# Palindromic lacunas of the Thue-Morse word*

A. Blondin-Massé, S. Brlek†, S. Labbé

Laboratoire de Combinatoire et d'Informatique Mathématique,
Université du Québec à Montréal,
C. P. 8888 Succursale "Centre-Ville", Montréal (QC), CANADA H3C 3P8
blondin_masse.alexandre@courrier.uqam.ca,
Brlek.Srecko@uqam.ca,Labbe.Sebastien@courrier.uqam.ca

**Abstract**

The palindromic defect of a finite word $w$ has been defined by $D(w) = |w|+1-|\mathrm{Pal}(w)|$ where $\mathrm{Pal}(w)$ is the set of its palindromic factors. In this paper we study the problem of computing the palindromic defect of finite and infinite words. Moreover, we describe completely the lacunas (the positions where the longest palindromic suffix is not uni-occurrent) in the Thue-Morse word, showing that there exist infinite words with infinitely many palindromes but infinite defect.

**Keywords** Palindromic complexity, defect, Thue-Morse, lacunary words.

## 1 Introduction

Among the many ways of measuring the information content of a finite word, counting the number of distinct factors of given length occurring in it has been widely used and known as its complexity. A refinement of this notion amounts to restrict the factors to palindromes. The motivations for the study of palindromic complexity comes from many areas ranging from the study of Schrödinger operators in physics [1, 4, 12] to number theory [3] and combinatorics on words where it appears as a powerful tool for understanding the local structure of words. It has been recently studied in various classes of infinite words, an account of which may be found in the survey provided by Allouche et al. [2].

In particular, the palindromic factors give an insight on the intrinsic structure, due to its connection with the usual complexity, of many classes of words. For instance, they completely characterize Sturmian words [15], and for the class of smooth words they provide a connection with the notion of recurrence [7, 8]. Droubay, Justin and Pirillo [10] noted that the palindrome complexity $|\mathrm{Pal}(w)|$ of a word $w$ is bounded by $|w| + 1$, and that finite Sturmian (and even episturmian) words realize the upper bound. Moreover they show that the palindrome

*with the support of NSERC (Canada)

†Corresponding author.

complexity is computed by a linear algorithm listing the longest palindromic suffixes that are uni-occurrent.

Words that realize that bound are called full in Brlek et al. [6]. In that paper it is also shown that there exist periodic full words, and an optimal algorithm is provided to check if an infinite periodic word is full or not. Moreover, a characterization by means of a rational language is given for the language $L_P$ of words whose palindromic factors belong to a fixed and finite set $P$ of palindromes. A finite automaton recognizing $L_P$ is then easy to obtain, and consequently, if there exists a recurrent infinite word having $P$ for palindromic factors, then there exist a periodic one sharing exactly the same palindromic factors. Furthermore, by using a representation of periodic words by circular words, a geometric characterization of those having an infinite set of palindromic factors, as well as those having a finite one, is furnished. More precisely, an infinite periodic word $\mathbf{w} = w^{\omega}$, where $w$ is primitive, has an infinite set of palindromes if and only if $w$ is the product of two palindromes. Consequently the periodic words having a finite set of palindromes are the words whose smallest periodic pattern is *asymmetric*. An enumeration formula for asymmetric words was also provided. Still in [6], the authors considered the defect $D(w) = |w| + 1 - |\mathrm{Pal}(w)|$ of finite words, and showed in particular that for any $k \in \mathbb{N}$ there exist periodic words having finite defect $k$.

In this paper, we study in more details the palindromic defect of infinite words. For that purpose, in Section 2 we recall from Lothaire's book [13] the basic terminology on words. Section 3 is devoted to the study of conjugate morphisms and contains some technical lemmas about their fixed points and palindromic structure. In Section 4 we borrow from Brlek et al. [6] the necessary definitions and results about the palindromic defect, full words and constructions on periodic infinite words. In section 5, we study the lacunas of words that are not full, and show in Section 6 that the Thue-Morse word $\mathbf{t}$ is lacunary. Moreover we give an explicit formula for lacunas of $\mathbf{t}$.

## 2 Definitions and notation

We borrow from M. Lothaire [14] the basic terminology about words. In what follows, $\Sigma$ is a finite *alphabet* whose elements are called *letters*. By *word* we mean a finite sequence of letters $w : [0..n-1] \longrightarrow \Sigma$, where $n \in \mathbb{N}$. The length of $w$ is $|w| = n$ and $w[i]$ or $w_i$ denote its $i$-th letter. The set of $n$-length words over $\Sigma$ is denoted $\Sigma^n$. By convention, the *empty* word is denoted $\varepsilon$ and its length is 0. The free monoid generated by $\Sigma$ is defined by $\Sigma^* = \bigcup_{n\geq 0} \Sigma^n$. The set of right infinite words is denoted by $\Sigma^{\omega}$ and we set $\Sigma^{\infty} = \Sigma^* \cup \Sigma^{\omega}$. Given a word $w \in \Sigma^{\infty}$, a *factor* $f$ of $w$ is a word $f \in \Sigma^*$ satisfying

$$\exists x \in \Sigma^*, \exists y \in \Sigma^{\infty}, w = xfy.$$

If $x = \varepsilon$ (resp. $y = \varepsilon$ ) then $f$ is called *prefix* (resp. *suffix*). The set of all factors of $w$ is denoted by $\mathrm{Fact}(w)$, those of length $n$ is $\mathrm{Fact}_n(w) = \mathrm{Fact}(w) \cap \Sigma^n$, and $\mathrm{Pref}(w)$ is the set of all prefixes of $w$. The number of occurrences of a factor $f \in \Sigma^*$ is $|w|_f$. A *period* of a word $w$ is an integer $p < |w|$ such that $w[i] = w[i+p]$, for all $i < |w| - p$. An infinite word $w$ is *recurrent* if it satisfies the condition $u \in \mathrm{Fact}(w) \implies |w|_u = \infty$ . If $w = pu$, with $|w| = n$ and $|p| = k$, then $p^{-1}w = w[k..n-1] = u$ is the word obtained by erasing $p$. A word is said to be *primitive* if it is not a power of another word. Two words $u$ and $v$ are *conjugate* when there are words $x, y$ such that $u = xy$ and $v = yx$. The conjugacy class of a word $w$ is denoted by $(w)$; note that the length is invariant under conjugacy. Moreover, the *reversal* of $u = u_0 u_1 \cdots u_{n-1} \in \Sigma^n$ is the word $\widetilde{u} = u_{n-1} u_{n-2} \cdots u_0$. A *palindrome* is a word $p$ such that $p = \widetilde{p}$, and for a language $L \subseteq \Sigma^\infty$, the set of its palindromic factors is denoted by $\mathrm{Pal}(L)$. Every word contains palindromes, the letters and $\varepsilon$ being necessarily part of them. This justifies the introduction of the function $\mathrm{LPS} : \Sigma^* \longrightarrow \Sigma^*$ which associates to any word $w$ its longest palindromic suffix $\mathrm{LPS}(w)$.

A *morphism* is a function $\varphi : \Sigma^* \longrightarrow \Sigma^*$ compatible with concatenation, that is, such that $\varphi(uv) = \varphi(u)\varphi(v)$ for all $u, v \in \Sigma^*$. A morphism $\varphi$ is primitive if $\forall \alpha \in \Sigma, \varphi^k(\alpha)$ contains each letter of $\Sigma$ for some $k$. For $\alpha \in \Sigma$ we call *$\varphi$-block* (block for short if no confusion arises) a factor of the form $\varphi(\alpha)$. A morphism is called *uniform* when the blocks have equal lengths. Clearly, every infinite fixed point of a primitive morphism is recurrent, and there exist recurrent but non-periodic words, the Thue-Morse word $\mathbf{t}$ [18], and the Sturmian words being some of these. The *mirror-image* of a morphism $\varphi$, denoted by $\widetilde{\varphi}$, is the morphism such that $\widetilde{\varphi}(\alpha) = \widetilde{\varphi(\alpha)}$ for all $\alpha \in \Sigma$. It is easy to check that

$$\widetilde{\varphi(w)} = \widetilde{\varphi}(\widetilde{w}) \quad \text{for all} \quad w \in \Sigma^*; \quad \widetilde{\varphi \circ \mu} = \widetilde{\varphi} \circ \widetilde{\mu}. \tag{1}$$

Recall from Lothaire [14] (Section 2.3.4) that $\varphi$ is *right conjugate* of $\varphi'$, noted $\varphi \triangleleft \varphi'$, if there exists $u \in \Sigma^*$ such that

$$\varphi(\alpha)u = u\varphi'(\alpha), \quad \text{for all } \alpha \in \Sigma, \tag{2}$$

or equivalently that $\varphi(x)u = u\varphi'(x)$, for all words $x \in \Sigma^*$. Clearly, this relation is not symmetric so that we say that two morphisms $\varphi$ and $\varphi'$ are *conjugate*, noted $\varphi \bowtie \varphi'$, if $\varphi \triangleleft \varphi'$ or $\varphi' \triangleleft \varphi$. It is easy to see that conjugacy of morphisms is an equivalence relation.

# 3 Preliminaries

From Lothaire [13] we borrow the following useful result: if $w = xy = yz$, then, for some $u, v$, and some $i \geq 0$ we have

$$x = uv, y = (uv)^i u, z = vu; \tag{3}$$

We start by extending Lemma 5 of [6].

**Lemma 1** *Assume that $w = xy = yz$. Let $u, v$ and $i \geq 0$ be such that Eq. 3 holds. Then the following conditions are equivalent :*

(i) $x = \widetilde{z}$;

(ii) *$u$ and $v$ are palindromes;*

(iii) *$w$ is a palindrome;*

(iv) *$xyz$ is a palindrome.*

*Moreover, if one of the equivalent conditions above holds then*

(v) *$y$ is a palindrome.*

*Proof.* To establish the equivalences we proceed as follows. (i) $\Longrightarrow$ (ii) Since $uv = x = \widetilde{z} = \widetilde{u}\,\widetilde{v}$, we have $u = \widetilde{u}$ and $v = \widetilde{v}$. (ii) $\Longrightarrow$ (iii) We have $w = xy = (uv)^{i+1}u$. Then, $\widetilde{w} = \widetilde{u}(\widetilde{v}\,\widetilde{u})^{i+1} = u(vu)^{i+1} = (uv)^{i+1}u = w$. (iii) $\Longrightarrow$ (iv) We have $xy = w = \widetilde{w} = \widetilde{z}\,\widetilde{y}$. Then, $y = \widetilde{y}$, $x = \widetilde{z}$ and $\widetilde{xyz} = \widetilde{z}\,\widetilde{y}\,\widetilde{x} = xyz$. (iv) $\Longrightarrow$ (i) Since $|x| = |z|$ we have $x = \widetilde{z}$. ∎

Note that condition (v) above is not equivalent to the conditions (i)-(iv) in Lemma 1, as shown by the following example. Let $y = aba$, $x = abaab$ and $z = ababa$. The equality $xy = yz$ holds, $y$ is a palindrome but $x \neq \widetilde{z}$. In fact, the crucial point is the fact that $|y| < |x|$, which corresponds to $i = 0$ in Eq. (3). However, if $|y| \geq |x|$, that is if $i > 0$, then (v) $\Longrightarrow$ (ii): indeed, if $y$ is a palindrome, then $(uv)^i u = y = \widetilde{y} = \widetilde{u}(\widetilde{v}\,\widetilde{u})^i = (\widetilde{u}\,\widetilde{v})^i\widetilde{u}$; hence, $u = \widetilde{u}$ and $v = \widetilde{v}$ since $i > 0$. We have thus proved:

**Lemma 2** *Assume that $w = xy = yz$ with $|y| \geq |x|$. Then, conditions* (i)-(v) *in Lemma 1 are equivalent.*

**Lemma 3** *Let $\varphi$ and $\varphi'$ be two morphisms on $\Sigma = \{a, b\}$ such that $\varphi \triangleleft \varphi'$. i.e. $\varphi(\alpha)u = u\varphi'(\alpha)$ for all $\alpha \in \Sigma$. Let $p = |\varphi(a)|$ and $q = |\varphi(b)|$. If $|u| \geq p + q - \gcd(p, q)$, then $\varphi$ and $\varphi'$ both have periodic fixed points.*

*Proof.* Assume that there exists $u \in \Sigma^*$ such that $\varphi(a)u = u\varphi'(a)$ and $\varphi(b)u = u\varphi'(b)$. From Equations (3), there exist $x, z \in \Sigma^*$, $i \in \mathbb{N}$ such that $\varphi(a) = xz$, $u = (xz)^i x$ and $\varphi'(a) = zx$. Also, there exist $w, y \in \Sigma^*$, $j \in \mathbb{N}$ such that $\varphi(b) = wy$, $u = (wy)^j w$ and $\varphi'(b) = yw$. Then, $(xz)^i x = u = (wy)^j w$ and hence, $p = |\varphi(a)| = |xz|$ and $q = |\varphi(b)| = |wy|$ are periods of $u$. The theorem of Fine and Wilf [14] says that $\gcd(p, q)$ is a period of $u$. It follows that $\varphi(a)$ and $\varphi(b)$ are powers of the same word since they are prefixes of $u$, and $\varphi'(a)$ and $\varphi'(b)$ also since they are suffixes of $u$. ∎

The next Lemma extends the Lemma 2.3.17 of Lothaire [14].

**Lemma 4** *Let $\varphi$, $\varphi'$, $\mu$, $\mu'$ be morphisms. Then the following conditions hold:*

(i) *if $\varphi \triangleleft \varphi'$ and $\mu \triangleleft \mu'$, then $\varphi \circ \mu \triangleleft \varphi' \circ \mu'$.*

(ii) *if $\varphi \triangleleft \varphi'$ and $\mu \triangleleft \mu'$, then $\varphi \circ \mu' \bowtie \varphi' \circ \mu$.*

(iii) *if $\varphi \bowtie \varphi'$ and $\mu \bowtie \mu'$, then $\varphi \circ \mu \bowtie \varphi' \circ \mu'$.*

*Proof.* (i) There exist $u, v \in \Sigma^*$ such that $\varphi(\alpha)u = u\varphi'(\alpha)$ and $\mu(\alpha)v = v\mu'(\alpha)$, for all $\alpha \in \Sigma$. We compute

$$\begin{aligned}\varphi \circ \mu(\alpha) \cdot u\varphi'(v) &= \varphi \circ \mu(\alpha) \cdot \varphi(v)u \\ &= \varphi(\mu(\alpha) \cdot v)u \\ &= u\varphi'(v \cdot \mu'(\alpha)) \\ &= u\varphi'(v) \cdot \varphi' \circ \mu'(\alpha).\end{aligned}$$

(ii) There exist $u, v \in \Sigma^*$ such that $\varphi(\alpha)u = u\varphi'(\alpha)$ and $\mu(\alpha)v = v\mu'(\alpha)$, for all $\alpha \in \Sigma$. If $|u| \leq |\varphi(v)|$, then $u^{-1}\varphi(v) = \varphi'(v)u^{-1}$ and we obtain

$$\begin{aligned}u^{-1}\varphi(v) \cdot \varphi \circ \mu'(\alpha) &= u^{-1}\varphi(v \cdot \mu'(\alpha)) \\ &= \varphi'(\mu(\alpha) \cdot v)u^{-1} \\ &= \varphi' \circ \mu(\alpha) \cdot \varphi'(v)u^{-1} \\ &= \varphi' \circ \mu(\alpha) \cdot u^{-1}\varphi(v),\end{aligned}$$

that is, $\varphi' \circ \mu \triangleleft \varphi \circ \mu'$. If $|u| \geq |\varphi(v)|$, then

$$\begin{aligned}\varphi \circ \mu'(\alpha) \cdot \varphi(v)^{-1}u &= \varphi(v)^{-1}\varphi(v) \cdot \varphi \circ \mu'(\alpha) \cdot uu^{-1} \cdot \varphi(v)^{-1}u \\ &= \varphi(v)^{-1} \cdot \varphi(v\mu'(\alpha))u \cdot u^{-1}\varphi(v)^{-1}u \\ &= \varphi(v)^{-1} \cdot u\varphi'(\mu(\alpha)v) \cdot (\varphi(v)u)^{-1}u \\ &= \varphi(v)^{-1}u \cdot \varphi' \circ \mu(\alpha) \cdot \varphi'(v)(u\varphi'(v))^{-1}u \\ &= \varphi(v)^{-1}u \cdot \varphi' \circ \mu(\alpha),\end{aligned}$$

that is, $\varphi \circ \mu' \triangleleft \varphi' \circ \mu$. (iii) The result follows from (i) and (ii). ∎

**Lemma 5** *Let $\varphi$ and $\varphi'$ be two uniform morphisms. Then $\varphi \triangleleft \varphi'$ if and only if there exists $x$, and for each $\alpha \in \Sigma$ there exists $z_\alpha$ such that*

$$\varphi(\alpha) = xz_\alpha \text{ and } \varphi'(\alpha) = z_\alpha x. \tag{4}$$

*Proof.* ($\Rightarrow$) Assume that there exists $u \in \Sigma^*$ such that for all $\alpha \in \Sigma, \varphi(\alpha)u = u\varphi'(\alpha)$. From Equations (3), there exist $x_\alpha, z_\alpha \in \Sigma^*$, $i_\alpha \in \mathbb{N}$ such that $\varphi(\alpha) = x_\alpha z_\alpha$, $u = (x_\alpha z_\alpha)^{i_\alpha} x_\alpha$ and $\varphi'(\alpha) = z_\alpha x_\alpha$. To show that the choice of $x_\alpha$ is independent from $\alpha$, we proceed as follows. Assume that there are two such letters, say $\alpha$ and $\alpha'$. We have for some integers $i$ and $i'$,

$$(x_\alpha z_\alpha)^i x_\alpha = u = (x_{\alpha'} z_{\alpha'})^{i'} x_{\alpha'}.$$

Since the morphisms are uniform, it follows that $|x_\alpha z_\alpha| = |x_{\alpha'} z_{\alpha'}|$, so that $i = i'$, $|x_\alpha| = |x_{\alpha'}|$, and hence $x_\alpha = x_{\alpha'}$.

($\Leftarrow$) We have $\varphi(\alpha)x = xz_\alpha x = x\varphi'(\alpha)$. ∎

The lemma above does not hold for some non-uniform morphisms. Consider the conjugate morphisms $\varphi_1 : a \mapsto abaab, b \mapsto ab$, $\varphi_2 : a \mapsto baaba, b \mapsto ba$, $\varphi_3 : a \mapsto aabab, b \mapsto ab$ and $\varphi_4 : a \mapsto ababa, b \mapsto ba$. The pairs $(\varphi_1, \varphi_3)$ and $(\varphi_3, \varphi_4)$ both satisfy Equation 4, but the pair $(\varphi_1, \varphi_4)$ does not. That is why the following result and especially its corollary are more general than Lemma 3 in [2].

**Proposition 1** *Let $\varphi$ and $\varphi'$ be two primitive morphisms, and let* $\mathbf{u} = \varphi(\mathbf{u})$*,* $\mathbf{v} = \varphi'(\mathbf{v})$ *be two respective fixed points. The following properties hold :*

(i) *If* $\varphi \triangleleft \varphi'$*, then* $\mathrm{Fact}(\mathbf{u}) = \mathrm{Fact}(\mathbf{v})$.

(ii) *If* $\varphi' = \widetilde{\varphi}$*, then* $\mathrm{Pal}(\mathbf{u}) = \mathrm{Pal}(\mathbf{v})$.

*Proof.* (i) By hypothesis and from Lemma 4 (i), for all $k$ there exists $w$ such that $\varphi^k(\alpha)w = w\varphi'^k(\alpha)$ for all $\alpha \in \Sigma$. First, we show that $\varphi^k(\alpha)w$ is a factor of both $\mathbf{u}$ and $\mathbf{v}$. Since $\varphi$ is primitive, there exist $x \in \Sigma^*, \mathbf{z} \in \Sigma^\omega$ such that $\mathbf{u} = x\alpha\mathbf{z}$. We have

$$\mathbf{u} = \varphi^k(x) \cdot \varphi^k(\alpha) \cdot \varphi^k(\mathbf{z}) = \varphi^k(x) \cdot \varphi^k(\alpha) \cdot w{\varphi'}^k(\mathbf{z}).$$

Therefore, $\varphi^k(\alpha)w$ is factor of $\mathbf{u}$. Similarly, there exist $x \in \Sigma^*, \mathbf{z} \in \Sigma^\omega$ such that $\mathbf{v} = x\alpha\mathbf{z}$. Then we have

$$w\mathbf{v} = w{\varphi'}^k(x) \cdot {\varphi'}^k(\alpha) \cdot {\varphi'}^k(\mathbf{z}) = \varphi^k(x)w \cdot {\varphi'}^k(\alpha) \cdot {\varphi'}^k(\mathbf{z}).$$

We can choose $x$ long enough in order to satisfy $|\varphi^k(x)| \geq |w|$. Therefore, $w{\varphi'}^k(\alpha)$ is factor of $\mathbf{v}$. Finally, let $f \in \mathrm{Fact}(\mathbf{u})$. Then there exists $n$ such that $f \in \varphi^n(\mathbf{u}_0)$. Therefore, $f \in \mathrm{Fact}(\mathbf{v})$. By the same argument, factors of $\mathbf{v}$ occur in $\mathbf{u}$.

(ii) We have $f \in \mathrm{Fact}(\mathbf{u})$, that is, $f \in \mathrm{Fact}(\varphi^n(\alpha))$ for some $n \in \mathbb{N}$ and $\alpha \in \Sigma$, or equivalently $\widetilde{f} \in \mathrm{Fact}(\widetilde{\varphi^n(\alpha)}) = \mathrm{Fact}(\widetilde{\varphi}^n(\alpha))$, that is $\widetilde{f} \in \mathrm{Fact}(\mathbf{v})$. It follows that $\mathrm{Pal}(\mathbf{u}) = \mathrm{Pal}(\mathbf{v})$. ■

There are morphisms, like $\varphi : a \mapsto ba, b \mapsto ab$, having no fixed point. That is why the following result is useful.

**Corollary 1** *Let $\varphi$ and $\varphi'$ be two primitive morphisms and suppose there are integers $k, l$ such that $\varphi^k$ and $\varphi'^l$ have* $\mathbf{u} = \varphi^k(\mathbf{u})$ *and* $\mathbf{v} = \varphi'^l(\mathbf{v})$ *as fixed points. Then the following properties hold :*

(i) *if* $\varphi \bowtie \varphi'$*, then* $\mathrm{Fact}(\mathbf{u}) = \mathrm{Fact}(\mathbf{v})$.

(ii) *if* $\varphi' = \widetilde{\varphi}$*, then* $\mathrm{Pal}(\mathbf{u}) = \mathrm{Pal}(\mathbf{v})$.

*Proof.* If $\varphi \bowtie \varphi'$, then $\varphi^{kl} \bowtie \varphi'^{kl}$ from Lemma 4. If $\varphi' = \widetilde{\varphi}$, then $\varphi'^{kl} = \widetilde{\varphi}^{kl} = \widetilde{\varphi^{kl}}$. Since $\mathbf{u} = \varphi^{kl}(\mathbf{u})$ and $\mathbf{v} = \varphi'^{kl}(\mathbf{v})$, the result follows from Proposition 1. ■

By setting $\varphi' = \varphi$ in Proposition 1 one obtains

**Corollary 2** *Let $\varphi$ be a primitive morphism, and let* $\mathbf{u} = \varphi(\mathbf{u})$*,* $\mathbf{v} = \varphi(\mathbf{v})$ *be two fixed points. Then,* $\mathrm{Fact}(\mathbf{u}) = \mathrm{Fact}(\mathbf{v})$.

## 4 Defect

Recall that Droubay, Justin and Pirillo showed in [10] that for any finite word $w$ of length $n$, its palindromic complexity $|\mathrm{Pal}(w)|$ is bounded by $n+1$, and that Sturmian and also episturmian words realize the bound. Brlek *et al.* defined in [6], the *palindromic defect* of $w$ to be

$$D(w) = |w| + 1 - |\mathrm{Pal}(w)|. \tag{5}$$

If $D(w) = 0$, that is when $w$ contains a maximal number of palindromic factors, then $w$ is said *full*. The next statements follow easily from the definition.

**Lemma 6** *Let $u, v \in \Sigma^*$ be such that $u \in \mathrm{Fact}(v)$, and let $\alpha \in \Sigma$. Then we have*

(i) $D(u) = D(\widetilde{u})$;

(ii) $D(u) \leq D(u\alpha)$*, and* $D(u) \leq D(\alpha u)$;

(iii) $D(u) \leq D(v)$;

(iv) *if $v$ is full then $u$ is full.*

*Proof.* (i) By definition $\mathrm{Pal}(u) = \mathrm{Pal}(\widetilde{u})$. (ii) We have $\mathrm{Pal}(u) \subseteq \mathrm{Pal}(u\alpha)$, and therefore $|\mathrm{Pal}(u\alpha)| - |\mathrm{Pal}(u)| \leq 1 = |u\alpha| - |u|$. It follows that $|u| - |\mathrm{Pal}(u)| \leq |u\alpha| - |\mathrm{Pal}(u\alpha)|$, so that $D(u) \leq D(u\alpha)$. For the second part, we have $D(u) = D(\widetilde{u}) \leq D(\widetilde{u}\alpha) = D(\widetilde{\alpha u}) = D(\alpha u)$. (iii) By induction and (ii). (iv) Directly follows from (iii). ■

For words generated by conjugate morphisms, we have the following nice property.

**Proposition 2** *Let $\varphi$ and $\varphi'$ be two primitive morphisms and suppose there are integers $k$, $l$ such that $\varphi^k$ and $\varphi'^l$ have fixed point, namely $\mathbf{u} = \varphi^k(\mathbf{u})$ and $\mathbf{v} = \varphi'^l(\mathbf{v})$. If $\varphi$ and $\varphi'$ are conjugate, then $D(\mathbf{u}) = D(\mathbf{v})$.*

*Proof.* From Corollary 1, we have $\mathrm{Fact}(\mathbf{u}) = \mathrm{Fact}(\mathbf{v})$. Hence,

$$D(\mathbf{u}) = \sup\{D(w) \mid w \in \mathrm{Fact}(\mathbf{u})\} = \sup\{D(w) \mid w \in \mathrm{Fact}(\mathbf{v})\} = D(\mathbf{v}). \blacksquare$$

Observe that the defect is easily computed with the help of the LPS function. It goes like this:

**Algorithm 1**
**Input :** $w \in \Sigma^*$ ; **Initialization :** $D := 0$ ;
**if** $w \neq \varepsilon$ **then**
**for** $i := 1$ **to** $|w|$ **do**
    $s := \mathrm{LPS}(w[1..i])$ ;
    **if** $s$ is uni-occurrent in $w[1..i]$ **then** $H[i] := |s|$
    **else** $H[i] := 0 \,; D := D + 1$ ;
    **end if**
**end for**
**end if**
Return $D, H$

This algorithm also computes the function $H : \Sigma^+ \longrightarrow \mathbb{N}$ defined by

$$H(w)[i] = \begin{cases} |\mathrm{LPS}(w[0..i])| & \text{if } \mathrm{LPS}(w[0..i]) \text{ is uni-occurent,} \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

for $i$ such that $0 \leq i \leq |w|-1$. For instance, for $w = bbaabbabaaba$, we have the following table

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $w$ | $b$ | $b$ | $a$ | $a$ | $b$ | $b$ | $a$ | $b$ | $a$ | $a$ | $b$ | $a$ |
| $H$ | 1 | 2 | 1 | 2 | 4 | 6 | 4 | 3 | 3 | 0 | 0 | 6 |
| $D$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 |

where the defect is given by the number of 0's in the table listing the values of $H$. In this case we have $D(w) = 2$.

The sequence of values of $H$ may or not contain null values. When $H$ vanishes in some position $k$, the word $w$ is said *lacunary* and the position $k$ is called a *lacuna*. Moreover, we say that $w$ is *end-lacunary* if $k+1 = |w|$. The word in the example above is lacunary and contains two lacunas, namely at positions 9 and 10. Note that $D$ and $H$ may be viewed, by virtue of condition (ii) in Lemma 5, as functions $D, H : \mathbb{N} \times \Sigma^* \longrightarrow \mathbb{N}$, $D$ increasing with respect to $\mathbb{N}$.

It is clear that if $k$ is a lacuna of some word $u$, that is $H(k,u) = 0$, then for all $v \in \Sigma^*$, we have $H(k, uv) = 0$. This means that lacunas are preserved by suffix concatenation. This is no longer true for prefix concatenation as the next example shows.

**Example**. Let $w = u \cdot v = aabca \cdot acbcaacbaa$. Then we have the following table where the lacunas of $v$ are $9, 10$, while the lacunas of $uv$ are $4, 5$.

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $w$ | $a$ | $a$ | $b$ | $c$ | $a$ | $a$ | $c$ | $b$ | $c$ | $a$ | $a$ | $c$ | $b$ | $a$ | $a$ |
| $H$ | 1 | 2 | 1 | 1 | 0 | 0 | 4 | 6 | 3 | 5 | 7 | 9 | 11 | 13 | 15 |

One might ask whether a relation exists or not between $D(uv)$ and $D(u)+D(v)$. We have

$$2 = D(aabca \cdot acbcaacbaa) \leq D(aabca) + D(acbcaacbaa) = 1 + 2.$$

However, $D(aabca \cdot acbcaacb) = 2$ while $D(aabca) + D(acbcaacb) = 1 + 0 = 1$.

The notion of palindromic defect is extended to infinite words $\mathbf{w} \in \Sigma^\omega$ by setting $D(\mathbf{w})$ to be the maximum of the defect of its factors: it may be finite or infinite. Hence $D(\mathbf{w})$ is also equal to the maximum of the defect of the finite prefixes of $\mathbf{w}$.

# 5 Periodic words

We know from [6] that there exist periodic infinite words with a finite number of palindromes, and consequently there exist infinite words with infinite defect (equivalently finite words with arbitrarily large defects).

A periodic word may not have both an infinite defect and an infinite palindromic complexity. This is a direct consequence of the following results [6].

**Proposition 3** (Theorem 4 [6]) *Let $w$ be a primitive word. Then we have the following equivalent conditions:*

(i) *$w$ is the product of two palindromes;*

(ii) *$|\mathrm{Pal}(w^\omega)|$ is infinite.*

A word satisfying the equivalent conditions of the theorem is called *symmetric*, and *asymmetric* otherwise.

**Theorem 1** (Theorem 6 [6]) *Let $w = uv$, with $|u| \geq |v|$ and $u, v \in \mathrm{Pal}(\Sigma^*)$, be a primitive symmetric word. Then the defect of $\mathbf{w} = w^\omega$ is bounded by the defect of its prefix of length $|uv| + \lfloor \frac{|u|-|v|}{3} \rfloor$.*

We exhibit now a few examples showing that the bound is not optimal, but is not far from being sharp. For instance, let $w = u \cdot v = aabaa \cdot bab$. Then, $|uv| + \lfloor \frac{|u|-|v|}{3} \rfloor = |uv|$. It is easy now to check that $D(aabaa \cdot bab) = 0$. More generally, for any integer $k > 1$,

$$D((aab^k aa \cdot bab)^\omega) = D((aab^k aa \cdot bab) \cdot aab^k) = 0,$$

showing that there exist an infinite family of full infinite periodic words, a result in the spirit of that of Droubay, Justin and Pirillo on Sturmian words [10]. If $w = u \cdot v = a^2 bacaba^2 \cdot c$, we have $\lfloor \frac{|u|-|v|}{3} \rfloor = 2$. But

$$D(w^\omega) = D((aabacabaa \cdot c) \cdot a).$$

Consider now the word $w = u \cdot v = a^{k+1} ba^k ca^k ba^{k+1} \cdot c$, then $\lfloor \frac{|u|-|v|}{3} \rfloor = \lfloor \frac{4k+4}{3} \rfloor$, We have

$$D(w) = 1 \quad \text{and} \quad D(w^\omega) = D(w \cdot a^{k+1} ba^k) = D(w \cdot a^k) = k + 1.$$

On the other hand, for $w = u \cdot \epsilon = abbabaababba$, we have $D(w) = 0$, but $D(u \cdot abba) = 2$. More generally, for $w = ab^k a \cdot b^{k-1} aab^{k-1} \cdot ab^k a$ which is palindromic ($v = \epsilon$) we have for $k > 1$

$$D(w) = 0 \quad \text{while} \quad D(w \cdot ab^{k-1}) = k \quad \text{and} \quad D(w^\omega) = k,$$

so that $|w| = |u| = 4k + 4$ and $|ab^{k-1}| = k$. Note that $D((cabac \cdot faf)^\omega)$=1, and, on a 2-letter alphabet we also have $D((aababbaabbabaa)^\omega) = 1$. These examples solve the problem of constructing periodic sequences having finite defect value.

Moreover, by taking a convenient periodic sequence having defect value 0, preceded by $w$ one obtains a non periodic sequence (ultimately periodic indeed) having defect $k$, so that we have $D(ww \cdot a^\omega) = k$. By choosing a convenient Sturmian infinite word $\mathbf{v}$ (they all are full), one easily shows that $D(ww.\mathbf{v}) = k$ as well. Summarizing the preceding discussion we have

**Proposition 4** *There exist an infinite class of infinite words having a fixed defect value $k \geq 0$.*

## 6 The Thue-Morse word

In this section we fix the alphabet $\Sigma = \{a, b\}$. Recall that the Thue-Morse word $\mathbf{t}$ is the fixed point starting with $a$ of the morphism $\mu$ defined by

$$\mu(a) = ab \quad ; \quad \mu(b) = ba$$

Note that $\mathbf{t}$ is also a fixed point of the morphism $\theta = \mu^2$, that is

$$\mathbf{t} = \theta^\omega(a) = abbabaabbaababba\cdots$$

where $\theta : \Sigma^* \to \Sigma^*$ is defined by $\theta : a \mapsto abba, b \mapsto baab$. While the complexity function $F_{\mathbf{t}}(n) = |\mathrm{Fact}_n(\mathbf{t})|$ has been established independently by Brlek [5] and de Luca and Varricchio [16], the recent interest in palindromic complexity — see Allouche and al. [2] for a detailed account — begs for a description of the palindromic structure of $\mathbf{t}$. For instance, the existence of infinitely many palindromic prefixes in an infinite word $\mathbf{u}$, implies that $\mathbf{u}$ is recurrent. This fact was previously observed in [7] and used in [8] (Prop. 15) in order to show that some infinite words, namely the so-called *smooth infinite words* are recurrent. Since it does not appear explicitely as a statement we provide it here.

**Proposition 5** *If an infinite word* $\mathbf{u}$ *has infinitely many palindromic prefixes, then the set of factors* $\mathrm{Fact}(\mathbf{u})$ *is closed under reversal and* $\mathbf{u}$ *is recurrent.*

*Proof.* Let $f$ be a finite factor of $\mathbf{u}$. Then $\mathbf{u} = xf\mathbf{v}$ for some $x \in \Sigma^*$ and $\mathbf{v} \in \Sigma^\omega$. By hypothesis there exists a palindromic prefix starting with $xf$, hence containing $\widetilde{xf}$, showing that $\mathrm{Fact}(\mathbf{u})$ is closed under reversal. For the recurrence property, an extra step is necessary and we proceed by contradiction. Assume that $f$ is a non recurrent factor of $\mathbf{u}$. Then we can choose $x$ and $\mathbf{v}$ such that $f \notin \mathrm{Fact}(\mathbf{v})$. Let $p$ and $q$ be palindromic prefixes of $\mathbf{u}$ such that $|p| \geq |xf|$ and $|q| \geq 2|p|$. Then $p$ contains both $f$ and $\widetilde{f}$. Moreover $p$ is a suffix of $q$ and, since $|q| \geq 2|p| \geq 2|xf|$, $p$ is a factor of $\mathbf{v}$, so that $f$ and $\widetilde{f}$ are factors of $\mathbf{v}$. Contradiction. ∎

It is easy to see that $\mathbf{t}$ has an infinite palindromic complexity since for all $n \in \mathbb{N}$, $\mathbf{t}_n = \theta^n(a)$ is a palindrome, so that $\mathbf{t}$ has infinitely many palindromic prefixes. However $\mathbf{t}$ is not full as can be seen in the table below, where there are lacunas at positions 8 and 9.

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $w$ | $a$ | $b$ | $b$ | $a$ | $b$ | $a$ | $a$ | $b$ | $b$ | $a$ | $a$ | $b$ | $a$ | $b$ | $b$ | $a$ |
| $H$ | 1 | 1 | 2 | 4 | 3 | 3 | 2 | 4 | 0 | 0 | 6 | 8 | 10 | 12 | 14 | 16 |

Our goal now is to give a complete characterization of the lacunas in $\mathbf{t}$, and for that purpose the following definition, equivalent to that of Mignosi [17](see also Frid [11]), is useful.

**Definition 1** *We say that* $v$ *is an* ancestor of $w$ with respect to $\theta$ *if there exist a proper block-prefix* $x$ *and a proper block-suffix* $y$ *such that* $xwy = \theta(v)$. *We denote by* $\mathrm{Anc}(w)$ *the set of ancestors of* $w$. *Moreover, we say that* $w$ *is* centered (with respect to $v$) *if* $|x| = |y|$.

Let $v$ be an ancestor of $w$ such that $xwy = \theta(v)$ as above. Let $p, s \in \Sigma^*$ such that $w = ps$. Then we say that there is a *bar line* between $p$ and $s$, written $p|_\theta s$ (or simply $p|s$ when the context is clear), if there exist $v_1, v_2 \in \Sigma^*$ such that $v = v_1 v_2$, $xp = \theta(v_1)$ and $sy = \theta(v_2)$. We start by listing some useful properties of $\mathbf{t}$.

**Lemma 7** *Let $u \in \mathrm{Pal}(\mathbf{t})$. Then the following conditions hold:*

(i) *if $|u| \geq 4$, then $|u|$ is even;*

(ii) *if $|u| \geq 4$, then all ancestors of $u$ are palindromes and $u$ is centered;*

(iii) *if $|u| = 4$, then $u$ has exactly two ancestors;*

(iv) *if $|u| > 4$, then $u$ has a unique ancestor.*

*Proof.* (i) Assume by contradiction that there exists a factor $u$ of odd length such that $|u| \geq 4$. Then $u = zv\tilde{z}$ for some word $z$ and some palindrome $v$ of length 5. But the only possibilities for $v$ over $\Sigma = \{a, b\}$ are

$$aaaaa, aabaa, ababa, abbba, baaab, babab, bbabb, bbbbb$$

which are not factors of $\mathbf{t}$.

(ii) Let $v$ be an ancestor of $u$. Then $xuy = \theta(v)$ for some words $x$ and $y$, with $|x|, |y| \leq 3$. Since $|u| \geq 4$, there exist letters $\alpha$ and $\beta$ and a word $z$ such that $u = z\alpha\beta\beta\alpha\tilde{z}$. Then there are four possible bar lines :

$$\text{(a) } xz|\alpha\beta\beta\alpha\tilde{z}y, \quad \text{(b) } xz\alpha|\beta\beta\alpha\tilde{z}y, \quad \text{(c) } xz\alpha\beta|\beta\alpha\tilde{z}y, \quad \text{(d) } xz\alpha\beta\beta|\alpha\tilde{z}y$$

Cases (b) and (d) cannot happen. Otherwise, we would respectively have a block beginning with $\beta\beta$ and a block ending with $\beta\beta$, which is impossible. For the two remaining cases, we have $|xz| \equiv |\tilde{z}y| \bmod 4$. Therefore $|x| = |y|$, since $|x|, |y| \leq 3$.

Now, let $xp$ be the prefix of length 4 of $\theta(v)$ and $\tilde{p}y$ be the suffix of length 4 of $\theta(v)$, where $p$ is some non empty word. The fact that both $xp$ and $\tilde{p}y$ are Thue-Morse blocks implies that they are palindromes as well. Thus $\tilde{p}y = \tilde{y}p$. Moreover, since $p$ is non empty, $xp$ and $\tilde{y}p$ must correspond to the same block, so that $x = \tilde{y}$. Hence, $\theta(v) = xu\tilde{x}$ is a palindrome. Finally, $\theta(v) = \widetilde{\theta(v)} = \tilde{\theta}(\tilde{v}) = \theta(\tilde{v})$. But $\theta$ is injective, so $v = \tilde{v}$ is a palindrome.

(iii) The only palindromes of length 4 are *abba* and *baab*. It is easy to see that $\mathrm{Anc}(abba) = \{a, bb\}$ and $\mathrm{Anc}(baab) = \{b, aa\}$.

(iv) Let $\alpha, \beta, \gamma \in \Sigma$ and $z \in \Sigma^*$ such that $u = z\gamma\alpha\beta\beta\alpha\gamma\tilde{z}$. From (ii) we know that $|\mathrm{Anc}(u)| \leq 2$ (see cases (a) and (c)). Now, if $\gamma = \alpha$, then we have $xz|\alpha\alpha\beta\beta\alpha\alpha\tilde{z}\tilde{x}$, which is impossible, since no block begins with $\alpha\alpha$. If $\gamma = \beta$, then $xz\beta\alpha\beta|\beta\alpha\beta\tilde{z}\tilde{x}$, which is also impossible, since no block begins with $\beta\alpha\beta$. In both cases $|\mathrm{Anc}(u)| = 1$. ■

It is now easy to compute the palindromic complexity $P_{\mathbf{t}}(n) = |\mathrm{Pal}(\mathbf{t}) \cap \mathrm{Fact}_n(\mathbf{t})|$ of the Thue-Morse using Lemma 7.

**Proposition 6** *The palindromic complexity $P_{\mathbf{t}}(n)$ of $\mathbf{t}$ satisfies the following recurrence :*

(i) $P_{\mathbf{t}}(0) = 1$, $P_{\mathbf{t}}(1) = P_{\mathbf{t}}(2) = P_{\mathbf{t}}(3) = P_{\mathbf{t}}(4) = 2$,

(ii) *for all* $n \geq 2$, $P_{\mathbf{t}}(2n+1) = 0$, *and*

(iii) *for all* $n \geq 2$, $P_{\mathbf{t}}(4n) = P_{\mathbf{t}}(4n-2) = P_{\mathbf{t}}(n) + P_{\mathbf{t}}(n+1)$.

*Proof.* These results follow from Theorem 9 of [2], but for sake of completeness, we give here a direct proof. First, by inspection, (i) is satisfied. Part (ii) follows from Lemma 7(i). To prove (iii), assume that $n \geq 2$. We first show that $P_{\mathbf{t}}(4n) = P_{\mathbf{t}}(4n-2)$. Let $p \in \mathrm{Pal}_{\mathbf{t}}(4n-2)$. Then by Lemma 7 there exist a unique $u$ and a unique $x$ of length $|x| \in \{1,3\}$ such that $\theta(u) = \widetilde{x}px$. Let $f : \mathrm{Pal}_{\mathbf{t}}(4n-2) \to \mathrm{Pal}_{\mathbf{t}}(4n)$, be the function defined by $p \mapsto x_0 p x_0$ where $x_0$ is the first letter of $x$. Then $f$ is a bijection.

It remains to show that $P_{\mathbf{t}}(4n) = P_{\mathbf{t}}(n) + P_{\mathbf{t}}(n+1)$. Let $p \in \mathrm{Pal}_{\mathbf{t}}(4n)$. Then by Lemma 7, there exist a unique $x$ and a unique palindrome $u$ such that $\widetilde{x}px = \theta(u)$, where $p$ is centered, $|x| \in \{0,2\}$ and $|u| \in \{n, n+1\}$. Let $anc : \mathrm{Pal}_{\mathbf{t}}(4n) \to \mathrm{Pal}_{\mathbf{t}}(n) \cup \mathrm{Pal}_{\mathbf{t}}(n+1)$ be the function defined by $p \mapsto u$. We first show that $anc$ is a injection. Indeed, let $p, q \in \mathrm{Pal}_{\mathbf{t}}(4n)$ and assume that $anc(p) = anc(q)$. Then $p$ and $q$ are both centered in $\theta(anc(p)) = \theta(anc(q))$ so that $p = q$. To show that $anc$ is surjective, let $u \in \mathrm{Pal}_{\mathbf{t}}(n) \cup \mathrm{Pal}_{\mathbf{t}}(n+1)$. If $|u| = n$, then $u = anc(\theta(u))$. On the other hand, if $|u| = n+1$, then $u = anc(\widetilde{x}^{-1}\theta(u)x^{-1})$. The result follows. ■

A closed formula for $P_{\mathbf{t}}(n)$ is then easily obtained by induction :

$$P_{\mathbf{t}}(n) = \begin{cases} 1 & \text{if } n = 0, \\ 2 & \text{if } 1 \leq n \leq 4, \\ 0 & \text{if } n \text{ is odd and } n \geq 5, \\ 4 & \text{if } n \text{ is even and } 4^k + 2 \leq n \leq 3 \cdot 4^k, \text{ for } k \geq 1, \\ 2 & \text{if } n \text{ is even and } 3 \cdot 4^k + 2 \leq n \leq 4^{k+1}, \text{ for } k \geq 1. \end{cases} \tag{7}$$

**Lemma 8** *Let $w \neq \varepsilon$ be a prefix of $\mathbf{t}$ and $u$ be a palindromic suffix of $w$. Moreover, suppose that $u$ is the ancestor of a word $v \neq \varepsilon$ such that $\theta(u) = \tilde{x}vx$, for some word $x$, $|x| \leq 3$. Then $u = \mathrm{LPS}(w)$ if and only if $v = \mathrm{LPS}(\theta(w)x^{-1})$.*

*Proof.* The overall situation is depicted in Figure 1.

($\Rightarrow$) By contradiction, assume that $v'$ is a palindromic suffix of $\theta(w)x^{-1}$ and $|v'| > |v|$. Since $4 \mid |\tilde{x}vx|$, we have that $|v|$ is even so that $|v| \geq 2$ and $|v'| \geq 3$. Moreover, the case $|v'| = 3$ is impossible : otherwise, we would have $v' = \alpha^3$ for some letter $\alpha$. Therefore, $|v'| \geq 4$. From Lemma 7 (ii), $v'$ is also centered and has a palindromic ancestor $u'$ which is a suffix of $w$. This means that $\theta(u') = \tilde{x}v'x$ is a (palindromic) suffix of $\theta(w)$. But $|v'| > |v|$ so that $|\theta(u')| = |\tilde{x}v'x| > |\tilde{x}vx| = |\theta(u)|$. Hence $|u'| > |u|$, contradicting the assumption that $u = \mathrm{LPS}(w)$.

($\Leftarrow$) Again by contradiction, assume that $u'$ is a palindromic suffix of $w$, $|u'| > |u|$. Then $\theta(u')$ is a palindromic suffix of $\theta(w)$. Moreover, there exists a

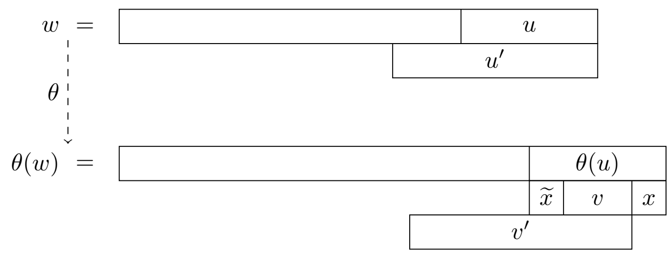


Figure 1: Schematic representation of the proof of Lemma 8

palindrome $v'$ such that $\theta(u') = \tilde{x}v'x$. But $|u'| > |u|$, so that $|\tilde{x}v'x| = |\theta(u')| > |\theta(u)| = |\tilde{x}vx|$. Hence $|v'| > |v|$. This contradicts $v = \mathrm{LPS}(\theta(w)x^{-1})$. ■

**Lemma 9** *Let $w$ be a prefix of $\mathbf{t}$. Then we have*

(i) $|\mathrm{LPS}(w)| = 1$ *if and only if* $|w| = 1$ *or* $|w| = 2$*,*

(ii) $|\mathrm{LPS}(w)| = 3$ *if and only if* $|w| = 5$ *or* $|w| = 6$*.*

*Proof.* (i) ($\Leftarrow$) Clearly, $\mathrm{LPS}(a) = \mathrm{LPS}(ab) = 1$. ($\Rightarrow$) We show that if $|w| \geq 3$, then $|\mathrm{LPS}(w)| \geq 2$. For $|w| = 3$, it is true since $\mathrm{LPS}(abb) = 2$. Now, assume that $|w| \geq 4$ and let $\alpha$ be the last letter of $w$. Since $\mathbf{t}$ is overlap-free, one of the words in $\{\alpha\alpha, \alpha\beta\alpha, \alpha\beta\beta\alpha\}$ is a suffix of $w$, where $\beta \neq \alpha$ is a letter. Hence $|\mathrm{LPS}(w)| \geq 2$.

(ii) ($\Leftarrow$) It is easy to see that $\mathrm{LPS}(abbab) = \mathrm{LPS}(abbaba) = 3$. ($\Rightarrow$) Let $y$ be the ancestor of $w$ such that $wx = \theta(y)$, for some word $x$, $|x| \leq 3$. By inspection, if $|y| \leq 3$, the only possibilities satisfying $|\mathrm{LPS}(w)| = 3$ are $|w| = 5$ or $|w| = 6$. Now, assume that $|y| \geq 4$. Let $u = \mathrm{LPS}(w)$. Then $u \in \{aba, bab\}$. Moreover $\mathrm{Anc}(aba) = \{ab, ba\} = \mathrm{Anc}(bab)$. Hence, either $ab$ or $ba$ is a suffix of $y$, so that there exists a word $s \in \{bab, baab, aba, abba\}$ that is a suffix of $y$. But $\tilde{x}^{-1}\theta(s)x^{-1}$ is a palindromic suffix of $w$ as well, and $|\tilde{x}^{-1}\theta(s)x^{-1}| \geq 6$, contradicting the hypothesis. ■

**Lemma 10** *Let $w$ be a prefix of $\mathbf{t}$ such that $|w| \geq 8$ and let $x$ be a suffix of $\theta(w)$ such that $|x| \leq 3$. Then $w$ is end-lacunary if and only if $\theta(w)x^{-1}$ is end-lacunary.*

*Proof.* ($\Rightarrow$) Let $u = \mathrm{LPS}(w)$. Since $w$ is end-lacunary, $u$ is not uni-occurrent in $w$ so that $w = yuzu$ for some word $y$ and some non empty word $z$. But $\theta(u)$ is a palindrome, which means that there exists a palindrome $v$ such that $\theta(u) = \tilde{x}vx$. Then we have the situation depicted in the following diagram.
By Lemma 8, $v = \mathrm{LPS}(\theta(w)x^{-1})$. Hence, $\theta(w)x^{-1}$ is end-lacunary, since $v$ is not uniocurrent in $\theta(w)x^{-1} = \theta(y)\tilde{x}vx\theta(z)\tilde{x}v$.

($\Leftarrow$) Let $v = \mathrm{LPS}(\theta(w)x^{-1})$. First, assume that $|v| > 4$. From Lemmas 7(ii) and 7(iv), $v$ is centered with respect to a unique palindromic ancestor $u$, i.e. $\tilde{x}vx = \theta(u)$. Therefore, since $\theta(w)x^{-1}$ is end-lacunary, there exist some word $y$

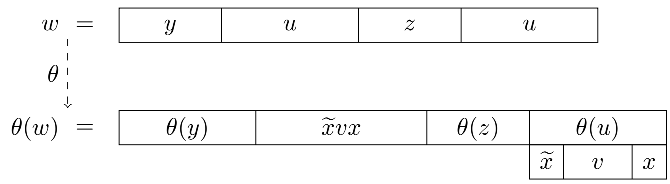


Figure 2: Schematic representation of the proof of Lemma 10

and some non empty word $z$ such that $\theta(w) = \theta(y)\tilde{x}vx\theta(z)\tilde{x}vx$ and $w = yuzu$. Hence $u$ is not uni-occurrent in $w$. But, from Lemma 8, $u = \mathrm{LPS}(w)$. This means that $w$ is end-lacunary.

Now, assume that $|v| \leq 4$. From Lemma 9, we have either $|v| = 2$ or $|v| = 4$. Then, $v \in \{aa, bb, abba, baab\}$. Moreover, if $u$ is an ancestor of $v$, then $u \in U = \{a, b, aa, bb\}$. From Lemma 8, $u = \mathrm{LPS}(w)$. Since $|w| \geq 8$, factors of $U$ have already appeared in $w$. Hence, $u$ is not uni-occurent in $w$ and $w$ is end-lacunary. ■

**Remark 1** *Lemma 10 can be restated as follows. Let $i \in \mathbb{N}$. Then $i$ is a lacuna of $\mathbf{t}$ if and only if all integers in $[4i..4i+3]$ are lacunas of $\mathbf{t}$. In particular, if $i, j \in \mathbb{N}$ and $i \leq j$, then all integers in $[i..j]$ are lacunas if and only if all integers in $[4i..4j+3]$ are lacunas as well.*

An explicit description of the lacunas is given now. For $n \in \mathbb{N}^+$, let $L(n)$ be the index where the $n$-th interval of lacunas start and $\ell(n)$ be its length.

**Theorem 2** *The sequences $L$ and $\ell$ satisfy the following recurrences :*

(i) $L(1) = 8$ *and* $L(2) = 24$,
(ii) $L(n) = 4L(n-2)$, *for* $n \geq 3$.
(iii) $\ell(1) = \ell(2) = 2$
(iv) $\ell(n) = 4\ell(n-2)$, *for* $n \geq 3$.

*Proof.* An easy proof by induction shows that the sequence $L$ is increasing and that the intervals described by $L$ and $\ell$ are pairwise non overlapping.

On the other hand, if we consider the prefix of length 32 of $\mathbf{t}$, the only lacunas are 8, 9, 24, 25, so that $L(1) = 8$, $L(2) = 24$ and $\ell(1) = \ell(2) = 2$. Now, for any prefix of $\mathbf{t}$ of length at least 32, Lemma 10 applies. Therefore, by Remark 1, $[L(i)..L(i)+\ell(i)-1]$ is an interval of lacunas if and only if $[4L(i)..4L(i)+4\ell(i)-1]$ is also an interval of lacunas. ■

Closed formulas for $L$ and $\ell$ are easily obtained :

$$L(n) = \begin{cases} 2^{n+2}, & \text{if } n \text{ is odd,} \\ 2^{n+2} + 2^{n+1}, & \text{if } n \text{ is even.} \end{cases}$$

and

$$\ell(n) = \begin{cases} 2^{n}, & \text{if } n \text{ is odd,} \\ 2^{n-1}, & \text{if } n \text{ is even.} \end{cases}$$

The first intervals where lacunas occur are

$$[8..9], [24..25], [32..39], [96..103], [128..159], [384..415], \ldots$$

## 7 Concluding remarks

Now, recall from [11] that a morphism $\varphi$ is *marked* if for every $\alpha, \beta \in \Sigma$ such that $\alpha \neq \beta$, the first letters and the last letters of $\varphi(\alpha)$ and $\varphi(\beta)$ are different. From this definition, we see that some results of section 6 can easily be extended to uniform marked morphisms $\varphi$ such that $\varphi(\alpha)$ is a palindrome, for every $\alpha \in \Sigma$. Obviously, Lemma 7 would have to be stated rather differently, but the essential part is the uniqueness of the ancestor, which is a simple matter for uniform marked morphisms (see Remark 3 of [11]). Moreover, it is possible to provide an algorithm to decide whether a uniform marked morphism generates full words or not : it suffices to generalize Lemma 10 accordingly. Finally, it would be easy to deduce from these Lemmas that a fixed point of a uniform marked morphism is either full or has an infinite defect. More precisely we state the following conjecture.

**Conjecture 1** *Let $\mathbf{u}$ be the fixed point $\mathbf{u} = \varphi(\mathbf{u})$ of a primitive morphism $\varphi$. If the defect is such that $0 < D(\mathbf{u}) < \infty$, then $\mathbf{u}$ is periodic.*

We conclude by suggesting some open problems :

1. It would be interesting to extend the optimal algorithm deduced from Theorem 6 of [6] (restated in Theorem 1) for computing the defect of an infinite periodic word to fixed points of primitive morphisms. For this purpose, recall that in [12], Hof et al. introduced morphisms of class $P$, i.e. morphisms such that there exist palindromes $p$ and $q_\alpha$ satisfying $\varphi(\alpha) = pq_\alpha$ for every $\alpha \in \Sigma$. They also conjectured that if a fixed point $\mathbf{u}$ of a primitive morphism has infinitely many palindromes, then there exists a morphism $\varphi$ such that either $\varphi$ or $\widetilde{\varphi}$ is of class $P$ and $\mathbf{u} = \varphi(\mathbf{u})$. Recently, a constructive proof has been provided for binary alphabets by Tan [19]. Let $\mathbf{u} = \varphi(\mathbf{u})$ be a fixed point of a morphism $\varphi$. Does there exist an algorithm for deciding whether $\mathbf{u}$ is full or not ? At first, it would be interesting to provide one for morphisms in class $P$. Indeed, from Proposition 2, the algorithm could be extended to any morphism having a conjugate of class $P$, and assuming that the Hof-Knill-Simon conjecture is true (which is the case for the binary alphabet), we would have an algorithm for any primitive morphism.

2. Is there a better algorithm than Algorithm 1 to compute the defect of a finite word ? In other words, is Algorithm 1 optimal ?

3. Let $f(n, k, d)$ be the number of words of length $n$ over a $k$-letter alphabet having $d$ lacunas. Is it possible to compute $f(n, k, d)$ in an efficient way ?

## References


[1] J.P. Allouche, *Schrödinger operators with Rudin-Shapiro potentials are not palindromic*, J. Math. Phys. **38** (1997) 1843–1848.

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

[3] J.P. Allouche and J. Shallit, *Sums of digits, overlaps, and palindromes*, Disc. Math. and Theoret. Comput. Sci. **4** (2000) 1–10.

[4] M. Baake, *A note on palindromicity*, Lett. Math. Phys. **49** (1999) 217–227.

[5] S. Brlek, *Enumeration of factors in the Thue-Morse word*, Disc. Appl. Math. **24** (1989) 83–96.

[6] S. Brlek, S. Hamel, M. Nivat and C. Reutenauer, *On the Palindromic Complexity of Infinite Words*, in J. Berstel, J. Karhumäki, D. Perrin, Eds, Combinatorics on Words with Applications, Int. J. of Found. Comput. Sci. **15**: 2 (2004) 293–306

[7] S. Brlek and A. Ladouceur, *A note on differentiable palindromes*, Theoret. Comput. Sci. **302**:1-3 (2003) 167–178.

[8] S. Brlek, D. Jamet and G. Paquin, *Smooth Words on 2-letter alphabets having same parity*, Theoret. Comput. Sci. **393**:1-3 (2008) 166–181.

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

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

[11] A. Frid, *Applying a uniform marked morphism to a word*, Disc. Math. and Theoret. Comput. Sci. **3** (1999) 125–140.

[12] A. Hof, O. Knill and B. Simon, *Singular continuous spectrum for palindromic Schrödinger operators*, Commun. Math. Phys. **174** (1995) 149–159.

[13] M. Lothaire, Combinatorics on words, Addison-Wesley, 1983.

[14] M. Lothaire, Algebraic Combinatorics on words, Cambridge University Press, 2002.

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

[16] A. de Luca and S. Varricchio, *Some combinatorial properties of the Thue-Morse sequence and a problem in semigroups*, Theoret. Comput. Sci. **63**:3 (1989) 333–348.

[17] F. Mignosi and P. Séébold, *If a D0L language is k-power free then it is circular*, Proc. ICALP'93, LNCS **700** (1993) 507–518.

[18] M. Morse and G. Hedlund, *Symbolic Dynamics*, Amer. J. Math. **60** (1938) 815–866.

[19] B. Tan, *Mirror substitutions and palindromic sequences*, Theoret. Comput. Sci. **389**:1-2 (2007), 118–124.