property (M_1); then w is periodic by Lemma 6.2. This contradicts the fact that, by Morse and Hedlund, w is not periodic. □

Lemma 6.4. *If w satisfies (MH_3), then it has a unique factorization $uxy\tilde{u}$, $\{x, y\} = \{a, b\}$.*

Proof. If w satisfies (MH_3), then w is the upper or lower Christoffel word as in Figure 1, right part. We see that these two words are of the form $uabv$, $ubav$. By central symmetry under the point represented on the figure, we see that $\tilde{u} = v$. Hence the two words are $uab\tilde{u}$ and $uba\tilde{u}$.

Now, suppose by contradiction that w has two such factorizations. Then w has two partial central symmetries (partial meaning here that these symmetries are defined for all but a finite number of letters of w). Since the product of two symmetries is a translation, we see that w is preserved by some partial translation; hence w is ultimately periodic on the left and on the right. Thus, some Sturmian sequence is periodic, a contradiction. This proves unicity. □

Lemma 6.5. *If w satisfies (MH_4), then it has at least two factorizations $w = u\,xy\,\tilde{u}$, with $\{x, y\} = \{a, b\}$.*

We need to define finite Christoffel words. These are defined similarly to infinite Christoffel words, but by replacing an infinite line by a finite segment from (0, 0) to (p, q), where p, q are relatively prime nonnegative integers. See [BL] or [BS] for a formal definition, and Figure 2 where is drawn a lower Christoffel word.

Proof. According to [MH] Th 5.2 (see also [H] Th 2.5), a two-sided infinite word w having property (MH_4) is either ${}^{\infty}a\,b\,a^{\infty}$ or ${}^{\infty}b\,a\,b^{\infty}$, or is of the following form: there exists some finite lower Christoffel word amb such that $w = {}^{\infty}(bma)\,bmb\,(amb)^{\infty}$ or $w = {}^{\infty}(amb)\,ama\,(bma)^{\infty}$. Now, m is a palindrome, as is well-known. Hence w has the two required factorizations. □

Remark. The proof of Lemma 6.1 shows that if w has property (M_1), then $N(w)$ is $\leq$ to the period of w minus 2. Actually, we have equality. Indeed,it follows from the theory of Sturmian sequences (see [BS]) that in the case $(M_1) \Leftrightarrow (MH_1)$, w is equal to ${}^{\infty}u^{\infty}$, for some lower finite Christoffel word u; moreover (disregarding the cases $w = {}^{\infty}a^{\infty}$ or ${}^{\infty}b^{\infty}$), $u = amb$, for some palindrome m. Hence $w = {}^{\infty}(amb)ambamb(amb)^{\infty}$ and this factorization shows that $N(w) \geq |m| =$ period minus 2.

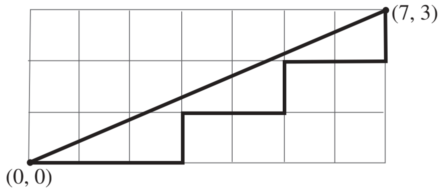

Fig. 2. A lower finite Christoffel word

7. Two-sided continued fractions

We consider a two-sided infinite sequence $A = \dots a_{-1}a_0a_1a_2\dots$, where the a_i are positive integers. Following [P] and [CF] define $\lambda_i(A) = a_i + [0, a_{i+1}, a_{i+2}, \dots] + [0, a_{i-1}, a_{i-2}, \dots]$, where $[a, b, c, \dots]$ denotes the real number whose continued fraction expansion is $a, b, c, \dots$ If B is an ordinary sequence $b_0, b_1, b_2, \dots$, we denote by $[B]$ the real number whose expansion into continued fraction is B.

Examples. $A = \cdots 111 \cdots$, $\lambda_i(A) = 1 + 2\frac{\sqrt{5}-1}{2} = \sqrt{5}$ for every i.
$B = \cdots 222 \cdots$, $\lambda_i(A) = \sqrt{8}$ for every i.
Define $M(A) = \sup_i \lambda_i(A)$.

Theorem 7.1. *(Markoff) (i) If $M(A) \leq 3$, then $a_i = 1$ or 2 for each i and the maximal blocks of 1 and 2 in A are of even length.*
(ii) Assuming that the maximal blocks of 1 and 2 in A are of even length, let w be the two-sided infinite word mapped onto A under the substitution $1 \mapsto 11, 2 \mapsto 22$. Then: $M(A) \leq 3 \Leftrightarrow w$ satisfies property (M).

Proof. (i) Suppose that $M(A) \leq 3$. Since $a_i < \lambda_i(A)$, we deduce that $a_i = 1$ or 2. Suppose that A has a block of 1's of odd length n. By part (i) of Lemma 8.1 in the Appendix, $n \geq 3$. Then A has the factor 21^n2, which we extend to the left. By Lemma 8.1 (ii), A has not the factor 222111 and by (i), no isolated 2; hence A has the factor 1221^n2. There are two cases:

- $1^{n-1}22111^{n-2}2$ is a factor of A: this contradicts Lemma 8.1 (ii) since $n-1$ is even. Note indeed that $[B] < [C]$ is equivalent to $b_1 < c_1$, or $b_1 = c_1$ and $b_2 > c_2$, or $b_1 = c_1$ and $b_2 = c_2$ and $b_3 < c_3, \dots$
- A has a factor 21^p221^n2 where $p < n-1$ must be even by induction, hence $p \leq n-3$. Then A has the factor 21^p22111^{p+1} and this contradicts Lemma 8.1 (ii) since p is even.

If A has a factor 12^n1, n odd, we argue similarly by extending it to the right.
(ii) Suppose that $M(A) \leq 3$. A factor 21 in w corresponds to a factor 2211 in A. Using Lemma 8.1 (ii), we see that A satisfies property (M) of Section 3 for this occurrence of the factor 21. For the factor 12, we argue similarly since, for the mirror sequence $\tilde{A}$, $\lambda_i(A) = \lambda_{-i}(\tilde{A})$, hence $M(A) = M(\tilde{A})$.

Conversely, suppose that w satisfies (M). Let $i \in \mathbb{Z}$. Then we must be in one of the 6 situations described in Lemma 8.1 (ii), (iii), (iv). Cases (iii) and (iv) imply that $\lambda_i(A) < 3$. For (ii) we note that if $A = \tilde{B}2\underline{2}11C$, then property (M) on w implies that either $B = C$ and $\lambda_i(A) = 3$, or $[B] < [C]$ since the blocks are of even length, hence $\lambda_i(A) < 3$. Thus $M(A) \leq 3$. □

We may now translate the classification of Section 6 in analytical terms.

Theorem 7.2. *Let w be a two-sided infinite word on $\{1, 2\}$ and A obtained from w under the substitution $\sigma : 1 \mapsto 11, 2 \mapsto 22$. Then w is balanced if and only if $M(A) \leq 3$. Suppose now that w is balanced; then:*

(i) *w has property $(M_1) \Leftrightarrow M(A) < 3$;*
(ii) *w has property $(M_2) \Leftrightarrow M(A) = 3$ and $\lambda_i(A) < 3$ for any i;*
(iii) *w has property $(M_3) \Leftrightarrow \lambda_i(A) = 3$ for a unique i;*
(iv) *w has property $(M_4) \Leftrightarrow \lambda_i(A) = 3$ for at least two values i.*

Proof. The first assertion is a consequence of Th. 3.1 and Th. 7.1. We assume now that w is balanced. If w satisfies property (M_1), then w is periodic. Hence A is periodic and $\lambda_i(A)$ takes only finitely many values; so we have to show that none of them is equal to 3. But this is handled exactly as at the end of the proof of Th. 7.1, knowing by (MH_1) that w has no factorization $w = u21\tilde{u}$ or $w = u12\tilde{u}$. Assume now that w satisfies (M_2). Then the same argument shows that $\lambda_i(A) \neq 3$ for any i. We show that however $M(A) = 3$. Indeed, w has factorizations $w = u\tilde{x}21xv$ or $w = u\tilde{x}12xv$ with arbitrarily long finite words x. Thus, without loss of generality, we may assume that for each n in $\mathbb{N}$, we can choose x_n such that $w = u_n\tilde{x}_n21x_nv_n$ where each x_n is a proper prefix of x_{n+1} (by finiteness of the alphabet $\{1, 2\}$). Then x_n tends to a right infinite word x. Then for each n we have a corresponding factorization $A = U_n\tilde{X}_n2211X_nV_n$ (with $U_n = \sigma(u_n)$ etc...) and choosing for i_n the second 2, we obtain $\lambda_{i_n}(A) = 2 + [0, 1, 1, X_n, V_n] + [0, 2, X_n, \tilde{U}_n]$. By definition $X := \lim_{n\to\infty} X_n$ exists. Hence $\lim_{n\to\infty} \lambda_{i_n}(A) = 2 + [0, 1, 1, X] + [0, 2, X]$. We now apply Lemma 8.1 (ii) to conclude that $\lim_{n\to\infty} \lambda_{i_n}(A) = \lambda_j(\tilde{X}2\underline{2}11X) = 3$, where j corresponds to the underlined 2. Thus $M(A) = 3$.

We now assume that w satisfies (M_3). Then w has a unique factorization $w = u21\tilde{u}$ or $u12\tilde{u}$. Lemma 8.1 shows that there is a unique i such that $\lambda_i(A) = 3$. Finally, if w has property (M_4), then w has two such factorizations and we argue similarly. □

8. Appendix: technical results

Lemma 8.1. *Let A be a two-sided infinite word on $\{1, 2, 3, \dots\}$ and i the index corresponding to the underlined letter.*

(i) *If $A = \dots 1\underline{2}1 \dots$ or $\dots 21\underline{2} \dots$, then $\lambda_i(A) > 3$.*
(ii) *If $A = \tilde{B}2\underline{2}11C$ or $A = \tilde{C}11\underline{2}2B$, then: $\lambda_i(A) \leq 3 \Leftrightarrow [B] \leq [C]$ and: $\lambda_i(A) = 3 \Leftrightarrow B = C$.*
(iii) *If $A = \dots 22\underline{1}1 \dots$ or $A = \dots 1\underline{1}22 \dots$, then $\lambda_i(A) < 3$.*
(iv) *If $A = \dots 1\underline{1}1 \dots$ or $A = \dots 2\underline{2}2 \dots$ then $\lambda_i(A) < 3$.*

These results are all implicit in Markoff's work; see also [D], [CF]. They are all obtained by straightforward computations, and we omit the proofs.

References

[A] Arnoux, P.: Sturmian sequences. In: N. Pytheas-Fogg, *Substitutions in Dynamics, Arithmetics and Combinatorics*, Lecture Notes in Mathematics 1794, Springer, Berlin Heidelberg, 2002, pp. 143–198

[AS] Allouche, J.-P., Shallit, J.: Automatic sequences. Cambridge: Cambridge University Press, 2003

[B] Bernoulli, J.: Sur une nouvelle espèce de calcul. Recueil pour les Astronomes I, 255–284, Berlin, 1772

[BL] Borel, J.-P., Laubie, F.: Quelques mots sur la droite projective réelle. Journal de Thérorie des Nombres de Bordeaux **5**, 23–51 (1993)

[BS] Berstel, J., Séébold, P.: Sturmian words, in: M. Lothaire, Algebraic Combinatorics on words. Cambridge: Cambridge University Press, 2002, Chapter 2

[C] Cassaigne, J.: Limit values of the recurrence quotient of Sturmian sequences. Theoretical Computer Science **218**, 3–12 (1999)

[CF] Cusick, T.W., Flahive, M.E.: The Markoff and Lagrange spectra. American Mathematical Society, 1989

[C1] Christoffel, E.B.: Observatio arithmetica. Annali di Matematica **6**, 145–152 (1875)

[C2] Christoffel, E.B.: Lehrsätze über arithmetische Eigenschaften du Irrationnalzahlen. Annali di Matematica Pura ed Applicata, Series II 15, 1888, pp. 253–276

[D] Dickson, L.E.: Studies in the theory of numbers. Chelsea, New York, 1930 (second edition 1957)

[H] Heinis, A.: Arithmetics and Combinatorics of Words of Low Complexity. Ph D, University of Leiden, 2001

[M1] Markoff, A.A.: Sur les formes quadratiques binaires indéfinies. Math. Ann. **15**, 381–496 (1879)

[M2] Markoff, A.A.: Sur les formes quadratiques binaires indéfinies (second mémoire). Math. Ann. **17**, 379–399 (1880)

[M3] Markoff, A.A.: Sur une question de Jean Bernouilli. Math. Ann. **19**, 27–36 (1882)

[MH] Morse, M., Hedlund, G.A.: Symbolic dynamics II: Sturmian Trajectories. Amer. J. Math. **62**, 1–42 (1940)

[P] Perron, O.: Über die Approximationen irrationaler Zahlen durch rationale. Sitzungsberichte der Heidelberger Akademie der Wissenschaften, Abhandlungen, **8**, 1921

[S] Smith, H.J.S.: Note on continued fractions. Messenger Math. **6**, 1–14 (1876)

[Y] Yasutomi, S.-I.: The continued fraction expansion of α with $\mu(\alpha) = 3$. Acta Arithmetica **84**, 337–374 (1998)