Theoretical Computer Science 262 (2001) 269–284
Theoretical Computer Science
www.elsevier.com/locate/tcs

Recurrence and periodicity in infinite words from local periods

J.-P. Duval[a, *], F. Mignosi[b], A. Restivo[b]

[a] LIFAR-ABISS, Facult des Sciences, Universit de Rouen, F-76821 Mont-Saint-Aignan Cedex, France
[b] Dipartemento di Mathematica ed Applicazioni, Universit di Palermo via Archirafi, 34, Palermo, Italy

Received October 1999; revised April 2000; accepted May 2000
Communicated by D. Perrin

Abstract

We study recurrence and periodicity in infinite words by using local conditions. In particular, we give a characterization of recurrent, periodic and eventually periodic infinite words in terms of local periods. © 2001 Elsevier Science B.V. All rights reserved.

1. Introduction

The study of periodicity is a central topic in the combinatorics on words and presents some important applications in algebra, in formal language theory and in string searching algorithms.

M.P. Schutzenberger wrote in [12]: "Periodicity is an important property of words that is often used in applications of combinatorics on words. The main result concerning it are the theorem of Fine and Wilf and the Critical Factorization Theorem ..".

The Critical Factorization Theorem is a theorem that relates local periods to the global period of a finite word.

Among the applications of the Critical Factorization Theorem we just recall a famous string matching algorithm (cf. [2]).

The search for connections between "local" and "global" regularities of some objects is present in several fields of Mathematics, Physics and Computer Science (cf. for instance [1, 17, 19, 20, 12, Chapter 8; 4–7, 13, Chapter 8, 11]).

In this paper we characterize recurrent, periodic and eventually periodic infinite words in terms of local periods, where the definition of local period is the same as that of

* Corresponding author. Tel.: +33-235-146610; fax: +33-235-146763.
E-mail addresses: jean-pierre.duval@dir.univ-rouen.fr (J.-P. Duval), mignosi@altair.math.unipa.it (F. Mignosi), restivo@altair.math.unipa.it (A. Restivo).

PII: S0304-3975(00)00204-8

the one used in the Critical Factorization Theorem. Indeed the Critical Factorization Theorem will be one of the main tools used to acheive this result.

A similar characterization has been given in [16] settling an open question of J. Shallit, by using a different definition of local periodicity.

The paper is organized as follows: In the next section we introduce some notations and some basic results. In Section 3 we give the definition of local periods and we give a new proof of the Critical Factorization Theorem. The main results of the paper are given in Section 4. In Section 5 we give a new characterization of eventually periodic on-sided infinite words.

2. Preliminaries

Let us start with some basic definitions.

Let $w = a_1 a_2 \cdots a_n$ be a word of length n over the alphabet A. Any word of the form $a_i \cdots a_j$ with $1 \leqslant i \leqslant j \leqslant n$ is said to be a *factor* or a *block* of w.

A positive integer p is a *period* of w if for any integer i, $1 \leqslant i \leqslant n - p$ $a_{i+p} = a_i$. It is easy to see that p is a period of w if and only if for any integers i, j, $1 \leqslant i, j \leqslant n$ $i \equiv j$ (mod p) implies $a_i = a_j$. The smallest period p of w is called *the* period of w and it is denoted by $p(w)$.

From the definition it follows that, if v is a factor of w, then $p(v) \leqslant p(w)$.

The positive rational number $|w|/p(w)$ is called the *order* of w and it is denoted by $ord(w)$. If u is the prefix of length $p(w)$ of w, we can write $w = u^\rho$ where $\rho = ord(w)$, and we say that w is a *rational power* of u. Remark that a rational power u^ρ is defined only if $|u|\rho$ is an integer. For instance, $p(abaaba) = 3$, $ord(abaaba) = 2$ and the word $abaaba$ can be uniquely written $abaaba = (aba)^2$. As another example, $p(ababaaba) = 5$, $ord(ababaaba) = 1.6$ and the word $ababaaba$ can be written in a unique way as $ababaaba = (ababa)^{1.6}$.

A word v that is both a prefix and a suffix of another word w, with $v \neq w$, is called a *border* of w. It is easy to see that $|w| - |v|$ is a period of w and, conversely, if $p \leqslant |w|$ is a period of w then the prefix v of w of length $|w| - p$ is a border of w. The empty string ε is a border of any string w. If there exists a nonempty border v of w then w is called *bordered* otherwise it is called *unbordered.*

It is easy to verify that a word is unbordered if and only if $ord(w) = 1$, or, equivalently, if and only if $|w| = p(w)$.

We now state three lemmas that will be often used in the sequel.

Lemma 1. *Let w be a word having two periods p and q, with $q < p \leqslant |w|$. Then the suffix and the prefix of w of length $|w| - q$ have both period $p - q$.*

Proof. We prove only that the prefix of w of length $|w| - q$ has period $p - q$, the proof for the suffix beeing analogous.

Since $|w| - q \geqslant p - q$, we have to prove that

$$a_{i+p-q} = a_i, \quad i = 1, \ldots, n - p.$$

Let i be such that, $1 \leqslant i \leqslant n - p$. Thus, $1 \leqslant i + p - q \leqslant n - q$. Since w has period p, one has that $a_i = a_{i+p}$. Since w has period q and $1 \leqslant i + p - q \leqslant n - q$, one has that $a_{i+p-q} = a_{i+p}$. $\square$

Lemma 2. *Let u, v, w be words such that uv and vw have period p and $|v| \geqslant p$. Then the word uvw has period p.*

Proof. Let $uvw = a_1 \cdots a_n$, $u = a_1 \cdots a_l$, $v = a_{l+1} \cdots a_j$, $w = a_{j+1} \cdots a_n$. By hypothesis $j - l \geqslant p$. Let i be an integer $1 \leqslant i \leqslant n - p$. We have to prove that $a_i = a_{i+p}$. If $i \leqslant j - p$, since uv has period p, $a_i = a_{i+p}$. If $i > j - p$, since $j - l \geqslant p$, $i \geqslant l + 1$. Since vw has period p, $a_i = a_{i+p}$. $\square$

Lemma 3. *Suppose that w has period q and that there exists a factor v of w with $|v| \geqslant q$ that has period r, where r divides q. Then w has period r.*

Proof. Let $w = a_1 \cdots a_n$ and let $v = a_h \cdots a_k$ with $i \leqslant h < k \leqslant n$ and $k - h + 1 \geqslant q$.

Let us suppose that $i \equiv j \pmod r$. We have to prove that $a_i = a_j$. Since, by hypothesis $k - h + 1 \geqslant q$, for any integers i, j there exist i', j' with $h \leqslant i', j' \leqslant k$ such that $i \equiv i' \pmod q$ and $j \equiv j' \pmod q$.

Since $i \equiv j \pmod r$ and since r divides q, $i' \equiv j' \pmod r$.

Word w has period q and, so, $a_i = a_{i'}$ and $a_j = a_{j'}$. Word v has period r and, so, $a_{i'} = a_{j'}$ and the lemma is proved. $\square$

An important result on periodicity, which will be used in the sequel, is the Theorem of Fine and Wilf (cf. [9],[10]).

Theorem 1 (Fine and Wilf [19]). *Let w be a word having periods p and q, with $q \leqslant p$. If $|w| \geqslant p + q - \gcd(p, q)$, then w has also period* $\gcd(p, q)$.

3. Local versus global periodicities

In this section we introduce a notion of local period in terms of *repetitions*. In the general case, cf. [13, Chapter 8], the order of a repetition occurring in a word w can be an arbitrary rational number ρ. Moreover the repetition is in general referred to as "point" of the word w and it is important to consider the relative positions of the repetitions and that of the point of the word w at which the repetition is detected.

The main results of this paper refer to *central* repetitions *of order 2*. Similar results, which take into account repetitions that are not central and have order different from 2, can be found in [16] (cf. also [13, Chapter 8]).

In order to give the formal definitions, we first introduce the notion of pointed word. This is the appropriate notion to define local properties of a word.

Let $w = a_1 a_2 \cdots a_n$ be a word over the alphabet A. A *pointed word* is a pair $(a_1 \cdots a_i, a_{i+1} \cdots a_n)$, $1 \leqslant i < n$. The pointed word is also denoted by (w, i) and we refer to (w, i) as the word w at the *point* (or the *position*) i.

Let (x, y) be a pair of words. The pair (x, y) *matches* the pointed word (w, i), or simply matches the word w at the point i, if

$$A^* x \cap A^* a_1 \cdots a_i \neq \emptyset$$

and

$$y A^* \cap a_{i+1} \cdots a_n A^* \neq \emptyset.$$

Remark that the word $z = xy$ is not in general a factor of the word w and that the pair (x, y) specifies the relative position of the word z and the point i.

A word w contains a *repetition* of order ρ having as *center* the point (or position) i, or shortly a *central repetition* of order ρ at the point i, if there exists a non empty word z of order $ord(z) = \rho$ and a factorization $z = xy$, with $|x| = |y|$, such that the pair (x, y) matches w at the point i. This means that the point i is *central* with respect to the repetition z. The word z is called a central repetition of (w, i) and must have even length. This central repetition is *proper* (or *internal*) if x is a suffix of $a_1 \cdots a_i$ and y is a prefix of $a_{i+1} \cdots a_n$. It is *left external* if $a_1 \cdots a_i$ is a proper suffix of x. It is *right external* if $a_{i+1} \cdots a_n$ is a proper prefix of y.

Central repetitions of order 2 play an important role in this theory. By definition, a central repetition of order 2 at the point i of w is a word z of the form $z = x^2$ such that the pair (x, x) matches w at the point i. We say that w has a *square* having its center in the position i.

Example 1. Given the word

$$w = abaababaabaaba$$

the pointed word $(w, 8)$ is the pair

$$(abaababa, abaaba).$$

The pair (aba, aba) matches the pointed word $(w, 8)$ and, so, the word $abaaba$ is a central repetition of w at the point 8. It has order 2 and period 3. In the point 8 of w there is another central repetition of order 2, or a square having its center in it. It is the word aa and it has period 1. Both these repetitions are proper. The pointed word $(w, 7)$ is the pair

$$(abaabab, aabaaba).$$

The word $aabaababaabaabab$ is a central repetition of $(w, 7)$ of order 2 and period 8. It is both left and right external. Since the pair $(abab, aaba)$ matches w at the point 7, the word $ababaaba$ is a proper central repetition of $(w, 7)$ of order 1.6 and period 5.

As shown in the previous example a word can have at a given point different central repetitions of same order. We are interested, for a given order, to detect the central repetition of minimal period. This leads to the notion of minimal central repetition and of central local period.

For any real $\alpha > 1$, $c_\alpha(w,i)$ denotes the *central local period* (of order α) of the pointed word (w,i), defined by

$$c_\alpha(w,i) = \min\{p(z) \mid z \text{ is a central repetition of } (w,i) \text{ of order } \geqslant \alpha\}.$$

The central repetition z of (w,i) such that $p(z) = c_\alpha(w,i)$ is called the *minimal* central repetition (of order α) of w at the point i.

It is immediate to verify that, if $\alpha < \beta$, then $c_\alpha(w,i) \leqslant c_\beta(w,i)$ and that for any α and any $i \geqslant 1$, $c_\alpha(w,i) \leqslant p(w)$.

In the special case $\alpha = 2$ one has that

$$c_2(w,i) = \min\{|x| \mid x \neq \varepsilon \text{ and } x^2 \text{ is a square having its center in the position } i\}.$$

Example 2.

$w = abaababaabaaba$

	a	*b*	*a*	*a*	*b*	*a*	*b*	*a*	*a*	*b*	*a*	*a*	*b*	*a*
i		1	2	3	4	5	6	7	8	9	10	11	12	13
$c_2(w,i)$		2	3	1	5	2	2	8	1	3	3	1	3	2
$c_{1.6}(w,i)$		2	3	1	5	2	2	5	1	3	3	1	3	2

We denote by $P_\alpha(w)$ the maximum of the central local periods (of order α) of w:

$$P_\alpha(w) = \max\{c_\alpha(w,i) \mid 1 \leqslant i < |w|\}.$$

A point (or position) i is *critical* if $c_\alpha(w,i) = P_\alpha(w)$. We denote by $C_\alpha(w)$ the set of critical points of w:

$$C_\alpha(w) = \{i \mid 1 \leqslant i < |w| \text{ and } c_\alpha(w,i) = P_\alpha(w)\}.$$

We denote further by $Z_\alpha(w)$ and $S_\alpha(w)$ the minimum and the maximum, respectively, of the critical, points:

$$Z_\alpha(w) = \min C_\alpha(w),$$

$$S_\alpha(w) = \max C_\alpha(w).$$

Example 2 (*continued*)**.** For $w = abaababaabaaba$,
$P_2(w) = 8$, $C_2(w) = \{7\}$, $Z_2(w) = S_2(w) = 7$,
$P_{1.6}(w) = 5$, $C_{1.6}(w) = \{4, 7\}$, $Z_{1.6}(w) = 4$, $S_{1.6}(w) = 7$.

Remark that the notion of critical point introduced in this paper slightly differs from that used in the literature (cf. [12]), where a critical point i usually denotes a position where the local period of order 2, $c_2(w,i)$, is equal to the global period $p(w)$. The difference is motivated by the fact that we here take in account also repetitions of an arbitrary order $\alpha > 1$.

It is easy to verify that $c_\alpha(w,i) \leqslant p(w)$ for $\alpha > 1$ and $i = 1,\dots,|w|-1$, i.e. the central local periods are smaller than or equal to the period. On the other hand, if α is sufficiently large, i.e. $\alpha \geqslant 2|w|$, it is possible to prove that $c_\alpha(w,i) = p(w)$ for all i, as stated in particular in next proposition.

Proposition 1. *Let $k = \lceil \alpha/2 \rceil$. If the period of w is smaller than or equal to k then in every position i, one has $c_\alpha(w,i) = p(w)$. Hence, if $\alpha \geqslant 2|w|$ then every position is critical of order α.*

Proof. Let i, $1 \leqslant i < |w|$ be a position in w. If the central repetition of order α at the point i is both left and right external then $c_\alpha(w,i)$ is also a period of w and, consequently, $p(w) \leqslant c_\alpha(w,i)$ and, by previous remark, the thesis follows.

Suppose now that the central repetition of order α at the point i is either left or right internal or both. Suppose that it is left internal. We claim that $c_\alpha(w,i)$ divides $p(w)$. Indeed if $c_\alpha(w,i) = 1$ there is nothing to prove. Suppose that $c_\alpha(w,i) > 1$. In this case the part v of the central repetition of order α at the point i that is at the left of point i has length, by hypothesis, $\geqslant 2p(w)$. Factor v has period $p(w)$ and $c_\alpha(w,i)$. We can apply the Theorem of Fine and Wilf and obtain that it has period $d = \gcd(p(w), c_\alpha(w,i))$. But d cannot properly divide $c_\alpha(w,i))$ by the minimality of $c_\alpha(w,i))$. Hence $d = c_\alpha(w,i))$ and the claim is proved. We can now apply Lemma 3 and obtain that $c_\alpha(w,i)$ is also a period of w and, consequently, $p(w) \leqslant c_\alpha(w,i)$ and, by previous remark, the thesis follows. $\square$

The Critical Factorization Theorem in particular states that for $\alpha = 2$ there exists at least a point such that the central local period detected at this point coincides with the (global) period of the word, i.e. there exists an integer j, $1 \leqslant j < |w|$, such that $c_2(w,j) = p(w)$.

An important step in the proof of the Critical Factorization Theorem is the following proposition.

Proposition 2. *If $z = x^2$ is the square of minimal length having its center in position j of w, $1 \leqslant j < |w|$, then x is unbordered.*

Proof. If there exists a nonempty border t of x, i.e. t is both prefix and suffix of x, then t^2 is a square having its center in the position j of w that is shorter than x^2, contradicting the definition of x. $\square$

Theorem 2 (Critical Factorization Theorem). *Let w be a word having length $|w| \geqslant 2$.*

In every sequence of $l \geqslant \max(1, p(w) - 1)$ consecutive positions there is a critical one and, moreover, $P_2(w) = p(w)$.

Proof. The proof is by induction on $P_2(w)$.

Suppose that $P_2(w)=1$. Since for all natural number i, $1\leqslant i<|w|$, $c_2(w, i)=1$, then $a_i=a_{i+1}$. If $a=a_1$ and $n=|w|$, then w is of the form $w=a^n$, $p(w)=1=P_2(w)$ and all positions are critical.

Let us suppose that the statement of the proposition holds true for all words w' such that $P_2(w')\leqslant k-1$, $k>1$.

Let w be a word having $P_2(w)=k$

We prove the following properties:

(i) If j is a critical position and $j+1,\dots,j+l$ are not critical then $P_2(w)>l+1$.
(ii) If j is a critical position and $j-l,\dots,j-1$ are not critical then $P_2(w)>l+1$.
(iii) Every sequence of at least $P_2(w)-1$ consecutive positions contains a critical one.
(iv) $p(w)=P_2(w)$.

In order to prove (i) let us consider the word $u=a_{j+1}\cdots a_{j+l}a_{j+l+1}$. Since any central repetition at point $j+i$ of w is a repetition having its center at point i of u one has

$$c_2(u,i)\leqslant c_2(w,j+i) \quad i=1,\dots,l.$$

Since no position $j+i$ of w, with $i=1,\dots,l$, is a critical position, one has that

$$c_2(u,j+i) < k \quad i=1,\dots,l.$$

As a consequence, $c_2(u,i)<k$ for $i=1,\dots,l$, and then $P_2(u)<k$. By inductive hypothesis $p(u)=P_2(u)<k$.

Let $z=x^2$ be the square of minimal length having its center at position j of w. Since by hypothesis position j is critical, one has that $c_2(w,j)=P_2(w)=k$, and $|x|=k$.

By Proposition 2 the word x is unbordered. If x is a prefix of the word $u=a_{j+1}\cdots a_{j+l}a_{j+l+1}$ then $p(x)\leqslant p(u)<k$, that is a contradiction. Hence, u is a proper prefix of x and, consequently, $P_2(w)=k=|x|>|u|=l+1$.

The proof of (ii) is analogous by taking $u=a_{j-l}\cdots a_{j-1}$.

Property (iii) is an immediate consequence of properties (i) and (ii).

Let us now prove property (iv).

Let us recall that for any position j of w one has $c_2(w,j)\leqslant p(w)$, and, so, $P_2(w)\leqslant p(w)$.

Let i be a position such that $1\leqslant i<i+P_2(w)\leqslant|w|$. By property (iii) there exists a critical position j in the set $\{i,\dots,i+P_2(w)-1\}$. There exists then a square x^2 having its center at position j with $|x|=P_2(w)$. Note that $a_i\cdots a_{i+P_2(w)}$ is a factor of x^2, and, consequently, $a_i=a_{i+P_2(w)}$. Therefore, $P_2(w)$ is a period of w and then $P_2(w)\geqslant p(w)$. It follows that $P_2(w)=p(w)$.

The statement of the theorem is a consequence of previous properties. $\square$

Corollary 1. *Let w be a word of length $|w|\geqslant 2$ and $p(w)>1$. We have that $Z_2(w)<P_2(w)$, i.e., the central repetition at point $Z_2(w)$ is left external. We have also that $|w|-P_2(w)<S_2(w)$, i.e., the central repetition at point $S_2(w)$ is right external.*

Corollary 2. *Let $w = a_1 \cdots a_n$, be a word of length n. Given i, j, $1 \leqslant i < j < n$, if $c_2(w,h) < c_2(w,j)$ for any h such that $i \leqslant h < j$, then $c_2(w,j) > j - i + 1$.*

Proof. Let $v = a_i \cdots a_j$, $c_2(v,h) \leqslant c_2(w,h) < c_2(w,j)$ for $i \leqslant h < j$. According to Theorem 2 we have that $p(v) < c_2(w,j)$.

Let u^2 be the square of length $2c_2(w,j)$ having its center at position j of w. According to Proposition 2, u is an unbordered word. Hence u cannot be a suffix of v longer than $p(v)$. Therefore v is a proper suffix of u and $|u| = c_2(w,j) > j - i + 1$. □

In Example 2, $P_2(w)$ is, according to the theorem, exactly the period of w. Moreover, the unique critical point of w is 7 and it satisfies the conditions of Theorem 2. The same example shows that the theorem does not hold true for $\alpha = 1.6$. Indeed $P_{1.6}(w) = 5 \neq p(w) = 8$. The following example shows that the value $\alpha = 2$ is tight.

Example 3. For any real number $\varepsilon > 0$, consider the word $y_m = ba^{m-1}ba^mb$, with m a positive integer such that $2m/(m+1) \geqslant 2 - \varepsilon$. The unique critical point of order 2 is the point $m+1$, corresponding to the pair $(ba^{m-1}b, a^mb)$. The minimal central repetition of order 2 at such a point is the word $a^mba^{m-1}ba^mba^{m-1}b$, which has period $2m+1$, according to the Critical Factorization Theorem. However, the minimal central repetition of order $2-\varepsilon$ at the same point is the word $u = a^{m-1}ba^m$. Indeed, u has period $m+1$ and order $2m/(m+1) \geqslant 2 - \varepsilon$. It is easy to verify that such a point is also a critical point of order $2-\varepsilon$, and then

$$P_{2-\varepsilon}(ba^{m-1}ba^mb) = m + 1 \neq p(ba^{m-1}ba^mb) = 2m + 1.$$

Indeed, the word $a^mba^mba^m$, $m \geqslant 1$, has period $m+1$ and exactly four critical points, $m, m+1, 2m+1$ and $2m+2$, corresponding to the pairs (a^m, ba^mba^m), (a^mb, a^mba^m), (a^mba^m, ba^m) and (a^mba^mb, a^m), respectively.

4. Recurrent and periodic infinite words

In this section we will consider applications of the results of previous section to the case of one-sided and two-sided infinite words.

One-sided infinite words are written as $x = x_0x_1, \ldots, x_i \in A$, where the index of the first letter of x is zero.

Two-sided infinite words are written as $x = \cdots x_{-1}x_0x_1, \ldots, x_i \in A$, and non positive positions naturally occurs in this case.

In the sequel, for "infinite words" without specifying, we will mean one-sided infinite word.

A one-sided infinite word $x = x_0x_1 \cdots$, is *periodic* if there exists a positive integer p such that $x_i = x_{i+p}$, for all $i \in \mathbb{N}$. The smallest p satisfying previous condition is called the period of x.

A one-sided infinite word $x = x_0x_1, \ldots,$ is *eventually periodic* if there exist two positive integers k, p such that $x_i = x_{i+p}$, for all $i \geqslant k$. An infinite word is *aperiodic* if it is not eventually periodic.

A one-sided infinite word x is *recurrent* if any factor occurring in x has an infinite number of occurrences.

Remark that a one-sided infinite word is periodic if and only if it is recurrent and eventually periodic.

A two-sided infinite word $x = \cdots x_{-1}x_0x_1 \cdots$ is *periodic* if there exists a positive integer p such that $x_i = x_{i+p}$, for all $i \in \mathbb{Z}$. The smallest p satisfying previous condition is called the period of x.

Concerning the notion of local period, the definitions of previous sections extend to one-sided and two-sided infinite words but there are some natural differences.

In the case of one-sided infinite words, for any order α, there could exist integers j such that there are no central repetitions of order α at position j. In this case the value of $c_\alpha(x, j)$ is $+\infty$.

Remark further that any central repetition cannot be right external.

As an example consider the one-sided word $x = x_0x_1x_2x_3 \cdots$ with $x_i \in \{a, b\}$ defined by $x_0 = a$ and for any $i \geqslant 1$, $x_i = b$ (i.e. $x = abbbbbbbb \cdots$). At the position 0 of x, for any $\alpha > 1$ there exists no central repetition of order α, and, consequently, $c_\alpha(x, 0) = +\infty$.

A more sophisticated example is the following one.

Example 4. Let f be the infinite word of Fibonacci, that is the limit of the sequence f_n, $n > 0$, $n \in \mathbb{N}$, of Fibonacci words defined inductively as $f_1 = a$, $f_2 = ab$ and $f_{n+1} = f_nf_{n-1}$ for $n \geqslant 2$. One has

$$f = abaababaabaababaababaabaababaabaababaababaabaababaababa \cdots.$$

For any position j, $c_2(f, j)$ is finite and the square of minimal length having its center in position j is external if and only if $j = F_n - 2$ for some Fibonacci number F_n as proved in next proposition.

Proposition 3. *In the infinite word of Fibonacci f, there exists a square having its center in any position and the square of minimal length having its center in position j is external if and only if $j = F_n - 2$ for some Fibonacci number F_n. Moreover, when $j = F_n - 2$, the minimal length $c_2(f, j)$ of a square having its center in position j is F_n.*

Proof. Recall that if F_n, $n \in \mathbb{N}$, is the sequence of Fibonacci numbers defined by $F_0 = 1$, $F_1 = 1$, $F_{n+1} = F_n + F_{n-1}$ for $n \geqslant 1$, one has that $|f_n| = F_n$ for any $n > 0$.

We will prove that for any $n \geqslant 2$ and for any position $j \leqslant F_n - 2$ one has that there exists a square having its center in j and the square of minimal length having its center in position j is external if and only if $j = F_k - 2$ for some Fibonacci number F_k, $k \leqslant n$. Moreover, when $j = F_k - 2$, the value of $c_2(f, j)$ is F_k.

The proof of this is by the induction on n. The base of the induction is easily verified for $n=2,3$.

Let us suppose previous statement is true for $n>3$ and let us prove it for $n+1$. By inductive hypothesis the statement is true for any j up to F_n-2.

We have that $f_{n+1}f_{n+1}$ is a prefix of f and $f_{n+1}=f_nf_{n-1}=f_{n-1}f_{n-2}f_{n-2}f_{n-3}$. Hence, f_{n-2},f_{n-2} matches position F_n-1. Moreover, since f_{n-2} is a prefix of $f_{n-3}f_{n+1}$, then in any position j, $F_n-1\leqslant j\leqslant F_n+F_{n-2}$ one has that there is a repetition of order 2 of length $2F_{n-2}$ centred in position j, i.e. $c_2(f,j)\leqslant F_{n-2}$.

Let us consider now j such that $F_n+F_{n-2}\leqslant j\leqslant F_n+F_{n-1}-2=F_{n+1}-2$. These positions belong to the occurrence of f_{n-1} of the prefix $f_{n+1}f_{n+1}=f_nf_{n-1}f_{n+1}$ of f. Since $f_{n-1}f_{n-1}$ is also a prefix of f and $f_nf_{n-1}f_{n-1}$ a prefix of $f_nf_{n-1}f_{n+1}$, the inductive hypothesis give us the information that in any position j with $F_n+F_{n-2}\leqslant j\leqslant F_{n+1}-2$ there exists an internal square having its center in j, with the exceptions of the position $\underline{j}=F_n+F_{n-1}-2=F_{n+1}-2$. Moreover, again by inductive hypothesis, we know that in both positions there is "almost" a square centred in it. More precisely there would be a square if the $(j+1)$th letter of f, would be equal to the $(\underline{j}-F_{n-1}+1)$th letter of f. And it is not difficult to prove that it is false. By inductive hypothesis there is no centred square in $\underline{j}$ of length $\leqslant 2F_{n-1}$. If there was an internal centred square, by a result in [18], it must have as length a fibonacci number, i.e F_n, because $F_{n+1}>j+1$. But again it is not difficult to prove that the $(\underline{j}+1)$th letter of f is different from the $(\underline{j}-F_n+1)$th letter of f, and, so, in $\underline{j}=F_{n+1}-2$ there are no internal centred squares. But, since $f_{n+1}f_{n+1}$ is a prefix of f, it is easy to see that $c_2(f,\underline{j})=F_{n+1}$, and this concludes the proof. $\square$

Also in the case of two-sided infinite words, "a fortiori" there could exist integers j such that there are no central repetitions of order α at position j, i.e. $c_\alpha(x,j)=+\infty$. However, in the case of two-sided infinite words all the central repetitions at every position j such that $c_\alpha(x,j)$ is finite, are internal. As an example consider the two-sided infinite word $y=\cdots y_{-2}y_{-1}y_0y_1y_2y_3\cdots$ with $y_i\in\{a,b\}$ defined by $y_0=a$ and for any $i\neq 0$, $y_i=b$ (i.e. $y=\cdots bbbbabbbb\cdots$). At the position 0 and at the position -1 of y, for any $\alpha>1$ there exist no central repetition of order α and for all other position there exists a square having its center in it.

The following proposition is an easy consequence of Lemma 3 and its proof is left to the reader.

Proposition 4. *Suppose that w is an infinite word that has period q and that there exists a factor v of w with $|v|\geqslant q$ that has period d, where d divides q. Then w has period d.*

The periodicity of an infinite word x strongly depends on the fact that the sequence $c_\alpha(x,j)$ of local periods is bounded or not. Let

$$M_\alpha(x)=\sup\{c_\alpha(x,j)\mid j\in\mathbb{N}\}.$$

The following theorem is a consequence of the Critical Factorization Theorem.

Theorem 3. *An infinite word x is periodic if and only if $M_2(x)$ is finite. Moreover, the period of x is equal to $M_2(x)$.*

Proof. If x is periodic then trivially in any position there exists a square having its center in it and the sequence of local periods $(c_2(x,i))_{i\in\mathbb{N}}$ is bounded by the period P of x, i.e. $M_2(x)\leqslant P$. If $M_2(x)<P$ then take a factor v of length $2P$ of x. Clearly, P is a period of v and $P_2(v)\leqslant M_2(x)<P$. By the Critical Factorization Theorem $P_2(v)$ is a period of v and $P_2(v)<P$. By the Theorem of Fine and Wilf v has also period $d=\gcd(P,P_2(v))<P$. Since d divides P, by Proposition 4, d is also a period of x, contradicting the minimality of P.

Let us prove the "if" part of the proposition. Let Z be a position where the sequence $(c_2(x,i))_{i\in\mathbb{N}}$ reach its maximum $M_2(x)$. We have to prove that for any $i\in\mathbb{N}$ $x_i=x_{i+M_2(x)}$. Take the factor $v=x_r\cdots x_s$ of x where

$$r=min\{i,Z-M_2(x)\} \quad \text{and} \quad s=\max\{i+M_2(x),Z+M_2(x)\}.$$

In previous definition, in the case of one-sided infinite words, if $r<0$ we consider v defined as $v=x_0\cdots x_s$.

It is easy to see that position Z is also a critical position for v and its central local period is again $M_2(x)$. This implies that $M_2(x)=P_2(v)$. Hence, by the Critical Factorization Theorem, $p(v)=P_2(v)=M_2(x)$ and, so, $x_i=x_{i+M_2(x)}$. $\square$

The proof of the following theorem is analogous to the proof of Theorem 3 and it is left to the reader.

Theorem 4. *A two-sided infinite word x is periodic if and only if $M_2(x)$ is finite. Moreover, the period of x is equal to $M_2(x)$.*

In the Theorem 3 the constant 2 is tight. Indeed, for any $\varepsilon>0$, we can construct one- and two-sided infinite words that are nonperiodic and in any position have a central repetition of order $2-\varepsilon$, as shown in the next example.

Example 5. For any $\varepsilon>0$ let m be a positive integer such that for any $n\geqslant m$, $\varepsilon>2/n$. Let v_n be the finite word defined by $v_n=a^{n^2}b^n$, and let y_m the infinite word obtained concatenating $v_m,v_{m+1},v_{m+2},\ldots$.

In any position of y_m there is a central repetition of order $2-\varepsilon$. Indeed, if the square aa or the square bb are not central in position j then either j is a position between the concatenation of v_n and v_{n+1} for some n or j is the position between the sequence of a's and b's inside a word v_n for some n.

In the first case the pair $(a^{n^2-n}b^n,a^{n^2})$ matches position j and in the second case the pair $(b^{n-1}a^{n^2},b^na^{n^2-1})$ matches position j. Both $a^{n^2-n}b^na^{n^2}$ and $b^{n-1}a^{n^2}b^na^{n^2-1}$ have period n^2+n. The first has length $2n^2$ and the second $2n^2+2n-2 \geqslant 2n^2$. In both cases the order α of this repetition is greater than or equal to

$$\frac{2n^2}{n^2+n} = \frac{2n}{n+1} = 2 - \frac{2}{n+1} > 2 - \varepsilon.$$

One can define a two-sided infinite word $x_m = \cdots x_{-1}x_0x_1\cdots$ starting from the previously defined one-sided infinite word $y_m = y_0y_1\cdots$ by the rule $x_i = y_{|i|}$. It is easy to check that also in any position of x there is a central repetition of order $2-\varepsilon$.

Recall that a one-sided infinite word x is said to be *recurrent* if any factor occurring in x has an infinite number of occurrences. We have that x is recurrent if and only if any prefix of x has a second occurrence in x.

Theorem 5. *Let x be a one-sided infinite word. If x is recurrent then in any position there is a central repetition of order α, for any α such that $1<\alpha\leqslant 2$. Conversely, for any $\alpha\geqslant 2$, if in any position there is a central repetition of order α, then x is recurrent. In particular, x is recurrent if and only if in any position there is a central repetition of order 2.*

Proof. Let us suppose that x is recurrent. If we prove that in every position j there exists a square having its center in it then, a fortiori, there is a central repetition of order α for any $\alpha<2$. Let $k>0$ be the position where the prefix $x_0\cdots x_j$ occurs for the second time, i.e. $x_0\cdots x_j = x_k\cdots x_{k+j}$. If we set $v = x_{j+1}\cdots x_{k+j}$ then it is not difficult to see that (v,v) matches position j and, so, $z=v^2$ is a square having its center in position j.

Suppose now that in every position of x there exists a central repetition of order $\alpha\geqslant 2$. In particular, in every position of x there exists a square having its center in it. If the sequence of central local periods is bounded then, by Theorem 3, x is periodic and, so, recurrent. If the sequence of central local periods is not bounded then there exists a sequence $(j_i)_{i\in\mathbb{N}}$ of positions such that for any i, $c_2(x,j_i) > c_2(x,h)$ for any $h<j_i$. For any i consider the finite word $v = x_0\cdots x_{j_i+c_2(x,j_i)}$. It is not difficult to prove that position j_i is the least critical position for v and its central local period is again $c_2(x,j_i)$, i.e. $j_i = Z_2(v)$ and $c_2(x,j_i) = c_2(v,j_i)$. By Corollary 1 the minimal square $z=u^2$ having its center in j_i is left external in v. This means that the prefix $x_0\cdots x_{j_i}$ is a suffix of u and then it is also suffix of $x_0\cdots x_{j_i+c_2(x,j_i)}$, i.e. the prefix $x_0\cdots x_{j_i}$ occurs a second time in x. Since the sequence $(j_i)_{i\in\mathbb{N}}$ of positions is an increasing sequence, we found a sequence of prefixes of x of increasing length that have a second occurrence in x. This fact easily implies that any prefix of x has a second occurrence, i.e. x is recurrent. $\square$

The number 2 is tight in both directions of previous proposition. For any $\varepsilon>0$ it is known that there exists a one-sided recurrent infinite word x that is $(1+\varepsilon)$-power free (cf. [16, 3]). For $\alpha\geqslant 2+2\varepsilon$, the word x has no central repetition of order α in any position.

Conversely, for fixed m, the word $y = a^m baaaaa \cdots$, i.e. the word $y = y_0 y_1 \cdots$ with $y_i = a$ if $i \neq m$ and $y_m = b$ has in every position a central repetition of order $2 - (1/m)$ and it is not recurrent.

For two-way infinite words there are no similar characterizations. Any recurrent two-sided square-free infinite word has obviously no square having its center in any position. Notice that in a square-free recurrent one-sided infinite word, in any position the minimal central repetition of order 2 exists (according to Theorem 5) and it must be left external.

5. Eventually periodic infinite words

In this section we give a new characterization of eventually periodic one-sided infinite words.

This characterization property is similar to the characterization property of one-sided recurrent infinite words described in Theorem 5. Now we require something less (at any "large enough" position j there is a central repetition of order 2) and something more (the minimal central repetition of order 2 at position j is internal if j is "large enough").

Remark further that here we do not explicitly require, as in Theorem 3, that the sequence $c_2(x, j)$, $j \in \mathbb{N}$, is bounded. This condition is actually obtained (cf. Lemma 4 below) as a consequence of the existence of an internal repetition for any large enough position j.

Theorem 6. *A one-sided infinite word $x = x_0 x_1 x_2 \cdots$ is eventually periodic if and only if there exists a number k such that for any $j \geqslant k$ there exists a suffix of $x_0 \cdots x_j$ that is also a prefix $x_{j+1} x_{j+2} \cdots$, i.e. at any position $j \geqslant k$ there exists a proper central repetition of order* 2.

The proof of this theorem is based on the following lemma.

Lemma 4. *If there exists a number k such that in any position $j \geqslant k$ there exists a proper central repetition of order* 2, *then the sequence of local periods $(c_2(x, j))_{j \geqslant k}$ is bounded.*

Proof. The proof is by contradiction. Let us suppose that the sequence of central local periods at positions $j \geqslant k$ is not bounded. By hypothesis $(j - c_2(x, j)) > 0$ for any $j \geqslant k$.

Let j_1 be such that $(j_1 - c_2(x, j_1))$ assumes the minimal value between all $j \geqslant k$ and let j_2 be the least position greater than j_1 such that $c_2(x, j_2) > c_2(x, j_1)$.

Consider the word $v = x_{j_1 - c_2(x, j_1) + 1} x_{j_1 - c_2(x - j_1) + 2} \cdots x_{j_2 + c_2(x - j_2)}$.

In the sequel of this proof, for simplicity, with abuse of notation, we will denote by t the position $t - (j_1 - c_2(x, j_1))$ of v.

It is easy to verify that $c_2(x, j_1) = c_2(v, j_1)$.

Remark also that the minimal square having its center in position j_2 of v is not right external.

This square cannot be left external by the minimality of j_1 and the fact that, for any position t of v, $c_2(x,t) \geqslant c_2(v,t)$.

Since this square is not left external then $c_2(v,j_2) = c_2(x,j_2) > c_2(x,j_1) \geqslant c_2(v,j_1)$.

Since for any position t of v, $c_2(x,t) \geqslant c_2(v,t)$, and since j_2 is the least position greater than j_1 such that $c_2(x,j_2) > c_2(x,j_1) = c_2(v,j_1)$, one has that for any position t of v with $j_1 \leqslant t < j_2$, $c_2(v,t) \leqslant c_2(v,j_1)$. If $j_1 - c_2(x,j_1) + 1 \leqslant t < j_1$, we also have, by construction of v, that $c_2(v,t) \leqslant c_2(v,j_1)$.

By Corollary 2, $c_2(v,j_2) > j_2 - (j_1 - c_2(x,j_1) + 1) + 1$, i.e. $j_2 - c_2(v,j_2) < j_1 - c_2(x,j_1)$, contradicting the minimality of j_1. □

Proof of Theorem 6. If x is eventually periodic then we can write $x = wy$, where y is a one-sided infinite word that is periodic with period P. Hence, if we set $k = |w| + P$, it is easy to check that at any position $j \geqslant k$ of x there exists a central repetition of order 2 that is internal.

Let us prove the "if" part. Let us write $x = uy$, where $|u| = k$. Let us consider a position i of y and the corresponding position $i + k$ of x. Since y is a suffix of x one has that in any position $i \geqslant 0$ there exists a central repetition of order 2 and $c_2(y,i) \leqslant c_2(x,i+k)$. Hence, by Lemma 4 the sequence of local periods $(c_2(y,i))_{i \in \mathbb{N}}$ is bounded and, by Theorem 3, y is periodic. □

By the fact that an infinite word is periodic if and only if it is recurrent and eventually periodic, and by Theorem 5, one has

Corollary 3. *A one-sided infinite word x is periodic if and only if at any position there is a central repetition of order 2 and this repetition is external only for finitely many positions.*

Remark 1. Notice that, in previous theorem, one cannot bound the period of y as function of k, as shown by the one-sided infinite word $y_n = ba^n ba^n ba^n \cdots$ where n is any positive natural number. In this word the number k is 2 and this word has period n.

Example 5 (*continued*)**.** The word y_1 previously defined with $m = 1$ shows that the number 2 is tight in Lemma 4. Indeed for any ε it is easy to see that there exists a constant $k(\varepsilon)$ such that at any position $j \geqslant k(\varepsilon)$ there exists a central repetition of order $2 - \varepsilon$, but the sequence of local periods at positions $j \geqslant k(\varepsilon)$ is not bounded.

Same word y_1 show also that the constant 2 is tight in Theorem 6, because y_1 is not eventually periodic.

A more sophisticated example, showing that the constant 2 is tight in Theorem 6, is given by the infinite word of Fibonacci f defined in Example 4. The word f is not eventually periodic but it is possible to prove, with same techniques used in Proposition 3, that for any ε there exists a constant $k(\varepsilon)$ such that at any position

$j \geqslant k(\varepsilon)$ there exists a central repetition of order $2-\varepsilon$, and that the sequence of local periods at positions $j \geqslant k(\varepsilon)$ is bounded.

An analogous of Theorem 6 does not hold for two-sided infinite words, as shown by the next example.

Example 6. For any $\alpha > 1$ one can construct a non periodic two-sided infinite word $x_\alpha = \cdots x_{-1}x_0x_1\cdots$ such that at any position there exists a central repetition of order α.

Let us consider the sequence of all integer numbers

$0, -1, 1, -2, 2, -3, 3, \ldots -i, i, \ldots$.

Our construction inductively determines letters in the word x_α in order to have a central repetition at position n_j, where $n_j = (-1)^j \lceil (j/2) \rceil$ is the jth element of previous sequence.

Firstly let $k = \lceil (\alpha/2) \rceil$ be the smallest integer greater than or equal to $\alpha/2$ and set $x_{-k} = x_{-k+1} = \cdots = x_0 = x_1 = \cdots = x_{k-1} = a$ and $x_k = b$.

By construction at position 0 there is a central repetition of order α.

Suppose that we have fixed letters from position s_j up to position t_j such that at all positions $n_0, \ldots n_j$ there exists a central repetition of order α that is internal to the word $x_{s_j} \cdots x_{t_j}$. Since position n_{j+1} is adjacent to position n_{j-1} then $s_j \leqslant n_{j+1} \leqslant t_j$.

Let us denote $u = x_{s_j}x_{s_j+1}\cdots x_{n_{j+1}}$ and $v = x_{n_{j+1}+1}\cdots x_{t_j}$.

Let us suppose that $w = uv$ has period P and that a is the letter such that wa has period strictly greater than P. Set $x_{t_j+1} = a$.

Let us now assign letters from the position $s_{j+1} = n_{j+1} - k|vau|$ to the position $t_{j+1} = n_{j+1} + 1 + k|vau|$ so that

$$(vau)^k = x_{s_{j+1}}x_{s_{j+1}+1}\cdots x_{n_{j+1}} \quad \text{and} \quad (vau)^k = x_{n_{j+1}+1}\cdots x_{t_{j+1}}.$$

Notice that this assigment is compatible with the previous assigment and that at position n_{j+1} there exists a central repetition of order α. Notice further that, since $x_{s_{j+1}}x_{s_{j+1}+1}\cdots x_{t_{j+1}}$ has wa as factor, its period is strictly greater than the period P of w. Using this property it is not difficult to prove that the infinite word x_α is non periodic.

References

[1] D. Breslauer, T. Jiang, Z. Jiang, Rotations of periodic strings and short superstring, J. Algorithms 24 (2) (1997) 340–353.

[2] C. Crochemore, D. Perrin, Two-way string-matching, J. ACM 38 (3) (1991) 651–675.

[3] J.D. Currie, R.O. Shelton, Cantor sets and Dejean's conjecture, J. Automata, Languages and Combinatorics I 1996 (2) 113–127.

[4] N. Dolbilin, J. Lagarias, M. Senechal, Multiregular point systems, Discrete Comput. Geom. 4 (1998) 477–498.

[5] J.P. Duval, Périodes et répétitions des mots du monoïde libre, Theoret. Comput. Sci. 9 (1979) 17–26.

[6] J.P. Duval, Une caractérisation des fonctions périodiques, C.R. Acad Sci. Paris 289 (1979) 185–187.

[7] J.P. Duval, Contribution à la combinatoire du monoïde libre. These d'Etat, Université de Rouen, 1980.

[8] J.P. Duval, Périodes locales et propagation de périodes dans un mot, Theoret. Comput. Sci. 204 (1–2) (1998) 87–98.
[9] N.J. Fine, H.S. Wilf, Uniqueness theorem for periodic functions, Proc. Amer. Math. Soc. 16 (1965) 109–114.
[10] R. Giancarlo, F. Mignosi, Generalizations of the Periodicity Theorem of Fine and Wilf. CAAP94, Lecture Notes in Computer Science, vol. 787, Springer, Berlin, 1994, pp. 130–141.
[11] A. Lepisto, Relations between local and global periodicity of words, in: J. Wiedermann, P. vam Emde Boas, M. Nielsen (Eds.), Proc. ICALP'99, Lecture Notes in Computer Science, vol. 1644, Springer, Berlin, 1999, pp. 534–543.
[12] M. Lothaire, in: Combinatorics on Words, Encyclopedia of Mathematics and its Applications, vol. 17, Addison-Wesley, Reading, MA, 1883.
[13] M. Lothaire, Algebraic Combinatorics on Words, Cambridge University Press, Cambridge, to appear. A preliminary version is actually available at URL http://www-igm.univ-mlv.fr/ $\sim$ berstel/ Lothaire/index.html.
[14] R.C. Lyndon, P.E. Schupp, Combinatorial Group Theory, Springer, Berlin, 1977.
[15] R.C. Lyndon, M.P. Schutzenberger, The equation $a^m = b^n c^p$ in a free group, Michigan Math. J. 9 (4) (1962) 289–298.
[16] F. Mignosi, A. Restivo, S. Salemi, Periodicity and the golden ratio, Theoret. Comput. Sci. 204 (1–2) (1998) 87–98.
[17] C. Radin, Global order from local sources, Bull. Amer. Math. Soc. 25 (2) (1991) 335–364.
[18] P. Séébold, propriétés combinatoires des mots infinis engendrés par certains morphismes, Thèse de doctorat, Technical Report L.I.T.P., 1985, pp. 85–14.
[19] M. Senechal, Quasicrystals and Geometry, Cambridge University Press, Cambridge, 1995.
[20] L. Vuillon, Local configurations in a discrete plane, Bull. Belg. Math. Soc. 6 (1999) 625–636.