RAIRO-Theor. Inf. Appl. **42** (2008) 467–475
DOI: 10.1051/ita:2008010

Available online at:
www.rairo-ita.org

FROM BI-IDEALS TO PERIODICITY

Jānis Buls[1] and Aivars Lorencs[2]

Abstract. The necessary and sufficient conditions are extracted for periodicity of bi-ideals. They cover infinitely and finitely generated bi-ideals.

Mathematics Subject Classification. 68R15, 94A55, 68Q15.

1. Introduction

The periodicities are fundamental objects, due to their primary importance in word combinatorics [8,9] as well as in various applications. The study of periodicities is motivated by the needs of molecular biology [6] and computer science. Particularly, we mention here such fields as string matching algorithms [4], text compression [13] and cryptography [11].

In different areas of mathematics, people consider a lot of hierarchies which are typically used to classify some objects according to their complexity. Here we deal with the hierarchy

$$\mathfrak{B} \supset \mathfrak{P}, \qquad \text{where}$$

$\mathfrak{B}$ is the class of bi-ideals,
$\mathfrak{P}$ is the class of periodic words.
This hierarchy comes from combinatorics on words, where these classes are being investigated intensively (*cf.* [2,8–10]). Bi-ideal sequences have been considered, with different names, by several authors in algebra and combinatorics [1,3,7,12,14].

Every bi-ideal x is the limit of some bi-ideal sequence (v_i). This bi-ideal sequence can be represented uniquely by the sequence (u_i), where $v_0 = u_0$ and

Keywords and phrases. Periodic words, bi-ideals, the sequence generates the bi-ideal, finitely generated bi-ideals.

[1] Department of Mathematics, University of Latvia, Raiņa bulvāris 19, Rīga, 1586, Latvia; buls@fmf.lu.lv; web site: http://home.lanet.lv/~buls

[2] Institute of Electronics and Computer Science, Dzērbenes street 14, Rīga, 1006, Latvia; lorencs@edi.lv

$\forall i \geq 0\ v_{i+1} = v_i u_{i+1} v_i$. We characterize the periodic words through this representation. At first we give an exhaustive description (Th. 3.7) of periodicity for all classes of bi-ideals. Then for periodic bi-ideals we demonstrate if every u_i appears infinitely often then every u_i is a power of the certain word. This leads to the effective method for finitely generated bi-ideals to check whether the bi-ideals are periodic.

2. Preliminaries

In this section we present most of the notations and terminology used in this paper. Our terminology is more or less standard (*cf.* [10]) so that a specialist reader may wish to consult this section only if need arise.

Let A be a finite non-empty set and A^* the free monoid generated by A. The set A is also called an *alphabet*, its elements *letters* and those of A^* *finite words*. The role of the identity element is performed by the *empty word* which is denoted by λ. We set $A^+ = A^* \backslash \{\lambda\}$.

A word $w \in A^+$ can be written uniquely as a sequence of letters as $w = w_1 w_2 \dots w_l$, with $w_i \in A$, $1 \leq i \leq l$, $l > 0$. The integer l is called the *length* of w and denoted $|w|$. The length of λ is 0. We set $w^0 = \lambda$ and $\forall i\ w^{i+1} = w^i w$;

$$w^+ = \bigcup_{i=1}^{\infty} \{w^i\}, \qquad w^* = w^+ \cup \{\lambda\}.$$

A positive integer p is called a *period* of $w = w_1 w_2 \dots w_l$ if the following condition is satisfied:

$$1 \leq i \leq l - p \ \Rightarrow\ w_i = w_{i+p}.$$

We recall the important periodicity theorem due to Fine and Wilf [5]:

Theorem 2.1. *Let w be a word having periods p and q and denote by* $\gcd(p,q)$ *the greatest common divisor of p and q. If $|w| \geq p + q - \gcd(p,q)$, then w has also the period* $\gcd(p,q)$.

The word $w' \in A^*$ is called a *factor* (or *subword*) of $w \in A^*$ if there exist $u, v \in A^*$ such that $w = uw'v$. The word u (respectively v) is called a *prefix* (respectively a *suffix*) of w. The ordered triple (u, w', v) is called an *occurrence* of w' in w. The factor w' is called a *proper* factor if $w \neq w'$. We denote respectively by $F(w)$, $\mathrm{Pref}(w)$ and $\mathrm{Suff}(w)$ the sets of w factors, prefixes and suffixes.

An (indexed) infinite word x on the alphabet A is any total map $x : \mathbb{N} \to A$. We set for any $i \geq 0$, $x_i = x(i)$ and write

$$x = (x_i) = x_0 x_1 \dots x_n \dots$$

The set of all the infinite words over A is denoted by A^ω.

The word $w' \in A^*$ is a *factor* of $x \in A^\omega$ if there exist $u \in A^*$, $y \in A^\omega$ such that $x = uw'y$. The word u (respectively y) is called a *prefix* (respectively a *suffix*) of x. We denote respectively by $F(x)$, $\mathrm{Pref}(x)$ and $\mathrm{Suff}(x)$ the sets of x factors, prefixes and suffixes. For any $0 \le m \le n$, both $x[m,n]$ and $x[m,n+1)$ denote a factor $x_m x_{m+1} \dots x_n$. The indexed word $x[m,n]$ is called an *occurrence* of w' in x if $w' = x[m,n]$. The suffix $x_n x_{n+1} \dots x_{n+i} \dots$ is denoted by $x[n,\infty)$.

If $v \in A^+$ we denote by v^ω the infinite word $v^\omega = vv\dots v\dots$ This word v^ω is called a *periodic* word. The *concatenation* of $u = u_1 u_2 \dots u_k \in A^*$ and $x \in A^\omega$ is the infinite word

$$ux = u_1 u_2 \dots u_k x_0 x_1 \dots x_n \dots$$

A word x is called *ultimately periodic* if there exist words $u \in A^*$, $v \in A^+$ such that $x = uv^\omega$. In this case, $|u|$ and $|v|$ are called, respectively, an *anti-period* and a *period* of x.

A sequence of words of A^*

$$v_0, v_1, \dots, v_n, \dots$$

is called a *bi-ideal sequence* if $\forall i \ge 0\ (v_{i+1} \in v_i A^* v_i)$. The term "a bi-ideal sequence" is due to the fact that $\forall i \ge 0\ (v_i A^* v_i)$ is a bi-ideal of A^*.

Corollary 2.2. *Let (v_n) be a bi-ideal sequence. Then*

$$v_m \in \mathrm{Pref}(v_n) \cap \mathrm{Suff}(v_n)$$

for all $m \le n$.

A bi-ideal sequence $v_0, v_1, \dots, v_n, \dots$ is called *proper* if $v_0 \ne \lambda$. In the following the term bi-ideal sequence will be referred only to proper bi-ideal sequences.

If $v_0, v_1, \dots, v_n, \dots$ is a bi-ideal sequence, then there exists a unique sequence of words

$$u_0, u_1, \dots, u_n, \dots$$

such that

$$v_0 = u_0, \quad \forall i \ge 0\ (v_{i+1} = v_i u_{i+1} v_i)\,.$$

Let us consider $u, v \in A^\infty = A^* \cup A^\omega$. Then $d(u,v) = 0$ if $u = v$, otherwise

$$d(u,v) = 2^{-n},$$

where

$$n = \max\{\, |w| \mid w \in \mathrm{Pref}(u) \cap \mathrm{Pref}(v)\}\,.$$

It is called a *prefix metric.*

Let $v_0, v_1, \dots, v_n \dots$ be an infinite bi-ideal sequence, where $v_0 = u_0$ and $\forall i \ge 0\ (v_{i+1} = v_i u_{i+1} v_i)$. Since for all $i \ge 0$ the word v_i is a prefix of the next word v_{i+1} the sequence (v_i) converges, with respect to the prefix metric, to the infinite word $x \in A^\omega$

$$x = v_0 (u_1 v_0)(u_2 v_1) \dots (u_n v_{n-1}) \dots$$

This word x is called a *bi-ideal.* We say the sequence (u_i) *generates* the bi-ideal x.

Convention. Let x be a bi-ideal generated by (u_i), then $x = \lim_{i\to\infty} v_i$, where $v_0 = u_0$ and $v_{i+1} = v_i u_{i+1} v_i$. We adopt this notational convention henceforth.

Let x be an infinite word. A factor u of x is called *recurrent* if it occurs infinitely often in x. The word x is called *recurrent* when any of its factors is recurrent.

Proposition 2.3. (see, *e.g.*, [10]) *A word x is recurrent if and only if it is a bi-ideal.*

Lemma 2.4. (see, *e.g.*, [10]) *Let $x \in A^\omega$ be an ultimately periodic word. If x is recurrent, then x is periodic.*

Due to this lemma we can restrict ourselves. Therefore we investigate only the periodicity of bi-ideals and say nothing about ultimate periodicity.

3. The periodicity of bi-ideals

The following three lemmas are very easy, but they turn out to be extremely useful:

Lemma 3.1. *If $x = w^\omega$ and T is the minimal period of the word x, then $T\backslash|w|$, i.e. T divides $|w|$.*

Proof. Let $n = T|w|$, then both T and $|w|$ are periods of the word $x[0,n)$. Hence (Th. 2.1) $t = \gcd(T, |w|)$ is a period of $x[0,n)$. Now we have

$$\forall i\ x[0,n) = x[ni, n(i+1)).$$

Therefore t is a period of x. Since T is the minimal period of the word x, then $t \geq T \geq \gcd(T,|w|) = t$. Hence $T = \gcd(T,|w|)$, thereby $T\backslash|w|$. □

Lemma 3.2. *If $x = w^\omega = uvy$ and $|w| = |v|$, then $vy = y = v^\omega$.*

Proof. Let $|w| = t$ and $|u| = k+1$, then $v = x_{k+1}x_{k+2}\ldots x_{k+t}$, since $|v| = |w|$. We have $\forall i\ x_{i+t} = x_i$, therefore

$$\forall j \in \overline{1,t}\ \forall s\ \ x_{k+j} = x_{k+j+st}.$$ □

Lemma 3.3. *If $\exists u \in A^+\ ux = x \in A^\omega$, then a word x is periodic with the minimal period $T\backslash|u|$.*

Proof. Let $u = a_1a_2\ldots a_{t-1}$, where $\forall j\ a_j \in A$, and $y = ux$, then $\forall i\ x_i = y_{i+t}$. Let

$$y = ux = x\,.$$

Hence

$$\forall i\ y_i = x_i = y_{i+t}\,.$$

This means that y is periodic with a period t. Since $y = x$, then x is periodic with a period t too. Let T is the minimal period of x, then by Lemma 3.1 $T\backslash t$, *i.e.* $T\backslash|u|$. □

Corollary 3.4. *Let $|v|$ be the minimal period of $x = v^\omega$.*

$$\text{If} \qquad v = x[k, k+|v|) \qquad \text{then} \qquad |v|\backslash k.$$

Proof. If, for any k, $v = x[k, k+|v|)$, then (see Lem. 3.2)

$$x = x[0,k)v^\omega = x[0,k)x.$$

Hence by Lemma 3.3 $|v|\backslash|x[0,k)| = k$. □

Lemma 3.5. *If exists n such that $v_n u \in v^*$ and $\forall i \in \mathbb{Z}_+ \ (u_{n+i} \in uv^*)$, then*

$$\forall i \in \mathbb{N}\ (v_{n+i} \in v^* v_n)\,.$$

Proof. If $i = 0$ then $v_{n+i} = v_n = \lambda v_n \in v^* v_n$.

Further, we shall prove the lemma by induction on i, *i.e.*, suppose that $v_{n+i} \in v^* v_n$, namely,

$$\exists k \in \mathbb{N}\ (v_{n+i} = v^k v_n)\,.$$

By assumption, $v_n u \in v^*$ and $u_{n+i+1} \in uv^*$, *i.e.*

$$\exists l \in \mathbb{N}\ (v_n u = v^l)\ \wedge\ \exists m \in \mathbb{N}\ (u_{n+i+1} = uv^m).$$

Hence

$$\begin{aligned} v_{n+i+1} &= v_{n+i}u_{n+i+1}v_{n+i} = (v^k v_n)(uv^m)(v^k v_n) \\ &= v^k(v_n u)v^{m+k}v_n = v^k v^l v^{m+k} v_n \in v^* v_n\,. \end{aligned}$$

We have completed the inductive step. □

Lemma 3.6. *If t is the period of the bi-ideal x and $|v_n| \geq t$, then*

$$\forall i \in \mathbb{Z}_+\ u_{n+1}x = u_{n+i}x\,.$$

Proof. We have $v_{n+i} = v_{n+i-1}u_{n+i}v_{n+i-1}$. Hence, if $i \in \mathbb{Z}_+$ then (Cor. 2.2)

$$\forall i \in \mathbb{Z}_+\ \exists v'_i\ v_{n+i} = v_n v'_i v_n\,.$$

Now, by definition of x

$$\begin{aligned} x &= v_n u_{n+1} v_n \ldots \\ x &= v_{n+i}u_{n+i+1}v_{n+i}\ldots = v_n v'_i v_n u_{n+i+1} v_n \ldots \end{aligned}$$

By assumption, x is periodic, therefore

$$x = v^\omega, \qquad \text{where} \qquad |v| = t\,.$$

Since $v \in \mathrm{Pref}(v_n)$ then by Lemma 3.2

$$\begin{aligned} x &= v_n u_{n+1} x\,, \\ x &= v_n u_{n+i+1} x\,. \end{aligned}$$

Hence $\forall i \in \mathbb{Z}_+ \; x = v_n u_{n+i} x$. Thus $\forall i \in \mathbb{Z}_+ \; u_{n+1} x = u_{n+i} x$. □

Theorem 3.7. *A bi-ideal x is periodic if and only if*

$$\exists n \in \mathbb{N}\, \exists u \exists v \, (v_n u \in v^* \;\wedge\; \forall i \in \mathbb{Z}_+ \; u_{n+i} \in uv^*)\,.$$

Proof. $\Rightarrow$ Let T be the minimal period of the word x, then $\exists n \in \mathbb{N}\; |v_n| \geq T$. Thus by Lemma 3.6

$$\forall i \in \mathbb{Z}_+ \quad u_{n+1} x = u_{n+i} x\,.$$

Let u be the longest word of the set $\bigcap_{i=1}^{\infty} \mathrm{Pref}(u_{n+i})$ then

$$\forall i \in \mathbb{Z}_+ \exists u_i' \, (u_{n+i} = u u_i')\,.$$

Particularly, $\exists k \; u_{n+k} = u$. This means that

$$\forall i \in \mathbb{Z}_+ \quad u u_i' x = u_{n+i} x = u_{n+k} x = ux\,.$$

Thus

$$\forall i \in \mathbb{Z}_+ \quad u_i' x = x\,.$$

Hence by Lemma 3.3

$$\forall i \in \mathbb{Z}_+ \quad T \backslash |u_i'|\,.$$

Thereby

$$\forall i \in \mathbb{Z}_+ \quad u_i' \in v^*\,,$$

where $v = x[0, T)$. Thus

$$\forall i \in \mathbb{Z}_+ \quad u_{n+i} = u u_i' \in uv^*\,.$$

Note

$$x = v_n u_{n+1} v_n \ldots = v_n u u_1' v_n \ldots$$

Since $u_1' \in v^*$ and $v \in \mathrm{Pref}(v_n)$, then [Lemma 3.2] $x = v_n u x$. Hence [Lem. 3.3] $v_n u \in v^*$.

$\Leftarrow$ By Lemma 3.5

$$\forall i \in \mathbb{N}\, \exists k_i \in \mathbb{N}\; v_{n+i} = v^{k_i} v_n\,.$$

Since $\lim_{k\to\infty} |v_k| = \infty$ then $\lim_{i\to\infty} k_i = \infty$. Thus

$$x = \lim_{k\to\infty} v_k = \lim_{i\to\infty} v_{n+i} = \lim_{i\to\infty} v^{k_i} v_n = v^{\omega}\,.$$ □

4. Powers

Observation. If all $u_i \in w^*$ for some word $w \neq \lambda$, then the bi-ideal generated by (u_i) is periodic.

The following example demonstrates the converse is not true in general.

Example 4.1. Let x be the bi-ideal generated by (u_i), where

$$\begin{array}{rrcl} & u_0 & = & 0, \\ & u_1 & = & 1, \\ \forall i > 1 & u_i & = & 00100\,. \end{array}$$

Then

$$\begin{array}{rcl} v_0 & = & 0, \\ v_1 & = & 010, \\ v_2 & = & 010\,00100\,010, \\ v_3 & = & 01000100010\,00100\,01000100010, \\ . & . & . \end{array}$$

and $x = \lim_{i\to\infty} v_i = (0100)^\omega$. Thus x is periodic.

Nevertheless, if every u_j appears infinitely often in (u_i), then the converse is valid.

Theorem 4.2. *Let (u_i) be a sequence of words, which contains every u_j infinitely often. The bi-ideal x generated by (u_i) is periodic if and only if*

$$\exists w \forall i \; u_i \in w^*\,.$$

Proof. $\Rightarrow$ Let x be a periodic bi-ideal, then by Theorem 3.7

$$\exists n \in \mathbb{N}\, \exists u \exists v\, (v_n u \in v^* \;\wedge\; \forall i \in \mathbb{Z}_+ \; u_{n+i} \in uv^*)\,.$$

Hence by Lemma 3.5 $|v|$ is the period of x. Therefore we can assume that $|v|$ is the minimal period of x and $|u| < |v|$. Since the sequence (u_i) contains every u_j infinitely often then by Theorem 3.7 $\forall i \in \mathbb{N}\, (u_i \in uv^*)$.

Now suppose that $u_i = u$ for all $i < m$ but $u_m = uv^k$, where $k > 0$. Then there exist $\alpha \in \mathbb{Z}_+$ and y such that

$$x = u^\alpha v^k y\,.$$

(i) If $u = \lambda$ then $\forall i \; u_i \in v^*$.

(ii) Otherwise $u \neq \lambda$. Then (Corollary 3.4) $|v| \backslash \alpha |u|$. Hence, there exists $\beta \in \mathbb{Z}_+$ such that $\alpha|u| = \beta|v|$. Thus $x = v^\omega = u^\omega$. Contradiction, since $|u| < |v|$ and $|v|$ is the minimal period of x.

$\Leftarrow$ See Observation. □

Now we turn our attention to the problem of effectiveness.

Definition 4.3. Assume that (u_i) generates a bi-ideal x. The bi-ideal x is called *finitely generated* if

$$\exists m\,\forall i\,\forall j\ (i \equiv j\ (\mathrm{mod}\, m) \Rightarrow u_i = u_j).$$

In this situation, we say that the m–tuple $(u_0, u_1, \ldots, u_{m-1})$ generates the bi-ideal x.

Theorem 4.4. *A bi-ideal x generated by $(u_0, u_1, \ldots, u_{m-1})$ is periodic if and only if*

$$\exists w \forall i \in \overline{0, m-1}\ \ u_i \in w^*.$$

Proof. As a corollary from Definition 4.3 and Theorem 4.2. □

This theorem gives a method to generate nonperiodic bi-ideals. Let

$$(u_0, u_1, \ldots, u_{m-1})$$

be any m-tuple chosen at random. Let v be any shortest word from the set

$$\{u_0, u_1, \ldots, u_{m-1}\}$$

and w be the shortest prefix of v such that $v \in w^+$. If there exists u_i such that $u_i \notin w^*$ then the bi-ideal generated by $(u_0, u_1, \ldots, u_{m-1})$ is not periodic. This can be easily checked by a deterministic algorithm.

Acknowledgements. The useful suggestions of two referees are gratefully acknowledged.

References

[1] D.B. Bean, A.E. Ehrenfeucht and G. McNulty. Avoidable patterns in strings of symbols. *Pacific J. Math.* **85** (1979) 261–294.

[2] J. Berstel, J. Karhumäki. *Combinatorics on Words — A Tutorial.* TUCS Technical Report (No. 530, June) (2003).

[3] M. Coudrain and M.P. Schützenberger. Une condition de finitude des monoïdes finiment engendrés. *C.R. Acad. Sci. Paris*, Sér. A **262** (1966) 1149–1151.

[4] M. Crochemore and W. Rytter. Squares, cubes, and time-space efficient string searchinng. *Algorithmica* **13** (1995) 405–425.

[5] N.J. Fine, H.S. Wilf. (1965) Uniqueness theorem for periodic functions. *Proc. Amer. Math. Soc.* **16** (1965) 109–114.

[6] D. Gusfield. *Algorithms on Strings, Trees, and Sequences.* Cambridge University Press (1997).

[7] N. Jacobson. *Structure of Rings.* American Mathematical Society, Providence, RI (1964).

[8] M. Lothaire. *Combinatorics on Words.* Encyclopedia of Mathematics and its Applications, Vol. 17. Addison-Wesley, Reading, Massachusetts (1983).

[9] M. Lothaire. *Algebraic Combinatorics on Words.* Encyclopedia of Mathematics and its Applications, Vol. 90. Cambridge University Press, Cambridge (2002).

[10] Aldo de Luca, Stefano Varricchio. *Finiteness and Regularity in Semigroups and Formal Languages.* Springer-Verlag, Berlin, Heidelberg (1999).

[11] R.A. Rueppel. *Analysis and Design of Stream Ciphers.* Springer-Verlag, Berlin (1986).

[12] I. Simon. Infinite words and a theorem of Hindman. *Rev. Math. Appl.* **9** (1988) 97–104.

[13] J.A. Storer. *Data compression: methods and theory.* Computer Science Press, Rockville, MD (1988).

[14] A.I. Zimin. Blocking sets of terms. Матем. сб., **119**, 363–375 (Russian) (1982).