From Bi-ideals to Periodicity

Jānis Buls, Aivars Lorencs
Department of Mathematics, University of Latvia, Raiņa bulvāris 19,
Rīga, LV-1586, Latvia

buls@fmf.lu.lv

28.07.06

Abstract

The necessary and sufficient conditions are extracted for periodicity of bi-ideals. By the way two proper subclasses of uniformly recurrent words are introduced.

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 [14] and cryptography [12].
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 hierarchy

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

$\mathfrak{B}$ — the class of bi-ideals,
$\mathfrak{R}_{\mathfrak{u}}$ — the class of uniformly recurrent words,
$\mathfrak{P}$ — the class of periodic words.
This hierarchy comes from combinatorics on words, where these classes are being investigated intensively (cf. [2, 8, 9, 10]). Bi-ideal sequences have been considered, with different names, by several authors in algebra and combinatorics [1, 3, 7, 13, 17].
We refine the hierarchy $\mathfrak{B} \supset \mathfrak{R}_{\mathfrak{u}} \supset \mathfrak{P}$ to the chain

$$\mathfrak{B} \supset \mathfrak{R}_{\mathfrak{u}} \supset \mathfrak{B}_{\mathfrak{b}} \supset \mathfrak{B}_{\mathfrak{f}} \supset \mathfrak{P}\,, \qquad \text{where}$$

$\mathfrak{B}_{\mathfrak{b}}$ — the class of bounded bi-ideals,
$\mathfrak{B}_{\mathfrak{f}}$ — the class of finitely generated bi-ideals. So we localize the class of uniformly recurrent words by means of bi-ideals. Corollary 7 gives one method how the words of $\mathfrak{B}_{\mathfrak{f}}$ can be generated.
At first we characterize periodic finitely generated bi-ideals: we give one necessary condition [Corollary 8] in prefix–suffix terms and demonstrate this is not sufficient [Example 12]. Then we turn our attention to factors and prove sufficient and necessary condition [Theorem 21], and demonstrate this is not necessary for bounded bi-ideals [Example 34]. Lastly we extract exhaustive description [Theorem 37] of periodicity for all class of bi-ideals (more complicated of course).

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 identity element is performed by *empty word* and denoted by λ. We set $A^+ = A^* \setminus \{\lambda\}$.

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

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

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

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

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

Theorem 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 $|w| \ge p + q - \gcd(p,q)$, 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 *proper* factor if $w \ne w'$. We denote respectively by $\mathrm{F}(w)$, $\mathrm{Pref}(w)$ 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 $i \ge 0$, $x_i = x(i)$ 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 $\mathrm{F}(x)$, $\mathrm{Pref}(x)$ 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*.

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 "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. *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 $v_0, v_1, \ldots, v_n, \ldots$ is called *proper* if $v_0 \neq \lambda$. In the following the term bi-ideal sequence will be referred only to proper bi-ideal sequences.
If $v_0, v_1, \ldots, v_n, \ldots$ is a bi-ideal sequence, then there exists a unique sequence of words

$$u_0, u_1, \ldots, u_n, \ldots$$

such that

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

3 The Class of Finitely Generated Bi-ideals

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

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

where n is the length of the maximal common prefix of u and v. It is called a *prefix metric.*
Let $v_0, v_1, \ldots, v_n \ldots$ be an infinite bi-ideal sequence, where $v_0 = u_0$ and $\forall i \geq 0\ (v_{i+1} = v_i u_{i+1} v_i)$. Since for all $i \geq 0$ 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 $x \in A^\omega$

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

This word x is called a *bi-ideal.* We say the sequence (u_i) *generates* the bi-ideal x.
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 3. (see, e.g., [10]) *A word x is recurrent if and only if it is a bi-ideal.*

Lemma 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 ultimately periodicity.

Definition 5. *Let (u_i) generates a bi-ideal x. The bi-ideal x is called* finitely generated *if $\exists m\, \forall i\, \forall j\ (i \equiv j \pmod m \Rightarrow u_i = u_j)$. We say in this situation m–tuple $(u_0, u_1, \ldots, u_{m-1})$ generates the bi-ideal x.*

Theorem 6. *If $\bigcup_{i=0}^{m-1} \mathrm{Pref}(u_i)$ or $\bigcup_{i=0}^{m-1} \mathrm{Suff}(u_i)$ has at least two words with the same length then a bi-ideal generated by $(u_0, u_1, \ldots, u_{m-1})$ is not periodic.*

Proof. Let $x \in A^\omega$ be a bi-ideal generated by $(u_0, u_1, \ldots, u_{m-1})$.
(i) Let $\bigcup_{i=0}^{m-1} \mathrm{Pref}(u_i)$ has at least two words with one and the same length. Then there exist u_i, u_j such that $ua \in \mathrm{Pref}(u_i)$, $ub \in \mathrm{Pref}(u_j)$, where $u \in A^*$, $a, b \in A$ and $a \neq b$.
Let $T_0 = |ua|$ and $t > T_0$. Then we can choose n so great that $|v_n| \geq t$, where $v_{n+1} = v_n u_i v_n$. Hence $v_n ua \in \mathrm{Pref}(v_{n+1})$. Therefore $a = x_s$, where $s = |v_n u|$.
Since the tuple $(u_0, u_1, \ldots, u_{m-1})$ generates the bi-ideal x then $\exists k > n\ v_{k+1} = v_k u_j v_k$. Hence $v_k ub \in \mathrm{Pref}(v_{k+1})$. Therefore $b = x_\sigma$, where $\sigma = |v_k u|$. Since $k > n$ then $v_n \in \mathrm{Suff}(v_k)$. Thus $x_{s-t} = x_{\sigma - t}$ but $x_s = a \neq b = x_\sigma$. That means that t is not a period of x.
(ii) Let $\bigcup_{i=0}^{m-1} \mathrm{Suff}(u_i)$ has at least two words with one and the same length. Then there exist u_i, u_j such that $au \in \mathrm{Suff}(u_i)$, $bu \in \mathrm{Suff}(u_i)$, where $u \in A^*$, $a, b \in A$ and $a \neq b$.
Let $T_0 = |au|$ and $t > T_0$. Then we can choose n so great that $|v_n| \geq t$, where $v_{n+1} = v_n u_i v_n$. Hence there exists v' such that $v_n u_i = v' au$. Therefore $a = x_s$, where $s = |v'|$.

Since the tuple $(u_0, u_1, \ldots, u_{m-1})$ generates the bi-ideal x then $\exists k > n\ v_{k+1} = v_k u_j v_k$. Hence there exists v'' such that $v_k u_j = v'' bu$. Therefore $b = x_\sigma$, where $\sigma = |v''|$. Since $k > n$ then $v_n \in \mathrm{Pref}(v_k)$. Thus $x_{s+t} = x_{\sigma+t}$ but $x_s = a \neq b = x_\sigma$. That means that t is not a period of x.
(iii) Let us suppose that T is a period of x. Then $\forall n \in \mathbb{Z}_+\ nT$ is a period too. This denies (i) and (ii) as well.

Corollary 7. *Let A be an alphabet and every letter $a \in A$ is chosen with one and the same probability $p(a) = \frac{1}{|A|}$. Let p be a probability that a bi-ideal generated by $(u_0, u_1, \ldots, u_m)$ is ultimately periodic. If $\forall i\ |u_i| \geq n$ then $p \leq \frac{1}{|A|^{mn}}$.*

Remarks. (i) Let $A = \{0, 1\}$ and $m = n = 10$ then probability $p \leq \frac{1}{2^{100}}$. This is practically negligible value.
(ii) Let a tuple $(u_0, u_1, \ldots, u_m)$ has been generated. Let u be the longest word of this tuple. There is only one dubious situation by Theorem 6 if we like a bi-ideal that is not periodic. This happens if all words of the tuple $(u_0, u_1, \ldots, u_m)$ are prefixes and suffixes of u. This can be easy verified by deterministic algorithm. Thus we have indeed practical method how to generate a bi-ideal that is not periodic.

Corollary 8. *If a bi-ideal generated by $(u_0, u_1, \ldots, u_{m-1})$ is periodic then*

$$\forall i\ \forall j\ \ (u_i \in \mathrm{Pref}(u_j) \cap \mathrm{Suff}(u_j) \vee u_j \in \mathrm{Pref}(u_i) \cap \mathrm{Suff}(u_i))\,.$$

Corollary 9. *The class of periodic words $\mathfrak{P}$ is the proper subclass of the class of finitely generated bi-ideals $\mathfrak{B}_\mathfrak{f}$.*

The following two lemmata are very easy, but those turn out to be extremely useful:

Lemma 10. *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 [Theorem 1] $t = \gcd(T, |w|)$ is a period of $x[0, n)$. Now we have

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

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 11. *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

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

Example 12. *The bi-ideal generated by $(0, 010)$ is not periodic.*

Proof. (i) Let x be the bi-ideal generated by $(0, 010)$, and

$$\begin{array}{rcl} w_0 & = & 0, \\ w_1 & = & 0\ 010\ 0, \\ w_2 & = & 00100\ 0\ 00100, \\ . & . & . \\ w_{2n} & = & w_{2n-1} 0 w_{2n-1}, \\ w_{2n+1} & = & w_{2n} 010 w_{2n}, \\ . & . & . \end{array}$$

in other words $x = \lim_{k\to\infty} w_k$.
Let t be a period of x. Then $t > 3$, otherwise the period of w_2 must be less than or equal to 3. Contradiction. So we have a word w such that $|w| = t > 3$ and $x = w^\omega$.
(ii) Now choose n so large odd number that $t < |w_n|$. Then

$$x \quad = \quad w_n 0 w_n \dots$$

and

$$w_n \quad = \quad (uv)^s u,$$

where $s \geq 1$, $uv = w$ and $u \neq \lambda$. (If $u = \lambda$ or $v = \lambda$, then t divides $|w_n|$. We shall analyse this situation later.) From Lemma 11 we conclude that

$$x \quad = \quad w_n 0 w^\omega = w_n 0 (uv)^\omega = w_n 0 x\,.$$

Thus

$$(uv)^s uvuvu \dots \quad = \quad (uv)^s u 0 uvuv \dots$$

Hence

$$vuvu \dots \quad = \quad 0uvuv \dots$$

Since $u \neq \lambda$ and $u \in \mathrm{Pref}(x)$, then $u = 0u'$. Hence

$$vu \dots \quad = \quad 00u'v \dots$$

Thus, if $|v| \geq 2$, then $v = 00v'$.
(iii) Note that

$$w_n 0 x = x = w_n 0 w_n 010 w_n \dots$$

Therefore $x = w_n 010 w_n \dots$ and

$$(uv)^\omega = x = (uv)^s u 010 \dots$$

Hence $v = 01v''$ but $v = 00v'$. Contradiction.
(iv) It remains to check that $|v| \leq 1$. Note

$$u010 \dots = uvu \dots$$

Hence, if $|v| = 1$, we can conclude that the first letter of u is 1. Contradiction! Otherwise $v = \lambda$, then $u = 01u''$. Again contradiction, since $w_1 = 00100$, therefore the first two letters of u must be 00. Finally, if $u = \lambda$, then it remains to interchange u with v in the last two sentences of the proof.
Now turn our attention to Corollary 8. We have proved that condition

$$\forall i\, \forall j \;\; (u_i \in \mathrm{Pref}(u_j) \cap \mathrm{Suff}(u_j) \vee u_j \in \mathrm{Pref}(u_i) \cap \mathrm{Suff}(u_i))$$

is necessary for periodicity of finitely generated bi-ideals. Nevertheless Example 12 demonstrates that this condition is not sufficient.
The following lemma is crucial:

Lemma 13. *If a bi-ideal x generated by $(u_0, u_1, \dots, u_{m-1})$ is periodic, then*

$$\forall i\, \forall j\; u_i x = u_j x\,.$$

Proof. (i) Since x is a bi-ideal generated by $(u_0, u_1, \ldots, u_{m-1})$, then

$$x = \lim_{k\to\infty} v_k\,;$$

$$\text{where} \qquad v_0 = u_0, \quad v_{k+1} = v_k u_{k+1} v_k,$$
$$\text{and} \qquad u_{k+1} = u_i, \quad \text{if} \quad k+1 \equiv i (\operatorname{mod} m)\,.$$

Let t be a period of x and choose n so large that $t < |v_n|$. For every $i \in \overline{0, m-1}$ we can find $s_i > n$ such that

$$v_{s_i+1} = v_{s_i} u_i v_{s_i}\,.$$

Hence, by Corollary 2,

$$\forall i\, \exists v'_i\ v_{s_i} = v_n v'_i v_n\,.$$

Therefore

$$x = v_{s_i} u_i v_{s_i} \ldots = v_n v'_i v_n u_i v_n \ldots$$

(ii) We suppose that x is periodic, thereby

$$x = v^\omega,\ \text{where } v = x[0, t)\,.$$

Note $v \in \mathrm{Pref}(v_n)$, therefore [Lemma 11]

$$x = v_n v'_i x = v_n v'_i v_n u_i x\,.$$

Hence, $\forall i\ x = v_n u_i x$, thereby $\forall i\, \forall j\ u_i x = u_j x$.

Examples 14.

(i) First, we reexamine Example 12 in light of the above lemma. Let us suppose that a bi-ideal x generated by $(0, 010)$ is periodic then $0x = 010x$. This contradicts the fact that the first letter of x is not 1 but 0. The same arguments show that a bi-ideal generated by $(010, 0)$ is not periodic too.
(ii) Both bi-ideals generated by $((01)^{n-1}0, (01)^n 0)$ and $((01)^n 0, (01)^{n-1}0)$are not periodic. Indeed, if we suppose that a bi-ideal x generated by $((01)^{n-1}0, (01)^n 0)$ is periodic then by Lemma 13

$$(01)^{n-1}0x = (01)^n 0x = (01)^{n-1}010x\,.$$

Hence $x = 10x$. This contradicts the fact that the first letter of x is not 1 but 0.
The same arguments show that a bi-ideal generated by $((01)^n 0, (01)^{n-1}0)$is not periodic too.
We now present some useful observations concerning the periodicity. We start with the following lemma.

Lemma 15. *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 $u = a_1 a_2 \ldots a_{t-1}$, 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 10 $T \backslash t$, i.e. $T \backslash |u|$.

Proposition 16. *A word $x \in A^\omega$ is periodic if and only if*

$$\exists u \in A^+\ ux = x \in A^\omega.$$

Proof. $\Rightarrow$ If x is periodic then $\exists u \in A^+ \; x = u^\omega$. Hence $x = uu^\omega = ux$.
$\Leftarrow$ Lemma 15.

Corollary 17. *Let* $u \in \{u_0, u_1, \ldots, u_{m-1}\}$ *and* $|u| = \max\{|u_0|, |u_1|, \ldots, |u_{m-1}|\}$. *If T is the minimal period of a periodic bi-ideal x generated by* $(u_0, u_1, \ldots, u_{m-1})$ *then* $T < |u|$.

Proof. If only $u \neq \lambda$ then $u = u_0$ and $x = u^\omega$. Hence the minimal period $T < |u|$.
Otherwise, there exists $v \in \{u_0, u_1, \ldots, u_{m-1}\}$ such that $u \neq v \neq \lambda$. Now by Lemma 13 $ux = vx$. Hence

$$\exists v' \neq u \;\; u = vv' \,.$$

Thus $vv'x = vx$, therefore $v'x = x$. Since $0 < |v'|$ then $T \backslash |v'|$. This follows immediately from Lemma 15. Thereby $T < |v'| < |u|$.

Proposition 18. (see, e.g., [10]) *Let $u, v \in A^+$ be such that $uv = vu$. Then there exists $w \in A^+$ such that $u, v \in w^+$.*

Lemma 19. *Let $u, v \in A^+$ be such that $u^k v = vu^k$ for any positive integer k. Then there exists $w \in A^+$ such that $u, v \in w^+$.*

Proof. If $k = 1$ then it is Proposition 18. Now we assume that $k > 1$; by Proposition 18

$$\exists x \in A^+ \; (u^k, v \in x^+).$$

(i) If $|x| > |u^{k-1}|$, then $x = u^k$, because $|x^2| > |u^k|$ and $u^k \in x^+$. Hence $x \in u^+$, therefore $v \in u^+$. Thus $\exists w \in A^+ \; (u, v \in w^+)$; here $w = u$.
(ii) If $|x| \leq |u^{k-1}|$, then $l = \gcd(|x|, |u|)$ is period of u^k by Theorem 1. Let

$$w \in \mathrm{Pref}(u) \wedge |w| = l$$

then $u \in w^+$. Since $u^k \in x^+$ then $u^k = x^m$ for any m. Hence $x^m \in w^+$. Since $|w| \backslash |x|$ then $x \in w^+$. Therefore $v \in w^+$.

Theorem 20. *If x is a periodic bi-ideal generated by (u_0, u_1) then*

$$\exists w \; u_0, u_1 \in w^* .$$

Proof. Obviously, if $u_0 = u_1$ then a bi-ideal generated by (u_0, u_1) is periodic. Now suppose that $u_0 \neq u_1$. Then by Lemma 13 $u_0 x = u_1 x$.
(i) Let $u_0 \in \mathrm{Pref}(u_1)$, then $u_1 = u_0 u$, where $u \neq \lambda$, and $u_0 x = u_1 x = u_0 u x$, therefore $x = ux$. Thus

$$\begin{aligned} x &= u_0 u_1 \ldots = u_0 u_0 u \ldots \\ x &= ux = u u_0 u_0 \ldots \end{aligned}$$

Hence $u_0^2 u = u u_0^2$, and by Lemma 19 $\exists w \; u_0, u_1 \in w^*$.
(ii) Let $u_1 \in \mathrm{Pref}(u_0)$, then $u_0 = u_1 u$, where $u \neq \lambda$, and $u_0 x = u_1 u x$. Since $u_0 x = u_1 x$ then $u_1 u x = u_1 x$, therefore $x = ux$. Thus

$$\begin{aligned} x &= u_0 \ldots = u_1 u \ldots \\ x &= ux = u u_0 \ldots = u u_1 u \ldots \end{aligned}$$

Hence $u_1 u = u u_1$, and by Proposition 18 $\exists w \; u, u_1 \in w^*$. Since $u_0 = u_1 u$ then $u_0 \in w^*$.

Theorem 21. *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. Since x is a bi-ideal generated by $(u_0, u_1, \ldots, u_{m-1})$, then

$$x = \lim_{k\to\infty} v_k\,;$$

$$\text{where} \qquad v_0 = u_0, \quad v_{k+1} = v_k u_{k+1} v_k,$$
$$\text{and} \qquad u_{k+1} = u_i, \quad \text{if} \quad k+1 \equiv i (\bmod m)\,.$$

$\Rightarrow$ We have $u_0 x = u_1 x = \ldots = u_{m-1} x$ by Lemma 13.
(i) First, we shall prove that $\exists w\ \ u_0, u_1 \in w^*$.
a) If $u_1 = \lambda$ or $u_1 = u_0$, then $w = u_0$.
Now we shall consider the situation $\lambda \neq u_1 \neq u_0$.
b) Let $u_0 \in \mathrm{Pref}(u_1)$ then

$$\begin{aligned} u_1 &= u_0 u, \quad \text{where} \quad u \neq \lambda, \quad \text{and} \\ u_0 x &= u_1 x = u_0 u x, \text{ therefore } x = ux\,. \\ x &= u_0 u_1 \ldots = u_0 u_0 u \ldots \\ x &= ux = u u_0 u_0 \ldots \end{aligned}$$

Hence $u_0^2 u = u u_0^2$, and by Lemma 19 $\exists w\ u_0, u \in w^*$. Since $u_1 = u_0 u$ then $u_1 \in w^*$.
c) Let $u_1 \in \mathrm{Pref}(u_0)$ then

$$\begin{aligned} u_0 &= u_1 u, \quad \text{where} \quad u \neq \lambda, \quad \text{and} \\ u_1 x &= u_0 x = u_1 u x, \text{ therefore } x = ux\,. \\ x &= u_0 \ldots = u_1 u \ldots \\ x &= ux = u u_0 \ldots = u u_1 \ldots \end{aligned}$$

Hence $u_1 u = u u_1$, thereby [Proposition 18] $\exists w\ u, u_1 \in w^*$. Since $u_0 = u_1 u$, then $u_0 \in w^*$.
(ii) Further, we shall prove the theorem by induction on n, i.e., suppose that $\exists v \forall i \in \overline{0,n}\quad u_i \in v^*$. Let $u_n \neq u_{n+1} \neq \lambda$, otherwise $u_{n+1} \in v^*$.
a) Let $u_n \in \mathrm{Pref}(u_{n+1})$ then

$$\begin{aligned} u_{n+1} &= u_n u, \quad \text{where} \quad u \neq \lambda, \quad \text{and} \\ u_n x &= u_{n+1} x = u_n u x, \text{ therefore } x = ux\,. \\ x &= v_n u_{n+1} \ldots = v_n u_n u \ldots \\ x &= ux = u v_n u_n \ldots \end{aligned}$$

Hence $v_n u_n u = u v_n u_n$. We have by induction $\exists k\ v^k = v_n u_n$ and $k \geq 1$, since $u_0 \neq \lambda$. Thus $v^k u = u v^k$, and by Lemma 19 $\exists w\ \ v, u \in w^*$.
Thereby $v \in w^*$, and by induction $\forall i \in \overline{0,n}\ \ u_i \in v^*$. Hence $\forall i \in \overline{0,n} u_i \in w^*$. Since $u_{n+1} = u_n u$ and $u_n, u \in w^*$, then $u_{n+1} \in w^*$.
b) Let $u_{n+1} \in \mathrm{Pref}(u_n)$ then

$$\begin{aligned} u_n &= u_{n+1} u, \quad \text{where} \quad u \neq \lambda, \quad \text{and} \\ u_{n+1} x &= u_n x = u_{n+1} u x, \text{ therefore } x = ux\,. \\ x &= v_{n-1} u_n \ldots = v_{n-1} u_{n+1} u \ldots \\ x &= ux = u v_{n-1} u_{n+1} \ldots \end{aligned}$$

Hence $v_{n-1}u_{n+1}u = uv_{n-1}u_{n+1}$, therefore by Proposition 18 $\exists w_0\ v_{n-1}u_{n+1},u \in w_0^*$. We have by induction $\exists k\ v^k = v_{n-1},\ k \geq 1$, since $u_0 \neq \lambda$. Thus

$$|v_{n-1}u_n| = |v_{n-1}u_{n+1}u| > |v_{n-1}| + |u| \geq |v| + |w_0|$$

and $v_{n-1}u_n \in v^*$, $v_{n-1}u_n = v_{n-1}u_{n+1}u \in w_0^*$. This means that both $|v|$ and $|w_0|$ are periods of $v_{n-1}u_n$. Now by Theorem 1 $l = \gcd(|v|, |w_0|)$ is the period of $v_{n-1}u_n$. Let

$$w \in \mathrm{Pref}(v_{n-1})\ \wedge\ |w| = l$$

then $v_{n-1}u_n \in w^+$ because $l \backslash |v_{n-1}u_n|$.
The word $v_{n-1}u_n = u_{i_1}u_{i_2}\dots u_{i_\varkappa}$, where all $u_{i_s} \in \{u_0, u_1, \dots, u_n\}$, besides,

$$\forall j \in \overline{0,n}\ \exists \nu \in \overline{1,\varkappa}\quad u_j = u_{i_\nu}.$$

Since $\forall i \in \overline{0,n}\ u_i \in v^*$ then $\forall i \in \overline{0,n}\ l\backslash|u_i|$. Thus $\forall i \in \overline{0,n}\ u_i \in w^*$.
Finally, $u_n = u_{n+1}u$ and $u_n \in w^*$, therefore l is the period of u_{n+1}. Since $u \in w_0^*$ then $l\backslash|u|$. Hence $l\backslash|u_{n+1}|$. Thus $u_{n+1} \in w^*$.
This completes the induction.
$\Rightarrow$ Since $\forall n\ v_n \in w^*$ then $x = w^\omega$.

4 The Class of Bounded Bi-ideals

Definition 22. *Let (u_i) generates a bi-ideal x. The bi-ideal x is called* bounded *if $\exists l\, \forall i\ |u_i| \leq l$.*

Proposition 23. *The class of finitely generated bi-ideals $\mathfrak{B}_\mathfrak{f}$ is the proper subclass of the class of bounded bi-ideals $\mathfrak{B}_\mathfrak{b}$.*

Proof. Note $\mathrm{card}\{(u_i) \mid \forall i\ u_i \in \{0,1\}\} = \mathfrak{c}$ — the cardinality of the set of real numbers. Let (u_i), (v_i) be two different sequences of letters in the alphabet $\{0,1\}$ that generate bi-ideals (x_i), (y_i) respectively. Since $(x_i) \neq (y_i)$ then $\mathrm{card}\{\,(x_i) \mid$ there is a sequence (u_i) of letters in the alphabet $\{0,1\}$ that generate a bi-ideal $(x_i)\} = \mathrm{card}\{(u_i) \mid \forall i\ u_i \in \{0,1\}\} = \mathfrak{c}$.
Let $\mathfrak{U}_m = \{(u_0, u_1, \dots, u_{m-1}) \mid \forall i\ u_i \in \{0,1\}^*\}$ then $\mathrm{card} \bigcup_{m=1}^{\infty} \mathfrak{U}_m = \aleph_0$, where $\aleph_0$ is the first infinite cardinality. Therefore the cardinality of the set of all finitely generated bi-ideals in the alphabet $\{0,1\}$ is equal to $\aleph_0$. Since $\aleph_0 < \mathfrak{c}$ then $\mathfrak{B}_\mathfrak{f} \subset \mathfrak{B}_\mathfrak{b}$.
Let $w = u_1w_1v_1 = u_2w_2v_2$. We define *a meet* $w_1 \cap w_2$ as follows. If there exists an occurrence (u_3, w_3, v_3) of w_3 in word w such that $w = u_3w_3v_3$, where $|u_3| = \max(|u_1|, |u_2|)$, $|v_3| = \max(|v_1|, |v_2|)$, then $w_1 \cap w_2 = w_3$. Otherwise, $w_1 \cap w_2 = \lambda$.

Lemma 24. *Let (u_i) generates a bi-ideal x, $v_0 = u_0$, $\forall i\,(v_{i+1} = v_iu_{i+1}v_i)$ and $\forall i\,|u_i| \leq l$. If $v \in \mathrm{F}(x)$ and $|v| = 2|v_m| + l$ for some m, then $v_m \in \mathrm{F}(v)$.*

Proof. Since $v \in \mathrm{F}(x)$ then $v \in \mathrm{F}(v_n)$ but $v \notin \mathrm{F}(v_{n-1})$ for some n. Moreover, $v_{n-1}u_nv_{n-1} = v_n = v'vv''$ for some v', v''.
Since $v \notin \mathrm{F}(v_{n-1})$ then $u_n \cap v \neq \lambda$. Hence, $|v_{n-1} \cap v| \geq |v_m|$ because of $|u_n| \leq l$. Therefore $v_m \in \mathrm{F}(v)$ by Corollary 2.

Definition 25. *It is said a factor u of an infinite word x occurs syndetically in x if there exists an integer k such that in any factor of x of length k there is at least one occurrence of u. A word x is called uniformly recurrent, or with bounded gaps, when all its factors occur syndetically in x.*

Proposition 26. *Bounded bi-ideals are uniformly recurrent.*

Proof. Let x be a bounded bi-ideal generated by (u_i) then there exists l such that $\forall i\,|u_i| \le l$. Let $u \in \mathrm{F}(x)$, $v_0 = u_0$ and $\forall i\,(v_{i+1} = v_i u_{i+1} v_i)$ then there exists m such that $u \in \mathrm{F}(v_m)$.
Let $v \in \mathrm{F}(x)$ and $|v| = 2|v_m| + l$ then $v_m \in \mathrm{F}(v)$ by Lemma 24. Therefore $u \in \mathrm{F}(v)$. So the factor u of x occurs syndetically in x.

Let $\phi : A^* \to A^*$ be a nonerasing morphism (namely, $\phi(A^+) \subseteq A^+$) such that there exists a letter $a \in A$ such that

$$\phi(a) = au\,, \qquad \text{with} \qquad u \in A^+.$$

For all $n \ge 0$ one has

$$\phi^{n+1}(a) = \phi^n(au) = \phi^n(a)\phi^n(u)\,,$$

so that $\phi^n(a)$ is a proper prefix of $\phi^{n+1}(a)$. Thus the sequence $(\phi^n(a))$ converges to a limit denoted by $\phi^\omega(a)$, that is,

$$\phi^\omega(a) = \lim_{n\to\infty} \phi^n(a)\,.$$

One says that $x = \phi^\omega(a)$ is the infinite word obtained by *iterating the morphism* ϕ on the letter a. Moreover, one has $x = \phi(x)$, that is, x is a fixed point for ϕ.
Very famous infinite word is Thue–Morse word t on two letters

$$t = 0110\,1001\,1001\,0110\ldots$$

t can be introduced by iterating, on the letter 0, the morphism

$$\tau : \{0,1\}^* \to \{0,1\}^*, \quad \text{defined as} \quad \tau(0) = 01,\ \tau(1) = 10\,.$$

The word t was introduced by Thue in two papers [15, 16] of 1906 and 1912 and, subsequently, rediscovered by Morse [11] and several other authors. Thue–Morse word is uniformly recurrent (see, e.g., [10]).

Definition 27. *A factor u of a word $x \in A^\infty$ is called an overlapping factor of x if $u = avava$, with $a \in A$ and $v \in A^*$. We say that x is overlap-free, if x does not contain overlapping factors.*

Corollary 28. *Let $y \in A^\infty$. If $x\backslash y$ and $x = uvuvu$, where $u \ne \lambda$, then both x and y contain an overlapping factor.*

Proof. Since $u \ne \lambda$, then exist a letter $a \in A$ and a word $w \in A^*$ such that $u = aw$. Hence $x = (aw)v(aw)v(aw) = a(wv)a(wv)aw$. Thus x contains the overlapping factor $a(wv)a(wv)a$.

Proposition 29. (see, e.g., [10]) *The Thue — Morse word t is overlap-free.*

Lemma 30. *If $\overset{\infty}{\forall} i\ u_i = \lambda$ then a bi-ideal x generated by (u_i) is periodic.*

Proof. Let $v_0 = u_0$, $v_{i+1} = v_i u_{i+1} v_i$. Since $\overset{\infty}{\forall}\ i\ \ u_i = \lambda$ then $\exists n \forall i > n\ \ u_i = \lambda$. Hence $v_{n+1} = v_n u_{n+1} v_n = v_n^2$.
Further, we shall prove the lemma by induction on j, i.e., suppose that $v_{n+j} = v_n^k$, where $k = 2^j$, then $v_{n+j+1} = v_{n+j} u_{n+j+1} v_{n+j} = v_{n+j}^2 = v_n^{2k}$, where $2k = 2 \cdot 2^j = 2^{j+1}$.
Thus $x = \lim\limits_{i\to\infty} v_i = \lim\limits_{k\to\infty} v_n^k = v_n^\omega$.

Lemma 31. *If A is a finite alphabet then every bounded bi-ideal $x \in A^\omega$ contains an overlapping factor.*

Proof. Let $x \in A^\omega$ be a bi-ideal generated by the sequence $u_0, u_1, \ldots, u_n, \ldots$.
(i) If $\overset{\infty}{\forall}\ i\ \ u_i = \lambda$ then x is periodic [Lemma 30]. Therefore [Corollary 28] x contains an overlapping factor.
(ii) If $\overset{\infty}{\exists}\, i\ u_i \ne \lambda$, then $\exists i \exists j\ (i < j \wedge u_i = u_j)$, because A is finite and $\exists l \forall i\ |u_i| \le l$. Since x is the bi-ideal generated by the sequence $u_0, u_1, \ldots, u_n, \ldots$ then $x = \lim\limits_{n\to\infty} v_n$, where $v_0 = u_0$ and $v_{n+1} = v_n u_{n+1} v_n$. Hence by Corollary 2 $v_j = v_{j-1} u_j v_{j-1} = v' u_i v_{i-1} u_j v_{i-1} u_i v''$. Thus $u_i v_{i-1} u_j v_{i-1} u_i \backslash x$, therefore [Corollary 28] x contains an overlapping factor.

Proposition 32. *The Thue–Morse word t is not a bounded bi-ideal.*

Proof. Let us suppose that t is a bounded bi-ideal generated by (u_i) then by Lemma 31 t contains an overlapping factor. This is contradiction [Proposition 29].

Theorem 33. *The class of bounded bi-ideals $\mathfrak{B}_\mathfrak{b}$ is the proper subclass of the class of uniformly recurrent words $\mathfrak{R}_\mathfrak{u}$, that is, $\mathfrak{B}_\mathfrak{b} \subseteq \mathfrak{R}_\mathfrak{u}$ and $\mathfrak{B}_\mathfrak{b} \neq \mathfrak{R}_\mathfrak{u}$.*

Proof. The class of bounded bi-ideals $\mathfrak{B}_\mathfrak{b}$ is the subclass of the class of uniformly recurrent words $\mathfrak{R}_\mathfrak{u}$ by Proposition 26. The Thue–Morse word is uniformly recurrent as mentioned above. Therefore $\mathfrak{B}_\mathfrak{b} \neq \mathfrak{R}_\mathfrak{u}$ by Proposition 32.

Example 34. *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 $x = \lim_{i\to\infty} v_i$. Thus x is the bounded bi-ideal, besides $x = (0100)^\omega$. This demonstrates that straightforward generalization of Theorem 21 for bounded bi-ideals is not valid.

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.

Lemma 35. *If $v_n u \in v^*$ 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

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

This completes the induction.

Lemma 36. *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 [Corollary 2]

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

Now, by definition of x

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

By assumption, x is periodic, therefore

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

Since $v \in \mathrm{Pref}(v_n)$ then by Lemma 11

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

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

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

$\Rightarrow$ Let T be the minimal period of the word x, then $\exists n \in \mathbb{N} \; |v_n| \geq T$. Thus by Lemma 36

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

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

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

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

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

Thus

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

Hence by Lemma 15

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

Thereby

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

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

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

Note

$$x = v_n u_{n+1} v_n \dots = v_n u u'_1 v_n \dots$$

Since $u'_1 \in v^*$ and $v \in \mathrm{Pref}(v_n)$, then [Lemma 11] $x = v_n ux$. Hence [Lemmma 15] $v_n u \in v^*$.
$\Leftarrow$ By Lemma 35

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

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

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

References

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

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

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

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

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

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

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

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

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

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

[11] M. Morse. (1921) *Recurrent geodesics on a surface of negative curvature.* Trans. Amer. Math. Soc. **22**, 84–110.

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

[13] I. Simon. (1988) *Infinite Words and a Theorem of Hindman.* Rev. Mat. Apl. **9**, 97–104.

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

[15] A. Thue (1906) *Über unendliche Zeichenreihen.* Norske Vidensk. Selsk. Skrifter. I. Mat.–Nat. K1., Christiania Nr 7, 1–22.

[16] A. Thue (1912) *Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen.* Norske Vidensk. Selsk. Skrifter. I. Mat.–Nat. K1., Christiania Nr 10, 1–67.

[17] А. И. Зимин. (1982) *Блокирующие множества термов.* [*Blocking Sets of Terms.*] Матем. сб., т.**119**, № 3, с. 363–375. (Russian)