Palindromes and local periodicity ⋆

A. Blondin Massé, S. Brlek ∗, A. Garon, 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

Abstract

In this paper we consider several types of equations on words, motivated by the attempt of characterizing the class of polyominoes that tile the plane by translation in two distinct ways. Words coding the boundary of these polyominoes satisfy an equation whose solutions are in bijection with a subset of the solutions of equations of the form $AB\widetilde{A}\widetilde{B} \equiv XY\widetilde{X}\widetilde{Y}$. It turns out that the solutions are strongly related to local periodicity involving palindromes and conjugate words.

Keywords Palindromes, periodicity, Fine and Wilf, polyominoes, tilings.

1 Introduction

A large part of combinatorics on words concerns the study of regularities and equations involving them. By describing their combinatorial properties, one is often led to structures that are useful for obtaining faster algorithms for several decidable problems. Their relevance is justified by numerous applications, among which the widely studied pattern matching in strings and more recently the advances in the genoma sequence analysis provided a broad range of algorithms. In addition the combinatorial properties often provide complexity bounds for both the space representation of objects and the execution time of the algorithms.

Let us give some examples. The well-known "lemma of three squares" [13] states that the equation $x^2 = y^2v = z^2uv$, where z is primitive, implies a condition on the

⋆ with the support of NSERC (Canada)

∗ Corresponding author.
Email addresses: alexandre.blondin.masse@gmail.com (A. Blondin Massé), Brlek.Srecko@uqam.ca (S. Brlek), garon.ariane@gmail.com (A. Garon), Labbe.Sebastien@courrier.uqam.ca (S. Labbé).

length of z, namely that $|z| < |x|/2$. In other words, it shows that such patterns are rather constrained instead of being freely organized. In turn this result is used for providing upper bounds on the number of squares leading to tighter bounds on square detection algorithms. The problem of overlapping palindromes appears in number theory, when trying to superpose Christoffel words [14] or Beatty sequences [16,17]. Indeed, a Christoffel word is characterized by a central palindrome. Motivations for the study of palindromic complexity emerge from many areas among which the study of Schrödinger operators in physics [3,10], number theory [2], discrete geometry [8], 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, a partial account of which may be found in the survey provided by Allouche et al. [1]. Another example stems from discrete geometry on square lattices, where polyominoes are conveniently represented by words on the 4-letter alphabet $\mathcal{E} = \{a, b, \overline{a}, \overline{b}\}$. A nice application is the detection of digital convexity. Indeed, one needs a combination of the computation of the Lyndon factorization followed by checking if all factors are Christoffel words, that is palindrome words extended by a letter at each end and satisfying an arithmetical condition [8].

In this paper, we study equations that yield periodic words. In the preliminary Section 2 we state some useful lemmas, where the key argument for obtaining periodicity is conjugacy. Then, in Section 3 we consider systems of two equations and show how palindromes propagate in a word. We provide a direct and much simpler proof of a result obtained by Labbé [11], based on a simple but efficient property (Lemma 7). By considering equations involving three palindromes, we obtain results in the spirit of the three squares lemma, which may be seen as extending the results of Paquin [14] for Christoffel words. More precisely, given three fixed palindromes in some overlapping configuration, the relative distances between each pair are constrained in order to apply Fine and Wilf's theorem and obtain periodicity. Finally, we consider equations of type $AB\widetilde{A}\widetilde{B} \equiv XY\widetilde{X}\widetilde{Y}$ on circular words, which are linked with the representation of a tile yielding tesselations of the plane with translated copies of it. Again, constraints on the way these factors overlap yield periodicity, leading to the discovery of an infinite class of polyominoes (called *double squares* in [15]) that tile the plane by translation in two distinct ways as depicted below.

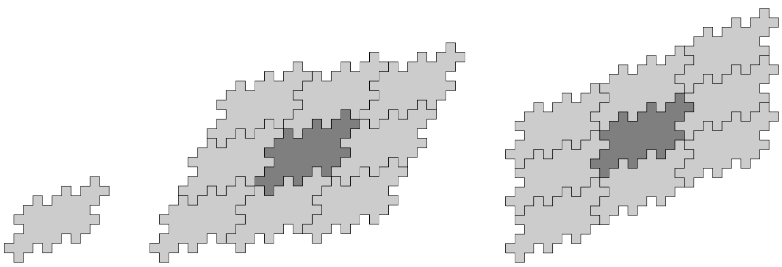


Fig. 1. A double square and its two associated tilings.

2 Preliminaries

All the basic terminology about words is taken from M. Lothaire [12]. In what follows, Σ 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 Σ is denoted Σ^n. By convention, the *empty* word is denoted ε and its length is 0. The free monoid generated by Σ is defined by $\Sigma^* = \bigcup_{n\geq 0} \Sigma^n$. The set of right infinite words is denoted by Σ^ω and we set $\Sigma^\infty = \Sigma^* \cup \Sigma^\omega$. Given a word $w \in \Sigma^\infty$, a *factor* u of w is a word $u \in \Sigma^*$ such that $w = xuy$, with $x \in \Sigma^*$ and $y \in \Sigma^\infty$. If $x = \varepsilon$ (resp. $y = \varepsilon$) then u is called *prefix* (resp. *suffix*). If $w = xu$, with $|w| = n$ and $|x| = k$, then $x^{-1}w = w[k..n-1] = u$ is the word obtained by erasing from w its prefix x. 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 $u \in \Sigma^*$ is $|w|_u$. A *period* of a word w is an integer $p < |w|$ such that $w[i] = w[i+p]$, for all $i < |w| - p$. If p is a period of w, then w is *periodic* with period p. An important result about periods is due to Fine and Wilf.

Theorem 1 (Fine et Wilf) *Let w be a word having p and q for periods. If $|w| \geq p+q-\gcd(p,q)$, then $\gcd(p,q)$ is also a period of w.*

A word is said to be *primitive* if it is not a power of another word. Two words u and v are *conjugate*, written $u \equiv v$, when there are words x,y such that $u = xy$ and $v = yx$. Moreover, we say that *u is conjugate to v with delay $|x|$*.

Lemma 2 *If u is conjugate to v with delay d_1 and v is conjugate to w with delay d_2, then u is conjugate to w with delay $d_1 + d_2 \bmod |u|$.*

Periodicity is related to conjugacy as the next propositions show.

Proposition 3 *If $w = ABC = CBA$, then $|AB|$, $|CB|$ and $\gcd(|AB|,|CB|)$ are periods of w.*

Proof. By hypothesis, we have

$$AB \cdot CB = ABC \cdot B = CBA \cdot B = CB \cdot AB$$

which is of the form $xy = yx$, and therefore there exist a word p, and two integers i and j such that $x = p^i$ et $y = p^j$ ([12], Prop. 1.3.2). Hence $AB = p^i$ and $BC = p^j$. It follows that $|p|$ divides both $|AB|$ and $|BC|$, so that $|p|$ also divides $\gcd(|AB|,|CB|)$. $\square$

Observe that if $A = \varepsilon$ in the proposition above, then w is conjugate to itself with delay $d = |B|$, so that $\gcd(d,|w|)$ is a period of w. Moreover, the statement above may be refined to take into account the fact that $|p|$ is a period of w. In fact a more general result holds. Indeed, let $\sigma : \{A,B,C\} \longrightarrow \{A,B,C\}$ be a bijection (or permutation). Then we have the following property showing that if the factors A,B and C have two occurrences in $w = ABC$ then periodicity appears.

Proposition 4 *Let $w = ABC = \sigma(A)\cdot\sigma(B)\cdot\sigma(C)$, where $\sigma \neq \mathrm{Id}$. If A, B and $C \neq \varepsilon$, then the following conditions hold*

(i) *if σ is cyclic then there exist $p \in \Sigma^*$ and $k > 1 \in \mathbb{N}$ such that $ABC = p^k$;*
(ii) *otherwise, either ABC or AB or BC is periodic.*

Proof. (i) In this case we have either $ABC = BCA$ or $ABC = CAB$. In both cases, it is an equation of the form $xy = yx$, so that Proposition 3 applies.

(ii) $ABC = CBA$ implies the first result, while $AB\cdot C = BA\cdot C$ implies the second and $A\cdot BC = A\cdot CB$ the third. □

The *reversal* of $u = u_0u_1\cdots u_{n-1} \in \Sigma^n$ is the word $\widetilde{u} = u_{n-1}u_{n-2}\cdots u_0$, and a *palindrome* is a word p such that $p = \widetilde{p}$. For a language $L \subseteq \Sigma^\infty$, the set of the palindromic factors of its elements is denoted by $\mathrm{Pal}(L)$. Every word contains palindromes, the letters and ε being necessarily part of them. This justifies the introduction of the function $\mathrm{LPS}(w)$ which associates to any word w its longest palindromic suffix. We recall from [5] a useful combinatorial property.

Proposition 5 (Blondin Massé et al. [5]) *Assume that $w = xy = yz$ with $|y| \neq 0$. Then for some u,v, and some $i \geq 0$ we have from* [12]

$$x = uv, y = (uv)^i u, z = vu; \tag{1}$$

and 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.*

An interesting consequence is the following result.

Corollary 6 (Blondin Massé et al. [7]) *Assume that $w = xp = qz$ where p and q are palindromes such that $|q| > |x|$. Then w has period $|x| + |z|$, and $x\widetilde{z}$ is a product of two palindromes.*

This basic property is best possible. Indeed, it says that when two distinct palindromes overlap, even for one letter, then a period appears in the word.

3 Equations involving palindromes

A very special case of this problem already appears in the literature. Indeed, Christoffel words (finite Sturmian words) have the property that the central word is a palindrome, and their superposition is possible under some arithmetic constraints. For more details see for instance Simpson [16,17] or more recently Paquin [14]. The following results are excerpts of the last author Master thesis [11]. The presentation given here is much simpler and is based on the following very useful lemma, which may be considered as a special case of Proposition 5. It is however independent since it does not rely on periodicity.

Lemma 7 *Let* $w = \widetilde{y}p = qy$ *where* $p, q \in \mathrm{Pal}(\Sigma)$. *Then* $w \in \mathrm{Pal}(\Sigma)$.

Proof. We have two cases to consider according to the length of y. If $|y| \geq |p|$ then $q = \widetilde{p}$ and $w = \widetilde{p}sp$ where s is the intersection of $\widetilde{y}$ and y. Since s is a palindrome the result follows. If $|y| < |p|$, we can write $q = \widetilde{y}q_1$ and $p = p_1y$. Then $\widetilde{y}p = qy$ implies $\widetilde{y}p_1y = \widetilde{y}q_1y$ and $q_1 = p_1$. On the other hand, taking the reversal of w we have $py = \widetilde{y}q$. Substituting $p = p_1y$ and $q = \widetilde{y}q_1$, we have $p_1 \cdot yy = \widetilde{y}\widetilde{y}q_1 = \widetilde{y}\widetilde{y} \cdot p_1$, which implies that $w \in \mathrm{Pal}(\Sigma)$ by Prop. 5 (i),(iii). □

Proposition 8 *Let* $x, y, p, s \in \Sigma^*$ *with* $|p| + |s| > 0$, *and let* $v, w \in \mathrm{Pal}(\Sigma)$. *Then the following conditions hold :*

(i) *if* $xs = pv$ *and* $\widetilde{x}s = pw$, *then* x *is a palindrome;*
(ii) *if* $xs = py$ *and* $\widetilde{x}s = p\widetilde{y}$, *then* x *and* y *are palindrome;*
(iii) *if* $sx = vp$ *and* $s\widetilde{x} = wp$, *then* x *is a palindrome.*

Proof. (i) We write $\widetilde{s}xs = \widetilde{s} \cdot pv = w\widetilde{p} \cdot s$. Then Lemma 7 applies with $y = \widetilde{p}s$ so that $\widetilde{s}xs$ is a palindrome, and x as well.
(ii) We write $\widetilde{s}xs = \widetilde{s} \cdot py = y\widetilde{p} \cdot s$. If $|y| > 0$, since this equality as the form $uy = y\widetilde{u}$ with $u = \widetilde{s}p$, by Proposition 5 (i) and (iii) $\widetilde{s}xs$ is a palindrome and so is x. If $|y| = \varepsilon$, then $xs = p = \widetilde{x}s$ and x is a palindrome.
(iii) Taking the reversal of the equations, we obtain from (i) the result. □

Corollary 9 *Let* $x, y, p, s \in \Sigma^*$, *where with* $|p| \neq |s|$, *and* $v, w \in \mathrm{Pal}(\Sigma)$. *Then the following conditions hold :*

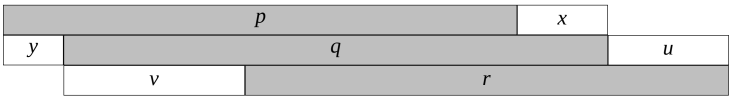
(i) *if* $sx = pv$ *and* $s\widetilde{x} = pw$, *then* x *is a palindrome;*
(ii) *if* $sx = py$ *and* $s\widetilde{x} = p\widetilde{y}$, *then* x *is a palindrome and* y *is a palindrome.*

Proof. (i) If $|p| > |s|$, then there exists $p' \neq \varepsilon$ such that $x = p'v$ and $\widetilde{x} = p'w$. From Proposition 8 (i), we conclude that x is a palindrome. If $|p| < |s|$, then there exists $s' \neq \varepsilon$ such that $s'x = v$ and $s'\widetilde{x} = w$ and x is a palindrome from Proposition 8(iii).
(ii) If $|p| > |s|$, then there exists $p' \neq \varepsilon$ such that $x = p'y$ and $\widetilde{x} = p'\widetilde{y}$. The result follows from Lemma 8(ii). If $|p| < |s|$, then there exists $s' \neq \varepsilon$ such that $s'x = y$ and $s'\widetilde{x} = \widetilde{y}$. The result follows from Proposition 8(ii). □

We consider now the relative position of three palindromes in a word, and start with a property inducing periodicity in overlapping palindromes.

Proposition 10 *Let $w = pxu = yqu = yvr$ where p,q and r are palindromes and $|q| > |x|$ and $|q| > |v|$. Let $a = |x| + |y|$ and $b = |u| + |v|$. If $|q| \geq a + b - \gcd(a,b)$, then $\gcd(a,b)$ is a period of w.*

Proof. The relative positions of p,q and r is as follows.



Applying Corollary 6 to the equations $px = yq$ and $qu = vr$, we obtain that yq is periodic with period $a = |x| + |y|$ and that yq is periodic with period $b = |u| + |v|$. In particular, q has both periods. Since $|q| \geq a + b - \gcd(a,b)$, then the Fine and Wilf's theorem tells us that q has period $\gcd(a,b)$ too. Finally, since $|q| \geq a + b - \gcd(a,b)$ implies that $|q| \geq a$ and $|q| \geq b$, then yq and qu both have period $\gcd(a,b)$ and $w = yqu$ is periodic with period $\gcd(a,b)$. □

As a direct consequence of this Proposition, we know a little bit more when q and r are the longest palindromic suffixes.

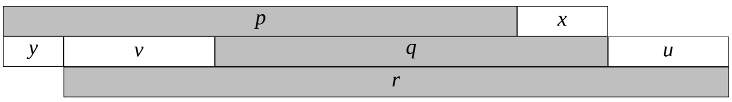
Corollary 11 *If $q = \mathrm{LPS}(px)$ or $r = \mathrm{LPS}(qu)$, with $|x| > 0$ and $|v| > 0$. Then $|q| \geq a + b - 1$, implies that $\gcd(a,b) \neq 1$.*

Proof. If $\gcd(a,b) = 1$, then $w = \alpha^{|w|}$ by Proposition 10. Therefore $\mathrm{LPS}(qu) = vr \neq r$ since $|v| > 0$, a contradiction. □

The next proposition is similar but deals with another configuration.

Proposition 12 *Let $w = pxu = yvqu = yr$ where p, q and r are palindromes such that $|q| > |x|$ and $|r| > |xu|$. Let $a = |x| + |y| + |u|$ and $b = |y| + |v| + |x|$. If $|yvq| > a + b - \gcd(a,b)$, then $\gcd(a,b)$ is a period of w.*

Proof. The relative positions of p,q and r is depicted below.



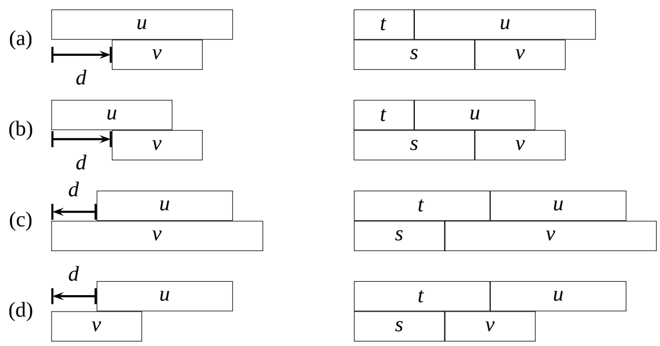
Applying Corollary 6 to the equations $w = pxu = yr$ and $px = yvq$, we obtain that w is periodic with period a and yvq is periodic with period b. If $|yvq| > a + b - \gcd(a,b)$, then from Fine and Wilf's theorem, we obtain that $\gcd(a,b)$ is a period of yvq. Finally, since $|yvq| > a$, the period $\gcd(a,b)$ can be extended over all w and the result follows. □

4 Equations on circular words

Words may overlap in many ways depending on their relative positions, and to deal with all the possible configurations we need a convenient relation. Let $u, v \in \Sigma^*$ and $d \in \mathbb{Z}$, we say that *u overlaps v with delay d* if

$$-|v| < d < |u|$$

and if there exist $s, t \in \Sigma^*$ such that $d = |s| - |t|$ and tu and sv are comparable for the prefix order. In the pictures below, u overlaps v with delay $d = |s| - |t|$ positive in both (a) and (b) and negative in both (c) and (d).



Examples. The word `cheval` overlaps `chevalet` with delay 0. The word `chevalet` overlaps `valet` with delay 3. The word `superduper` overlaps `up` with delay $\{1, 6\}$.

The relation $\mathcal{R} = \{(u, v, d) \in \Sigma^* \times \Sigma^* \times \mathbb{Z} \mid u$ overlaps v with delay $d\}$ satifies some properties that are easy to verify (see [11]). If $(u, v, d) \in \mathcal{R}$ and $(u, v, d') \in \mathcal{R}$ with $d \neq d'$, we will abuse the notation and simply write $(u, v, \{d, d'\}) \in \mathcal{R}$. For instance, the word `perdu` overlaps `superduper` with delay $\{-2, -7\}$.

Lemma 13 [11](4.2) *Let u, $v \in \Sigma^*$ and $d \in \mathbb{Z}$. Then the following conditions are equivalent:*

(i) $(u, v, d) \in \mathcal{R}$;
(ii) $(v, u, -d) \in \mathcal{R}$;
(iii) $(\widetilde{u}, \widetilde{v}, |u| - |v| - d) \in \mathcal{R}$
(iv) $(\widetilde{v}, \widetilde{u}, |v| - |u| + d) \in \mathcal{R}$

Lemma 14 *Let u, v, w, $z \in \Sigma^*$ and $d \in \mathbb{Z}$ such that $(uv, wz, d) \in \mathcal{R}$. Then:*

(i) $(uv, w, d) \in \mathcal{R}$;
(ii) $(u, wz, d) \in \mathcal{R}$ *if* $d < |u|$;
(iii) $(v, wz, d - |u|) \in \mathcal{R}$ *if* $-|w| < d - |u|$.

Lemma 15 *Let u, v, $w \in \Sigma^*$ and d_1, $d_2 \in \mathbb{Z}$ such that $(u, v, d_1) \in \mathcal{R}$ and $(v, w, d_2) \in \mathcal{R}$. If $d_1 d_2 \geq 0$ and $|d_1 + d_2| < |u|$, then $(u, w, d_1 + d_2) \in \mathcal{R}$.*

An immediate consequence of Lemmas 2 and 14 follows.

Lemma 16 *Let $u,v,w,x,y,z \in \Sigma^*$ be such that $|u| = |v|$. Suppose that uy is conjugate to vx with delay d_1 and vw is conjugate to uz with delay d_2. If $d_1d_2 > 0$ and $d_1+d_2 < |u|$, then d_1+d_2 is a period of both u and v.*

Proposition 17 *Let $u,v,x,y \in \Sigma^*$, and $k, \ell, n \in \mathbb{Z}$ such that $|u| = |v| = k$, $|x| = |y| = \ell$ and $n = k+\ell$. Suppose that uy is conjugate to vx with delay d_1 and that $v\widetilde{y}$ is conjugate to $u\widetilde{x}$ with delay d_2. Let $d = d_1+d_2$. Then we have:*

(i) *if $d < k$, then d is a period of u and v;*
(ii) *if $2n-d < k$, then $2n-d$ is a period of u and v;*
(iii) *if $d < \ell$, then d is a period of x and y;*
(iv) *if $2n-d < \ell$, then $2n-d$ is a period of x and y.*

Proof. (i) Follows directly from Lemma 16. (ii) Follows from Lemma 16, since vx is conjugate to uy with delay $n-d_1$ and $u\widetilde{x}$ is conjugate to $v\widetilde{y}$ with delay $n-d_2$. Similarly, (iii) and (iv) both follow from Lemma 16 since yu is conjugate to xv with delay d_1 from Lemma 2 and $x\widetilde{u}$ is conjugate to $y\widetilde{v}$ with delay d_2 from Lemma 13. □

Now, we consider equations of type $AB\widetilde{A}\widetilde{B} \equiv XY\widetilde{X}\widetilde{Y}$, or equivalently, words $w \in \Sigma^*$ satisfying the following properties:

(i) $w = AB\widetilde{A}\widetilde{B}$, with $A,B \in \Sigma^*$
(ii) $wp = pXY\widetilde{X}\widetilde{Y}$ or $sw = XY\widetilde{X}\widetilde{Y}s$

where p and s are respectively a prefix and a suffix of w. Without loss of generality, the case $sw = XY\widetilde{X}\widetilde{Y}s$ may be dropped since it amounts to a renaming of the variables. An example of the situation $wp = pXY\widetilde{X}\widetilde{Y}$ is depicted in Figure 2.

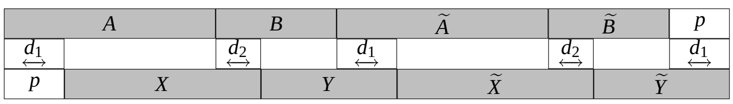


Fig. 2. Equation $AB\widetilde{A}\widetilde{B} \equiv XY\widetilde{X}\widetilde{Y}$.

Let d_1 be the delay between A and X, and d_2 the one between B and Y. We will suppose in addition that $|A| > |d_1| \geq 0$ and $|B| > |d_2| \geq 0$. Clearly, the overlapping of $AB\widetilde{A}\widetilde{B}$ and $XY\widetilde{X}\widetilde{Y}$ is completely determined by A, B, $|X|$ and d_1. Indeed, by construction, the delay d_2 between B and Y is defined by $d_2 = d_1 + |X| - |A|$ and $|Y| = |A| + |B| - |X|$ since $|AB| = |XY|$. We are interested in the particular cases where $d_1d_2 \geq 0$. Without loss of generality, we may restrict our study to the case $d_1, d_2 \geq 0$, since the case $d_1, d_2 \leq 0$ is obtained from the first one by taking the mirror image of the equation $AB\widetilde{A}\widetilde{B} \equiv XY\widetilde{X}\widetilde{Y}$. The overall situation is completely described by the following set of eight overlapping relations:

$$(R)\quad \begin{array}{ll} \text{(a)} & (A,X,d_1) \in \mathcal{R} \text{ and } (\widetilde{A},\widetilde{X},d_1) \in \mathcal{R}, \\ \text{(b)} & (X,B,|X|-d_2) \in \mathcal{R} \text{ and } (\widetilde{X},\widetilde{B},|X|-d_2) \in \mathcal{R}, \\ \text{(c)} & (B,Y,d_2) \in \mathcal{R} \text{ and } (\widetilde{B},\widetilde{Y},d_2) \in \mathcal{R}, \\ \text{(d)} & (Y,\widetilde{A},|Y|-d_1) \in \mathcal{R} \text{ and } (\widetilde{Y},A,|Y|-d_1) \in \mathcal{R}. \end{array}$$

For practical reasons, it is convenient to work with this decomposition instead of the general equation $AB\widetilde{A}\widetilde{B} \equiv XY\widetilde{X}\widetilde{Y}$. We set $\delta = d_1 + d_2$ and start with a general result on periodicity of A, B, X and Y.

Proposition 18 *If the length of A (resp. B, X, Y) is strictly greater than δ then A (resp. B, X, Y) has period δ.*

Proof. First, note that we have $|X| = |A| + d_2 - d_1$. By the overlapping relations decomposition, $(A,X,d_1) \in \mathcal{R}$ and $(\widetilde{A},\widetilde{X},d_1) \in \mathcal{R}$. By Lemma 13,

$$(\widetilde{A},\widetilde{X},d_1) \in \mathcal{R} \Longleftrightarrow (X,A,|X|-|A|+d_1) \in \mathcal{R} \Longleftrightarrow (X,A,d_2) \in \mathcal{R}.$$

Then by Lemma 15, $\delta = d_1 + d_2$ is a period of A if $|A| > \delta$ and a period of X if $|X| > \delta$. The argument is exactly the same for B and Y. $\square$

Now, let A_1, A_2 be the prefix and the suffix of A of length d_1 and B_1, B_2 be the prefix and the suffix of B of length d_2, that is

$$A = A_1 u = v A_2, \quad \text{and} \quad B = B_1 s = t B_2.$$

Observe that with those notations, we can write X and Y as

$$Y = \widetilde{A_1} u' = v' \widetilde{A_2} \quad \text{and} \quad X = B_2 s' = t' B_1,$$

as shown in Figure 3.

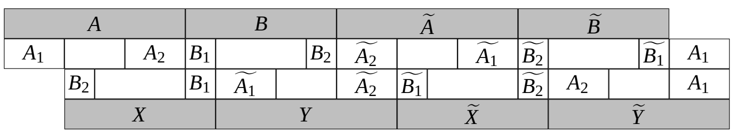


Fig. 3. Finer factorization of $AB\widetilde{A}\widetilde{B}$

Recall that the fractional power of a word $w \in \Sigma^*$ is defined as $w^r = w^{\lfloor r \rfloor} p$ where r is a rational such that $r|w| \in \mathbb{N}$ and p is the prefix of w of length $(r - \lfloor r \rfloor)|w|$. Then using Proposition 18, we can write

$$
\begin{aligned}
A &= (A_1B_2)^{\frac{|A|}{\delta}} \quad \text{and} \quad \widetilde{A} = (\widetilde{A_2}\widetilde{B_1})^{\frac{|A|}{\delta}};\\
B &= (B_1\widetilde{A_1})^{\frac{|B|}{\delta}} \quad \text{and} \quad \widetilde{B} = (\widetilde{B_2}A_2)^{\frac{|B|}{\delta}};\\
X &= (B_2A_1)^{\frac{|X|}{\delta}} \quad \text{and} \quad \widetilde{X} = (\widetilde{B_1}\widetilde{A_2})^{\frac{|X|}{\delta}};\\
Y &= (\widetilde{A_1}B_1)^{\frac{|Y|}{\delta}} \quad \text{and} \quad \widetilde{B} = (A_2\widetilde{B_2})^{\frac{|Y|}{\delta}};
\end{aligned}
$$

to obtain the following proposition.

Proposition 19 *Let A, B, X, $Y \in \Sigma^*$, such that $AB\widetilde{A}\widetilde{B} \equiv XY\widetilde{X}\widetilde{Y}$. Let A_1, A_2, B_1 and B_2 be as above. Let $d_1 = |A_1| = |A_2|$, $d_2 = |B_1| = |B_2|$ and $\delta = d_1 + d_2$. Then the following properties hold:*

(i) *if $|A| \geq \delta$, then A_1B_2 is conjuguate to A_2B_1 with delay $|A| + d_2 \mod \delta$,*
(ii) *if $|B| \geq \delta$, then $A_2\widetilde{B_2}$ is conjuguate to $A_1\widetilde{B_1}$ with delay $|B| + d_1 \mod \delta$.*

Proof. (i) Suppose that $|A| \equiv k \mod \delta$. Since $A = (A_1B_2)^{\frac{|A|}{\delta}}$, $\widetilde{A} = (\widetilde{A_2}\widetilde{B_1})^{\frac{|A|}{\delta}}$ and $|A| \geq \delta$, then A_1B_2 is conjugate to B_1A_2 with delay k. We also have that B_1A_2 is conjuguate to A_2B_1 with delay d_2. Then by Lemma 2, we conclude that A_1B_2 is conjuguate to A_2B_1 with delay $(k + d_2) \bmod \delta = (|A| + d_2) \mod \delta$.

(ii) Suppose that $|B| \equiv l \mod \delta$. Since $B = (B_1\widetilde{A_1})^{\frac{|B|}{\delta}}$, $\widetilde{B} = (\widetilde{B_2}A_2)^{\frac{|B|}{\delta}}$ and $|B| \geq \delta$, then $B_1\widetilde{A_1}$ is conjuguate to $\widetilde{A_2}B_2$ with delay l. Since $|A_2| = d_1$, we also have that $\widetilde{A_2}B_2$ is conjuguate to $B_2\widetilde{A_2}$ with delay d_1. By Lemma 2, we obtain that $B_1\widetilde{A_1}$ is conjuguate to $B_2\widetilde{A_2}$ with delay $(l + d_1) \bmod \delta$. Finally, applying Lemma 13, we conclude that $A_2\widetilde{B_2}$ is conjuguate to $A_1\widetilde{B_1}$ with delay $(l + d_1) \bmod \delta = (|A| + d_1) \mod \delta$. □

Several results follow from this proposition since it makes the conditions of Proposition 17 satisfied. Due to lack of space we only mention some of them, that we used in connection with the problem of determining a class of tilings of the plane.

Corollary 20 *If $|A| + d_2 \equiv 0 \mod \delta$ or $|B| + d_1 \equiv 0 \mod \delta$, then $A_1 = A_2$ and $B_1 = B_2$.*

Corollary 21 *If $|A|$, $|B| \geq \delta$, $|A| + d_2 \equiv 0 \mod \delta$ and $|B| + d_1 \equiv 1 \mod \delta$, then $AB\widetilde{A}\widetilde{B} \equiv XY\widetilde{X}\widetilde{Y} = \alpha^n$.*

Corollary 22 *Assume that $d_1 = d_2 = d$, i.e. $|A| = |B| = |X| = |Y| = \ell$ and $\delta = 2d$. Let $\ell \geq \delta$ and $\ell \equiv n \mod \delta$. If $2|d - n| \leq d$ then A_1, A_2, B_1 and B_2 have period $2|d - n|$.*

Characterization of double squares. A polyomino P is represented by a word $b(P)$ on the 4-letter alphabet $\mathcal{E} = \{a, b, \overline{a}, \overline{b}\}$ encoding its contour. There is a natural involution defined on $\mathcal{E}$, by $\overline{\cdot} : a \mapsto \overline{a}; b \mapsto \overline{b}$, consisting in swapping the elementary directions. The composition $\widehat{\cdot} = \overline{\cdot} \circ \widetilde{\cdot}$ is an antimorphism interpreted as follows: if $w \in \mathcal{E}^*$ is a path, then $\widehat{w}$ is the same path traversed in the opposite direction. Deciding if a polyomino P tiles the plane by translation amounts to check if the circular word $b(P)$ can be factorized as $b(P) \equiv ABC\widehat{A}\widehat{B}\widehat{C}$, where at most one variable is empty [4]. P is called a *square* if one variable is empty, otherwise it is called an *hexagon*. Moreover, although some polyominoes admit $O(|b(P)|)$ distinct hexagon factorizations, it has been conjectured in [15] that squares admit at most two distinct factorizations. Squares having exactly two distinct factorizations yield distinct tilings as illustrated in Figure 1.

All results of Section 4 obtained from equations of type $AB\widetilde{A}\widetilde{B} \equiv XY\widetilde{X}\widetilde{Y}$, are also valid for equations of the form $AB\widehat{A}\widehat{B} \equiv XY\widehat{X}\widehat{Y}$ by replacing $\widetilde{\cdot}$ with $\widehat{\cdot}$ everywhere. Since the problem of characterizing double squares requires solving equations of this last form, it provides additional constraints for finding them. For instance, it has been conjectured in [15] that the boundary of prime double squares always satisfies the equation

$$AB\widehat{A}\widehat{B} \equiv XY\widehat{X}\widehat{Y}, \quad \text{for } A,B,X,Y \in \mathrm{Pal}(\mathcal{E}^*). \tag{2}$$

In this specific case, by combining the palindromic conditions with the $\widehat{\cdot}$ version of Propostion 19, we obtain the following proposition:

Proposition 23 *Assume that* $|A|$, $|B|$, $|X|$ *and* $|Y|$ *are greater than* δ *and let* $k = |A| + d_2 \mod \delta$, $l = |B| + d_1 \mod \delta$. *Then* $k + l \neq d_1 + d_2$.

Incidentally, an interesting first step in tackling the general problem is to describe all polyominoes whose boundary words satisfy Equation (2). If we fix $|A| = |B| = |X| = |Y| = 43$, and consider all possible delays d such that $1 \leq d \leq 42$ in Equation (2), then prime double squares are obtained only when

$$d = 1,2,4,5,6,9,14,17,26,29,34,37,38,39,41,42.$$

Also, note that each double square obtained with delay d is isometric to the one obtained with delay $43 - d$. The eight polyominoes in question are illustrated in Figure 4. A description of two infinite families of double squares can be found in [6].

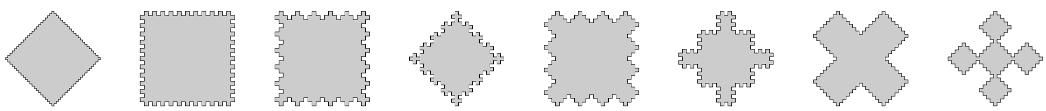


Fig. 4. Prime double squares with $|A| = |B| = |X| = |Y| = 43$.

References

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

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

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

[4] D. Beauquier, M. Nivat, On Translating one Polyomino to Tile the Plane, Discrete Comput. Geom. 6 (1991) 575–592.

[5] A. Blondin Massé, S. Brlek, S. Labbé, Palindromic lacunas of the Thue-Morse word, Proc. GASCOM 2008 (16-20 June 2008, Bibbiena, Arezzo-Italia) 53–67

[6] A. Blondin Massé, S. Brlek, A. Garon, S. Labbé, Christoffel and Fibonacci tiles, Proc. DGCI 2009 (Septeber 30-October 2, 2009, Montreal,Canada) LNCS (To appear)

[7] A. Blondin Massé, S. Brlek, A. Frosini, S. Labbé, S. Rinaldi, Reconstructing words from a fixed palindromic length sequence, Proc. TCS 2008, 5th IFIP Int. Conf. on Theoretical Computer Science (Springer) IFIP 273 (2008) 101–114.

[8] S. Brlek, J.-O. Lachaud, X. Provençal, C. Reutenauer, Lyndon + Christoffel = Digitally Convex, Pattern Recognition (2009) In press, available at doi:10.1016/j.patcog.2008.11.010

[9] S. Brlek, X. Provençal, J.-M Fédou, On the Tiling by Translation Problem, Discrete Applied Mathematics **157**(3) (2009) 464–475

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

[11] S. Labbé, Propriétés combinatoires des f-palindromes, Mémoire de maîtrise en Mathématiques, Montréal, UQAM, 2008, 115 p.

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

[13] M. Lothaire, Applied combinatorics on words, Cambridge University Press, 2002.

[14] G. Paquin, Mots équilibrés et mots lisses, thèse de doctorat en mathématiques, Université du Québec à Montréal, D1706 2008.

[15] X. Provençal, Combinatoires des mots, géométrie discrète et pavages, Ph.D. thesis, Université du Québec à Montréal, D1715 2008.

[16] R. Simpson, Disjoint covering systems of rational Beatty sequences, Discr. Math., 185 vol. 92 (1991) 361–369.

[17] R. Simpson, Disjoint beatty sequences, Electronic journal of combinatorial number theory, vol. 4 (2004) p A12.