Publ. Math. Debrecen
77/3-4 (2010), 313–330

Semiperiodic vectors and the context-freeness of $Q_n = Q \cap (ab^*)^n$

By LÁSZLÓ KÁSZONYI (Szombathely)

This paper is dedicated to Professor P. Dömösi

Abstract. It is conjectured in DÖMÖSI *et al.* LNCS 710, pp. 194–203, that if Q denotes the set of all primitive words over a given alphabet X containing the letters a and b, then the languages $Q_n = Q \cap (ab^*)^n$ are context-free for all positive numbers n. In this paper we classify the elements of Q_n, in order to get a new method for constructing elements of Q_n.

1. Introduction

Let Q be the set of all primitive words over a fixed alphabet X. In the papers [2], [3] and [4] the still unsolved problem was investigated: whether the whole set Q is context-free or not.(See also: [10].) The simplest idea to show that Q is not context-free would be to use one of the pumping lemmata for context-free languages. This approach fails, because Q has seemingly context-free properties (DÖMÖSI *et al.* [4]). Another idea would be the investigation of context-freeness of the intersection of Q with a regular language L: if Q would be context-free then $Q \cap L$ would be context-free as well. In papers [4], [15], [17] and [16] we investigated the context-freeness of languages $Q_n = Q \cap (ab^*)^n$ for some natural numbers n. Our results suggest that Q_n is context-free for all natural numbers n. The sharpest result considering this conjecture was the following:

Mathematics Subject Classification: 20M35, 68A40, 57Q45.
Key words and phrases: primitive words, bounded languages, **DLI**-languages, context-freeness.

Theorem 1 (Kászonyi, Katsura [17]). *Let $n = p_1^{f_1} \cdots p_k^{f_k}$ where $p_1, \ldots, p_k$ are distinct prime numbers and $f_1, \ldots, f_k$ are positive integers. Assume that*

$$\sum_{i=1}^{k} 1/p_i < 4/5. \tag{1}$$

Then the language $Q \cap (ab^*)^n$ is context-free.

In order to get new constructions for grammars generating subclasses of Q_n without using the condition in Theorem 1, we develop the small theory of so called semiperiodic vectors. (See: Section 4.)

2. Definitions

Let X be a fixed alphabet, having at least two letters. A *primitive word* (over X) is a nonempty word not of the form w^m for any (nonempty) word w and integer $m \geq 2$. The set of all primitive words over X will be denoted by Q. Let $a, b \in X$, $a \neq b$, $n \in \{1, 2, \ldots\}$, and W be an arbitrary subset of the language $(ab^*)^n$. For $w \in W$ let $w = ab^{e_0} \cdots ab^{e_{n-1}}$ and denote the set of all vectors of the form $e(w) = (e_0, \ldots, e_{n-1})$ by $E(W)$. The index-set $\underline{n} = \{0, \ldots, (n-1)\}$ will be considered as a "cyclically ordered" set, i.e. the "open intervalls" (i, j) of $\underline{n}$ are given by $(i, j) = \{k | i < k < j\}$ if $i < j$ and by $(i, j) = \{k | k < j \text{ or } k > i\}$ if $i > j$. We will use the notations $[i, j), (i, j]$ and $[i, j]$ for the "semi closed" and "closed" intervalls defined in the usual manner: $[i, j) = \{i\} \cup (i, j), (i, j] = (i, j) \cup \{j\}$ and $[i, j] = \{i\} \cup (i, j) \cup \{j\}$. The addition and multiplication in $\underline{n}$ are meant as (mod n)-operations.

We say that the pairs of indices $\{i, j\}$ and $\{k, l\}$ are *crossing* if $k \in (i, j)$ and $l \in (j, i)$ or if $l \in (i, j)$ and $k \in (j, i)$. The subsets R and T of $\underline{n}$ are said to be *non-crossing* sets, if there exist two elements i and j of $\underline{n}$ such that $R \subseteq [i, j)$ and $T \subseteq [j, i)$ holds. For the expression "non-crossing" we will use the abbreviation n.c. If there are given more than two subsets of $\underline{n}$, then for the expression *pairwise non-crossing* we will use the abbreviation p.n.c.

In general, a language $L \subseteq \Sigma^*$, is a *bounded language* if and only if there exist non-empty words $w_0, \ldots, w_{m-1}$ such that $L \subseteq w_0^* \ldots w_{m-1}^*$. The words $w_0, \ldots, w_{m-1}$ are said to be the corresponding words of language L. Note that for a word $w \in \Sigma^*$ we use w^* as a short-hand notation for $\{w\}^*$.

Obviously, $Q_n = Q \cap (ab^*)^n$ is a bounded language. A necessary and sufficient condition for a bounded language to be context-free was given by Ginsburg:

Theorem 2 (Ginsburg [5])**.** *Let L be a bounded language over the alphabet Σ. Language L is context-free if and only if set*

$$E(L) = \{(e_0, \dots, e_{m-1}) \in \mathbb{N}^m \mid w_0^{e_0} \dots w_{m-1}^{e_{m-1}} \in L\}, \tag{2}$$

where the words $w_0, \dots, w_{m-1}$ are the corresponding words of L, is a finite union of stratified linear sets.

Definition 1. A set $F \subseteq \mathbb{N}^m$ where $\mathbb{N} = \{0, 1, \dots\}$ and $m \geq 1$ is called a *stratified linear set* if and only if either $F = \emptyset$ or there exist $r \geq 1$ and $v_0, \dots, v_r \in \mathbb{N}^m$ such that

(1). $F = \{v_0 + \sum_{i=1}^{r} k_i v_i \mid k_i \geq 0\}$

and for the vector set $P = \{v_i \mid 1 \leq i \leq r\}$

(2). every $v \in P$ has at most two nonzero components, and

(3). there exist no natural numbers i, j, k, l, with $0 \leq i < j < k < l \leq m-1$, and no vectors $u = (u_0, \dots, u_{m-1})$ and $x = (x_0, \dots, x_{m-1})$ from P such that $u_i x_j u_k x_l \neq 0$.

The vector v_0 and the vector-set P appearing in (1) are often called *preperiod* and the set of *periods* of F, respectively.

Often the set $E(L)$ is **D**efined by **L**inear **I**nequalities, and the problem is to check stratifiedness. Define the concept of **DLI**-sets as follows:

Definition 2. The set

$$E(\Theta, \delta, \epsilon) = \bigcap_{I \in \Theta} \{(e_0, \dots, e_{m-1}) \in \mathbb{N}^m \mid \epsilon(I) \sum_{i \in I} \delta_i e_i \geq 0\} \tag{3}$$

is a **DLI**-set where

(1). Θ is a system of index-sets, (i.e., of subsets of $\underline{m}$). Θ is considered as a multi-set i.e., elements of Θ may have multiplicity greater then one.

(2). $\delta = (\delta_0, \dots, \delta_{m-1})$ is a fixed vector of *signs* i.e., for $i = 0, \dots m-1$ $\delta_i \in \{-1, 0, 1\}$.

(3). ϵ is a function from Θ into the set $\{-1, 1\}$.

Definition 3. A bounded language L is a **DLI**-*language* if the set

$$E(L) = \{(e_0, \dots, e_{m-1}) \in \mathbb{N}^m \mid w_0^{e_0}, \dots, w_{m-1}^{e_{m-1}} \in L\}$$

of the corresponding exponent-vectors is a **DLI**-set. (Here $w_0, \dots,$ and w_{m-1} are the corresponding words.)

DLI-languages are often used as examples or counterexamples for context-free languages. In such cases we have to decide whether or not a given **DLI**-language is context-free. The following "Flip-Flop-Theorem" gives a necessary and sufficient condition for a **DLI**-set to be stratified semilinear.

Theorem 3 (Flip-Flop theorem, Kászonyi [12]). *Let the set E be a* **DLI**-*set with respect to the sign-vector $\delta = (\delta_0, \ldots, \delta_{m-1})$, index-set-system Θ, and function ϵ:*

$$E = E(\Theta, \delta, \epsilon) = \bigcap_{I \in \Theta} \{(e_0, \ldots, e_{m-1}) \in \mathbb{N}^m \mid \epsilon(I) \sum_{i \in I} \delta_i e_i \geq 0\} \tag{4}$$

E is stratified semilinear if and only if for every $e \in E$ there exists a hypergraph H, having the following properties:

(i). *The vertices of H are the vertices of a convex m-polygon, indexed by the elements of a cyclically ordered set $\underline{m}$ according to their cyclical order.*

(ii). *The edges of H are one- or two-element subsets of the vertex-set $\mathbf{V}(H)$ of H.*

(iii). *If $\{i, j\}$ is a two-element edge of H, then the signs associated with the end-points i and j are opposite, i.e., $\delta_i = -\delta_j \neq$.*

(iv). *The edge f is* forbidden *if there exists an index-set $I \in \Theta$ such that $f \cap I = \{i\}$ and $\epsilon(I) = -\delta_i$. Hypergraph H doesn't contain forbidden edges.*

(v). *The edges of H are non-crossing.*

(vi). *The degree of each vertex i is e_i.*

Using the Flip-Flop Theorem some lemmata may be proved guaranteeing the stratified semilinearity of a **DLI**-set in the case that the index set system Θ possesses some special properties.

(See: [10], [11], [12], [13] and [14]). We will apply the so called Flip-Flop lemma.

Lemma 1 (Flip-Flop lemma, Holzer–Kászonyi [11]). *Let the set E be a* **DLI**-*set with respect to the sign vector $\delta = (\delta_0, \ldots, \delta_{m-1})$, index set system Θ, and function ϵ:*

$$E = E(\Theta, \delta, \epsilon) = \bigcap_{I \in \Theta} \{(e_0, \ldots, e_{m-1}) \in \mathbb{N}^m \mid \epsilon(I) \sum_{i \in I} \delta_i e_i \neq 0\} \tag{5}$$

If Θ consists of pairwise non-crossing sets then the set E is stratified semilinear.

3. Boxes and differences

In the sequel we adopt some concepts and results from [17]. (Definitions 4–6, Lemmata 2–4 and 5.) In this section n denotes an integer with $n = p_1^{f_1} \dots p_k^{f_k}$ where $p_1, \dots, p_k$ are pairwise distinct prime numbers and $f_1, \dots, f_k$ are positive integers.

Definition 4. Let $\pi = \{q_1, \dots, q_r\}$ be a nonempty subset of $\{p_1, \dots, p_k\}$. A *π-scale* is a set $\{t_1/q_1, \dots, t_r/q_r\}$ where t_i is an integer relatively prime to p_i for each $i = 1, \dots, r$. For a π-scale $S = \{t_1/q_1, \dots, t_r/q_r\}$, $\xi \in \underline{n}$, we define a *π-box* by:

$$B = B(\xi; S) = \{\xi + \rho | \rho = \sum_{i=1}^{r} \rho_i t_i n/q_i, \rho_i \in \{0, 1\}\} \tag{6}$$

Definition 5. For a vector $e = (e_0, \dots, e_{n-1}) \in N^n$, the corresponding *difference* is:

$$\Delta_e(B) = \Delta_e(\xi; S) = \sum_{\xi+\rho \in B} (-1)^{\rho_1 + \dots + \rho_r} e_{\xi+\rho} \tag{7}$$

In other words, a difference defined for a vector e and a box B is a signed sum of such components of e whose indices belong to B, and if the index-pair (i, j) is an "edge" of box B then the corresponding members e_i and e_j of the sum have opposite signs.

Definition 6. For a π-scale S and $e \in N^n$, consider the subset $\Omega_e(S)$ of $\underline{n}$ defined by the rule $\Omega_e(S) = \{\xi \in \underline{n} \mid \Delta_e(\xi; S) \neq \emptyset\}$.

In the sequel we will investigate the question, whether or not $\Omega_e(S)$ is the empty set. The following lemma says that the answer to this question is independent of the choice of the scale S.

Lemma 2. *For π-scales S and S', $\Omega_e(S) \neq \emptyset$ if and only if $\Omega_e(S') \neq \emptyset$.*

Proof. Let $S = \{t_1/q_1, \dots, t_r/q_r\}$ and $S' = \{t'_1/q_1, \dots, t'_r/q_r\}$. For any i, there exists an s_i such that $s_i t'_i \equiv t_i \pmod{q_i}$. Hence

$$\Delta_e(\xi; t_1/q_1, \dots, t_r/q_r)$$
$$= \sum_{j_1=0}^{s_1-1} \cdots \sum_{j_r=0}^{s_r-1} \Delta_e(\xi + j_1 t'_1 n/q_1 + \dots + j_r t'_r n/q_r; t'_1 n/q_1, \dots, t'_r n/q_r).$$

Thus $\Omega_e(S') = \emptyset$ implies $\Omega_e(S) = \emptyset$. □

We will say that $\Omega_e(\pi) \neq \emptyset$ if $\Omega_e(S) \neq \emptyset$ for some (and thus any) π-scale S.

The following lemma asserts that for subsets π of the set $\{p_1, \dots, p_k\}$ the property $\Omega_e(\pi) \neq \emptyset$ is "hereditary".

Lemma 3. *Let $\emptyset \neq \pi' \subseteq \pi \subseteq \{p_1, \dots, p_k\}$. If $\Omega_e(\pi') = \emptyset$ then $\Omega_e(\pi) = \emptyset$.*

Proof. Let $\pi' = \{q'_1, \dots, q'_{r'}\}$ and $\pi = \{q'_1, \dots, q'_{r'}, q_1, \dots, q_r\}$. Then

$$\Delta_e(\xi; 1/q'_1, \dots, 1/q'_{r'}, 1/q_1, \dots, 1/q_r)$$
$$= \sum_{\rho_1=0}^{1} \cdots \sum_{\rho_r=0}^{1} (-1)^{\rho_1+\dots+\rho_r} \Delta_e(\xi + \rho_1 n/q_1 + \dots + \rho_r n/q_r; n/q'_1, \dots, n/q'_{r'}). \quad \square$$

Lemma 4. *For a π-scale S and $q \in \{p_1, \dots, p_k\} \setminus \pi$, the following conditions are equivalent:*

(1) $\Omega_e(\pi \cup \{q\}) = \emptyset$.

(2) *If $\xi \equiv \xi' \pmod{n/q}$ then $\Delta_e(\xi; S) = \Delta_e(\xi'; S)$.*

Proof. Note that $S \cup \{1/q\}$ is a $(\pi \cup \{q\})$-scale, and $\Delta_e(\xi; S \cup \{1/q\}) = \Delta_e(\xi; S) - \Delta_e(\xi + n/q; S)$ holds for any ξ. It follows that $\Omega_e(\pi \cup \{q\}) = \emptyset$ if and only if $\Delta_e(\xi; S) = \Delta_e(\xi + n/q; S)$ holds for any ξ. $\square$

Lemma 5. *For a π-scale S and $\{q_1, \dots, q_r\} \subseteq \{p_1, \dots, p_k\} \setminus \pi$, the following conditions are equivalent:*

(1) $\Omega_e(\pi \cup \{q_i\}) = \emptyset$ *for any $i = 1, \dots, r$.*

(2) *If $\xi \equiv \xi' \pmod{n/q_1 \dots q_r}$ then $\Delta_e(\xi; S) = \Delta_e(\xi'; S)$.*

Proof. The equivalence of (1) and (2) follows from Lemma 4. $\square$

Lemma 6. *Let $\pi = \{q_1, \dots, q_r\}$ and $S = \{t_1/q_1, \dots, t_r/q_r\}$ be any π-scale. Then for any $\xi \in \underline{n}$ and $s = \sum_{i=1}^{r} t_i n/q_i$*

$$e_\xi - e_{\xi+s} = \sum_{S' \subseteq S, S' \neq \emptyset} (-1)^{|S'|-1} \Delta_e(\xi; S') \tag{8}$$

holds.

Proof. The proof is by mathematical induction on the number of elements in S. For $|S| = 1$ equality (8) is trivially true. Assume that $|S| = r \geq 2$ and that (8) holds for any S with $|S| < r$. Let $S = \{t_1/q_1, \dots, t_{r-1}/q_{r-1}, t_r/q_r\}$, $S_{r-1} = S \setminus \{t_{r-1}/q_{r-1}\}$ and $S_r = S \setminus \{t_r/q_r\}$, consider

$$e_\xi - e_{\xi+s} = (e_\xi - e_{\xi+nt_r/p_r}) + (e_{\xi+nt_r/p_r} - e_{\xi+s}). \tag{9}$$

Here

$$e_{\xi+nt_r/p_r} - e_{\xi+s} = \sum_{S' \subseteq S_r, S' \neq \emptyset} (-1)^{|S'|-1} \Delta_e(\xi + nt_r/p_r; S') \tag{10}$$

holds by our hypothesis. It is easy to show that

$$\Delta_e(\xi; S' \cup \{t_r/p_r\}) = \Delta_e(\xi; S') - \Delta_e(\xi + nt_r/p_r; S'), \tag{11}$$

thus

$$\Delta_e(\xi + nt_r/p_r; S') = \Delta_e(\xi; S') - \Delta_e(\xi; S' \cup \{t_r/p_r\}). \tag{12}$$

Substituting (12) into (10) we have

$$e_{\xi+nt_r/p_r} - e_{\xi+s} = \sum_{S' \subseteq S_r, S' \neq \emptyset} (-1)^{|S'|-1} (\Delta_e(\xi; S') - \Delta_e(\xi; S' \cup \{t_r/p_r\}) \tag{13}$$

and

$$\begin{aligned}(e_\xi - e_{\xi+nt_r/p_r}) + (e_{\xi+nt_r/p_r} - e_{\xi+s}) &= \Delta_e(\xi; \{t_r/p_r\}) + (e_{\xi+nt_r/p_r} - e_{\xi+s}) \\ &= \sum_{S' \subseteq S, t_r/p_r \in S'} (-1)^{|S'|-1} \Delta_e(\xi; S') + \sum_{S' \subseteq S, t_r/p_r \notin S', S' \neq \emptyset} (-1)^{|S'|-1} \Delta_e(\xi; S') \\ &= \sum_{S' \subseteq S, S' \neq \emptyset} (-1)^{|S'|-1} \Delta_e(\xi; S'). \end{aligned} \tag{14}$$

□

In the sequel we develop some kind of "Discrete Fourier Analysis", i.e., we will show that for any vector $e \in \mathbb{N}^n$ e is a sum of some periodic vectors, associated with π-scales in a special manner. (As before, $n = p_1^{f_1} \cdots p_k^{f_k}$.)

Definition 7. Let $n = p_1^{f_1} \cdots p_k^{f_k}$, and $e \in \mathbb{N}^n$. Consider the $\{p_1, \ldots, p_k\}$-scale $S = \{t_1/p_1, \ldots, t_k/p_k\}$, let $\theta = p_1^{f_1-1} \cdots p_k^{f_k-1}$. For any vector $e \in \mathbb{N}^n$ and $S' \subseteq S$, $(S' \neq \emptyset)$, we define the vector $\phi(S', e) = (\phi_0(S', e), \ldots, \phi_{n-1}(S', e))$ as follows:

(1). For $0 \leq r \leq \theta - 1$ let $\phi_r(S', e) = d_r(S')$, such that $d_r(S') \in Z$ and

$$\sum_{S' \subseteq S,\ (S' \neq \emptyset)} d_r(S') = e_r. \tag{15}$$

(2). Assume that for any $0 \leq i \leq n - 1 - \theta$ $\phi_i(S', e)$ is already defined. Then let

$$\phi_{i+\theta}(S', e) = \phi_i(S', e) + (-1)^{|S'|} \Delta_e(i; S'). \tag{16}$$

Lemma 7. *For any vector $e \in \mathbb{N}^n$ and $S' \subseteq S$, $(S' \neq \emptyset)$, let the vector $\phi(S', e) \in \mathbb{N}^n$, $\phi(S', e) = (\phi_0(S', e), \ldots, \phi_{n-1}(S', e))$ be the vector given in Definition 7. Then*

$$e = \sum_{S' \subseteq S, S' \neq \emptyset} \phi(S', e). \tag{17}$$

Proof. Let us choose the scale $S = \{t_1/p_1, \ldots, t_k/p_k\}$ such that

$$nt_1/p_1 +, \ldots, + nt_k/p_k = \theta. \tag{18}$$

Then

$$e_{i+\theta} = e_i - (e_i - e_{i+\theta}) = e_i - \sum_{S' \subseteq S, S' \neq \emptyset} (-1)^{|S'|-1} \Delta_e(i; S') \tag{19}$$

holds by (8). □

Lemma 8. *Let $S' \subseteq S$, $(S' \neq \emptyset, S' \neq S)$, where S' and S are π' resp. π-scales. For $e \in \mathbb{N}^n$ the vector $\phi(S', e)$ given in Definition 7 is a c-periodic function where*

$$c = n / \prod_{q \in \pi \setminus \pi'} q. \tag{20}$$

Proof. Let

$$e = \sum_{S' \subseteq S, S' \neq \emptyset} \phi(S', e) \tag{21}$$

and define the vector e' as follows:

$$e' = \sum_{S_0 \subseteq S',\, S_0 \neq \emptyset} \phi(S_0, e) \tag{22}$$

It is easy to see that if $q \in \pi \setminus \pi'$ then $\Omega_{e'}(\pi' \cup q) = \emptyset$. (Here π' corresponds to S'.) It follows by Lemma 5, that if $\xi \equiv \xi' \pmod{n / \prod_{q \in \pi \setminus \pi'} q}$ then $\Delta_{e'}(\xi; S) = \Delta_{e'}(\xi'; S)$. It means by the definition of $\phi(S', e)$ that it is a c-periodic vector. □

4. Semiperiodic vectors

In this section we examine some properties of so-called semiperiodic vectors playing a central rule in our investigations concerning the context-freeness of Q_n. We will make use of the so-called Chinese Reminder Theorem. (See e.g.: [20].)

Theorem 4 (Chinese Reminder Theorem). *Let $m_1, \ldots, m_k$ be pair-wise relatively prime numbers and $a_1, \ldots, a_k$ be integers. Then for $i = 1, \ldots, k$ the congruences*

$$x \equiv a_i \pmod{m_i} \tag{23}$$

have a common solution. If x_1 and x_2 are any two solutions of the system (23), *then*

$$x_1 \equiv x_2 \pmod{m_1 \cdots m_k} \tag{24}$$

holds.

Definition 8. The vector $(e_0, \ldots, e_{n-1}) \in \mathbb{N}^n$ is (r,s)-semiperiodic if for any $\xi \in \underline{n}$

$$e_\xi - e_{\xi+r} = e_{\xi+s} - e_{\xi+r+s} \tag{25}$$

holds.

We get an interesting subclass of (r,s)-semiperiodic vectors in the case of $n = rs$ $r, s \neq 1$ and $\gcd(r,s) = 1$. The following lemma gives the motivation of the notion '(r,s)-semiperiodic vector'.

Lemma 9. *Let $n = rs$, $r, s \neq 1$ and $\gcd(r,s) = 1$. The vector $e \in \mathbb{N}^n$ is an (r,s)-semiperiodic if and only if e may be written as the sum of an r-periodic and an s-periodic vector, i.e.,*

$$e = f + g, \tag{26}$$

where for $f = (f_0, \ldots, f_{n-1})$, $g = (g_0, \ldots, g_{n-1})$, and for any $\xi \in \underline{n}$,

$$f_\xi = f_{\xi+r} \tag{27}$$

and

$$g_\xi = g_{\xi+s} \tag{28}$$

holds. If

$$e = f' + g', \tag{29}$$

is any other decomposition of e into an r-periodic and an s-periodic vector, then $f' = f + c$ and $g' = g - c$ holds for some constant vector c.

Proof. Let i be any element of $\underline{n}$. By the Chinese Reminder Theorem, there are integers α $(0 \leq \alpha \leq s-1)$ and β $(0 \leq \beta \leq r-1)$ such that

$$i = \alpha r + \beta s. \tag{30}$$

Let f and g defined by

$$f_i = -e_0 + c + e_{\beta s} \tag{31}$$

$$g_i = -c + e_{\alpha r} \tag{32}$$

where c is a fixed integer. Here

$$i + r = (\alpha + 1)r + \beta s \tag{33}$$

and

$$i + s = \alpha r + (\beta + 1)s, \tag{34}$$

thus

$$f_{i+r} = f_i = -e_0 + c + e_{\beta s} \tag{35}$$

$$g_{i+s} = g_i = -c + e_{\alpha r}. \tag{36}$$

We have to prove that $e = f + g$. The proof is by mathematical induction on the number $\alpha + \beta$. If $\beta = 0$ then

$$f_i + g_i = (-e_0 + c + e_0) + (-c + e_{\alpha r}) = e_i \tag{37}$$

Similarly follows that $f_i + g_i = e_i$ whenever $\alpha = 0$. Let $\alpha > 0$ and $\beta > 0$, assume that for $j = \alpha' r + \beta' s \quad f_j + g_j = e_j$ holds when $\alpha' + \beta' < \alpha + \beta$. Vector e is semiperiodic hence

$$e_{\alpha r+\beta s} - e_{(\alpha-1)r+\beta s} = e_{\alpha r+(\beta-1)s} - e_{(\alpha-1)r+(\beta-1)s} \tag{38}$$

thus

$$e_{\alpha r+\beta s} = e_{(\alpha-1)r+\beta s} + e_{\alpha r+(\beta-1)s} - e_{(\alpha-1)r+(\beta-1)s} \tag{39}$$

Here

$$e_{(\alpha-1)r+\beta s} = f_{(\alpha-1)r+\beta s} + g_{(\alpha-1)r+\beta s}$$

$$e_{\alpha r+(\beta-1)s} = f_{\alpha r+(\beta-1)s} + g_{\alpha r+(\beta-1)s}$$

$$e_{(\alpha-1)r+(\beta-1)s} = f_{(\alpha-1)r+(\beta-1)s} + g_{(\alpha-1)r+(\beta-1)s} \tag{40}$$

by our hypothesis. Vector f is r-periodic, thus

$$f_{\alpha r+\beta s} = f_{(\alpha-1)r+\beta s} + f_{\alpha r+(\beta-1)s} - f_{(\alpha-1)r+(\beta-1)s}. \tag{41}$$

Similarly, g is s-periodic, thus

$$g_{\alpha r+\beta s} = g_{(\alpha-1)r+\beta s} + g_{\alpha r+(\beta-1)s} - g_{(\alpha-1)r+(\beta-1)s} \tag{42}$$

It follows by 39, 40, 41 and 42 that

$$e_{\alpha r+\beta s} = f_{\alpha r+\beta s} + g_{\alpha r+\beta s} \tag{43}$$

Let $e = f' + g'$ any decomposition of e into an r-periodic and an s-periodic vector. It is easy to check that if $f'_0 = f_0$ then $f' = f$ and $g' = g$.

Let us assume that vector e is of the form $e = f + g$, where f and g are r, resp. s-periodic functions. It follows that

$$e_\xi - e_{\xi+r} = f_\xi + g_\xi - f_{\xi+r} - g_{\xi+r}. \tag{44}$$

Vector f is r-periodic, thus $f_\xi = f_{\xi+r}$. Here

$$\begin{aligned} e_\xi - e_{\xi+r} &= g_\xi - g_{\xi+r} = g_\xi - g_{\xi+r} + f_{\xi+s} - f_{\xi+s} \\ &= g_{\xi+s} - g_{\xi+r+s} + f_{\xi+s} - f_{\xi+r+s} = e_{\xi+s} - e_{\xi+r+s}, \end{aligned} \tag{45}$$

it means that vector e is (r, s)-semiperiodic. □

In the sequel we define the general concept of semiperiodic vectors and investigate some of their properties.

Definition 9. Let $n = n_1\theta$, where $n_1 = p_1^{f_1} \cdots p_k^{f_k}$, consequently, $\theta = p_1^{f_1-1} \cdots p_k^{f_k-1}$. Vector $e \in \mathbb{N}^n$ is called *semiperiodic*, if for any decomposition $n = rs\theta$ of n where $r, s \neq n_1$ and $\gcd(r, s) = 1$, vector e is $(r, s\theta)$-periodic.

For the sake of simplicity, in the following lemma we assume that n is square-free.

Lemma 10. *Let $n = p_1 \cdots p_k$ and $\pi = \{p_1, \ldots, p_k\}$. The vector $e \in \mathbb{N}^n$ is semiperiodic if and only if e is of the form*

$$e = \sum_{p\in\pi} e^{(p)} \tag{46}$$

where for $p \in \pi$ $e^{(p)}$ is a p-periodic vector.

Proof. *Step 1, sufficiency.* Let e be of the form 46, $n = rs$, $r, s \neq 1$. We have to prove that e is (r, s)-semiperiodic. Let us denote the set of prime divisors of r by π_r, and by π_s that of s. Consider the following decomposition of vector e:

$$e = e^{(r)} + e^{(s)}, \tag{47}$$

where

$$e^{(r)} = \sum_{q\in\pi_r} e^{(q)} \tag{48}$$

and

$$e^{(s)} = \sum_{q\in\pi_s} e^{(q)}. \tag{49}$$

It is easy to show that $e^{(r)}$ is an r-periodic and $e^{(s)}$ is an s-periodic vector, thus vector e is (r, s)-semiperiodic by Lemma 9.

Step 2, necessity. Assume now that the vector $e \in \mathbb{N}^n$ is semiperiodic. We have to prove that $e = e^{(p_1)} + \cdots + e^{(p_k)}$, where for $i = 1, \ldots, k$ $e^{(p_i)}$ is a p_i-periodic vector. Especially, for any $p \in \{p_1, \ldots, p_k\}$ e is $(p, n/p)$-semiperiodic, thus in Definition 17 for any $i \in \underline{n}$, $\Delta_e(i; S') = 0$ holds if $|S'| \geq 2$. Let

$$e = \sum_{S' \subseteq S,\, S' \neq \emptyset} \phi(S', e) = \sum_{S' \subseteq S, |S'|=1} \phi(S', e) + \sum_{S' \subseteq S, |S'| \geq 2} \phi(S', e). \tag{50}$$

Here

$$\sum_{S' \subseteq S, |S'| \geq 2} \phi(S', e) = C \tag{51}$$

holds for some constant $C \in \mathbf{Z}$, and every summand in the sum

$$\sum_{S' \subseteq S, |S'|=1} \phi(S', e). \tag{52}$$

is p-periodic for $S' = \{p\}$. (See: Lemma 8) $\square$

Theorem 5. *Let $n = p_1^{f_1} \cdots p_k^{f_k}$ and $\pi = \{p_1, \ldots, p_k\}$. The vector $e \in \mathbb{N}^n$ is semiperiodic if and only if e is of the form*

$$e = \sum_{p \in \pi} e^{(p)} \tag{53}$$

where for $p \in \pi$ and $\theta = p_1^{f_1-1} \cdots p_k^{f_k-1}$ $e^{(p)}$ is a $p\theta$-periodic vector.

Proof. For $j = 0, \ldots, \theta - 1$ and $n_1 = p_1 \cdots p_k$ let us define the vectors $f^j = (f_0^j, \ldots, f_{n_1-1}^j)$ as follows:

$$f_k^j = e_{j+k\theta} \quad k = 0, \ldots, n_1 - 1 \tag{54}$$

It is easy to show that e is semiperiodic if and only if for any $j = 0, \ldots, \theta - 1$ vector f^j is semiperiodic as well.

$$f^j = \sum_{p \in \pi} f^{(j,p)}, \tag{55}$$

where $f^{(j,p)}$ is p-periodic. Let $e^{(p)} = (e_0^{(p)}, \ldots, e_n^{(p)})$ defined as

$$e_{j+k\theta}^{(p)} = f_k^{(j,p)} \quad j = 0, \ldots, \theta - 1 \; k = 0, \ldots, n_1 - 1 \tag{56}$$

Vector $f^{(j,p)}$ is p-periodic, thus $e^{(p)}$ is $p\theta$-periodic. $\square$

5. The classification of $E(Q_n)$

In the sequel we will define a classification of $E(Q_n)$. It is conjectured that each class of this classification is a stratified semilinear **DLI**-set.

Definition 10. For $\alpha = 1, \dots, k-1$ let E_α defined by

$$E_\alpha = \{e \in \mathbb{N}^n \mid \Omega_e(\pi) \neq \emptyset,\ \text{if } |\pi| = \alpha,\ \Omega_e(\pi') = \emptyset,\ \text{if } |\pi'| = \alpha + 1\}, \tag{57}$$

and for $\alpha = k$ by

$$E_k = \{e \in \mathbb{N} \mid \Omega_e(\pi) \neq \emptyset,\ |\pi| = k, \}, \tag{58}$$

Proving the context-freeness of Q_6, Katsura Masashi introduced two special types of **DLI**-sets. Here we generalize this constructions.

Theorem 6. *Let E_α given in Definition 10. Then $E_k \cap E(Q_n)$ and $E_1 \cap E(Q_n)$ are stratified semilinear sets.*

Proof. $E_k \cap E(Q_n)$ is a stratified semilinear set by the Flip-Flop lemma. Let $e \in E_1 \cap E(Q_n)$. Let $\pi = \{p_1, \dots, p_k\}$, assume that $p_1 < \cdots < p_k$. By Lemma 5, vector e is $(p\theta, n/p)$-semiperiodic for any $p \in \pi$. It means that

$$e_\xi - e_{\xi+n/p} = e_{\xi+p\theta} - e_{\xi+n/p+p\theta}. \tag{59}$$

Case 1, $k \geq 3$. $e \in E(Q_n)$, thus there is a $\xi \in \underline{n}$, for which $e_\xi - e_{\xi+n/p_k} \neq 0$ holds. We may assume that $\xi = n - 1 - n/p_k$, thus

$$e_{n-1-n/p_k} - e_{n-1} \neq 0. \tag{60}$$

Similarly, there is a $\xi_1 \in \underline{n}$ such that

$$e_{\xi_1} - e_{\xi_1+n/p_1} \neq 0. \tag{61}$$

Using(59) we can choose an m_1 such that

$${\xi_1}' = \xi_1 + m_1 p_1 \theta \in [0, p_1\theta), \tag{62}$$

and

$$e_{\xi_1'} - e_{\xi_1'+n/p_1} = e_{\xi_1} - e_{\xi_1+n/p_1} \neq 0. \tag{63}$$

In the same way, let $\xi_2' \in [\xi_1', \xi_1' + p_2\theta)$ such that

$$e_{\xi_2'} - e_{\xi_2'+n/p_2} \neq 0. \tag{64}$$

In general, let us assume that ξ_i' is already given ($i \in \{1, \dots, k-2\}$). We define

$\xi'_{i+1} \in [0, p_{i+1}\theta)$ such that

$$e_{\xi'_{i+1}} - e_{\xi'_{i+1}+n/p_{i+1}} \neq 0. \tag{65}$$

We show that the set system

$$\Theta = \{\{\xi'_i, \xi'_i + n/p_i\} \mid i = 1, \ldots, k\} \tag{66}$$

consists of pair-wise non-crossing elements. It is enough to prove that for $i = 1, \ldots, n-2$ $\{\xi'_i, \xi'_i + n/p_i\}$ and $\{\xi'_{i+1}, \xi'_{i+1} + n/p_{i+1}\}$ further $\{\xi'_1, \xi'_1 + n/p_1\}$ and $\{\xi'_k, \xi'_k + n/p_k\}$ are non-crossing sets. Indeed, $d_1 = (\xi_i + n/p_i) - (\xi_i + n/p_{i+1}) = n/p_i - n/p_{i+1}$ is divisible by $p_k\theta$, thus $p_i\theta < p_k\theta < d_1$, it means that $\{\xi'_i, \xi'_i + n/p_i\}$ and $\{\xi'_{i+1}, \xi'_{i+1} + n/p_{i+1}\}$ are really non-crossing sets. Similarly, $d_2 = (n - n/p_k) - (n/p_1)$ is divisible by $p_2\theta$, thus $p_1\theta < p_2\theta < d_2$, hence $\{\xi'_1, \xi'_1 + n/p_1\}$ and $\{\xi'_k, \xi'_k + n/p_k\}$ are non-crossing sets.

Case 2.1, $k = 2$, $p_1 = 2$, $p_2 = 3$ Vector e belongs to $E(Q_n)$, thus there are $\xi_1 \in \underline{n}$, $\xi_2 \in \underline{n}$ such that

$$e_{\xi_1} - e_{\xi_1+n/2} \neq 0 \tag{67}$$

and

$$e_{\xi_2} - e_{\xi_2+n/3} \neq 0. \tag{68}$$

It is easy to shaw that for any $j \in \underline{n}$

$$e_{\xi_1+jn/3} - e_{\xi_1+n/2+jn/3} \neq 0 \tag{69}$$

thus there is a j_0 such that one of the following relations is valid:

$$\xi_1 + j_0 n/3 \in [\xi_2, \xi_2 - \theta) \tag{70}$$

or

$$\xi_1 + j_0 n/3 \in [\xi_2, \xi_2 + \theta). \tag{71}$$

$$\xi_1 + n/2 + j_0 n/3 \in [\xi_2, \xi_2 - \theta). \tag{72}$$

$$\xi_1 + n/2 + j_0 n/3 \in [\xi_2, \xi_2 + \theta). \tag{73}$$

Without loss of generality we may assume that (70) holds. If

$$\eta = \xi_1 + j_0 n/3 \in (\xi_2, \xi_2 - \theta) \tag{74}$$

(i.e., $\eta \neq \xi_2$,) then

$$e_\eta - e_{\eta+n/2} \neq 0 \tag{75}$$

and

$$e_{\xi_2} - e_{\xi_2+n/3} \neq 0. \tag{76}$$

Here the index sets $\{\eta, \eta + n/2\}$ and $\{\xi_2, \xi_2 + n/3\}$ are non-crossing sets, thus vector e is contained in the stratified semilinear **DLI**-set

$$\{e' \mid e'_\eta - e'_{\eta+n/2} \neq 0, e'_{\xi_2} - e'_{\xi_2+n/3} \neq 0.\} \tag{77}$$

If $\eta = \xi_1 + j_0 n/3 = \xi_2$ then let us consider the **DLI**-set

$$D_1 = \{e' \mid e'_\eta - e'_{\eta+n/2} \neq 0, (e'_\eta - e'_{\eta+n/2}) + (e'_{\eta+n/2+n/3} - e'_{(\eta+n/3)+n/3}) \neq 0.\} \tag{78}$$

Assume that $e \in D_1$ but $e \notin E(Q_n)$. It means that e is either a quadrat or a cube. $e_{\xi_2} - e_{\xi_2+t_1 n/2} \neq 0$, thus e may not be a quadrat. If e is a cube, then

$$\begin{aligned} e_\eta - e_{\eta+n/3} &= (e_\eta - e_{\eta+n/2}) + (e_{\eta+n/2} - e_{\eta+n/3}) \\ &= (e_\eta - e_{\eta+n/2}) + (e_{\eta+n/2+n/3} - e_{\eta+n/3+n/3}) = 0 \end{aligned}$$

holds for any $\eta \in \underline{n}$, a contradiction by (78) .

Case 2.2, $k = 2$, $r = 1$, $p_1 p_2 > 6$ Let us assume that $p_1 < p_2$ and that $t_1 = t_2 = 1$. If the index sets $I_1 = \{\xi_1, \xi_1 + n/p_1\}$ and $I_2 = \{\xi_2, \xi_2 + n/p_2\}$ are non-crossing sets then vector e is contained in the stratified semilinear vector set

$$\{e' \mid e'_{\xi_1} - e'_{\xi_1+n/p_1} \neq 0,\ e'_{\xi_2} - e'_{\xi_2+n/p_2} \neq 0\}. \tag{79}$$

Assume that I_1 and I_2 are crossing sets, i.e., $\xi_1 \in [\xi_2, \xi_2 + n/p_2]$ or $\xi_1 + n/p_1 \in [\xi_2, \xi_2 + n/p_2]$. Without loss of generality may be assumed that $\xi_1 \in [\xi_2, \xi_2 + n/p_2]$ Vector e is semiperiodic thus

$$0 \neq e_{\xi_1} - e_{\xi_1+n/p_1} = e_{\xi_1+n/p_2} - e_{\xi_1+n/p_1+n/p_2}. \tag{80}$$

We will show that the index sets $I'_1 = \{\xi_1 + n/p_2, \xi_1 + n/p_1 + n/p_2\}$ and $I_2 = \{\xi_2, \xi_2 + n/p_2\}$ are non-crossing sets. We have to prove that

$$n/p_1 + 2n/p_2 < n, \tag{81}$$

i.e., that

$$1/p_1 + 2/p_2 < 1. \tag{82}$$

Indeed, $p_2 \geq 5$ and $p_1 \geq 2$ thus

$$0 < p_2 - 4 = 2(p_2 - 2) - p_2 \leq p_1(p_2 - 2) - p_2 = p_1 p_2 - 2p_1 - p_2, \tag{83}$$

and(82) holds.(Note that $p_1 p_2 > 6$.) □

It is conjectured that $E_\alpha \cap E(Q_n)$ is stratified semilinear for any $\alpha(2 \leq \alpha \leq k-1)$. In Lemma 11 we investigate the structure of E_α in this case.

Lemma 11. *For α $(2 \leq \alpha \leq k-1)$, let E_α given in Definition 10. Let us consider the vector $e \in \mathbb{N}^n$, which may be written in the form*

$$e = \sum_{S' \subseteq S,\, S' \neq \emptyset} \phi(S', e). \tag{84}$$

(See: Lemma 7) Vector e belongs to E_α if and only if

(i) *In case of $|S'| > \alpha$, $\phi(S', e)$ is a θ-periodic vector.*

(ii) *For any $\alpha' \leq \alpha$ there is an S', such that $|S'| = \alpha'$, and vector $\phi(S', e)$ is not θ-periodic, but it is a c-periodic function where*

$$c = n / \prod_{q \in \pi \setminus \pi'} q. \tag{85}$$

Proof. By the definition of E_α, $\Delta_e(\xi, S') = 0$ holds for any $e \in E_\alpha$ and $\xi \in \underline{n}$ whenever $|S'| > \alpha$. Further $\Delta_e(\xi_0, S') \neq 0$ holds for some $\xi_0 \in \underline{n}$ if $|S'| \leq \alpha$. It means that for $\phi(S', e) = (\phi_0(S', e), \ldots, \phi_{n-1}(S', e))$

$$\phi_{\xi+\theta}(S', e) = \phi_\xi(S', e) + (-1)^{|S'|} \Delta_e(\xi; S') = \phi_\xi(S', e) \tag{86}$$

holds if $|S'| > \alpha$ and $\phi(S', e)$ is c-periodic by Lemma 8, whenever $|S'| \leq \alpha$. Conversely, assume that for $e \in \mathbb{N}^n$,(86) and

$$\phi_{\xi_0+\theta}(S', e) = \phi_{\xi_0}(S', e) + (-1)^{|S'|} \Delta_e(\xi_0; S') \neq \phi_{\xi_0}(S', e). \tag{87}$$

holds. Then in case(i) $\Delta_e(\xi, S') = 0$ holds for any $\xi \in \underline{n}$ and in case(ii) there is a $\xi_0 \in \underline{n}$ with $\Delta_e(\xi_0; S') \neq 0$. It follows that $e \in E_\alpha$. $\square$

6. Examples

Theorem 17 allows us to construct semiperiodic vectors.

Example 1. Let $n = 15 = 3 \cdot 5$, further define vectors f, g and e as follows:

i= 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
f= 1 3 2—1 3 2—1 3 2—1 3 2— 13 2
g= 0 0 1 1 2— 00 1 1 2— 0 0 1 1 2
e= 1 3 3 2 5 2 1 4 3 3 3 2 2 4 4

Here f is a 3-periodic and g is a 5-periodic vector and row e is the sum of the previous two. For the components of e holds:

$$e_0 - e_5 = e_3 - e_8 = e_6 - e_{11} = e_9 - e_{14} = e_{12} - e_2 = -1,$$

$$e_1 - e_6 = e_4 - e_9 = e_7 - e_{12} = e_{10} - e_0 = e_{13} - e_3 = 2,$$

further

$$e_2 - e_7 = e_5 - e_{10} = e_8 - e_{13} = e_{11} - e_1 = e_{14} - e_4 = -1,$$

thus vector e is $(3, 5)$-semiperiodic.

The following example shows, how to construct elements of E_α, using Lemma 11.

Example 2. Let $n = 2 \cdot 3 \cdot 5$ and $\alpha = 2$ and

e_1	=0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1
e_2	=0 0 1 1 2 0 0 1 1 2 0 0 1 1 2 0 0 1 1 2 0 0 1 1 2 0 0 1 1 2
e_3	=1 3 2 1 3 2 1 3 2 1 3 2 1 3 2 1 3 2 1 3 2 1 3 2 1 3 2 1 3 2
e_4	=1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0

e =2 4 3 3 5 3 1 5 3 4 4 3 2 5 4 2 3 4 2 6 3 2 4 4 3 4 2 3 4 5.

Here vector $e = e_1 + e_2 + e_3 + e_4$ given above belongs to E_2.

References

[1] J. Dassow and Gh. Păun, Regulated Rewriting in Formal Language Theory, volume 18 of *EATCS Monographs in Theoretical Computer Science*, *Springer, Berlin*, 1989.

[2] P. Dömösi, S. Horváth and M. Ito, Formal languages and primitive words, *Publ. Math. Debrecen* **42** (1993), 315–321.

[3] P. Dömösi, S. Horváth, M. Ito, L. Kászonyi and M. Katsura, Some combinatorial properties of words, and the Chomsky-hierarchy, In Proceedings of the 2nd International Colloquium on Words, Languages and Combinatorics, Kyoto, Japan, 1992, *World Scientific, Singapore.*

[4] P. Dömösi, S. Horváth, M. Ito, L. Kászonyi and M. Katsura, Formal languages consisting of primitive words, Proceedings of the 9th International Conference on Fundamentals of Computation Theory, number 710 in LNCS, pages 194–203, Szeged, Hungary, (Z. Ésik, ed.), *Springer*, 1993.

[5] S. Ginsburg, The Mathematical Theory of Context-Free Languages, *McGraw-Hill, New York*, 1966.

[6] M. Ito and M. Katsura, Context-free languages consisting of non-primitive words, *Int. J. Comput. Math.* **40** (1991), 157–167.

[7] Gh. Păun, Morphisms and primitivity, *EATCS Bull., Techical Contributions* (1997), 85–88.

[8] A. Salomaa, From Parikh vectors to GO territories, *EATCS Bull., Formal Language Column* (1995), 89–95.

[9] J. Berstel and L. Boasson, The set of Lyndon words is not context-free, *EATCS Bull., Techical Contributions* **63** (1997), 139–140.

[10] L. Kászonyi, On a class of stratified linear sets, *Pure Math. Appl. Ser. A* **6**(2) (1995), 203–210.

[11] L. Kászonyi and M. Holzer, A generalization of the flip-flop lemma, *Publ. Math. Debrecen* **54** (1999), 203–210, AFL'96, Salgótarján (1996), supplementum.

[12] L. Kászonyi, On Bounded Context-Free Languages, Proceedings of the First Symposium on Algebra, Languages and Computation, (T. Imaoka and C. Nehaniv, eds.), *University of Aizu, Japan*, 1997.

[13] L. Kászonyi, A pumping lemma for **DLI**-languages, *Discrete Math.* **258** (2002), 105–122.

[14] L. Kászonyi, How to Generate Binary Codes Using Context-Free Grammars, Words, Semigroups, Transductions, Festschrift in Honor of Gabriel Thierrin, (M. Ito, G. Paun, Sheng Yu, eds.), *World Scientific*, 2001, 289–301.

[15] L. Kászonyi and M. Katsura, On an algorithm concerning the languages $Q\cap(ab^*)^n$, *Pure Appl. Math.* **10**(3) (1999), 313–322.

[16] L. Kászonyi and M. Katsura, On the context-freeness of a class of primitive words, *Publ. Math. Debrecen* **51** (1997), 1–11.

[17] L. Kászonyi and M. Katsura, Some new results on the context-freeness of languages $Q\cap(ab^*)^n$, *Publ. Math. Debrecen* (1999), 885–894, In: *Proc. of the 8th Conf. Automata and Formal Languages (AFL'96),* Salgótarján, Hungary, 29 Juli 2 August.

[18] L. Kászonyi, On **DLI**-sets of Katsura type, *AFL'02, Debrecen* (2002).

[19] A. Mateescu, Gh. Păun, G. Rozenberg and A. Salomaa, Parikh prime words and GO-like territories, *J.UCS* **1**(12) (1995), 790–810.

[20] Van Niven and H. S. Zuckerman, An Introduction to the Theory of Numbers, Third Edition, *John Wiley and Sons, Inc. New York.*

LÁSZLÓ KÁSZONYI
DEPARTMENT OF INFORMATICS
UNIVERSITY OF WEST HUNGARY
KÁROLYI GÁSPÁR TÉR 4
H-9700 SZOMBATHELY
HUNGARY

E-mail: kaszonyi@ttmk.nyme.hu

(Received July 30, 2009; revised May 29, 2010)