RAIRO-Theor. Inf. Appl. 48 (2014) 267–280
DOI: 10.1051/ita/2014009

Available online at:
www.rairo-ita.org

ROOT CLUSTERING OF WORDS *

Gerhard Lischke[1]

Abstract. Six kinds of both of primitivity and periodicity of words, introduced by Ito and Lischke [M. Ito and G. Lischke, *Math. Log. Quart.* **53** (2007) 91–106; Corrigendum in *Math. Log. Quart.* **53** (2007) 642–643], give rise to defining six kinds of roots of a nonempty word. For $1 \leq k \leq 6$, a k-root word is a word which has exactly k different roots, and a k-cluster is a set of k-root words u where the roots of u fulfil a given prefix relationship. We show that out of the 89 different clusters that can be considered at all, in fact only 30 exist, and we give their quasi-lexicographically smallest elements. Also we give a sufficient condition for words to belong to the only existing 6-cluster. These words are also called Lohmann words. Further we show that, with the exception of a single cluster, each of the existing clusters contains either only periodic words, or only primitive words.

Mathematics Subject Classification. 68Q45, 68R15.

1. Preliminaries

In the algebraic theory of codes and formal languages, the set Q of all primitive words over some alphabet X has received special interest (see, for instance, [6,8,9]). A nonempty word is primitive if and only if it is not periodic. And a word p is periodic if and only if it is a concatenation of two or more copies of the same word v,

Keywords and phrases. Periodicity of words, primitivity of words, roots of words, classification of words.

* *This work was supported by the Fellowship Program of the Japan Society for Promotion of Science (JSPS) under No. S-11717. Parts of it have been presented at the Second Russian Finnish Symposium on Discrete Mathematics in Turku 2012 under the title* k-root words and their clusters *and at the 22nd Theorietag Automata and Formal Languages in Prague under the title* Lohmann words and more clusters of words.

[1] Fakultät für Mathematik und Informatik, Friedrich-Schiller-Universität Jena, Ernst-Abbe-Platz 1-4, 07743 Jena, Germany. gerhard.lischke@uni-jena.de

$p = v^n$, $n \geq 2$.[2] Instead of the usual concatenation of words, Ito and Lischke in [1] took into consideration a more general kind of concatenation (concatenation with overlaps or folding) and generalized both the notions of primitivity and periodicity to six kinds of both of primitivity and periodicity. Whereas the shortest word v such that $p = v^n$ for some natural number n is denoted as the root of p, our generalizations give rise to define five further kinds of roots of a nonempty word. These are the topic of this paper.

First, let us repeat the most important notions about words.

X should be a fixed alphabet, which means, it is a finite and nonempty set of symbols, also called letters. Further, we assume that it is a nontrivial alphabet, which means that it has at least two symbols which we will denote by a and b, $a \neq b$ (otherwise, all of our results become trivial or meaningless), and that we have a fixed order on X, $a < b$. X^* is the free monoid generated by X or the set of all words over X. The number of letters of a word p, with their multiplicities, is the length of the word p, denoted by $|p|$. The order on X will be quasi-lexicographically extended to an order on X^*, this means, we order the words first by their lengths and lexicographically for words of equal length. If $|p| = 0$, then p is the empty word, denoted by e (in other papers also by ϵ or λ). The set of words of length n over X is denoted by X^n. Then $X^* = \bigcup_{n \in \mathbb{N}} X^n$ and $X^0 = \{e\}$. For the set of nonempty words over X we will use the notation $X^+ = X^* \setminus \{e\}$.

The concatenation of two words $p = x_1 x_2 \dots x_m$ and $q = y_1 y_2 \dots y_n$, $x_i, y_j \in X$, is the word $pq = x_1 x_2 \dots x_m y_1 y_2 \dots y_n$ which has the length $|pq| = |p| + |q|$. The powers of a word $p \in X^*$ are defined inductively: $p^0 = e$, and $p^n = p^{n-1} p$ for $n \geq 1$.

For words $p, q \in X^*$, p is a prefix of q, in symbols $p \sqsubseteq q$, if there exists $r \in X^*$ such that $q = pr$. p is a strict prefix of q, in symbols $p \sqsubset q$, if $p \sqsubseteq q$ and $p \neq q$. $Pr(q) =_{Df} \{p : p \sqsubset q\}$ is the set of all strict prefixes of q (including e if $q \neq e$). p is a suffix of q, if there exists $r \in X^*$ such that $q = rp$.

For sets A, B, $A \subseteq B$ denotes their inclusion and $A \subset B$ denotes their strict inclusion.

2. Folding of words

Instead of the usual concatenation of words, we also consider the concatenation with overlaps or folding operation: For $p, q \in X^*$,

$$p \otimes q =_{Df} \begin{cases} \{p\} & \text{if } q = e \\ \{w_1 w_2 w_3 : \, w_3 \neq e \, \wedge \, w_1 w_2 = p \, \wedge \, w_2 w_3 = q\} & \text{otherwise,} \end{cases}$$

$$p^{\otimes 0} =_{Df} \{e\}, \; p^{\otimes n} =_{Df} \bigcup \left\{ w \otimes p : w \in p^{\otimes n-1} \right\} \text{ for } n \geq 1.$$

For sets $A, B \subseteq X^*$, $A \otimes B =_{Df} \bigcup \{p \otimes q : p \in A \wedge q \in B\}$.

[2] Remark that in some papers, the term *periodic* has a more general meaning which for instance in [1–3] is covered by the term *semi-periodic*.

The following example illustrates this operation.
Let $p = aabaa$. Then $p \otimes p = p^{\otimes 2} = \{aabaaaabaa, aabaaabaa, aabaabaa\}$.

This operation $\otimes$ does not fulfill the associative law. If for instance $p = aa$, $q = b$ and $r = aba$ then $p \otimes (q \otimes r) = \{aababa\} \subset (p \otimes q) \otimes r = \{aababa, aaba\}$ and $(r \otimes q) \otimes p = \{ababaa\} \subset r \otimes (q \otimes p) = \{ababaa, abaa\}$. For each $p \in X^+$, we have $p^{\otimes 2} \subset p \otimes (p \otimes p)$, but $p^{\otimes 2} \not\subseteq (p \otimes p) \otimes p = p^{\otimes 3}$ and $p^{\otimes k} = (\ldots((p \otimes p) \otimes p) \ldots \otimes p)$ for $k > 3$.

Remark that our definition of $\otimes$ in [1–3] was more general than the present one but all results and proofs there remain unchanged under the new definition. The reason for the new definition is that for the former one we have $p^{\otimes k} \subset p^{\otimes k+1}$ for each $k \geq 2$ and each word $p \in X^+$ which is no more true for the present definition (and this makes the formulation and proof of the forthcoming Lem. 4.11 easier).

If $u \in p^{\otimes m}$ then we can say that any letter in u is covered by the word p, and if $m \geq 2$ and p is the shortest word with this property then we will call it the hyperroot of u and show that it cannot be longer than $\frac{2|u|}{3}$ (Lem. 4.11). Based on the definitions of concatenation with or without overlaps in combination with prefixes, six sets of words were defined which are periodic in different senses, and also six sets of words were defined which are primitive in different senses. We don't want to repeat these definitions, and refer the interested reader to [1–3]. These sets are deeply investigated in [1].

3. Roots and their clusters

Based on the definitions of different kinds of periodicity resp. primitivity of words, we define the following kinds of roots of nonempty words.

Definition 3.1. Let $u \in X^+$.

The shortest word v such that there exists a natural number n with $u = v^n$ is called the root of u, denoted by $root(u)$.

The shortest word v such that there exists a natural number n with $u \in v^n \cdot Pr(v)$ is called the strong root of u, denoted by $sroot(u)$.

The shortest word v such that there exists a natural number n with $u \in v^{\otimes n}$ is called the hyperroot of u, denoted by $hroot(u)$.

The shortest word v such that there exists a natural number n with $u \in \{v^n\} \otimes Pr(v)$ is called the super strong root of u, denoted by $ssroot(u)$.

The shortest word v such that there exists a natural number n with $u \in v^{\otimes n} \cdot Pr(v)$ is called the strong hyperroot of u, denoted by $shroot(u)$.

The shortest word v such that there exists a natural number n with $u \in v^{\otimes n} \otimes Pr(v)$ is called the hyperhyperroot of u, denoted by $hhroot(u)$.

root, *sroot*, *hroot*, *ssroot*, *shroot* and *hhroot* are word functions over X^+, *i.e.*, functions from X^+ to X^+. Generally, for word functions we define the following partial ordering, also denoted by $\sqsubseteq$. $dom(f)$ for a function f denotes the domain of f.

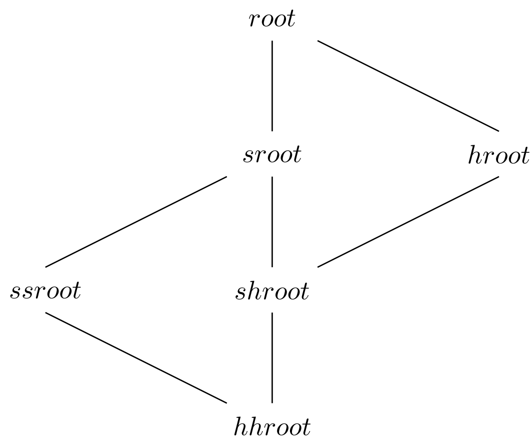

FIGURE 1. Partial ordering of the root-functions.

Definition 3.2. For word functions f and g having the same domain, $f \sqsubseteq g =_{Df} \forall u(u \in dom(f) \to f(u) \sqsubseteq g(u))$.

Theorem 3.3. *The partial ordering $\sqsubseteq$ for the functions from Definition 3.1 is given in Figure 1.*

The proof follows from the definition and can be done as a simple exercise. It is also contained in [1,3].

For most words u, some of the six roots coincide, and we have the question how many roots of u are different, and whether there exist words u such that all the six roots of u are different each other. This last question was raised in [1], and it was first assumed that they do not exist. But in 2010 Georg Lohmann discovered the first of such words [4].

Definition 3.4. Let $k \in \{1, 2, 3, 4, 5, 6\}$. A word $u \in X^+$ is called a k-root word if $\{root(u), sroot(u), hroot(u), ssroot(u), shroot(u), hhroot(u)\}$ has exactly k elements.

A 6-root word is also called a Lohmann word.

If u is a k-root word and $k \neq 1$ and $k \neq 6$, it is interesting to know which of the roots of u coincide and which do not. This is the question for classifying the k-root words. To answer this question we introduce the following notions.

Definition 3.5. Let $k \in \{1, 2, 3, 4, 5, 6\}$.

A k-cluster is a set of the form $[\alpha_{11} \ldots \alpha_{1i_1}/\alpha_{21} \ldots \alpha_{2i_2}/ \ldots /\alpha_{k1} \ldots \alpha_{ki_k}]$, where $i_1 + i_2 + \ldots + i_k = 6$, $\{\alpha_{11}, \ldots, \alpha_{1i_1}\} \cup \{\alpha_{21}, \ldots, \alpha_{2i_2}\} \cup \ldots \cup \{\alpha_{k1}, \ldots, \alpha_{ki_k}\} = \{hh, ss, sh, h, s, r\}$ and $[\alpha_{11} \ldots \alpha_{1i_1}/ \ldots /\alpha_{k1} \ldots \alpha_{ki_k}] =_{Df} \{u : u \in X^* \wedge \alpha_{11}root(u) = \ldots = \alpha_{1i_1}root(u) \sqsubset \alpha_{21}root(u) = \ldots = \alpha_{2i_2}root(u) \sqsubset \ldots \sqsubset \alpha_{k1}root(u) = \ldots = \alpha_{ki_k}root(u)\}$.

For $\alpha = r$, $\alpha root(u)$ means $root(u)$. Thus, for instance, $[\,hh\,sh\,h/ss/\,s\,r\,]$ denotes the set of all words u satisfying $hhroot(u) = shroot(u) = hroot(u) \sqsubset ssroot(u) \sqsubset sroot(u) = root(u)$.

We will say, that a cluster exists if it is not empty.

For instance, clusters of the form $[r/\ldots]$ or $[\ldots/hh\ldots]$ cannot exist.

For different clusters $\mathcal{C}$, $\mathcal{C}'$ we have $\mathcal{C} \cap \mathcal{C}' = \emptyset$, $\bigcup_{\substack{\mathcal{C} \text{ is a} \\ k\text{-cluster}}} \mathcal{C}$ is the set of all k-root words, and $\bigcup\{\mathcal{C} : \mathcal{C} \text{ is a } k\text{-cluster} \wedge k \in \{1,2,3,4,5,6\}\} = X^+$.

Thus the k-clusters give a more detailed characterization for all k-root words. Of course there is only one 1-cluster namely $[\,hh\,ss\,sh\,h\,s\,r\,]$. It contains all 1-root words. But there are periodic and primitive words in it, for instance $(ab)^n$ and ab^n for $n > 1$. Therefore it is also interesting to characterize these words.

Definition 3.6. A cluster $\mathcal{C}$ is called a primitive cluster if it contains only primitive words. A cluster $\mathcal{C}$ is called a periodic cluster if it contains only periodic words. A cluster is called a mixed cluster if it contains both primitive and periodic words.

In the next section we shall prove some lemmas which we will use for a complete overview of all k-clusters for $1 < k \leq 6$, and after that we make still some remarks for Lohmann words.

4. Some lemmas

Lemma 4.1. *For a nonempty word u, only one of the following relationships is possible (where we will use hh, ss, sh, h, s, r, respectively, as shorthand expressions instead of $hhroot(u)$, $ssroot(u)$, $shroot(u)$, $hroot(u)$, $sroot(u)$, $root(u)$, respectively):*

(1) $hh \sqsubseteq ss \sqsubseteq sh \sqsubseteq h \sqsubseteq s \sqsubseteq r$;
(2) $hh \sqsubseteq ss \sqsubseteq sh \sqsubseteq s \sqsubseteq h \sqsubseteq r$;
(3) $hh \sqsubseteq sh \sqsubseteq h \sqsubseteq ss \sqsubseteq s \sqsubseteq r$;
(4) $hh \sqsubseteq sh \sqsubseteq ss \sqsubseteq h \sqsubseteq s \sqsubseteq r$;
(5) $hh \sqsubseteq sh \sqsubseteq ss \sqsubseteq s \sqsubseteq h \sqsubseteq r$.

The proof follows immediately from Theorem 3.3 and Figure 1.

For a nonempty word u and a nonempty prefix v of u we define three parameters $s_{u,v}$, $t_{u,v}$, and $h_{u,v}$ which are natural numbers between 0 and $|u|-1$.

Assume, we have $u \in v^{\otimes n} \otimes Pr(v)$. Then $s_{u,v}$ is the sum of all lengths of overlaps occuring in this representation, $t_{u,v}$ is the length of the overlap between the last complete v and the coupled strict prefix of v (the $(n+1)$st v), and $h_{u,v}$ is the length of the overhanging part of the coupled $(n+1)$st v. More formally: Let $u \in v^{\otimes n} \otimes v'$ where $v' \in Pr(v)$. Then $s_{u,v} = n \cdot |v| + |v'| - |u|$, $t_{u,v} = |u'| + |v'| - |u|$ if $u \in u' \otimes v'$ and $u' \in v^{\otimes n}$, and $h_{u,v} = \begin{cases} 0 & \text{if } v' = e \\ |v| - |v'| & \text{if } v' \neq e \end{cases}$.

The parameters $s_{u,v}$, $t_{u,v}$, $h_{u,v}$ do not exist for each pair $[u,v]$ with $v \sqsubseteq u$. Let, for instance, $u = abaababa$. For $v \in \{a, ab, abaa\} \subseteq Pr(u)$, the parameters do not exist because there is no representation $u \in v^{\otimes n} \otimes Pr(v)$. For $v = aba$ we get $[s_{u,v}, t_{u,v}, h_{u,v}] = [1, 0, 0]$, for $v = abaab$ we get $[0, 0, 2]$, for $v = abaaba$ we get $[1, 1, 3]$, for $v = abaabab$ we get $[0, 0, 6]$ or $[2, 2, 4]$, and for $v = u$ we get $[0, 0, 0]$ or $[1, 1, 7]$ or $[3, 3, 5]$. This also illustrates that the parameters are not unique in every case.

Further, it is clear, that one of the parameters exists if and only if both of the others exist, that always $t_{u,v} \leq s_{u,v}$, and that $s_{u,v} = 0$ or $h_{u,v} = 0$ implies $t_{u,v} = 0$.

Lemma 4.2. *Let $u \in X^+$.*

$v = hhroot(u)$ iff v is the shortest prefix of u where $s_{u,v}$ exists.

$v = shroot(u)$ iff v is the shortest prefix of u where $s_{u,v}$ exists and $t_{u,v} = 0$.

$v = ssroot(u)$ iff v is the shortest prefix of u where $s_{u,v}$ exists and $s_{u,v} = t_{u,v}$.

$v = hroot(u)$ iff v is the shortest prefix of u where $s_{u,v}$ exists and $h_{u,v} = 0$.

$v = sroot(u)$ iff v is the shortest prefix of u where $s_{u,v}$ exists and $s_{u.v} = 0$.

$v = root(u)$ iff v is the shortest prefix of u where $s_{u,v}$ exists and $s_{u,v} = h_{u,v} = 0$.

The proof follows immediately from the definitions.

For the above example $u = abaababa$ we get $hhroot(u) = shroot(u) = hroot(u) = aba$, $ssroot(u) = sroot(u) = abaab$, and $root(u) = u$, and therefore it belongs to the cluster $[\,hh\,sh\,h\,/\,ss\,s\,/\,r\,]$.

Lemma 4.3. *If $shroot(u) = ssroot(u)$ for some $u \in X^+$, then also $sroot(u) = shroot(u) = ssroot(u)$, and therefore there is no 5-cluster with $sh = ss$.*

Proof. Let $v = shroot(u) = ssroot(u)$. By Lemma 4.2, v is the shortest prefix of u where $t_{u,v} = 0$ and at the same time it is also the shortest prefix of u where $s_{u,v} = t_{u,v}$. By Theorem 3.3 or Lemma 4.1, v is a prefix of $sroot(u)$. The latter is the shortest prefix of u where $s_{u,v} = 0$. This must be v itself. □

Lemma 4.4. *If $sroot(u) = hroot(u)$ for some $u \in X^+$, then also $root(u) = sroot(u) = hroot(u)$, and therefore there is no 5-cluster with $s = h$.*

The proof follows immediately in the same way.

Lemma 4.5. *If $ssroot(u) = hroot(u)$ for some $u \in X^+$, then also $root(u) = sroot(u) = ssroot(u) = hroot(u)$, and therefore there is no 4-cluster and no 5-cluster with $ss = h$.*

Proof. Let $v = ssroot(u) = hroot(u)$. By Lemma 4.2, v is the shortest prefix of u where $h_{u,v} = 0$ and at the same time it is also the shortest prefix of u where $s_{u,v} = t_{u,v}$. $h_{u,v} = 0$ implies $t_{u,v} = 0$ and thus we have $h_{u,v} = t_{u,v} = s_{u,v} = 0$. v is a prefix of $root(u)$ which by Lemma 4.2 is the shortest prefix of u where $s_{u,v} = h_{u,v} = 0$. This must be v itself, and $ssroot(u) = root(u)$ implies $sroot(u) = root(u)$. □

The next two lemmas are standard in the combinatorics of words, see for instance [6, 8]. Lemma 4.7 was first proved by Lyndon and Schützenberger [7] for elements of a free monoid, see also [3].

Lemma 4.6. *If a word $u \in X^+$ has a nonempty strict prefix which at the same time is also a suffix of u then $u = (\alpha\beta)^m\alpha$ for some $\alpha \in X^+$, $\beta \in X^*$ and $m \geq 1$, and the common prefix and suffix is $(\alpha\beta)^{m-1}\alpha$.*

Lemma 4.7. *If $pq = qp$ for nonempty words p and q, then p and q are powers of a common word.*

Lemma 4.8. *If $u \in X^+$ is in a k-cluster with $k > 1$ and $sroot(u) = root(u)$ then u is periodic.*

Proof. Let $v = hhroot(u)$, and assume $u = root(u) = sroot(u)$. Then $u \in v^{\otimes n} \otimes v'$ for some $n \geq 1$ and $v' \sqsubset v$. If u is in a k-cluster with $k > 1$ then $v \sqsubset u$. If $v' = e$ then v is also a suffix of u, otherwise v' is a nonempty suffix of u. Thus u has a common nonempty prefix and suffix. Therefore, by Lemma 4.6, $u = (\alpha\beta)^m\alpha$ for some $\alpha \neq e$ and $m \geq 1$. Then $sroot(u) \sqsubseteq \alpha\beta \sqsubset root(u)$, a contradiction. □

Lemma 4.9. *If $ssroot(u) \sqsubset root(u)$ for some $u \in X^+$, then u is primitive.*

Proof. Let $v = ssroot(u) \sqsubset p = root(u) \sqsubset u$ for a periodic word $u = p^r$, $r \geq 2$. Then $u \in \{v^n\} \otimes Pr(v)$ with $n \geq r$ because of $|v| < |p|$, and $u = v^n v''$ for a strict subword v'' of v, and $p = v^k v'$ for some $v' \sqsubset v$ and $k \geq 1$. It is $v' \neq e$ and $v'' \neq e$ because otherwise $root(u) \sqsubseteq v$, a contradiction. Now, v' is a nonempty strict prefix of p and at the same time a strict suffix of p. By Lemma 4.6, $v' = (\alpha\beta)^{m-1}\alpha$ and $p = (\alpha\beta)^m\alpha$ for some $\alpha \in X^+$, $\beta \in X^*$ and $m \geq 1$. It is $\beta \neq e$ because otherwise $root(u) \sqsubseteq \alpha \sqsubset p$. We have $p = v^k v' = (\alpha\beta)^m\alpha = v^k(\alpha\beta)^{m-1}\alpha$ which implies $v^k = \alpha\beta$, and because of $v' = (\alpha\beta)^{m-1}\alpha \sqsubset v \sqsubseteq v^k = \alpha\beta$ we get $m = 1$. Then $v' = \alpha$, $p = \alpha\beta\alpha = v^k\alpha$ and $u = p^r = (v^k\alpha)^r = v^n v''$, where $n \geq rk \geq 2k$. Note that $\alpha = v' \sqsubset v$, and omitting the prefix v^k we get $\alpha(v^k\alpha)^{r-1} = v^{n-k}v''$. If $n - k \geq 2$, this means $\alpha v \ldots = v\alpha \ldots$, and by Lemma 4.7, α and v and therefore also p are powers of a common word $x \sqsubseteq \alpha \sqsubset v \sqsubset p$, contradicting $p = root(u)$.

It remains the case that $n - k < 2$, which means $k = 1$, $n = 2$ and $r = 2$ because of $2 \leq 2k \leq rk \leq n < k+2$ in this case. Now we have $v' = \alpha$, $v = \alpha\beta$, $p = \alpha\beta\alpha$ and $u = p^2 = \alpha\beta\alpha\alpha\beta\alpha = v^2v'' = \alpha\beta\alpha\beta v''$ and therefore $|v''| = 2|\alpha|$ and $\alpha\beta\alpha = \beta v''$. If $v'' \sqsubset v$ then $\alpha\beta = \beta\alpha$ would follow and therefore by Lemma 4.7, α and β and therefore also p are powers of a common word, the same contradiction as before. Thus there must be a real overlap between v^2 and a strict prefix of v yielding to v'', which means that $0 < |v''| = 2|\alpha| \leq |v| - 2 = |\alpha| + |\beta| - 2$ and therefore $|\alpha| \leq |\beta| - 2$. But then α must be a strict prefix of β, $\beta = \alpha\alpha_1$ for some $\alpha_1 \in X^+$. Then from $\beta v'' = \alpha\beta\alpha$ follows $\alpha\alpha_1 v'' = \alpha\alpha\alpha_1\alpha$ and therefore $\alpha_1 v'' = \alpha\alpha_1\alpha = \beta\alpha$. Hence $\alpha_1 \sqsubset \beta$ and $\beta = \alpha_1\alpha_2$ for some $\alpha_2 \in X^+$. By Lemma 4.6 we get $\beta = (\gamma\delta)^l\gamma$, $\alpha_1 = (\gamma\delta)^{l-1}\gamma$, $\alpha = \gamma\delta$, $\alpha_2 = \delta\gamma$ for some $\gamma \in X^+$, $\delta \in X^*$ and $l \geq 1$. It is $\delta \neq e$ and $\delta \neq \gamma$ since otherwise $root(u) \sqsubset p$. We get $\beta v'' = \alpha_1\alpha_2 v'' = \alpha\beta\alpha = \alpha\alpha_1\alpha_2\alpha$. Hence $\alpha_2\alpha = v''$ since $|\alpha| = |\alpha_2|$. $v'' = \alpha_2\alpha = \delta\gamma\gamma\delta$ is a subword of $v = \alpha\beta = (\gamma\delta)^{l+1}\gamma$. Then we have four cases.

(i) $\delta\gamma\gamma\delta = \gamma\delta\gamma\delta$. Then $\delta\gamma = \gamma\delta$, and by Lemma 4.7, δ and γ and therefore also p are powers of a common word, a contradiction.
(ii) $\delta\gamma\gamma\delta = \gamma''\delta\gamma\delta\gamma'$ where $\gamma'\gamma'' = \gamma$. Then $\delta\gamma'\gamma'' = \gamma''\delta\gamma' = \gamma'\gamma''\delta$, which means $\delta\gamma = \gamma\delta$ with the same conclusion as in (i).
(iii) $\delta\gamma\gamma\delta = \delta\gamma\delta\gamma$, as in (i).
(iv) $\delta\gamma\gamma\delta = \delta''\gamma\delta\gamma\delta'$ where $\delta'\delta'' = \delta$, as in (ii).

All cases yield to a contradiction which proves the lemma. $\square$

A consequence of Lemmas 4.8 and 4.9 is

Lemma 4.10. *There is no cluster with* $ss \sqsubset s = r$.

Lemma 4.11. *If* p *is the shortest word such that* $u \in p^{\otimes m}$ *for some* $m \geq 2$ *then the greatest overlap in the representation of* $u \in p^{\otimes m}$ *is not longer than* $\frac{|p|}{2}$ *and therefore* $|p| \leq \frac{2|u|}{3}$.

Proof. If there is an overlap w in $p \otimes p$ then w is a common prefix and suffix of p and therefore by Lemma 4.6, $p = (\alpha\beta)^r\alpha$ for some $\alpha \in X^+$, $\beta \in X^*$ and $r \geq 1$, and $w = (\alpha\beta)^{r-1}\alpha$. Since p is the shortest word for the representation of u, $r = 1$ must follow. Therefore $w = \alpha$ and $|w| \leq \frac{|p|}{2}$. Then from $u \in p^{\otimes m}$ with $m \geq 2$ and w is the longest overlap in this representation, we get $|u| \geq m|p| - (m-1)|w| \geq \frac{m+1}{2}|p| \geq \frac{3}{2}|p|$ or $|p| \leq \frac{2}{3}|u|$. $\square$

Lemma 4.12. *If* $ssroot(u) \sqsubset hroot(u) \sqsubset root(u)$ *for some* $u \in X^+$, *then* $ssroot(u) = sroot(u)$, *and therefore there is no 6-cluster with* $ss \sqsubset h$.

Proof. Let $v = ssroot(u) \sqsubset p = hroot(u) \sqsubset root(u)$ and $ssroot(u) \sqsubset sroot(u)$. Then by Lemma 4.9, u is primitive and therefore $u = root(u) \in p^{\otimes m}$ and $u \in v^n \otimes v'$ for some $m \geq 2$, $n \geq 1$ and $v' \sqsubset v$. There must be an overlap between v^n and v' since otherwise $sroot(u) \sqsubseteq v = ssroot(u)$. This means, $u = v^n v''$ for a strict subword v'' of v, more exactly, $v = v_l v'' v_r$ with $v_l, v_r \neq e$. If $n = 1$, then $u = vv''$ with $|u| \leq 2|v| - 2$ and therefore $|v| \geq \frac{|u|}{2} + 1$ and $|p| \geq \frac{|u|}{2} + 2$. If m would be greater than 2 then by Lemma 4.11, $|u| \geq m|p| - (m-1) \cdot \frac{|p|}{2} = \frac{m+1}{2}|p| \geq 2|p| \geq |u| + 4$ which is not possible. Therefore $m = 2$ must follow and then we have $p = wqw$ and $u = wqwqw \in p^{\otimes 2}$. But then $ssroot(u) \sqsubseteq sroot(u) \sqsubseteq wq$ which means, $|v| \leq \frac{|u|}{2}$, a contradiction. Now we must have $u = v^n v'' \in p^{\otimes m}$ with $m, n \geq 2$, $v \sqsubset p$. Let $p = v^k q$ for some q with $|q| < |v|$ and $1 \leq k \leq n$. It is $q \neq e$ because otherwise $hroot(u) \sqsubseteq v$. It is $k < n$ because otherwise $|p| > \frac{2}{3}|u|$ contradicting Lemma 4.11. Then $q \sqsubset v$ because of $p = v^k q \sqsubset u = v^n v''$. Now we have $v = qq'$ for nonempty words q, q', and $p = (qq')^k q$. p is the hyperroot of u, and therefore $k = 1$ because otherwise $qq'q$ would be a shorter candidate for the hyperroot. Hence $u = (qq')^n v'' \in (qq'q)^{\otimes m}$ and v'', which is shorter than $v = qq'$, is a suffix of $q'q$. If $|v''| \leq |q|$ then v'' is a suffix of q and $qq'v'' = v''q'q$ must follow. This means, $v'' \sqsubseteq q$ and $sroot(u) \sqsubseteq qq' = ssroot(u)$, a contradiction. If $|q| < |v''| < |qq'|$ then remember that $v = qq' = v_l v'' v_r$ with $v_l, v_r \neq e$ and v_l is a suffix of v. From

$[\,hh\,/\,ss\,sh\,h\,s\,r\,]$	$+$	$[\,hh\,sh\,h\,/\,ss\,s\,r\,]$	$+$
$[\,hh\,ss\,/\,sh\,h\,s\,r\,]$	10	$[\,hh\,ss\,sh\,h\,/\,s\,r\,]$	3,5,10
$[\,hh\,sh\,/\,ss\,h\,s\,r\,]$	$+$	$[\,hh\,ss\,sh\,s\,/\,h\,r\,]$	$+$
$[\,hh\,ss\,sh\,/\,h\,s\,r\,]$	3,10	$[\,hh\,ss\,sh\,h\,s\,/\,r\,]$	4,5

FIGURE 2. 2-clusters.

$u = (v_l v'' v_r)^n v'' \in (v_l v'' v_r q)^{\otimes m}$ we get $v'' = v_1 q$ for some v_1 which is a strict nonempty suffix of q', and $v_l v_1 q v_r v_1 q = \dots v_l v_1 q v_r q$, therefore $q v_r v_1 = v_1 q v_r$ and v_l is a suffix of $v_l v_1$. By Lemma 4.7, $q v_r$ and v_1 are powers of a common primitive word x, $v_1 = x^\alpha$, $q v_r = x^\beta$ for integers $\alpha, \beta \geq 1$. Then $v = v_l v_1 q v_r = v_l x^{\alpha+\beta}$. Since v_l is a suffix of $v_l v_1$ it follows from Lemma 4.6 that $v_l = y v_1^s$ for some $s \geq 0$ and a nonempty suffix y of v_1. Since $v_1 = x^\alpha$ we can assume that $v_l = y x^t$ for some $t \geq 0$ and y is a nonempty suffix of x. Now we have $v = y x^c$ with $c = t + \alpha + \beta \geq 2$ and $p = vq = y x^c q$ with $e \sqsubset q \sqsubset x^\beta$. Let $q = x^\gamma q_1$ with $0 \leq \gamma < \beta$ and $q_1 \sqsubset x$. Then $p = y x^c q = y x^{c+\gamma} q_1$ has the hyperroot $hroot(p) \sqsubseteq y x q_1$ (since y is a suffix of x and $c \geq 2$) which is shorter than p. This is a contradiction with $hroot(u) = p$. □

5. k-CLUSTERS FOR $2 \leq k \leq 5$

By Lemma 4.1, altogether 89 k-clusters are possible: 1 for $k = 1$, 8 for $k = 2$, 23 for $k = 3$, 32 for $k = 4$, 20 for $k = 5$, and 5 for $k = 6$. For $k \in \{2, 3, 4, 5\}$ they are listed in Figures 2 to 5. The only 1-cluster is $[\,hh\,ss\,sh\,h\,s\,r\,]$. Most of the remaining clusters cannot exist by our Lemmas 4.3, 4.4, 4.5, 4.10, 4.12. In the tables in Figures 2 to 5, the numbers of the lemmas proving the non-existence of the clusters are indicated. Lemma 4.4, in this connection is unnecessary. The sign $+$ means that the existence of the corresponding cluster is sure. The 28 existing k-clusters for $2 \leq k \leq 5$ are also listed in Theorems 5.1 to 5.4 together with their smallest elements. The only existing 6-cluster will be considered in the next section.

Theorem 5.1. *There exist only the following 2-clusters which have the shown quasi-lexicographic smallest elements* (*according to the order* $<$ *on* $X = \{a, b\}$)*:*

$$aba \in [\,hh\,ss\,sh\,s\,/\,h\,r\,],$$

$$abaababaab \in [\,hh\,sh\,/\,ss\,h\,s\,r\,],$$

$$ababaababa \in [\,hh\,sh\,h\,/\,ss\,s\,r\,],$$

$$abaabababaabab \in [\,hh\,/\,ss\,sh\,h\,s\,r\,].$$

$[\,hh\,/\,ss\,/\,sh\,h\,s\,r\,]$	10	$[\,hh\,ss\,sh\,/\,s\,/\,h\,r\,]$	3
$[\,hh\,/\,ss\,sh\,/\,h\,s\,r\,]$	3,10	$[\,hh\,ss\,sh\,s\,/\,h\,/\,r\,]$	+
$[\,hh\,/\,ss\,sh\,h\,/\,s\,r\,]$	3,5,10	$[\,hh\,/\,sh\,/\,h\,ss\,s\,r\,]$	+
$[\,hh\,/\,ss\,sh\,h\,s\,/\,r\,]$	4,5	$[\,hh\,/\,sh\,h\,/\,ss\,s\,r\,]$	+
$[\,hh\,ss\,/\,sh\,/\,h\,s\,r\,]$	10	$[\,hh\,sh\,/\,h\,/\,ss\,s\,r\,]$	+
$[\,hh\,ss\,/\,sh\,h\,/\,s\,r\,]$	10,12	$[\,hh\,sh\,/\,h\,ss\,/\,s\,r\,]$	5,10
$[\,hh\,ss\,/\,sh\,h\,s\,/\,r\,]$	4,12	$[\,hh\,sh\,/\,h\,ss\,s\,/\,r\,]$	4,5
$[\,hh\,ss\,sh\,/\,h\,/\,s\,r\,]$	3,10,12	$[\,hh\,sh\,h\,/\,ss\,/\,s\,r\,]$	10
$[\,hh\,ss\,sh\,/\,h\,s\,/\,r\,]$	3,4,12	$[\,hh\,sh\,h\,/\,ss\,s\,/\,r\,]$	+
$[\,hh\,ss\,sh\,h\,/\,s\,/\,r\,]$	3,5	$[\,hh\,sh\,/\,ss\,/\,h\,s\,r\,]$	10
$[\,hh\,/\,ss\,sh\,s\,/\,h\,r\,]$	+	$[\,hh\,sh\,/\,ss\,s\,/\,h\,r\,]$	+
$[\,hh\,ss\,/\,sh\,s\,/\,h\,r\,]$	+		

Figure 3. 3-clusters.

Theorem 5.2. *There exist only the following 3-clusters which have the shown quasi-lexicographic smallest elements:*

$$\begin{aligned}
ababa &\in [\,hh\,ss\,sh\,s\,/\,h\,/\,r\,],\\
ababaa &\in [\,hh\,sh\,/\,ss\,s\,/\,h\,r\,],\\
abaabab &\in [\,hh\,ss\,/\,sh\,s\,/\,h\,r\,],\\
abaababa &\in [\,hh\,sh\,h\,/\,ss\,s\,/\,r\,],\\
abaababab &\in [\,hh\,/\,ss\,sh\,s\,/\,h\,r\,],\\
(abaabaab)^2 &\in [\,hh\,sh\,/\,h\,/\,ss\,s\,r\,],\\
(ababaabab)^2 &\in [\,hh\,/\,sh\,/\,h\,ss\,s\,r\,],\\
(abaababaabab)^2 &\in [\,hh\,/\,sh\,h\,/\,ss\,s\,r\,].
\end{aligned}$$

Theorem 5.3. *There exist only the following 4-clusters which have the shown quasi-lexicographic smallest elements:*

$$\begin{aligned}
abaabaabab &\in [\,hh\,ss\,/\,sh\,/\,s\,/\,h\,r\,],\\
ababaababaa &\in [\,hh\,sh\,/\,ss\,s\,/\,h\,/\,r\,],\\
ababaababab &\in [\,hh\,/\,ss\,/\,sh\,s\,/\,h\,r\,],\\
abababaabab &\in [\,hh\,/\,sh\,/\,ss\,s\,/\,h\,r\,],\\
aabaaabaabaa &\in [\,hh\,sh\,/\,h\,/\,ss\,s\,/\,r\,],\\
abaababaabaa &\in [\,hh\,sh\,/\,ss\,/\,s\,/\,h\,r\,],\\
abaababaabab &\in [\,hh\,/\,ss\,sh\,s\,/\,h\,/\,r\,],\\
ababaabababa &\in [\,hh\,sh\,h\,/\,ss\,/\,s\,/\,r\,],\\
abaababaabababaabab &\in [\,hh\,/\,sh\,h\,/\,ss\,s\,/\,r\,],\\
(ababaababaabab)^2 &\in [\,hh\,/\,sh\,/\,h\,/\,ss\,s\,r\,].
\end{aligned}$$

$[\,hh\,ss\,/\,sh\,h\,/\,s\,/\,r\,]$	12	$[\,hh\,sh\,/\,ss\,/\,h\,s\,/\,r\,]$	4,12
$[\,hh\,ss\,/\,sh\,/\,h\,s\,/\,r\,]$	4,12	$[\,hh\,sh\,/\,ss\,/\,h\,/\,s\,r\,]$	10,12
$[\,hh\,ss\,/\,sh\,/\,h\,/\,s\,r\,]$	10,12	$[\,hh\,sh\,/\,ss\,s\,/\,h\,/\,r\,]$	+
$[\,hh\,/\,ss\,sh\,/\,h\,s\,/\,r\,]$	3,4,12	$[\,hh\,sh\,/\,ss\,/\,s\,/\,h\,r\,]$	+
$[\,hh\,/\,ss\,sh\,/\,h\,/\,s\,r\,]$	3,10,12	$[\,hh\,/\,sh\,/\,ss\,s\,/\,h\,r\,]$	+
$[\,hh\,/\,ss\,/\,sh\,h\,/\,s\,r\,]$	10,12	$[\,hh\,ss\,sh\,/\,h\,/\,s\,/\,r\,]$	3,12
$[\,hh\,ss\,/\,sh\,s\,/\,h\,/\,r\,]$	12	$[\,hh\,/\,ss\,sh\,h\,/\,s\,/\,r\,]$	3,5
$[\,hh\,ss\,/\,sh\,/\,s\,/\,h\,r\,]$	+	$[\,hh\,/\,ss\,/\,sh\,h\,s\,/\,r\,]$	4,12
$[\,hh\,/\,ss\,sh\,/\,s\,/\,h\,r\,]$	3	$[\,hh\,/\,ss\,/\,sh\,/\,h\,s\,r\,]$	10
$[\,hh\,/\,ss\,/\,sh\,s\,/\,h\,r\,]$	+	$[\,hh\,ss\,sh\,/\,s\,/\,h\,/\,r\,]$	3,12
$[\,hh\,sh\,/\,h\,ss\,/\,s\,/\,r\,]$	5	$[\,hh\,/\,ss\,sh\,s\,/\,h\,/\,r\,]$	+
$[\,hh\,sh\,/\,h\,/\,ss\,s\,/\,r\,]$	+	$[\,hh\,sh\,h\,/\,ss\,/\,s\,/\,r\,]$	+
$[\,hh\,sh\,/\,h\,/\,ss\,/\,s\,r\,]$	10	$[\,hh\,/\,sh\,/\,h\,ss\,s\,/\,r\,]$	4,5
$[\,hh\,/\,sh\,h\,/\,ss\,s\,/\,r\,]$	+	$[\,hh\,/\,sh\,/\,h\,/\,ss\,s\,r\,]$	+
$[\,hh\,/\,sh\,h\,/\,ss\,/\,s\,r\,]$	10	$[\,hh\,sh\,ss\,/\,h\,/\,s\,/\,r\,]$	3,12
$[\,hh\,/\,sh\,/\,h\,ss\,/\,s\,r\,]$	5,10	$[\,hh\,/\,sh\,/\,ss\,/\,h\,s\,r\,]$	10

FIGURE 4. 4-clusters.

$[\,hh\,ss\,/\,sh\,/\,h\,/\,s\,/\,r\,]$	12	$[\,hh\,/\,sh\,h\,/\,ss\,/\,s\,/\,r\,]$	+
$[\,hh\,/\,ss\,sh\,/\,h\,/\,s\,/\,r\,]$	3,12	$[\,hh\,/\,sh\,/\,h\,ss\,/\,s\,/\,r\,]$	5
$[\,hh\,/\,ss\,/\,sh\,h\,/\,s\,/\,r\,]$	12	$[\,hh\,/\,sh\,/\,h\,/\,ss\,s\,/\,r\,]$	+
$[\,hh\,/\,ss\,/\,sh\,/\,h\,s\,/\,r\,]$	4,12	$[\,hh\,/\,sh\,/\,h\,/\,ss\,/\,s\,r\,]$	10
$[\,hh\,/\,ss\,/\,sh\,/\,h\,/\,s\,r\,]$	10,12	$[\,hh\,sh\,/\,ss\,/\,h\,/\,s\,/\,r\,]$	12
$[\,hh\,ss\,/\,sh\,/\,s\,/\,h\,/\,r\,]$	12	$[\,hh\,/\,sh\,/\,ss\,/\,h\,s\,/\,r\,]$	4,12
$[\,hh\,/\,ss\,sh\,/\,s\,/\,h\,/\,r\,]$	3,12	$[\,hh\,/\,sh\,/\,ss\,/\,h\,/\,s\,r\,]$	10,12
$[\,hh\,/\,ss\,/\,sh\,s\,/\,h\,/\,r\,]$	12	$[\,hh\,sh\,/\,ss\,/\,s\,/\,h\,/\,r\,]$	12
$[\,hh\,/\,ss\,/\,sh\,/\,s\,/\,h\,r\,]$	+	$[\,hh\,/\,sh\,/\,ss\,s\,/\,h\,/\,r\,]$	+
$[\,hh\,sh\,/\,h\,/\,ss\,/\,s\,/\,r\,]$	+	$[\,hh\,/\,sh\,/\,ss\,/\,s\,/\,h\,r\,]$	+

FIGURE 5. 5-clusters.

Theorem 5.4. *There exist only the following 5-clusters which have the shown quasi-lexicographic smallest elements:*

$$\begin{aligned} abaabaabaabababaab &\in [\,hh\,sh\,/\,h\,/\,ss\,/\,s\,/\,r\,], \\ ababaababaabababab &\in [\,hh\,/\,ss\,/\,sh\,/\,s\,/\,h\,r\,], \\ ababaabababaabab &\in [\,hh\,/\,sh\,/\,ss\,s\,/\,h\,/\,r\,], \\ abaababaabaababaabab &\in [\,hh\,/\,sh\,/\,ss\,/\,s\,/\,h\,r\,], \\ ababaababaabababaabab &\in [\,hh\,/\,sh\,/\,h\,/\,ss\,s\,/\,r\,], \\ abaababaababababaababaababaabab &\in [\,hh\,/\,sh\,h\,/\,ss\,/\,s\,/\,r\,]. \end{aligned}$$

Proofs. The membership of the elements is easy to see. That they are the appropriate smallest elements can be shown, for instance, by computer experiments [5].

In the computer programs [5], Lemma 4.2 was exploited. The non-existence of the remaining clusters follows by our lemmas as noted. □

6. Lohmann words

Up to the discovering of the first 6-root word by Georg Lohmann [4] it was unknown whether 6-clusters exist. On March 25, 2010 [4], he discovered the first 6-root word which was *ababaababababaababaababababaabab*. Later on he found with *ababaababaababaababababaabab* a smaller one. To give a sufficient condition for a word to be a Lohmann word, we introduce the following notions.

Definition 6.1. For finite sequences $(k_1,\dots,k_r)$ and $(t_1,\dots,t_m)$ of natural numbers, let $(k_1,\dots,k_r)\odot(t_1,\dots,t_m) =_{Df} (k_1,\dots,k_{r-1},k_r+t_1,t_2,\dots,t_m)$ and $(k_1,\dots,k_r)^{\odot 0} =_{Df} (0)$, $(k_1,\dots,k_r)^{\odot s} =_{Df} (k_1,\dots,k_r)^{\odot s-1}\odot(k_1,\dots,k_r)$ for $s\geq 1$.
Let $(k_1,\dots,k_r,t)$ be a sequence of natural numbers with $r\geq 2$, $2\leq k_1\leq k_i\leq 2k_1$ for each $i\in\{1,\dots,r\}$ and $0\leq t\leq k_1$, and let $s\geq 2$ and $(k_1,\dots,k_r,t)^{\odot s} = (k_1,\dots,k_r,k_{r+1},\dots,k_{s\cdot r},t)$. If $t\neq 0$ and $k_1\leq k' < k_1+t$ then $(k_1,\dots,k_{s\cdot r},k')$ is called an L-sequence with its producer $(k_1,\dots,k_r,t)$; if $t=0$ and k' with $\max\{k_1,\dots,k_r\} < k' \leq 2k_1$ exists then $(k_1,\dots,k_{s\cdot r-1},k')$ is called an L-sequence with its producer $(k_1,\dots,k_r,0)$.

Theorem 6.2. *Let v and w be words such that $e\sqsubset v\sqsubset w$, $wv\not\sqsubseteq p^l$ for some $p\sqsubset w$ and $l>1$, let $(k_1,\dots,k_n)$ be an L-sequence and k^+ the greatest number in this sequence, and let w' be a word such that $w^{k-1}v\sqsubset w'\sqsubset w^k v$ for some $k\geq 2$ with $k^+-k_1\leq k\leq k_1$ and $w^2\sqsubseteq w'$. Then $u=w^{k_1}vw^{k_2}v\dots w^{k_n}vw'$ is a Lohmann word.*

Proof. Let $(k_1,\dots,k_r,t)$ be the shortest producer of the L-sequence $(k_1,\dots,k_n)$. Then the proof is done by verifying the roots:

$$\begin{aligned}
hhroot(u) &= wv,\\
shroot(u) &= w^{k_1}v,\\
hroot(u) &= w^{k_1}vw',\\
ssroot(u) &= w^{k_1}vw^{k_2}v\dots w^{k_r}vw^t,\\
ssroot(u) &\sqsubset sroot(u)\sqsubseteq w^{k_1}vw^{k_2}v\dots w^{k_{n-1}}vw^{k_n-k_1},\\
root(u) &= u.
\end{aligned}$$
□

Examples. (2,3,0), (2,2,0), (2,2,1), (2,4,1), (2,4,2), (3,4,4,2) are producer of the L-sequences (2,3,2,4), (2,2,2,3), (2,2,3,2,2), (2,4,3,4,2), (2,4,4,4,3), (3,4,4,5,4,4,5,4,4,3), respectively. (2,3,0) is also a producer of (2,3,2,3,2,4). From (2,3,2,4) and (2,2,2,3) with $w=ab$, $v=a$, and $w'=(ab)^2$ we get the Lohmann words mentioned above before Definition 6.1.

The condition in Theorem 6.2 is not necessary for Lohmann words. For instance, (2,1,3,1,3,1,3,1,2,1) produces Lohmann words as in Theorem 6.2 but it is not an L-sequence according to Definition 6.1. In the same way, the sequence (2,1,2,1,4,1,2,1,3,1) produces words belonging to the cluster $[\,hh\,/\,sh\,h\,/\,ss\,/\,s\,/\,r\,]$. All Lohmann words belong to the 6-cluster $[\,hh\,/\,sh\,/\,h\,/\,ss\,/\,s\,/\,r\,]$. The four forther 6-clusters which are possible by Lemma 4.1 cannot exist by Lemma 4.12.

7. Final remarks

The periodic k-root words are also called strong k-root words.

The next two theorems follow from Section 5, Section 6 and Lemmas 4.8 and 4.9. Theorem 7.1 corrects Theorem 55 in [3].

Theorem 7.1. *The quasi-lexicographic smallest k-root words are a for $k = 1$, aba for $k = 2$, $ababa$ for $k = 3$, $abaabaabab$ for $k = 4$, $abaabaabaababaab$ for $k = 5$, and $ababaababaababaabababaabab$ for $k = 6$.*

The quasi-lexicographic smallest strong k-root words are aa for $k = 1$, $abaababaab$ for $k = 2$, $(abaabaab)^2$ for $k = 3$, and $(ababaababaabab)^2$ for $k = 4$. There exist no strong k-root words for $k = 5$ and $k = 6$.

Theorem 7.2. *There exists only one mixed cluster namely the 1-cluster $[\,hh\,ss\,sh\,h\,s\,r\,]$. The clusters $[\,hh\,sh\,/\,ss\,h\,s\,r\,]$, $[\,hh\,sh\,h\,/\,ss\,s\,r\,]$, $[\,hh\,/\,ss\,sh\,h\,s\,r\,]$, $[\,hh\,sh\,/\,h\,/\,ss\,s\,r\,]$, $[\,hh\,/\,sh\,/\,h\,ss\,s\,r\,]$, $[\,hh\,/\,sh\,h\,/\,ss\,s\,r\,]$ and $[\,hh\,/\,sh\,/\,h\,/\,ss\,s\,r\,]$ are periodic clusters, and all other existing clusters are primitive.*

The following theorem was first conjectured by Georg Lohmann in a weaker form.

Theorem 7.3. *For an arbitrary nonempty word p holds true that $\alpha root(p^d) = \alpha root(p^{d'})$ for all $d, d' \geq 2$ and $\alpha \in \{hh, sh, ss, h, s, r\}$ and therefore all periodic words which are powers of the same word are in the same cluster.*

Proof. Let p be primitive and $u = p^d$ periodic, $d \geq 2$. Then by Lemmas 4.9 and 4.1, $ssroot(u) = sroot(u) = root(u) = p$ and therefore also $ssroot(p^{d'}) = sroot(p^{d'}) = root(p^{d'}) = p$ for each $d' \geq 2$. Now we show that $\alpha root(p^{d'}) = \alpha root(p^2)$ for each $d' \geq 2$ and $\alpha \in \{h, sh, hh\}$ by induction over d'.

Step from 2 to 3. Let $\alpha root(p^2) = v$, $v \sqsubseteq p$.

Then $p^2 = v_{1,1}v_{1,2}v_{2,1}v_{2,2}\ldots v_{k,1}v_{k,2}\ldots v_{l,1}v_{l,2}v_{l+1,1}$ where $k < l$, $v_{1,1}v_{1,2} = v_{1,2}v_{2,1}v_{2,2} = \ldots = v_{k-1,2}v_{k,1}v_{k,2} = \ldots = v_{l-1,2}v_{l,1}v_{l,2} = v$ with $v_{l+1,1} = e$ if $\alpha = h$, $v_{l,2} = e$ and $v_{l+1,1} \sqsubset v$ if $\alpha = sh$, $v_{l,2}v_{l+1,1} \sqsubset v$ if $\alpha = hh$, and $p = v_{1,1}\ldots v_{k-1,2}v' = v''v_{k+1,1}\ldots v_{l,2}v_{l+1,1}$ with $v'v'' = v_{k,1}v_{k,2}$. Let $p_1 = v_{1,1}\ldots v'$ and $p_2 = v''\ldots v_{l+1,1}$, $p_1 = p_2 = p$. Then $p^3 = p_1p_2p_2 = p_1p_1p_2$ must be of the form $v_{1,1}v_{1,2}v_{2,1}v_{2,2}\ldots v_{m,1}v_{m,2}v_{m+1,1}$ where $m > l$, $v_{1,1}v_{1,2} = v_{1,2}v_{2,1}v_{2,2} = \ldots = v_{m-1,2}v_{m,1}v_{m,2} = v$ and $v_{m,2} = v_{l,2}$, $v_{m+1,1} = v_{l+1,1}$. Therefore $\alpha root(p^3) \sqsubseteq v$. If $\alpha root(p^3)$ would be shorter than v then removing of the middle p from an appropriate representation of ppp would yield to a shorter $\alpha root$ of p^2. Thus $\alpha root(p^3) = v$.

The step from $d \geq 2$ to $d+1$ is done analogously. $\square$

Acknowledgements. The author is very grateful to Masami Ito for the cooperation and for the hospitality during his stay in Kyoto, and to Georg Lohmann in Apolda for his interest and the assistence with his computer programs.

References

[1] M. Ito and G. Lischke, Generalized periodicity and primitivity for words. *Math. Log. Quart.* **53** (2007) 91–106; Corrigendum in *Math. Log. Quart.* **53** (2007) 642–643.
[2] G. Lischke, The primitivity distance of words, in *Automata, Formal Languages and Algebraic Systems*, edited by M. Ito, Y. Kobayashi and K. Shoji. World Scientific (2010) 125–137.
[3] G. Lischke, Primitive words and roots of words. *Acta Univ. Sapientiae, Informatica* **3** (2011) 5–34.
[4] G. Lohmann, e-mail to G. Lischke (2010).
[5] G. Lohmann, Program packet LIMA, Apolda (2010). Improvements 2012.
[6] M. Lothaire, *Combinatorics on Words*. Addison-Wesley, Reading Mass. (1983).
[7] R.C. Lyndon and M.P. Schützenberger, On the equation $a^M = b^N c^P$ in a free group. *Michigan Math. J.* **9** (1962) 289–298.
[8] H.J. Shyr, *Free Monoids and Languages*. Hon Min Book Company, Taichung (1991).
[9] S.S. Yu, *Languages and Codes*. Tsang Hai Book Publishing Co., Taichung (2005).

Communicated by Ch. Choffrut.
Received March 14, 2014. Accepted March 14, 2014.