

# TAMENESS OF THE PSEUDO-VARIETY $\mathbf{LSI}$

J. C. COSTA and M. L. TEIXEIRA<sup>1</sup>

*Centro de Matemática da Universidade do Minho*

*Campus de Gualtar, 4700-320 Braga, Portugal*

*Email: jcosta@math.uminho.pt, mlurdes@math.uminho.pt*

December 31, 2002

AMS Mathematics Subject Classification: 20M05, 20M07, 20M35

The notion of  $\kappa$ -tameness of a pseudovariety was introduced by Almeida and Steinberg and is a strong property which implies decidability of pseudovarieties. In this paper we prove that the pseudovariety  $\mathbf{LSI}$ , of local semilattices, is  $\kappa$ -tame.

*Keywords:* Semigroup; pseudovariety; implicit operation; tameness; local semilattice; infinite word; graph labelling.

## 1 Introduction

A class of finite semigroups  $\mathbf{V}$  is said to be decidable if the membership problem for  $\mathbf{V}$  has a solution, that is, if we can construct an algorithm to test whether a given semigroup lies in  $\mathbf{V}$ . Decidability of pseudovarieties is not preserved by some of the most common pseudovariety operators, such as semidirect product, Mal'cev product and join [1, 17]. In particular Rhodes [17] has exhibited a decidable pseudovariety  $\mathbf{V}$  such that the semidirect product  $\mathbf{SI} * \mathbf{V}$  is not decidable, where  $\mathbf{SI}$  is the pseudovariety of semilattices. This fact suggested the necessity to establish and to study stronger conditions which were expected to be important in proving decidability of pseudovarieties using these operators. In this context several concepts were introduced such as the notion of tameness, introduced by Almeida and Steinberg [7]. Tameness consists of a refinement of the concept of hyperdecidability previously introduced by Almeida [3].

If tameness is preserved by the most common operators is an open problem yet. That it is a property of the “usual” pseudovarieties is a conjecture. In general, proving tameness of a pseudovariety is a non trivial, but useful, exercise. Some pseudovarieties are already known to be tame such as, for example:  $\mathbf{G}$ , the pseudovariety of all finite groups (consequence of the interpretation of Ash’s results [10]);  $\mathbf{Ab}$  [6], the pseudovariety of all finite abelian groups;  $\mathbf{K}$  and  $\mathbf{D}$  [9], the pseudovarieties of semigroups whose idempotents are, respectively, left zeros and right zeros. For more examples and recent developments on the subject, the reader is referred to Almeida’s papers [4, 5].

Here we study the tameness of  $\mathbf{LSI}$ , the pseudovariety of finite semigroups  $S$  such that  $eSe \in \mathbf{SI}$ , for all idempotents  $e \in S$ . This problem was proposed by Almeida

---

<sup>1</sup>This work was supported, in part, by FCT through the *Centro de Matemática da Universidade do Minho*, and by the FCT and POCTI approved project POCTI/32817/MAT/2000 which is participated by the European Community Fund FEDER.

in [4]. Recall that the pseudovariety **LSI** is associated via Eilenberg's correspondence with the class of locally testable languages and that  $\mathbf{LSI} = \mathbf{SI} * \mathbf{D}$  [11, 14]. Proving the  $\kappa$ -tameness of **LSI** involves proving two properties: that the  $\kappa$ -word problem for **LSI** is decidable, and that **LSI** is  $\kappa$ -reducible. The first problem was already solved by the first author in [12]. So, this paper is essentially dedicated to the proof of the second problem. As an example of comparison, we note that the problem of proving the  $\kappa$ -tameness of the semidirect product  $\mathbf{SI} * \mathbf{K}$  is much more simple and was announced by the first author [13].

This work is organized as follows. After a section of preliminaries, where we introduce some notation and review some basic results on semigroups, pseudovarieties, words, graphs,  $\kappa$ -reducibility and  $\kappa$ -tameness, we present in Section 3 some properties about subpseudovarieties of **LSI** and their respective implicit operations. Section 4 is devoted to introduce some definitions on words and to establish some technical combinatoric results which are essential for our purposes. We next recall the result which establishes the decidability of the  $\kappa$ -word for **LSI** in Section 5. Finally, Section 6 gives the proof of the  $\kappa$ -reducibility of **LSI**.

## 2 Preliminaries

For general background and terminology from the classical theories of semigroups, formal languages, pseudovarieties and profinite semigroups the reader is referred to [2, 8, 15]. Here, we begin by presenting a brief description of the most relevant basic definitions, notation and results that will be needed in the following sections.

### 2.1 Words

A finite non-empty set  $A$  is called an *alphabet* and its elements are called *letters*. A (*finite*) *word* on  $A$  is a finite sequence  $w = (a_1, \dots, a_n)$  of elements of  $A$ , usually written  $w = a_1 \cdots a_n$ . The integer  $n$  is called the *length* of  $w$ . The empty sequence, called the *empty word*, is denoted by 1 and its length is 0. The length of a word  $w$  is denoted by  $|w|$ . The *product* of two words  $w = a_1 a_2 \cdots a_n$  and  $z = b_1 b_2 \cdots b_m$  is the word  $wz = a_1 a_2 \cdots a_n b_1 b_2 \cdots b_m$ . We denote by  $A^*$  the set of words on  $A$  and by  $A^+$  the set of non-empty words. The set  $A^*$  (resp.  $A^+$ ) endowed with the product is a monoid (resp. semigroup) whose identity is the empty word and is called the *free monoid* (resp. the *free semigroup*) generated by  $A$ .

A word  $w \in A^+$  is said to be *primitive* if it is not a power of another word; that is, if  $w = u^n$  for some  $u \in A^*$  and  $n \in \mathbb{N}$  implies  $w = u$  (and  $n = 1$ ). Two words  $w$  and  $z$  are said to be *conjugate* if there exist words  $u, v \in A^*$  such that  $w = uv$  and  $z = vu$ . We notice that, if  $w$  is a primitive word and  $z$  is a conjugate of  $w$ , then  $z$  is also primitive. Let an order be fixed for the letters of the alphabet  $A$ . A *Lyndon word* is a primitive word which is minimal in its conjugation class.

A *bi-infinite* (resp. *right-infinite*, *left-infinite*) word on  $A$  is a sequence  $w = (a_n)_n$  of letters of  $A$  indexed by  $\mathbb{Z}$  (resp.  $\mathbb{N}$ ,  $-\mathbb{N}$ ), also written

$$w = \cdots a_{-2} a_{-1} a_0 a_1 a_2 \cdots \quad (\text{resp. } w = a_1 a_2 \cdots, \quad w = \cdots a_{-2} a_{-1}).$$

We denote  $w(n) = a_n$  and say that  $a_n$  is the letter of  $w$  on position  $n$ . For integers  $i$  and  $j$  such that  $i < j$ , we denote

$$w[i, i[ = 1, \quad w[i, j[ = a_i \cdots a_{j-1}, \quad w[i, +\infty[ = a_i a_{i+1} \cdots .$$

Analogously one would define  $w]i, j]$ ,  $w]i, j]$ ,  $w] - \infty, j]$ , etc. When they make sense, these notations are used also for finite and infinite words.

A word  $x \in A^*$  is a *factor* of a word  $w$ , and  $w$  is an *extension* of  $x$ , if  $x = 1$  or  $x = w[i, j]$ , for some integers  $i$  and  $j$ . In such case,  $w[i, j]$  is said to be an *occurrence* of the factor  $x$  in  $w$ . We will denote by  $\text{Fact}(w)$  the set of all factors of  $w$  and, for a positive integer  $k$ , by  $\text{Fact}_k(w)$  the set of all factors of  $w$  of length  $k$ .

We say that two occurrences  $w[i, j]$  and  $w[k, l]$  of factors in a word  $w$  are *disjoint* if the integer intervals  $[i, j]$  and  $[k, l]$  are disjoint sets. For each pair of words  $w, x \in A^*$ , we denote by  $\text{occ}(x, w)$  the number of occurrences of  $x$  in  $w$ , and by  $\text{docc}(x, w)$  the maximal number of disjoint occurrences of  $x$  in  $w$ . Of course, we have  $\text{docc}(x, w) \leq \text{occ}(x, w)$ . For instance,  $\text{occ}(baab, baabaaba) = 2$  and  $\text{docc}(baab, baabaaba) = 1$ .

The sets of bi-infinite, right-infinite and left-infinite words on  $A$  will be denoted, respectively, by  $A^{\mathbb{Z}}$ ,  $A^{\mathbb{N}}$  and  $A^{-\mathbb{N}}$ . We denote also

$$A^{\infty} = A^* \cup A^{\mathbb{N}} \quad \text{and} \quad A^{-\infty} = A^* \cup A^{-\mathbb{N}}.$$

The product of two elements  $w$  and  $z$  of  $A^{\infty}$  is defined as follows: if  $w, z \in A^*$ , then  $wz$  is already defined; right-infinite words are left zeros; finally, if  $w = a_0 a_1 \cdots a_n$  is a finite word and  $z = b_1 b_2 \cdots$  is a right-infinite word then  $wz$  is the right-infinite word

$$wz = a_0 a_1 \cdots a_n b_1 b_2 \cdots .$$

The product of two elements of  $A^{-\infty}$  is defined symmetrically. Now, we observe that  $A^{\mathbb{N}}$ ,  $A^{-\mathbb{N}}$ ,  $A^{\infty}$  and  $A^{-\infty}$  are semigroups.

A word  $x \in A^*$  is a *prefix* of a word  $w \in A^{\infty}$ , and  $w$  is a *right-extension* of  $x$ , if there exists  $z \in A^{\infty}$  such that  $w = xz$ . Symmetrically,  $x \in A^*$  is a *suffix* of  $w \in A^{-\infty}$ , and  $w$  is a *left-extension* of  $x$ , if there exists  $z \in A^{-\infty}$  such that  $w = zx$ .

A right-infinite word of the form  $vu^{+\infty} = vuuu \cdots$ , with  $u \in A^+$  and  $v \in A^*$ , is said to be *ultimately periodic* and  $u$  is said to be a *period* of  $vu^{+\infty}$ . Dually, we will use the notation  $u^{-\infty}v$  to represent the (*ultimately periodic of period  $u$* ) left-infinite word  $u^{-\infty}v = \cdots uuv$ . As one can easily verify, a (right-infinite or left-infinite) ultimately periodic word has one and only one period which is a Lyndon word. An ultimately periodic word  $w \in A^{\mathbb{N}}$  (resp.  $w \in A^{-\mathbb{N}}$ ) which can be written in the form  $w = u^{+\infty}$  (resp.  $w = u^{-\infty}$ ) for some  $u \in A^+$ , is said to be *periodic*.

When writing a specific bi-infinite word, we need to specify which is the letter on position 0. We do this by putting a “.” on the left of the letter. For example,

$$w = \cdots aaa \cdot baaa \cdots$$

represents the bi-infinite word over  $\{a, b\}$  which contains exactly one occurrence of the letter  $b$ , being this occurrence on position 0. This word can also be written  $w = a^{-\infty} \cdot ba^{+\infty}$ .

More generally, given words  $x = (x_i)_{i \in -\mathbb{N}} \in A^{-\mathbb{N}}$  and  $y = (y_i)_{i \in \mathbb{N}} \in A^{\mathbb{N}}$ , we denote by  $x \cdot y$  the bi-infinite word  $w = \cdots x_{-2}x_{-1} \cdot y_1y_2 \cdots$ ; that is,  $w$  is the bi-infinite word such that  $w_i = x_i$  for  $i \leq -1$  and  $w_i = y_{i+1}$  for  $i \geq 0$ . A bi-infinite word  $w$  is said to be *ultimately periodic* if  $w = x \cdot y$  for some ultimately periodic words  $x \in A^{-\mathbb{N}}$  and  $y \in A^{\mathbb{N}}$ . The word  $w$  is said to be *periodic* if one can choose  $x = u^{-\infty}$  and  $y = u^{+\infty}$  for some  $u \in A^+$ . For instance, the word  $w = a^{-\infty} \cdot ba^{+\infty}$  is ultimately periodic and non-periodic.

For each  $n \in \mathbb{Z}$ , we define a *shift* function  $\sigma^n : A^{\mathbb{Z}} \rightarrow A^{\mathbb{Z}}$  by setting, for each  $x = (x_i)_{i \in \mathbb{Z}} \in A^{\mathbb{Z}}$ ,  $\sigma^n(x) = (x_{i+n})_{i \in \mathbb{Z}}$ . The element  $\sigma^n(x)$  is called a *shift* of  $x$ . For instance,  $\sigma^3(a^{-\infty} \cdot ba^{+\infty}) = a^{-\infty} baa \cdot a^{+\infty}$ . Notice that for  $n > 0$ , the function  $\sigma^n$  shifts letters  $n$  places to the left, while  $\sigma^{-n}$  shifts them  $n$  places to the right. It is well known that a bi-infinite word  $w$  is periodic if and only if  $w = \sigma^n(w)$ , for some  $n > 0$ .

We denote by  $\sim$  the equivalence relation on  $A^{\mathbb{Z}}$  given by

$$w \sim z \quad \text{if and only if} \quad \exists n \in \mathbb{Z}, z = \sigma^n(w).$$

Given a left-infinite word  $x \in A^{-\mathbb{N}}$  and a right-infinite word  $y \in A^{\mathbb{N}}$ , we denote by  $xy$  the  $\sim$ -class of the bi-infinite word  $x \cdot y$ . For instance,  $a^{-\infty}ba^{+\infty}$  is the  $\sim$ -class of all bi-infinite words over  $\{a, b\}$  which contain exactly one occurrence of the letter  $b$ . For  $u \in A^+$ , we will usually denote  $u^\infty = u^{-\infty}u^{+\infty}$ . Sometimes we will not distinguish a bi-infinite word from its  $\sim$ -class.

## 2.2 Semigroups, pseudovarieties and implicit operations

If  $S$  is a semigroup, we denote by  $S^1$  the smallest monoid containing  $S$ . Given an element  $s$  of a profinite semigroup the subsemigroup generated by  $s$  has a unique idempotent, denoted by  $s^\omega$ . The set of idempotents of a semigroup  $S$  is denoted by  $E(S)$ .

A proof of the following classical result can be found in [2, Proposition 3.7.1].

**Lemma 2.1** *Let  $S$  be a finite semigroup with  $k$  elements. Then, for any  $s_1, \dots, s_k \in S$ , there exist integers  $1 \leq i \leq j \leq k$ , such that,  $s_1 \cdots s_k = s_1 \cdots s_{i-1} (s_i \cdots s_j)^\omega s_{j+1} \cdots s_k$ .*  
□

A *morphism* of semigroups is an application  $\varphi : S \rightarrow T$  between semigroups  $S$  and  $T$  that respects the multiplication operation of the semigroups. A *relational morphism* of semigroups is a binary relation  $\theta : S \rightharpoonup T$  between semigroups  $S$  and  $T$  with domain  $S$  which is a subsemigroup of  $S \times T$ .

A *semigroup pseudovariety* is a class of finite semigroups closed under taking subsemigroups, homomorphic images and finite direct products. For a set  $A$  endowed with a topology, we say that a (topological) semigroup  $S$  is *A-generated* if there is a continuous function, from  $A$  to  $S$ , such that  $S$  is the smallest closed semigroup that contains the image of  $A$ . We consider finite sets endowed with discrete topology. For a pseudovariety  $\mathbf{V}$  and a set  $A$ , the  $A$ -generated elements of  $\mathbf{V}$  form a directed system and the respective projective limit will be denoted by  $\overline{\Omega}_A \mathbf{V}$ . In general, we say that a semigroup  $S$  is *pro- $\mathbf{V}$*  if it is a projective limit of semigroups of  $\mathbf{V}$ . In case  $\mathbf{V} = \mathbf{S}$ , where  $\mathbf{S}$  is the pseudovariety of all finite semigroups, we say that  $S$  is *profinite*.

The semigroup  $\overline{\Omega}_A \mathbf{V}$  is the free semigroup over  $A$  for the class of all pro- $\mathbf{V}$  semigroups, which means that, for each pro- $\mathbf{V}$  semigroup  $S$  and each continuous function  $\varphi : A \rightarrow S$ , there is a unique continuous morphism  $\overline{\varphi} : \overline{\Omega}_A \mathbf{V} \rightarrow S$  extending  $\varphi$ . This leads to a natural interpretation of elements of  $\overline{\Omega}_A \mathbf{V}$  as ( $A$ -ary) *implicit operations*: to each  $\pi \in \overline{\Omega}_A \mathbf{V}$  is associated an operation  $\pi_S : S^A \rightarrow S$  which maps  $\varphi \in S^A$  to  $\pi \overline{\varphi}$ . The subsemigroup of  $\overline{\Omega}_A \mathbf{V}$  generated by  $A$  is denoted by  $\Omega_A \mathbf{V}$  and is a dense subsemigroup of  $\overline{\Omega}_A \mathbf{V}$ , whose elements are called *explicit operations*. The following examples of implicit operations play an important role in the next sections: the binary implicit operation *multiplication*, denoted by  $_ \cdot _$ , interpreted as the semigroup multiplication on each profinite semigroup, and the unary implicit operation  $\omega$ -*power*, denoted by  $_ ^\omega$ , which, for a profinite semigroup  $S$ , to  $s \in S$  associates  $s^\omega$ . Note that the composition of implicit operations is an implicit operation.

We denote by  $q_{\mathbf{V}} : \overline{\Omega}_A \mathbf{S} \rightarrow \overline{\Omega}_A \mathbf{V}$  the continuous morphism mapping the generators of  $\overline{\Omega}_A \mathbf{S}$  to the generators of  $\overline{\Omega}_A \mathbf{V}$ .

A *pseudoidentity* is a formal equality  $\pi = \rho$  where  $\pi, \rho \in \overline{\Omega}_A \mathbf{S}$  with  $A$  finite. A finite semigroup  $S$  *satisfies* a pseudoidentity  $\pi = \rho$  if, for every continuous function  $\varphi : A \rightarrow S$ ,  $\pi \overline{\varphi} = \rho \overline{\varphi}$ . A pseudovariety  $\mathbf{V}$  satisfies a pseudoidentity  $\pi = \rho$  if every semigroup in  $\mathbf{V}$  satisfies  $\pi = \rho$ , which means that  $\pi q_{\mathbf{V}} = \rho q_{\mathbf{V}}$ . Reiterman's theorem [16] states that every pseudovariety  $\mathbf{V}$  is defined by a set  $\Sigma$  of pseudoidentities, in the sense that  $\mathbf{V}$  is the class of all semigroups that satisfy all the pseudoidentities of  $\Sigma$ , and in such case we write  $\mathbf{V} = \llbracket \Sigma \rrbracket$ .

Given  $\pi \in \overline{\Omega}_A \mathbf{S}$ , we say that  $\rho \in \overline{\Omega}_A \mathbf{S}$  is a factor (resp. a prefix, a suffix) of  $\pi$  if there are  $\pi_1, \pi_2 \in (\overline{\Omega}_A \mathbf{S})^1$  such that  $\pi = \pi_1 \rho \pi_2$  (resp.  $\pi = \rho \pi_2$ ,  $\pi = \pi_1 \rho$ ). A bi-infinite word  $w$  is a (bi-infinite) factor of  $\pi$  if every finite word which is a factor of  $w$  is also a factor of  $\pi$ . Notice that, a bi-infinite word  $w$  is a factor of  $\pi$  if and only if every shift of  $w$  is a factor of  $\pi$ .

## 2.3 Graphs

By a (*directed*) *graph*  $\Gamma$  we mean a partial algebra with a support set  $\mathcal{V}(\Gamma) \overset{\circ}{\cup} \mathcal{E}(\Gamma)$  with two sorts of elements, called *vertices* and *edges* respectively, and two binary operations  $\alpha, \omega : \mathcal{E}(\Gamma) \rightarrow \mathcal{V}(\Gamma)$  which define the *orientation* of the edges. For any edge  $e \in \mathcal{E}(\Gamma)$ ,  $e\alpha$  is the *beginning* of  $e$  and  $e\omega$  is the *end* of  $e$ .

Two edges  $e_1$  and  $e_2$  are said to be *consecutive* if  $e_1\omega = e_2\alpha$ . A (*directed*) *path* of  $\Gamma$  is a finite sequence  $e_1, e_2, \dots, e_n$  of consecutive edges of  $\Gamma$ . Given a sequence of edges  $e_1, \dots, e_n$ , if it is possible to invert the orientation of some edges in such a way to obtain a directed path, we say that  $e_1, \dots, e_n$  is a *non-oriented path*. A path  $e_1, \dots, e_n$  is a *circuit* if  $e_1\alpha = e_n\omega$ . By similarity, a non-oriented path is a *non-oriented circuit* if it is possible to invert the orientation of some edges in such a way to obtain a circuit.

A graph  $\Gamma$  is said to be *connected* if for every pair of vertices of  $\Gamma$  there is a non-oriented path between them. It is said to be *strongly connected* if, for any two of its vertices  $v_1$  and  $v_2$ , there is a path from  $v_1$  to  $v_2$  and a path from  $v_2$  to  $v_1$ . The *connected component* (resp. *strongly connected component*) of a vertex of  $\Gamma$  is the largest connected subgraph (resp. strongly connected subgraph) of  $\Gamma$  containing that vertex.

By a *labelling* of a graph  $\Gamma$  by a semigroup  $S$ , we mean a mapping  $\delta : \Gamma \longrightarrow S^1$  such that  $\mathcal{E}(\Gamma)\delta \subseteq S$ . Such labelling is *consistent* if  $\epsilon\alpha\delta \cdot \epsilon\delta = \epsilon\omega\delta$ , for all  $\epsilon \in \mathcal{E}(\Gamma)$ . The *label* of a path  $\mathbf{e}_1, \dots, \mathbf{e}_n$  is by definition  $\mathbf{e}_1\delta \cdots \mathbf{e}_n\delta$ . If  $\gamma : \mathcal{E}(\Gamma) \longrightarrow A^+$  is a function, the *label* of a non-oriented path  $\mathbf{e}_1, \dots, \mathbf{e}_n$  is the reduced form of the word  $(\mathbf{e}_1\gamma)^{\epsilon_1} \cdots (\mathbf{e}_n\gamma)^{\epsilon_n}$  in the free group generated by  $A$ , where  $\epsilon_i = 1$  if in the non-oriented path the edge  $\mathbf{e}_i$  is read in the direct way and  $\epsilon_i = -1$  otherwise. We say that the function  $\gamma$  *commutes* if the label of any non-oriented circuit is 1.

## 2.4 $\kappa$ -reducibility and $\kappa$ -tameness

An *implicit signature*  $\sigma$  is a set of implicit operations containing the multiplication. In this paper, we are particularly interested in the implicit signature  $\kappa = \{ \_ \cdot \_ , \_ ^\omega \}$ .

Every profinite semigroup can naturally be considered an algebra for the signature  $\sigma$ . Such algebras will be called  $\sigma$ -semigroups. For a pseudovariety  $\mathbf{V}$ , we denote by  $\Omega_A^\sigma \mathbf{V}$  the free  $\sigma$ -semigroup generated by  $A$  in the variety of  $\sigma$ -semigroups generated by  $\mathbf{V}$ , which as one can observe is the  $\sigma$ -subsemigroup of  $\overline{\Omega}_A \mathbf{V}$  generated by  $A$ . The elements of  $\Omega_A^\sigma \mathbf{S}$  are called  $\sigma$ -words or  $\sigma$ -terms. We denote by  $p_{\mathbf{V}} : \Omega_A^\sigma \mathbf{S} \longrightarrow \Omega_A^\sigma \mathbf{V}$  the morphism of  $\sigma$ -semigroups determined by the choice of generators. That is,  $p_{\mathbf{V}}$  is the restriction of  $q_{\mathbf{V}}$  to  $\Omega_A^\sigma \mathbf{S}$ . The  $\sigma$ -word problem for  $\mathbf{V}$  is the problem of deciding, for  $\sigma$ -words  $\pi$  and  $\rho$ , if  $\mathbf{V}$  satisfies the pseudoidentity  $\pi = \rho$ , that is, if  $\pi p_{\mathbf{V}} = \rho p_{\mathbf{V}}$ .

Let  $\gamma : \Gamma \longrightarrow S^1$  be a labelling of a finite graph  $\Gamma$  by a finite semigroup  $S$  and let  $\mu : S \rightrightarrows T$  be a relational morphism between semigroups  $S$  and  $T$ . The labelling  $\gamma$  is said to be  $\mu$ -inevitable if there is a consistent labelling  $\delta : \Gamma \longrightarrow T^1$  such that  $(\mathbf{g}\gamma, \mathbf{g}\delta) \in \mu \cup \{(1, 1)\}$  for every  $\mathbf{g} \in \Gamma$ . The labelling  $\gamma$  is said to be  $\mathbf{V}$ -inevitable if it is  $\mu$ -inevitable for every relational morphism  $\mu : S \rightrightarrows T$ , with  $T \in \mathbf{V}$ . Using a compactness result, Almeida [3] has shown the following result.

**Proposition 2.2** *A labelling  $\gamma$  of a finite graph  $\Gamma$  by a finite  $A$ -generated semigroup  $S$  is  $\mathbf{V}$ -inevitable if and only if there exists a labelling  $\delta$  of  $\Gamma$  by  $\overline{\Omega}_A \mathbf{S}$  such that  $\delta\psi = \gamma$ , where the homomorphism  $\psi : \overline{\Omega}_A \mathbf{S} \rightarrow S$  respects the choice of generators, and  $\delta q_{\mathbf{V}}$  is consistent.  $\square$*

The pseudovariety  $\mathbf{V}$  is said to be  $\sigma$ -reducible if every  $\mathbf{V}$ -inevitable labelling of a finite graph by a finite semigroup  $S$  is  $\mu_{\mathbf{V}}^\sigma$ -inevitable, where  $\mu_{\mathbf{V}}^\sigma = \varphi^{-1} p_{\mathbf{V}}$ , for  $\varphi : \Omega_A^\sigma \mathbf{S} \longrightarrow S$  the morphism of  $\sigma$ -semigroups which respects the choice of generators.

The following commutative diagram expresses the  $\sigma$ -reducibility of  $\mathbf{V}$ , which means that proving  $\sigma$ -reducibility of  $\mathbf{V}$  is proving the existence of  $\delta'$  in the conditions of the diagram.

$$\begin{array}{ccccc}
 & & \Omega_A^\sigma \mathbf{S} & & \\
 & & \uparrow \cap & & \\
 & & \overline{\Omega}_A \mathbf{S} & & \\
 & \delta' & \uparrow \psi & & \\
 \Gamma & \xrightarrow{\gamma} & S & \xrightarrow{\mu_{\mathbf{V}}} & \overline{\Omega}_A \mathbf{V} \supseteq \Omega_A^\sigma \mathbf{V} \\
 & & \downarrow \psi & & \\
 & & S & & \\
 & & \downarrow q_{\mathbf{V}} & & \\
 & & \Omega_A^\sigma \mathbf{V} & & \\
 & & \downarrow p_{\mathbf{V}} & & \\
 & & \Omega_A^\sigma \mathbf{V} & & 
 \end{array}$$

A pseudovariety  $\mathbf{V}$  is said to be  $\sigma$ -tame if it is recursively enumerable,  $\sigma$ -reducible and the  $\sigma$ -word problem for  $\mathbf{V}$  is decidable. Finally we say that a pseudovariety is *tame* if it is  $\sigma$ -tame with respect to a recursively enumerable implicit signature  $\sigma$  consisting of computable implicit operations.

### 3 Implicit operations on **LSI**

This section gathers some basic statements about the free pro- $\mathbf{V}$  semigroups, when  $\mathbf{V} \in \{\mathbf{K}, \mathbf{D}, \mathbf{LI}, \mathbf{LSI}\}$  which will be used in this work. Recall that  $\mathbf{K}, \mathbf{D} \subseteq \mathbf{LI} \subseteq \mathbf{LSI}$ ,

$$\mathbf{K} = \llbracket x^\omega y = x^\omega \rrbracket, \quad \mathbf{D} = \llbracket yx^\omega = x^\omega \rrbracket, \quad \mathbf{LI} = \llbracket x^\omega yz^\omega = x^\omega z^\omega \rrbracket$$

and

$$\mathbf{LSI} = \llbracket x^\omega yx^\omega yx^\omega = x^\omega yx^\omega, x^\omega yx^\omega zx^\omega = x^\omega zx^\omega yx^\omega \rrbracket.$$

Notice also that  $\mathbf{K} \vee \mathbf{D} = \mathbf{LI}$ . This means that a pseudoidentity  $\pi = \rho$  is satisfied by  $\mathbf{LI}$  if and only if it is satisfied by both  $\mathbf{K}$  and  $\mathbf{D}$ . The following properties are well known (see [2], for instance).

**Lemma 3.1** *Let  $\mathbf{V}$  be one of the pseudovarieties  $\mathbf{K}, \mathbf{D}$  or  $\mathbf{LI}$ . Then,  $\Omega_A \mathbf{V}$  is isomorphic to  $A^+$ ,  $E(\overline{\Omega}_A \mathbf{V})$  is an ideal consisting of the non-explicit operations of  $\overline{\Omega}_A \mathbf{V}$  and  $\overline{\Omega}_A \mathbf{K}, \overline{\Omega}_A \mathbf{D}$  and  $E(\overline{\Omega}_A \mathbf{LI})$  are isomorphic to, respectively,  $A^\infty$ ,  $A^{-\infty}$  and  $A^{\mathbb{N}} \times A^{-\mathbb{N}}$ .  $\square$*

In view of this lemma, we shall identify explicit operations (on  $\mathbf{K}, \mathbf{D}$  or  $\mathbf{LI}$ ) with finite words. Moreover, the idempotents of  $\overline{\Omega}_A \mathbf{K}$  and  $\overline{\Omega}_A \mathbf{D}$  may be identified, respectively, with right-infinite words and left-infinite words. Notice that in  $\overline{\Omega}_A \mathbf{K}$ , the right-infinite word  $yx^{+\infty}$ , where  $y \in A^*$  and  $x \in A^+$ , corresponds to the implicit operation  $yx^\omega$ . A dual remark holds for  $\mathbf{D}$ .

Let  $n$  be a positive integer. We denote by  $\equiv_n$  the congruence on  $\overline{\Omega}_A \mathbf{S}$  given, for every  $\pi, \rho \in \overline{\Omega}_A \mathbf{S}$ , by

$$\pi \equiv_n \rho \quad \text{if} \quad \pi \text{ and } \rho \text{ have the same prefix, suffix and factors of length } n.$$

We summarize in the next proposition some properties of the pseudoidentities satisfied by  $\mathbf{LSI}$ . This result is an immediate consequence of Theorem 3.3 in [12].

**Proposition 3.2** *Let  $\pi, \rho \in \overline{\Omega}_A \mathbf{S}$ . The following conditions are equivalent:*

- i)  $\pi q_{\mathbf{LSI}} = \rho q_{\mathbf{LSI}}$ ;
- ii)  $\pi \equiv_n \rho$  for every  $n \in \mathbb{N}$ ;
- iii)  $\pi q_{\mathbf{LI}} = \rho q_{\mathbf{LI}}$  and  $\pi$  and  $\rho$  have the same finite factors;
- iv)  $\pi q_{\mathbf{LI}} = \rho q_{\mathbf{LI}}$  and  $\pi$  and  $\rho$  have the same bi-infinite factors.

Moreover, if  $\pi$  and  $\rho$  are non-explicit, then a bi-infinite word  $\mathbf{w} \in A^{\mathbb{Z}}$  is a factor of  $\pi\rho$  if and only if  $\mathbf{w}$  is a factor of  $\pi$  or a factor of  $\rho$ , or  $\mathbf{w}$  is a shift of the bi-infinite word  $\pi q_{\mathbf{D}} \cdot \rho q_{\mathbf{K}}$ .  $\square$

Now we state the following simple, but fundamental, result.

**Lemma 3.3** *Let  $S \in \mathbf{S}$ ,  $\pi \in \overline{\Omega}_A \mathbf{S}$  be a non-explicit operation and  $n \in \mathbb{N}$ . Then, there exists a word  $u_\pi \in A^+$  such that  $u_\pi \equiv_n \pi$  and  $S$  satisfies  $\pi = u_\pi$ .*

**Proof.** Let  $(u_i)_i$  be a sequence of words converging to  $\pi$ . Since  $S$  is finite,  $S$  satisfies  $u_i = \pi$  for every  $i$  after a certain order. Moreover, taking  $i$  sufficiently large, the word  $u_\pi = u_i$  is such that  $u_\pi \equiv_n \pi$  and  $S$  satisfies  $\pi = u_\pi$ .  $\square$

A proof of the following useful result can be found in [9].

**Lemma 3.4** *Let  $\pi, \rho \in \overline{\Omega}_A \mathbf{S}$  be non-explicit operations such that  $\pi q_{\mathbf{D}} = \rho q_{\mathbf{D}}$ . Then there exist factorizations  $\pi = \pi_1 \pi_2$  and  $\rho = \rho_1 \pi_2$  where  $\pi_1, \pi_2, \rho_1$  are non-explicit operations. Moreover, if  $\pi q_{\mathbf{D}} = u^{-\infty} v$  where  $u \in A^+$  and  $v \in A^*$ , then one can choose  $\pi_2 = u^\omega v$ .  $\square$*

A dual result could be stated for the pseudovariety  $\mathbf{K}$ .

## 4 Some combinatorial results

In this section, we introduce some definitions on words and prove some results that will be used later.

### 4.1 Centers of bi-infinite words

Let  $u \in A^+$  and let  $x = u[l, r]$  be an occurrence of a factor  $x$  in  $u$ . An occurrence  $v = u[l', r]$ , with  $l' \leq l$ , of a factor  $v$  in  $u$  is said to be a *left-extension* of the occurrence  $u[l, r]$ . In this case, the word  $v$  itself is said to be a *left-extension (in  $u$ )* of the occurrence  $u[l, r]$ .

Let  $x \in A^+$  and let  $\overleftarrow{x} \in A^+$  be a left-extension of  $x$ . We say that an occurrence  $x = u[l, r]$  in a word  $u \in A^+$  is *allowed in  $u$  relatively to  $\overleftarrow{x}$*  if  $\overleftarrow{x}$  is a left-extension in  $u$  of the occurrence  $u[l, r]$ . For instance, let  $x = ab$  and let  $\overleftarrow{x} = aabab$ . Then

$$u = baab**ab**ababab$$

has two allowed occurrences of  $x$ :  $u[5, 6]$  and  $u[12, 13]$ ; and three occurrences of  $x$  not allowed:  $u[3, 4]$ ,  $u[7, 8]$  and  $u[10, 11]$ . If  $\overleftarrow{x} = bbab$  then  $x$  has no allowed occurrences in  $u$ .

Let  $\mathbf{w} \in A^{\mathbb{Z}}$  be a bi-infinite word. For every pair of integers  $p, q \in \mathbb{N}_0$ , the factor  $\mathbf{w}[-p, q]$  is said to be a *center* of  $\mathbf{w}$ . Let  $\mathbf{w}' \in A^{\mathbb{Z}}$  be another bi-infinite word such that  $\mathbf{w} \not\sim \mathbf{w}'$  (i.e.,  $\mathbf{w}'$  is not a shift of  $\mathbf{w}$ ). An occurrence of  $\mathbf{w}[-p, q]$  in  $\mathbf{w}'$ , when it exists, is of the form  $\mathbf{w}'[h - p, h + q]$  for some  $h \in \mathbb{Z}$ . Notice that in this case  $h$  is the position in  $\mathbf{w}'$  of the central letter  $\mathbf{w}(0)$  of  $\mathbf{w}$ . We then say that the center  $\mathbf{w}[-p, q]$  of  $\mathbf{w}$  *occurs in  $\mathbf{w}'$  on position  $h$* . Notice that, since by hypothesis  $\mathbf{w} \not\sim \mathbf{w}'$ , there is not a  $h \in \mathbb{N}_0$  such that  $\mathbf{w}[-p, q]$  occurs in  $\mathbf{w}'$  on position  $h$ , for every  $p$  and  $q$ .

We now assume that  $\mathbf{w}$  and  $\mathbf{w}'$  are non-aperiodic and prove the following statement.

$$\exists q \in \mathbb{N}_0 \forall h \in \mathbb{N}_0 \exists p \in \mathbb{N}_0, \mathbf{w}[-p, q] \text{ does not occur in } \mathbf{w}' \text{ on a position in } [-h, h]. \quad (1)$$

Suppose that (1) is not true. Hence,

$$\forall q \in \mathbb{N}_0 \exists h \in \mathbb{N}_0 \forall p \in \mathbb{N}_0, \mathbf{w}[-p, q] \text{ occurs in } \mathbf{w}' \text{ on a position in } [-h, h].$$

Since the integer interval  $[-h, h]$  is finite, we deduce that

$$\forall q \in \mathbb{N}_0 \exists h_q \in \mathbb{Z} \forall p \in \mathbb{N}_0, \mathbf{w}[-p, q] \text{ occurs in } \mathbf{w}' \text{ on position } h_q.$$

Let  $q \in \mathbb{N}_0$ . Then, for every integer  $i \leq q$ ,  $\mathbf{w}(i) = \mathbf{w}'(i + h_q)$ , so that

$$\mathbf{w}[-\infty, q] = \mathbf{w}'[-\infty, q + h_q].$$

Since  $\mathbf{w} \not\sim \mathbf{w}'$ , the equality  $h_q = h_{q'}$  holds only for a finite number of integers  $q' \in \mathbb{N}_0$ . Therefore,  $h_q \neq h_{q'}$  for some integer  $q' > q$ . On the other hand  $\mathbf{w}[-\infty, q] = \mathbf{w}'[-\infty, q + h_q] = \mathbf{w}'[-\infty, q' + h_{q'}]$ , whence  $\mathbf{w}[-\infty, q] = \mathbf{w}'[-\infty, q + h_q] = \mathbf{w}'[-\infty, q' + h_{q'}]$ . Assuming, without loss of generality, that  $h_q < h_{q'}$  and putting  $u = \mathbf{w}'[h_q, h_{q'}]$ , we deduce that  $\mathbf{w}[-\infty, q] = u^{-\infty}$ . Since  $q$  is arbitrary, this clearly implies that  $\mathbf{w}$  is periodic. This contradicts the hypothesis and so we conclude that (1) holds. Now, using the fact that  $\mathbf{w}[-P, Q]$  is a factor of  $\mathbf{w}[-p, q]$  for every  $q \geq Q$  and  $p \geq P$  and using symmetry, we easily deduce the more generic statement,

$$\begin{aligned} \exists Q \in \mathbb{N}_0 \forall q \geq Q \forall h \in \mathbb{N}_0 \exists P \in \mathbb{N}_0 \forall p \geq P, \mathbf{w}[-p, q] \text{ does not occur in } \mathbf{w}' \text{ and} \\ \mathbf{w}'[-p, q] \text{ does not occur in } \mathbf{w} \text{ on positions in the interval } [-h, h]. \end{aligned} \quad (2)$$

Fix two integers  $q, q' \geq Q$  and write  $c_i = \mathbf{w}[-i, q]$  and  $c'_i = \mathbf{w}'[-i, q']$  for every  $i \in \mathbb{N}_0$ . Let  $t$  and  $t'$  be two positive integers. Then, by (2), for  $h = \max(t, t') + \max(q, q')$  there exists  $P \in \mathbb{N}_0$  such that

$$\begin{aligned} \forall p \geq P, \quad c_p \text{ does not occur in } \mathbf{w}' \text{ and } c'_p \text{ does not occur in } \mathbf{w} \\ \text{on positions in the interval } [-h, h]. \end{aligned} \quad (3)$$

Let  $p \geq \max(P, t, t')$ , let  $T$  be a large integer (to be more precise, it suffices to take  $T \geq \max(q, q') + 2p$ ), and let  $\overleftarrow{c}_t = c_T$  and  $\overleftarrow{c}'_{t'} = c'_T$ . Our objective is to prove the following:

**Claim** *Any two allowed occurrences of  $c_t$  and  $c'_{t'}$  in a finite word  $u \in A^+$ , relatively to  $\overleftarrow{c}_t = c_T$  and  $\overleftarrow{c}'_{t'} = c'_T$  respectively, are disjoint.*

Indeed, suppose that  $c_t = u[r - t, r + q]$  and  $c'_{t'} = u[r' - t', r' + q']$  are two allowed occurrences in  $u$ . Since these occurrences are allowed, they can be extended, respectively, to occurrences  $c_T = u[r - T, r + q]$  and  $c'_T = u[r' - T, r' + q']$  of  $\overleftarrow{c}_t = c_T$  and  $\overleftarrow{c}'_{t'} = c'_T$  in  $u$ . Suppose without loss of generality that  $r + q \leq r' + q'$  and that the occurrences  $c_t = u[r - t, r + q]$  and  $c'_{t'} = u[r' - t', r' + q']$  are not disjoint. Then  $r - r' \leq q' - q$  and

$r + q \geq r' - t'$ . Hence  $-q - t' \leq r - r' \leq q' - q$  and we deduce from the definitions of  $p$  and  $T$  that

$$r - p \geq r' - t' - q - p \geq r' - q - 2p \geq r' - T.$$

Therefore, the occurrence  $c_p = u[r - p, r + q]$  is contained in the occurrence  $c'_T = u[r' - T, r' + q']$ . Since  $r - r' \in [-q - t', q' - q] \subseteq [-h, h]$ , this implies that  $c_p$  occurs in  $\mathbf{w}'$  on a position in the interval  $[-h, h]$ , which contradicts (3). Therefore, the occurrences  $c_t = u[r - t, r + q]$  and  $c'_{t'} = u[r' - t', r' + q']$  are disjoint, and the claim is true for every sufficiently large  $T$ .

Notice that, since  $\mathbf{w} \in A^{\mathbb{Z}}$  is non-periodic (so that  $\mathbf{w} \neq \sigma^m(\mathbf{w})$  for every integer  $m \neq 0$ ), the claim remains true for  $\mathbf{w}' = \mathbf{w}$ . That is, any two distinct allowed occurrences of  $c_t = \mathbf{w}[-t, q]$  in a finite word  $u$ , relatively to  $\overleftarrow{c}_t = c_T$ , are disjoint.

Since  $T$  is arbitrarily large, we could consider any finite set of bi-infinite words instead of only two and deduce the following result.

**Lemma 4.1** *Let  $B$  be a finite set of bi-infinite words and let, for each  $\mathbf{w} \in B$ ,  $t_{\mathbf{w}}$  be a positive integer. There exists a positive integer  $Q$  such that, given an integer  $q_{\mathbf{w}} \geq Q$  for each  $\mathbf{w} \in B$ , for every large integer  $T \geq \max\{t_{\mathbf{w}} \mid \mathbf{w} \in B\}$ , if  $\mathbf{w}, \mathbf{w}' \in B$  with  $\mathbf{w} \neq \sigma^m(\mathbf{w}')$  for every integer  $m \neq 0$  and there are two distinct occurrences of  $\mathbf{w}[-t_{\mathbf{w}}, q_{\mathbf{w}}]$  and  $\mathbf{w}'[-t_{\mathbf{w}'}, q_{\mathbf{w}'}]$  which are allowed in a finite word  $u \in A^+$ , relatively to  $\mathbf{w}[-T, q_{\mathbf{w}}]$  and  $\mathbf{w}'[-T, q_{\mathbf{w}'}]$  respectively, then these occurrences of  $\mathbf{w}[-t_{\mathbf{w}}, q_{\mathbf{w}}]$  and  $\mathbf{w}'[-t_{\mathbf{w}'}, q_{\mathbf{w}'}]$  are disjoint.  $\square$*

## 4.2 Marked factors

In this paper, we will need to fix and use several integers which depend one on the others. For our and the readers commodity, we will introduce now some of these constants.

- $k$  represents the number of elements of a finite  $A$ -generated semigroup  $S$  which will be introduced in the beginning of Section 6.
- $k' = 6k|A|^{3k}$ . (Notice that this constant is large enough to guarantee that, if  $x_1, \dots, x_r$  are all the factors of a word  $u \in A^+$  with  $2k \leq |x_i| < 3k$  and if  $\text{docc}(x_i, u) > k' - 2$  for every  $1 \leq i \leq r$ , then it is possible to choose one occurrence of each factor  $x_i$  such that these occurrences are pairwise disjoint. We will be interested in these factors  $x_i$  because they can be factored in the form  $x_i = x_{i,1}x_{i,2}x_{i,3}$  with  $|x_{i,1}| = |x_{i,3}| = k$  and  $|x_{i,2}| < k$ .)
- $k'' = [2k'(3k - 1)]^{|A|^{3k-1}+1}$ . (The definition of  $k''$  is motivated by Lemma 4.2 below.)

The definitions that follow, although given for the constants  $k$ ,  $k'$  and  $k''$  above, could be made for generic integers.

A finite word  $v$  is said to be  $k'$ -abundant if  $\text{docc}(y, v) \geq k'$  for all factors  $y$  of  $v$  with length  $3k - 1$ .

Let  $u = a_1a_2 \cdots a_n$  ( $a_i \in A$ ) be a finite word. A  $k''$ -neighborhood of an occurrence  $x = u[i, j]$  of a factor  $x$  in  $u$  is an occurrence  $v = u[i', j']$  extending  $u[i, j]$  (i.e., such that

$i' \leq i$  and  $j' \geq j$ ) and such that  $|v| \leq k''$ . An occurrence  $x = u[i, j]$  of a factor  $x$  of length  $3k - 1$  in  $u$  will be said to be *free* if there exists a  $k''$ -neighborhood  $v$  of  $u[i, j]$  such that  $v$  is  $k'$ -abundant. Notice that, in this case, every occurrence of a factor  $y$ , of length  $3k - 1$ , in the  $k''$ -neighborhood  $v$  is free. The occurrence  $x = u[i, j]$  (and the letters  $a_i, a_{i+1}, \dots, a_j$ ) will be said to be *marked* if  $u[i, j]$  is not free. Clearly, there is a unique factorization

$$u = x_0 w_1 x_1 w_2 \cdots w_r x_r \quad (4)$$

such that: if  $|u| < 3k - 1$ ,  $u = w_1$ ; otherwise,

- $r \geq 0$ ;
- $x_0, x_r \in A^*$ ,  $x_1, \dots, x_{r-1} \in A^+$ ;
- for each  $1 \leq i \leq r$ , the letters of  $w_i$  are marked;
- for each  $0 \leq j \leq r$ , the letters of  $x_j$  are not marked.

The factorization (4) is called the *marked factorization of  $u$  (for  $k$ )*. The factors  $w_1, \dots, w_r$  (resp.  $x_0, \dots, x_r$ ) are said to be the *marked factors* (resp. the *free factors*) of  $u$  (for  $k$ ).

In the following lemma, we prove that the marked factors of a word are of bounded length.

**Lemma 4.2** *Let  $u = x_0 w_1 x_1 w_2 \cdots w_r x_r$  be the marked factorization of a word  $u \in A^+$ . Then  $|w_i| < k''$  for every  $1 \leq i \leq r$ .*

**Proof.** The lemma is obvious when  $|u| < 3k - 1$ . So, we assume that  $|u| \geq 3k - 1$ . Let  $b = 2k'(3k - 1)$  and  $e = |A|^{3k-1} + 1$ , so that  $k'' = b^e$ . Suppose that  $|w_i| \geq k''$  for some integer  $1 \leq i \leq r$ . Then, since every letter of  $w_i$  is marked, there is a marked occurrence  $x_1$  of a factor of length  $3k - 1$  in  $w_i$ . Let  $v_1$  be an occurrence of length  $k''$  in  $w_i$  such that  $v_1$  is a  $k''$ -neighborhood of  $x_1$ . In particular, since  $x_1$  is marked,  $v_1$  is not  $k'$ -abundant. Hence,

$$\exists y_1 \in \text{Fact}_{3k-1}(v_1), \text{ docc}(y_1, v_1) < k'.$$

Notice that, therefore, the occurrences of  $y_1$  in  $v_1$  “use” less than  $b$  letters; that is, if

$$I = \bigcup_{\substack{[i,j] \subset \mathbb{N} \\ y_1 = v_1[i,j]}} [i, j]$$

then  $\#I < b$ . Hence, as  $|v_1| = k'' = b^e$ , there is an integer interval  $[h, g] \subseteq [1, |v_1|]$ , of amplitude  $b^{e-1}$ , that does not contain any element of  $I$ . In particular, this implies that

$$\exists v_2 \in \text{Fact}(v_1), \text{ occ}(y_1, v_2) = 0 \text{ and } |v_2| = b^{e-1}.$$

Now, since  $|v_2| \geq 2(3k - 1)$  and every letter of  $v_2$  is marked, there is in  $v_2$  an occurrence  $x_2$  of length  $3k - 1$  that is marked in  $u$ . In particular  $v_2$  is a  $k''$ -neighborhood of  $x_2$ , whence  $v_2$  is not  $k'$ -abundant. Hence,

$$\exists y_2 \in \text{Fact}_{3k-1}(v_2), \text{ docc}(y_2, v_2) < k'.$$

Therefore, as above, one can show that

$$\exists v_3 \in \text{Fact}(v_2), \text{occ}(y_2, v_3) = 0 \text{ and } |v_3| = b^{e-2}.$$

Notice that, as  $v_3$  is a factor of  $v_2$ ,  $y_1$  does not occur in  $v_3$ . Therefore  $v_3$  is a factor of  $w_i$  of length  $b^{e-2}$  with no occurrences of  $y_1$  and  $y_2$ .

Iterating this process, one deduces, after  $e - 1 = |A|^{3k-1}$  steps, the existence of distinct words  $y_1, \dots, y_{e-1} \in A^+$  of length  $3k - 1$  and a factor  $v_e$  of  $w_i$  such that

$$y_1, \dots, y_{e-1} \notin \text{Fact}(v_e) \text{ and } |v_e| = b.$$

Hence  $|v_e| \geq 3k - 1$  and so  $v_e$  has a factor of length  $3k - 1$ . This is absurd because  $y_1, \dots, y_{e-1}$  are  $e - 1$  distinct words on  $A$  of length  $3k - 1$  that are not factors of  $v_e$  and there exist exactly  $e - 1$  distinct words on  $A$  of length  $3k - 1$ . Therefore  $|w_i| < k''$ .  $\square$

The process of marking letters of a given word  $u \in A^+$ , described in this subsection, is a way to localize the factors of a given length  $3k - 1$  that (locally) have “few” occurrences: the definition of what “few” means is made in such a way that the marked factors have bounded lengths.

## 5 The $\kappa$ -word problem for **LSI**

A non-explicit  $\kappa$ -term  $\pi$  of the form

$$\pi = u_0 x_1^\omega u_1 x_2^\omega \cdots x_n^\omega u_n$$

with  $n \geq 1$ ,  $u_0, \dots, u_n \in A^*$  and  $x_1, \dots, x_n \in A^+$ , will be called a rank 1  $\kappa$ -term. It is clear that  $\pi q_{\mathbf{K}}$  is the right-infinite word  $u_0 x_1^{+\infty}$  and  $\pi q_{\mathbf{D}}$  is the left-infinite word  $x_n^{-\infty} u_n$ . The bi-infinite words of the form  $x_i^{-\infty}$  and  $x_j^{-\infty} u_j x_{j+1}^{+\infty}$ , with  $i = 1, \dots, n$  and  $j = 1, \dots, n - 1$ , are the bi-infinite factors of  $\pi$ .

The  $\kappa$ -word problem for **LSI** is decidable as shown by the first author [12]. Note that, by Proposition 3.2,  $\Omega_A \mathbf{LSI}$  is isomorphic to  $A^+$ . Moreover, if  $\pi \in \overline{\Omega}_A \mathbf{S}$  is a non-explicit operation and a pseudoidentity  $\pi = \rho$  holds in **LSI**, then  $\rho$  is non-explicit too. The decision criterion to test if two non-explicit  $\kappa$ -terms are equal over **LSI** is given in next proposition (Theorem 7.1 of [12]).

**Proposition 5.1** *Let  $\pi \in \Omega_A^{\kappa} \mathbf{S}$  be a non-explicit  $\kappa$ -term. Then, there is a rank 1  $\kappa$ -term  $\pi_1 = u_0 x_1^\omega u_1 x_2^\omega \cdots x_n^\omega u_n$  such that  $\pi q_{\mathbf{LSI}} = \pi_1 q_{\mathbf{LSI}}$ . Moreover, if  $\rho \in \Omega_A^{\kappa} \mathbf{S}$  is another non-explicit  $\kappa$ -term and  $\rho_1 = v_0 y_1^\omega v_1 y_2^\omega \cdots y_m^\omega v_m$  is a rank 1  $\kappa$ -term such that  $\rho q_{\mathbf{LSI}} = \rho_1 q_{\mathbf{LSI}}$ , then  $\pi q_{\mathbf{LSI}} = \rho q_{\mathbf{LSI}}$  if and only if  $u_0 x_1^{+\infty} = v_0 y_1^{+\infty}$ ,  $x_n^{-\infty} u_n = y_m^{-\infty} v_m$  and  $\pi_1$  and  $\rho_1$  have the same bi-infinite factors. Furthermore, the equality  $\pi q_{\mathbf{LSI}} = \rho q_{\mathbf{LSI}}$  is effectively decidable.*

$\square$

## 6 $\kappa$ -reducibility of **LSI**

In this section, we show that the pseudovariety **LSI** is  $\kappa$ -reducible.

Let  $\gamma$  be a **LSI**-inevitable labelling of a finite graph  $\Gamma$  by a finite  $A$ -generated semi-group  $S$  and let  $\psi$  be the unique homomorphism from  $\overline{\Omega}_A \mathbf{S}$  to  $S$  which respects the choice of generators. By Proposition 2.2, there is a labelling  $\delta$  of  $\Gamma$  by  $\overline{\Omega}_A \mathbf{S}$  such that  $\delta\psi = \gamma$  and  $\delta q_{\mathbf{LSI}}$  is consistent. Notice that since **K** and **D** are subpseudovarieties of **LSI**, this implies that  $\delta q_{\mathbf{K}}$  and  $\delta q_{\mathbf{D}}$  are both consistent. We have to construct a labelling  $\delta' : \Gamma \rightarrow \Omega_A^{\kappa} \mathbf{S}$  such that  $\delta'\psi = \gamma$  and  $\delta' p_{\mathbf{LSI}}$  is consistent.

### 6.1 Reductions on the graph $\Gamma$

We first reduce to the case in which all edges are labelled by non-explicit operations under  $\delta$ . Let  $\mathcal{V}_{\omega}$  be the set of vertices  $v \in \mathcal{V}(\Gamma)$  such that  $v\delta$  is a non-explicit operation. Let  $\phi$  be the equivalence relation on  $\mathcal{V}_{\omega}$  generated by the relation

$$\{(v, w) \mid v, w \in \mathcal{V}_{\omega} \text{ and there is an edge } e \text{ from } v \text{ to } w \text{ such that } e\delta \text{ is explicit}\}.$$

For each  $v \in \mathcal{V}_{\omega}$ , let

- $v\phi$  be the  $\phi$ -class of  $v$ ;
- $\mathcal{E}_{v\phi}$  be the subset of  $\mathcal{E}(\Gamma)$  consisting of all edges  $e$  such that  $e\delta$  is explicit and  $e\omega \in v\phi$  (and so also  $e\alpha \in v\phi$ );
- $\mathcal{E}_v$  be the set of edges  $e \in \mathcal{E}(\Gamma)$  such that  $e\omega = v$ ;
- $\mathcal{E}_{v,\omega}$  be the subset of  $\mathcal{E}_v$  consisting of all edges labelled under  $\delta$  by a non-explicit operation;
- $\mathcal{E}_{v,0} = \mathcal{E}_v \setminus \mathcal{E}_{v,\omega}$ .

Select a vertex  $v_0 \in \mathcal{V}_{\omega}$  such that  $\mathcal{E}_{v_0\phi}$  is non-empty. Let  $m_{v_0\phi}$  be the maximal length of labels of non-oriented paths, having no repeated edges, consisting of edges of  $\mathcal{E}_{v_0\phi}$ . Since  $v_0\delta$  is not explicit, there is a factorization

$$v_0\delta q_{\mathbf{D}} = z_{v_0\phi} s_{v_0} \tag{5}$$

where  $s_{v_0}$  is a word of length  $m_{v_0\phi}$  and  $z_{v_0\phi} \in \overline{\Omega}_A \mathbf{D}$  is a non-explicit operation. Let  $v \in v_0\phi$ . Select a non-oriented path from  $v_0$  to  $v$  with no repeated edges and consisting of edges of  $\mathcal{E}_{v_0\phi}$ . Let  $h$  be the label of this path and put  $s_v = s_{v_0}h$ . Since the length of  $h$  is at most  $m_{v_0\phi}$ , and since the action  $h$  on  $s_{v_0}$  is defined,  $s_v$  belongs to  $A^*$ . By consistency of  $\delta q_{\mathbf{D}}$ , we have

$$v\delta q_{\mathbf{D}} = z_{v_0\phi} s_v.$$

Moreover, for every edge  $e \in \mathcal{E}_{v,\omega}$ ,

$$e\delta q_{\mathbf{D}} = v\delta q_{\mathbf{D}} = z_{v_0\phi} s_v. \tag{6}$$

Now, we deduce from Lemma 3.4 the existence of non-explicit operations  $\pi_{v_0\phi}$ ,  $\rho_v$  and  $\rho_e$  such that

$$v\delta = \rho_v \cdot \pi_{v_0\phi} \cdot s_v, \quad (7)$$

$$e\delta = \rho_e \cdot \pi_{v_0\phi} \cdot s_v \quad \text{for any edge } e \in \mathcal{E}_{v,\omega}. \quad (8)$$

Suppose that the restriction of  $\delta$  to  $\mathcal{E}_{v_0\phi}$  is not a commuting labelling. Then, by Lemma 2.5 of [9],  $z_{v_0\phi} = u^{-\infty}v$ , for some explicit operations  $u \neq 1$  and  $v$ , and, by Lemma 3.4, one can choose  $\pi_{v_0\phi} = u^\omega v$ . In this case, we drop all the edges of  $\mathcal{E}_{v_0\phi}$  and let  $\gamma_1$  and  $\delta_1$  be the labellings of  $\Gamma' = \Gamma \setminus \mathcal{E}_{v_0\phi}$  which differ from  $\gamma$  and  $\delta$  on the following edges and vertices in which they are defined by:

- $v\gamma_1 = (\rho_v \cdot \pi_{v_0\phi})\psi$  and  $v\delta_1 = \rho_v \cdot \pi_{v_0\phi}$ , for each  $v \in v_0\phi$ ;
- $e\gamma_1 = (\rho_e \cdot \pi_{v_0\phi})\psi$  and  $e\delta_1 = \rho_e \cdot \pi_{v_0\phi}$ , for each edge  $e$  such that  $e\alpha \notin v_0\phi$  and  $e\omega \in v_0\phi$ ;
- $e\gamma_1 = (s_{e\alpha} \cdot \rho_e \cdot \pi_{v_0\phi})\psi$  and  $e\delta_1 = s_{e\alpha} \cdot \rho_e \cdot \pi_{v_0\phi}$ , for each edge  $e$  such that  $e\alpha \in v_0\phi$  and  $e\omega \in v_0\phi$ ;
- $e\gamma_1 = (s_{e\alpha} \cdot e\delta)\psi$  and  $e\delta_1 = s_{e\alpha} \cdot e\delta$ , for each edge  $e$  such that  $e\alpha \in v_0\phi$  and  $e\omega \notin v_0\phi$ .

In these conditions,  $\delta_1\psi = \gamma_1$  and, by Proposition 3.2,  $\delta_1q_{\mathbf{LSI}}$  is consistent. Suppose that we construct a labelling  $\delta'_1$  of  $\Gamma'$  by rank 1  $\kappa$ -terms satisfying conditions:

- c1)**  $\delta'_1\psi = \gamma_1$  and  $\delta'_1p_{\mathbf{LSI}}$  is consistent;
- c2)** for any  $g \in \Gamma'$ , if  $g\delta_1q_{\mathbf{D}} = u^{-\infty}v$ , where  $u \neq 1$  and  $v$  are finite words, then  $g\delta'_1p_{\mathbf{D}} = u^{-\infty}v$ .
- c3)** for any fixed integer  $M \geq m_{v_0\phi}$  and any  $g \in \Gamma'$ , if  $g\delta_1 = u\sigma$ , where  $u \in A^+$  and  $\sigma \in \overline{\Omega}_A\mathbf{S}$  with  $|u| \leq M$ , then  $g\delta'_1 = u\sigma'$  where  $\sigma' \in \Omega_A^\kappa\mathbf{S}$  is such that  $\sigma\psi = \sigma'\psi$ . Hence, in particular, for each edge  $e$  such that  $e\alpha \in v_0\phi$ ,  $e\delta'_1 = s_{e\alpha} \cdot \sigma_e$  for some  $\sigma_e \in \Omega_A^\kappa\mathbf{S}$ ;
- c4)** for any vertices  $v_1$  and  $v_2$ , if  $v_1\delta_1q_{\mathbf{LSI}} = v_2\delta_1q_{\mathbf{LSI}}$ , then  $v_1\delta'_1p_{\mathbf{LSI}} = v_2\delta'_1p_{\mathbf{LSI}}$ . Notice that in this case, we have  $v\delta'_1p_{\mathbf{LSI}} = v_0\delta'_1p_{\mathbf{LSI}}$ , for each  $v \in v_0\phi$ .

Therefore, if we define  $\delta'$  to coincide with  $\delta'_1$  on  $\Gamma'$  and let,

- $v\delta' = v\delta'_1 \cdot s_v$ , for each vertex  $v \in v_0\phi$ ;
- $e\delta' = e\delta'_1 \cdot s_{e\omega}$ , for each edge  $e$  such that  $e\alpha \notin v_0\phi$  and  $e\omega \in v_0\phi$ ;
- $e\delta' = \sigma_e \cdot s_{e\omega}$ , for each edge  $e$  such that  $e\alpha \in v_0\phi$  and  $e\omega \in v_0\phi$ ;
- $e\delta' = \sigma_e$ , for each edge  $e$  such that  $e\alpha \in v_0\phi$  and  $e\omega \notin v_0\phi$ ;
- $e\delta' = e\delta$ , for each edge  $e \in \mathcal{E}_{v_0\phi}$ ;

then  $\delta'$  is a labelling of  $\Gamma$  by  $\Omega_A^k \mathbf{S}$  such that  $\delta'\psi = \gamma$  and, by Proposition 5.1,  $\delta'p_{\mathbf{LSI}}$  is consistent.

The case in which the restriction of  $\delta$  to  $\mathcal{E}_{v_0\phi}$  is a commuting labelling can be treated analogously (in this case, condition c2) above is superfluous). Therefore, by induction on the number of  $\phi$ -classes  $v_0\phi$  such that  $\mathcal{E}_{v_0\phi}$  is non-empty, we may assume that all edges beginning in vertices labelled by non-explicit operations are labelled by non-explicit operations. Finally, we can easily eliminate edges labelled by explicit operations which begin in vertices labelled by explicit operations (and vertices labelled by explicit operations which are not the beginning of edges labelled by non-explicit operations). For that, it suffices to define the labels of these edges under  $\delta'$  to be equal to their labels under  $\delta$  by the imposition of condition c3) above which preserves prefixes of any fixed length of labels of vertices. Therefore, we may assume that all edges are labelled by non-explicit operations.

Finally, we reduce to the case in which all vertices are labelled by non-explicit operations. Suppose that  $v$  is a vertex labelled by an explicit operation and let  $e_1, \dots, e_n$  be all the edges beginning in  $v$ . Notice that, since all edges are labelled by non-explicit operations,  $v$  is not the end of an edge. Moreover, for each  $i \in \{1, \dots, n\}$ , we can write

$$e_i\delta = \pi_i\rho_i$$

for some non-explicit operations  $\pi_i$  and  $\rho_i$ . Drop vertex  $v$  and, for each  $i \in \{1, \dots, n\}$ , drop edge  $e_i$ , let  $v'_i$  be a new vertex, let  $e'_i$  be a new edge with beginning in  $v'_i$  and end in  $e_i\omega$  and let  $\gamma_1$  and  $\delta_1$  be the labellings of  $\Gamma' = (\Gamma \setminus \{v, e_1, \dots, e_n\}) \cup \{v'_i, e'_i \mid i = 1, \dots, n\}$  defined as follows:

- $\gamma_1$  and  $\delta_1$  coincide, respectively, with  $\gamma$  and  $\delta$  on  $\Gamma \setminus \{v, e_1, \dots, e_n\}$ ;
- $v'_i\gamma_1 = (v\delta \cdot \pi_i)\psi$ ,  $v'_i\delta_1 = v\delta \cdot \pi_i$ ,  $e'_i\gamma_1 = \rho_i\psi$  and  $e'_i\delta_1 = \rho_i$ , for each  $i \in \{1, \dots, n\}$ .

Since  $\delta\psi = \gamma$  and  $\delta q_{\mathbf{LSI}}$  is consistent, it is clear that  $\delta_1\psi = \gamma_1$  and  $\delta_1 q_{\mathbf{LSI}}$  is consistent. Suppose that there exists a labelling  $\delta'_1$  of  $\Gamma'$  by rank 1  $\kappa$ -terms satisfying conditions c1) and c3) above [in condition c3) we have to choose  $M \geq |v\delta|$  so that, for each  $i$ ,  $v'_i\delta'_1 = v\delta \cdot \sigma_i$  for some  $\sigma_i \in \Omega_A^k \mathbf{S}$ ]. Then, let  $\delta'$  coincide with  $\delta'_1$  on  $\Gamma \setminus \{v, e_1, \dots, e_n\}$ , let  $v\delta' = v\delta$  and let  $e_i\delta' = \sigma_i \cdot e'_i\delta'_1$  for all  $i$ . Therefore  $\delta'$  is a labelling of  $\Gamma$  by  $\Omega_A^k \mathbf{S}$  such that  $\delta'\psi = \gamma$  and  $\delta'p_{\mathbf{LSI}}$  is consistent. By induction on the number of vertices labelled by explicit operations under  $\delta$ , we may assume that all vertices are labelled by non-explicit operations.

We assume therefore that  $\Gamma$  is a finite graph all of whose vertices and edges are labelled by non-explicit operations, and the objective is to construct the labelling  $\delta'$  of  $\Gamma$  satisfying conditions c1) to c4) above with  $M$  large enough.

## 6.2 Reduction rule

Before giving explicitly the algorithm, let us give some details of the process. For a sufficiently large integer  $n$  (to be fixed below) and for each  $g \in \Gamma$  consider a word  $u_g\delta \in A^+$ , given by Lemma 3.3. We will apply to each of these words a process of

transformation described in sections 6.4 and 6.5 below. The  $\kappa$ -term obtained (in fact it is a rank 1  $\kappa$ -term), denoted  $\widehat{u_{\mathbf{g}\delta}}$ , will be chosen to be  $\mathbf{g}\delta'$ . The process of transformation will use a unique reduction rule which consists in the substitution of certain factors by certain rank 1  $\kappa$ -terms. We proceed to describe (and to justify the use of) this rule.

Since  $\psi$  is a homomorphism and so, in particular, commutes with  $\omega$ -powers, we deduce from Lemma 2.1 that for every word  $a_1 \cdots a_k \in A^+$  of length  $k$ ,

$$[a_1 \cdots a_k]\psi = [a_1 \cdots a_{i-1}(a_i \cdots a_j)^\omega a_{j+1} \cdots a_k]\psi$$

for some integers  $1 \leq i \leq j \leq k$ . Therefore, for each word  $x = a_1 \cdots a_k$  of length  $k$  we may fix a rank 1  $\kappa$ -term

$$\bar{x} = a_1 \cdots a_{i-1}(a_i \cdots a_j)^\omega a_{j+1} \cdots a_k$$

such that  $x\psi = \bar{x}\psi$ .

Now, suppose that  $x \in A^+$  is a word such that  $|x| > k$  and  $S$  satisfies  $x = x^\omega$ . Let  $y \in A^+$  be the unique primitive word such that  $x$  is a power of  $y$ . In this case we define

$$\bar{x} = y^\omega$$

and, so, also in this case the equality  $x\psi = \bar{x}\psi$  holds.

Let  $\bar{A}$  be the alphabet

$$\bar{A} = A \cup \{\bar{x} \mid x \in A^+ \text{ and } \bar{x} \text{ is defined}\}.$$

The reduction rule (to be applied on terms in the alphabet  $\bar{A}$ ) is the following:

$$R) t_1 x t_2 \rightarrow t_1 \bar{x} t_2, \quad \text{where } t_1, t_2 \in \bar{A}^*, x \in A^+ \text{ and } \bar{x} \text{ is defined.}$$

Notice that:

- rule  $R$ ) substitutes occurrences of certain factors  $x \in A^+$ , with  $|x| \geq k$ , by elements  $\bar{x}$  of the alphabet  $\bar{A}$ ; therefore, this is clearly a Noetherian system since rule  $R$ ) reduces the length of terms in the alphabet  $\bar{A}$ ;
- rule  $R$ ) apply explicit operations and rank 1  $\kappa$ -terms to rank 1  $\kappa$ -terms;
- $(t_1 x t_2)\psi = (t_1 \bar{x} t_2)\psi$ . (Notice that, since Lemma 3.3 states that  $S$  satisfies  $\mathbf{g}\delta = u_{\mathbf{g}\delta}$ , this equality will guarantee that  $\delta'\psi = \gamma$ .)

Notice that, since we want to guarantee the consistency of  $\delta'p_{\mathbf{LSI}}$ , the pseudoidentity  $\widehat{u_{\mathbf{e}\alpha\delta}} \cdot \widehat{u_{\mathbf{e}\delta}} = \widehat{u_{\mathbf{e}\omega\delta}}$  must hold in  $\mathbf{LSI}$  for every edge  $\mathbf{e}$  of  $\Gamma$ . Hence, if  $\widehat{u_{\mathbf{e}\alpha\delta}} \cdot \widehat{u_{\mathbf{e}\delta}}$  has a factor of the form  $\bar{x}y\bar{z}$  with  $x, z \in A^+$  and  $y \in A^*$ , then  $\bar{x}y\bar{z} = x_1 x_2^\omega x_3 y z_1 z_2^\omega z_3$  for some  $x_1, x_3, y, z_1, z_3 \in A^*$  and  $x_2, z_2 \in A^+$ , and, by Proposition 5.1,  $x_2^\omega x_3 y z_1 z_2^\omega$  must also occur in  $\widehat{u_{\mathbf{e}\omega\delta}}$ . For that, it suffices to guarantee that  $\bar{x}y\bar{z}$  is also a factor of  $\widehat{u_{\mathbf{e}\omega\delta}}$ . Reciprocally, we will guarantee that every factor of  $\widehat{u_{\mathbf{e}\omega\delta}}$  of the form  $\bar{x}y\bar{z}$ , with  $x, z \in A^+$  and  $y \in A^*$ , is also a factor of  $\widehat{u_{\mathbf{e}\alpha\delta}} \cdot \widehat{u_{\mathbf{e}\delta}}$ . Moreover, the substitutions will be made in such a way that the unique prefix (resp. suffix) of  $\widehat{u_{\mathbf{e}\alpha\delta}} \cdot \widehat{u_{\mathbf{e}\delta}}$  of the form  $y\bar{z}$  (resp.  $\bar{z}y$ ) with  $y \in A^*$  and  $z \in A^+$  is also a prefix (resp. suffix) of  $\widehat{u_{\mathbf{e}\omega\delta}}$ .

### 6.3 Factorizations of the words $u_{\mathbf{g}\delta}$

The objective of this subsection is, for each  $\mathbf{g} \in \Gamma$ , to identify in the word  $u_{\mathbf{g}\delta}$  the first and the last occurrences of factors on which rule  $R$ ) is going to be applied. These applications of rule  $R$ ) will determine the unique prefix (resp. suffix) of  $\widehat{u_{\mathbf{g}\delta}}$  of the form  $y\bar{z}$  (resp.  $\bar{z}y$ ) with  $y \in A^*$  and  $z \in A^+$ .

As noticed in the end of last section, we want to guarantee in particular that, for every edge  $\mathbf{e}$  of  $\Gamma$ ,  $\widehat{u_{\mathbf{e}\alpha\delta}} \cdot \widehat{u_{\mathbf{e}\delta}}$  and  $\widehat{u_{\mathbf{e}\omega\delta}}$  have the same factors of the form  $\bar{x}y\bar{z}$  with  $x, z \in A^+$  and  $y \in A^*$ . Since the last occurrence of a factor of the form  $\bar{x}$  in  $\widehat{u_{\mathbf{e}\alpha\delta}}$  and the first occurrence of a factor of the form  $\bar{z}$  in  $\widehat{u_{\mathbf{e}\delta}}$  originate a factor of the form  $\bar{x}y\bar{z}$  in  $\widehat{u_{\mathbf{e}\alpha\delta}} \cdot \widehat{u_{\mathbf{e}\delta}}$ , we must guarantee that  $\bar{x}y\bar{z}$  is also a factor of  $\widehat{u_{\mathbf{e}\omega\delta}}$  (the word  $xyz \in A^+$  will be called *the center of e*). On the other hand, since **LSI** verifies the pseudoidentity  $a^\omega b a^\omega b a^\omega = a^\omega b a^\omega$ , it may happen for instance that the word  $xyz$  has only one occurrence in  $u_{\mathbf{e}\alpha\delta} \cdot u_{\mathbf{e}\delta}$  and “many” occurrences in  $u_{\mathbf{e}\omega\delta}$  or vice-versa. This may constitute a problem since we want that  $\widehat{u_{\mathbf{e}\alpha\delta}} \cdot \widehat{u_{\mathbf{e}\delta}}$  and  $\widehat{u_{\mathbf{e}\omega\delta}}$  have the same factors of the form  $\bar{x}y\bar{z}$ . Therefore, we must localize all occurrences of the word  $xyz$  in  $u_{\mathbf{e}\alpha\delta} \cdot u_{\mathbf{e}\delta}$  and  $u_{\mathbf{e}\omega\delta}$  and substitute  $x$  by  $\bar{x}$  and  $z$  by  $\bar{z}$ . In fact, we may not be able to substitute all these occurrences since they may be not disjoint and rule  $R$ ) can only be applied on occurrences of factors of  $A^+$  which are disjoint. So these factors  $xyz \in A^+$  will be chosen to permit substitutions on certain occurrences which are conveniently “separated” by a process, that uses the results of Section 4.1, that we proceed to describe.

Let  $\mathbf{g} \in \Gamma$ . We denote by  $\mathbf{g}\theta$  the connected component of  $\mathbf{g}$  in the graph  $\Gamma$ , by  $\mathcal{V}_{\mathbf{g}\theta}$  the set of vertices of  $\mathbf{g}\theta$  and by  $\mathcal{E}_{\mathbf{g}\theta}$  the set of edges of  $\mathbf{g}\theta$ . We denote also

$$\mathbf{k}_{\mathbf{g}} = \mathbf{g}\delta q_{\mathbf{K}} \quad \text{and} \quad \mathbf{d}_{\mathbf{g}} = \mathbf{g}\delta q_{\mathbf{D}}.$$

Notice that  $\mathbf{k}_{\mathbf{g}} \in A^{\mathbb{N}}$ , that  $\mathbf{d}_{\mathbf{g}} \in A^{-\mathbb{N}}$  and, by consistency of  $\delta q_{\mathbf{D}}$ , that  $\mathbf{d}_{\mathbf{e}} = \mathbf{d}_{\mathbf{e}\omega}$  for any edge  $\mathbf{e}$ . Moreover, the consistency of  $\delta q_{\mathbf{K}}$  implies the equality  $\mathbf{k}_{\mathbf{v}} = \mathbf{k}_{\mathbf{v}'}$ , for all vertices  $\mathbf{v}$  and  $\mathbf{v}'$  in the same connected component. We then denote, for any vertex  $\mathbf{v}$ ,  $\mathbf{k}_{\mathbf{v}\theta} = \mathbf{k}_{\mathbf{v}}$ .

Let  $M$  be a large enough positive integer [cf. condition c3) on page 14]. For each vertex  $\mathbf{v}$  and for each edge  $\mathbf{e}$ , let

$$l_{\mathbf{v}} = \mathbf{k}_{\mathbf{v}\theta}[1, M + k], \tag{9}$$

$$l_{\mathbf{e}} = \mathbf{k}_{\mathbf{e}}[1, i_{\mathbf{e}}], \quad \text{where } i_{\mathbf{e}} \geq M + k \text{ is an integer defined below,} \tag{10}$$

$$r_{\mathbf{v}} = \mathbf{d}_{\mathbf{v}}[-i_{\mathbf{v}}, -1], \quad \text{where } i_{\mathbf{v}} \geq M + k \text{ is an integer defined below,} \tag{11}$$

$$r_{\mathbf{e}} = r_{\mathbf{e}\omega}. \tag{12}$$

Now, for each edge  $\mathbf{e}$ , let  $\mathbf{w}_{\mathbf{e}}$  be the bi-infinite word

$$\mathbf{w}_{\mathbf{e}} = \mathbf{d}_{\mathbf{e}\alpha} \cdot \mathbf{k}_{\mathbf{e}}$$

and let  $c_{\mathbf{e}} \in A^+$ , called the *center* of  $\mathbf{e}$ , be the following center of  $\mathbf{w}_{\mathbf{e}}$

$$c_{\mathbf{e}} = r_{\mathbf{e}\alpha} l_{\mathbf{e}} = \mathbf{w}_{\mathbf{e}}[-i_{\mathbf{e}\alpha}, i_{\mathbf{e}}].$$

We say also that  $c_{\mathbf{e}}$  is a center of  $\Gamma$ . The center  $c_{\mathbf{e}}$  is said to be *periodic* if and only if the bi-infinite word  $\mathbf{w}_{\mathbf{e}}$  is periodic.

For  $e \in \mathcal{E}(\Gamma)$  and  $v \in \mathcal{V}(\Gamma)$ , the integers  $i_e$  and  $i_v$  in (10) and (11), respectively, are chosen sufficiently large so that:

- (a) if  $e'$  is an edge such that  $\mathbf{w}_e \sim \mathbf{w}_{e'}$ , then  $c_e = c_{e'}$  (not necessarily  $r_{e\alpha} = r_{e'\alpha}$  and  $l_e = l_{e'}$ ).
- (b) If  $\mathbf{k}_e$  is ultimately periodic, say  $\mathbf{k}_e = vu^{+\infty}$  where  $u$  is the unique period of  $\mathbf{k}_e$  which is a Lyndon word, then  $l_e = vu^p$  for some positive integer  $p$  such that  $S$  satisfies  $u^\omega = u^p$ .
- (c) If  $\mathbf{d}_v$  is ultimately periodic, say  $\mathbf{d}_v = u^{-\infty}v$  where  $u$  is the unique period of  $\mathbf{d}_v$  which is a Lyndon word, then  $r_v = u^pv$  for some positive integer  $p$  such that  $S$  satisfies  $u^\omega = u^p$ .
- (d) For every  $g \in \Gamma$ , if  $\mathbf{w}_e$  is not a bi-infinite factor of  $g\delta$ , then  $c_e$  is not a factor of  $g\delta$ .
- (e)  $i_e \geq Q$ , where, for  $B = \{\mathbf{w}_e \mid e \in \mathcal{E}(\Gamma)\}$ ,  $Q$  is a positive integer in the conditions of Lemma 4.1.

Notice that  $i_e$  and  $i_{e\omega}$  may effectively be chosen satisfying these conditions since, for instance, if  $e$  and  $e'$  are edges such that  $\mathbf{w}_e \sim \mathbf{w}_{e'}$  and  $\mathbf{k}_e = vu^{+\infty}$  is ultimately periodic, then  $\mathbf{k}_{e'}$  is also ultimately periodic of the form  $\mathbf{k}_{e'} = v'u^{+\infty}$  and one of  $v$  or  $v'$  is a suffix of the other.

Notice also that, if  $e$  is an edge such that  $\mathbf{w}_e$  is ultimately periodic, say of the form  $u^{-\infty}vx^{+\infty}$  where  $u$  and  $x$  are Lyndon words, then  $c_e = u^pvx^q$  for some positive integers  $p$  and  $q$  such that  $S$  satisfies  $u^\omega = u^p$  and  $x^\omega = x^q$ . In particular: if  $\mathbf{w}_e$  is periodic, then  $u = x$  and  $v$  is a power of  $u$ , so that we can write  $c_e = u^hu^i u^h = u^{2h+i}$  for some integers  $h > 0$  and  $i \geq 0$  such that  $S$  satisfies  $u^\omega = u^h$  [one can choose, for instance  $h = \min(p, q)$ ]; if  $\mathbf{w}_e$  is non-periodic, then  $c_e = u^pvx^q$  is not a (non-trivial) power of a word.

Since, for each edge  $e$ , the word  $c_e$  is a center of the bi-infinite word  $\mathbf{w}_e$  and  $i_e \geq Q$ , Lemma 4.1 guarantees the existence of an integer  $T \geq \max\{i_v \mid v \in \mathcal{V}(\Gamma)\}$  such that, for every edges  $e$  and  $e'$ , with  $\mathbf{w}_e$  and  $\mathbf{w}_{e'}$  non-periodic and  $\mathbf{w}_e \not\sim \mathbf{w}_{e'}$ ,  $\overleftarrow{c}_e = \mathbf{w}_e[-T, i_e[$  and  $\overleftarrow{c}_{e'} = \mathbf{w}_{e'}[-T, i_{e'}[$ , and every word  $u \in A^+$ ,

- if two occurrences of  $c_e$  and  $c_{e'}$  are allowed in  $u$  relatively to  $\overleftarrow{c}_e$  and  $\overleftarrow{c}_{e'}$  respectively, then these occurrences of  $c_e$  and  $c_{e'}$  are disjoint;
- any two distinct allowed occurrences of  $c_e$  in  $u$ , relatively to  $\overleftarrow{c}_e$ , are disjoint.

For each edge  $e$  such that  $\mathbf{w}_e$  is non-periodic, the left-extension  $\overleftarrow{c}_e = \mathbf{w}_e[-T, i_e[$  of the center  $c_e$  is called the *extended-center* of  $e$ .

Finally, we fix the integer  $n$ , mentioned in the beginning of Section 6.2. Let first  $L$  be the maximal length of the centers and extended-centers of  $\Gamma$ . That is

$$L = \max\{|\overleftarrow{c}_e|, |c_f| \mid e, f \in \mathcal{E}(\Gamma) \text{ with } c_f \text{ periodic and } c_e \text{ non-periodic}\}.$$

Then we fix an integer  $n > 3L + k''$  and, for each  $\mathbf{g} \in \Gamma$ , we consider a word  $u_{\mathbf{g}\delta} \in A^+$ , given by Lemma 3.3.

Let  $\mathbf{g} \in \Gamma$ . We want to localize in  $u_{\mathbf{g}\delta}$  the allowed occurrences of the non-periodic centers of  $\Gamma$ . Since these occurrences are allowed in  $u_{\mathbf{g}\delta}$  if and only if they occur as suffixes of the respective extended-centers, the occurrences that occur “near” the beginning of  $u_{\mathbf{g}\delta}$  are not allowed. If  $\mathbf{g}$  is a vertex, this does not constitute a problem, but if  $\mathbf{g}$  is an edge  $\mathbf{e}$  this is not desirable since we want to guarantee that **LSI** satisfies  $\widehat{u_{\mathbf{e}\alpha\delta}} \cdot \widehat{u_{\mathbf{e}\delta}} = \widehat{u_{\mathbf{e}\omega\delta}}$ . We overcome this situation by localizing in  $u_{\mathbf{e}\alpha\delta} \cdot u_{\mathbf{e}\delta}$  the allowed occurrences of the non-periodic centers of  $\Gamma$ . Then, an occurrence of a center  $c_{\mathbf{f}}$  in  $u_{\mathbf{e}\alpha\delta} \cdot u_{\mathbf{e}\delta}$  which occurs in  $u_{\mathbf{e}\delta}$  will be said to be allowed in  $u_{\mathbf{e}\delta}$  if it is an allowed occurrence in  $u_{\mathbf{e}\alpha\delta} \cdot u_{\mathbf{e}\delta}$ .

Notice that  $l_{\mathbf{g}}$  is a prefix of  $u_{\mathbf{g}\delta}$  and  $r_{\mathbf{g}}$  is a suffix of  $u_{\mathbf{g}\delta}$ , so that we can write

$$u_{\mathbf{g}\delta} = l_{\mathbf{g}}u_{\mathbf{g}}r_{\mathbf{g}} \quad (13)$$

for some  $u_{\mathbf{g}} \in A^+$ . Now, we localize in  $u_{\mathbf{g}\delta}$  **every** allowed occurrences of the **non-periodic** centers  $c_{\mathbf{e}}$  of  $\Gamma$ . As observed above, the allowed occurrences of non-periodic centers are disjoint. Moreover, assuming that  $T$  was chosen large enough, these occurrences are also disjoint from the prefix  $l_{\mathbf{g}}$  of  $u_{\mathbf{g}\delta}$ . This is clear when  $\mathbf{g}$  is a vertex. When  $\mathbf{g}$  is an edge, then an occurrence of a center  $c_{\mathbf{e}}$  in  $u_{\mathbf{g}\delta}$  is allowed if the corresponding occurrence is allowed in  $u_{\mathbf{g}\alpha\delta} \cdot u_{\mathbf{g}\delta}$ . Now, as

$$u_{\mathbf{g}\alpha\delta} \cdot u_{\mathbf{g}\delta} = l_{\mathbf{g}\alpha}u_{\mathbf{g}\alpha}r_{\mathbf{g}\alpha} \cdot l_{\mathbf{g}}u_{\mathbf{g}}r_{\mathbf{g}} = l_{\mathbf{g}\alpha}u_{\mathbf{g}\alpha}c_{\mathbf{g}}u_{\mathbf{g}}r_{\mathbf{g}}$$

and the occurrence of  $c_{\mathbf{g}}$  signed in this factorization is allowed (since  $u_{\mathbf{g}\alpha\delta}$  has the same suffix of  $\mathbf{g}\alpha\delta$  of length  $n > T$ ), the occurrence of  $c_{\mathbf{e}}$  in  $u_{\mathbf{g}\delta}$  occurs in  $u_{\mathbf{g}}r_{\mathbf{g}}$  and so it is disjoint from  $l_{\mathbf{g}}$ . Analogously, we may admit that the allowed occurrences of centers are also disjoint from the suffix  $r_{\mathbf{g}}$  of  $u_{\mathbf{g}\delta}$ . Indeed, consider an allowed occurrence of a center  $c_{\mathbf{e}}$  in  $u_{\mathbf{g}\delta}$  and consider the word  $u_{\mathbf{g}\delta} \cdot l_{\mathbf{f}}$  where  $\mathbf{f}$  is an edge beginning in  $\mathbf{g}$  or in  $\mathbf{g}\omega$  whether  $\mathbf{g}$  is a vertex or not (we assume that every vertex of  $\Gamma$  is the beginning of some edge, by adding new vertices and edges if necessary). Then the corresponding occurrence of  $c_{\mathbf{e}}$  in  $u_{\mathbf{g}\delta} \cdot l_{\mathbf{f}}$  is allowed. Since  $u_{\mathbf{g}\delta} \cdot l_{\mathbf{f}} = l_{\mathbf{g}}u_{\mathbf{g}}r_{\mathbf{g}} \cdot l_{\mathbf{f}} = l_{\mathbf{g}}u_{\mathbf{g}}c_{\mathbf{f}}$  and the occurrence of the suffix  $c_{\mathbf{f}}$  is allowed in  $u_{\mathbf{g}\delta} \cdot l_{\mathbf{f}}$ , we deduce that the occurrence of  $c_{\mathbf{e}}$  in  $u_{\mathbf{g}\delta}$  occurs in  $l_{\mathbf{g}}u_{\mathbf{g}}$  and so it is disjoint from  $r_{\mathbf{g}}$ . We notice that these last comments show that, when  $\mathbf{g}$  is a vertex, the allowed occurrences of centers in  $u_{\mathbf{g}\delta}$  are the same as the allowed occurrences of centers in  $u_{\mathbf{g}\delta} \cdot u_{\mathbf{f}\delta}$  which occur in  $u_{\mathbf{g}\delta}$ , where  $\mathbf{f}$  is an edge beginning in  $\mathbf{g}$ . Therefore, the process of allowing occurrences of centers used for edges is not necessary for vertices.

Now, if  $c_{\mathbf{e}} = x^i$  is a periodic center of  $\Gamma$  that occurs in  $u_{\mathbf{g}\delta}$ , then, by condition (d) above,  $x^\omega$  is a bi-infinite factor of  $\mathbf{g}\delta$ . Therefore, since  $n > 3L$  and  $\mathbf{g}\delta$  and  $u_{\mathbf{g}\delta}$  have the same factors of length  $n$ ,  $x^j$  is a factor of  $u_{\mathbf{g}}$  for a sufficiently large integer  $j$  to guarantee that we can choose one occurrence of  $c_{\mathbf{e}}$  disjoint from the allowed occurrences of the other centers and from  $l_{\mathbf{g}}$  and  $r_{\mathbf{g}}$ .

The process above permits to obtain a factorization

$$u_{\mathbf{g}\delta} = l_{\mathbf{g}}u_{\mathbf{g},0}c_{\mathbf{g},1}u_{\mathbf{g},1}c_{\mathbf{g},2} \cdots c_{\mathbf{g},s_{\mathbf{g}}}u_{\mathbf{g},s_{\mathbf{g}}}r_{\mathbf{g}} \quad (14)$$

of  $u_{\mathbf{g}\delta}$  such that

- $s_g \geq 0$ ;
- $c_{g,1}, \dots, c_{g,s_g}$  are centers of  $\Gamma$ , called the *allowed centers of  $u_{g\delta}$* ;
- $u_{g,0}, u_{g,1}, \dots, u_{g,s_g} \in A^*$ ;
- if  $c_{g,i}$  is a non-periodic center, then the factorization (14) contains every allowed occurrence of  $c_{g,i}$ ;
- if  $c_{g,i}$  is a periodic center and  $c_{g,i}$  is a factor of  $u_{g\delta}$ , then the factorization (14) contains one occurrence of  $c_{g,i}$ .

We fix a factorization of type (14) and we say that this is the *factorization of  $u_{g\delta}$  for the centers of  $\Gamma$* .

#### 6.4 Transformations on the words $l_g$ , $r_g$ and $c_g$

Let  $v$  be a vertex. By (9),  $l_v = \mathbf{k}_{v\theta}[1, M+k]$ . Using the notation introduced in Section 6.2, let

$$\widehat{l}_v = \mathbf{k}_{v\theta}[1, M] \cdot \overline{\mathbf{k}_{v\theta}M, M+k}. \quad (15)$$

Moreover, by (11),  $r_v = \mathbf{d}_v[-i_v, -1]$ . If  $\mathbf{d}_v$  is ultimately periodic, say  $\mathbf{d}_v = u^{-\infty}v$  where  $u$  is a Lyndon word, then, by condition (c) on page 18,  $r_v = u^p v$  where  $p$  is a positive integer such that  $S$  satisfies  $u^\omega = u^p$ . In this case, let

$$\widehat{r}_v = \overline{u^p} \cdot v = u^\omega v. \quad (16)$$

If  $\mathbf{d}_v$  is not ultimately periodic, let

$$\widehat{r}_v = \overline{\mathbf{d}_v[-i_v, -i_v+k]} \cdot \mathbf{d}_v[-i_v+k, -1]. \quad (17)$$

Consider now an edge  $e$ . By (12),  $r_e = r_{e\omega}$ . We then let

$$\widehat{r}_e = \widehat{r}_{e\omega}. \quad (18)$$

On the other hand, by (10),  $l_e = \mathbf{k}_e[1, i_e]$ . If  $\mathbf{k}_e$  is ultimately periodic, say  $\mathbf{k}_e = vu^{+\infty}$  where  $u$  is a Lyndon word, then, by condition (b) on page 18,  $l_e = vu^p$  where  $p$  is a positive integer such that  $S$  satisfies  $u^\omega = u^p$ . In this case, let

$$\widehat{l}_e = v \cdot \overline{u^p} = vu^\omega. \quad (19)$$

If  $\mathbf{k}_e$  is not ultimately periodic, let

$$\widehat{l}_e = \mathbf{k}_e[1, i_e - k] \cdot \overline{\mathbf{k}_e[i_e - k, i_e]}. \quad (20)$$

Remember that  $c_e = r_{e\alpha} l_e$ . Hence, we define at last

$$\widehat{c}_e = \widehat{r}_{e\alpha} \cdot \widehat{l}_e. \quad (21)$$

Notice that  $\widehat{c}_e$  is a rank 1  $\kappa$ -term of the form  $\overline{xy\bar{z}} = x_1 x_2^\omega x_3 y z_1 z_2^\omega z_3$ , with  $x_1, x_3, y, z_1, z_3 \in A^*$  and  $x_2, z_2 \in A^+$ . Moreover, if the bi-infinite word  $\mathbf{w}_e$  is ultimately periodic, then  $x_1 = x_3 = z_1 = z_3 = 1$  and  $x_2^{-\infty} y z_2^{+\infty}$  is the  $\sim$ -class of  $\mathbf{w}_e$ .

## 6.5 Transformations on the words $u_{\mathbf{g},i}$

In this section, we describe a process to make correspond a  $\kappa$ -term  $\widehat{u_{\mathbf{g},i}}$  to each word  $u_{\mathbf{g},i}$  ( $i = 0, \dots, s_{\mathbf{g}}$ ) in factorization (14). The process of transformation is the following:

**First step** Consider the marked factorization, described in section 4.2,

$$u_{\mathbf{g},i} = x_0 \overline{w_1} x_1 w_2 \cdots w_r x_r$$

of  $u_{\mathbf{g},i}$ . If  $|u_{\mathbf{g},i}| < 3k - 1$ , then  $u_{\mathbf{g},i} = w_1$ . In this case, let  $\widehat{w}_1 = w_1$ . Suppose now that  $|u_{\mathbf{g},i}| \geq 3k - 1$ . Then, by definition of marked factors, it is clear that  $|w_j| \geq 2k$  for every  $j \in \{1, \dots, r\}$ , so that we can write  $w_j = w_{j,1} w_{j,2} w_{j,3}$  for some words  $w_{j,1}$ ,  $w_{j,2}$  and  $w_{j,3}$  with  $|w_{j,1}| = |w_{j,3}| = k$ . We define  $\widehat{w}_j = \overline{w_{j,1}} w_{j,2} \overline{w_{j,3}}$ .

We let  $u_{\mathbf{g},i}(1)$  be the following rank 1  $\kappa$ -term

$$u_{\mathbf{g},i}(1) = x_0 \widehat{w}_1 x_1 \widehat{w}_2 x_2 \cdots x_{r-1} \widehat{w}_r x_r.$$

**Second step** Define

$$\widehat{x}_0 = \begin{cases} x_0 & \text{if } |x_0| < k \\ \overline{x_{0,1}} x_{0,2} & \text{if } |x_0| \geq k \end{cases}$$

where  $x_{0,1}$  is the prefix of length  $k$  of  $x_0$  and  $x_0 = x_{0,1} x_{0,2}$ . Symmetrically, let

$$\widehat{x}_r = \begin{cases} x_r & \text{if } |x_r| < k \\ x_{r,1} \overline{x_{r,2}} & \text{if } |x_r| \geq k \end{cases}$$

where  $x_{r,2}$  is the suffix of length  $k$  of  $x_r$  and  $x_r = x_{r,1} x_{r,2}$ . Now, let

$$u_{\mathbf{g},i}(2) = \widehat{x}_0 \widehat{w}_1 x_1 \widehat{w}_2 x_2 \cdots x_{r-1} \widehat{w}_r \widehat{x}_r.$$

**Third step** Let  $y \in A^+$  be a factor of  $u_{\mathbf{g},i}$  such that  $2k \leq |y| < 3k$ . Two cases may arise.

**First case** Every extension of length  $3k - 1$  in  $u_{\mathbf{g},i}$ , of an occurrence of  $y$ , is a marked occurrence (see Section 4.2). In this case every occurrence of  $y$  in  $u_{\mathbf{g},i}$  is contained in the marked factors.

**Second case** There exists a free occurrence in  $u_{\mathbf{g},i}$  of an extension  $\tilde{y}$ , of length  $3k - 1$ , of an occurrence of  $y$ . In this case, by definition of free occurrence (cf. Section 4.2), there exists a  $k''$ -neighborhood  $v$  of  $\tilde{y}$  such that  $v$  is  $k'$ -abundant. In particular,  $\text{docc}(\tilde{y}, v) \geq k'$ . Furthermore, every occurrence of a factor  $z$  of length  $3k - 1$  in the  $k''$ -neighborhood  $v$  is free. This means that at least  $k' - 2$  of the disjoint occurrences of  $\tilde{y}$  in  $v$  occur disjoint of the marked factors. More precisely, there exists an integer  $0 \leq j \leq r$  such that  $\text{docc}(\tilde{y}, x_j) \geq k' - 2$ . Hence, since  $y$  is a factor of  $\tilde{y}$ ,  $\text{docc}(y, x_j) \geq k' - 2$ . In this case, we say that  $y$  has a free occurrence in  $u_{\mathbf{g},i}$ .

Consider the set  $F$  of all factors  $y$  of  $u_{\mathbf{g},i}$  such that  $2k \leq |y| < 3k$  and  $y$  has a free occurrence in  $u_{\mathbf{g},i}$ . By the second case above, for each  $y \in F$  there exists an integer  $0 \leq j \leq r$  such that  $\text{docc}(y, x_j) \geq k' - 2$ . Therefore, the choice of  $k'$  allows us to select an occurrence for each  $y \in F$  in such a way that these occurrences are pairwise disjoint. These occurrences are selected in the factors  $x_j$  and, if  $j = 0$  or  $j = r$ , then we can select them, respectively, in  $x_{0,2}$  and in  $x_{r,1}$ . Since  $2k \leq |y| < 3k$ , we can write

$$y = y_1 y_2 y_3$$

for some words  $y_1, y_2$  and  $y_3$  with  $|y_1| = |y_3| = k$ . We substitute in  $u_{\mathbf{g},i}$  (2) the selected occurrence of  $y$  by  $\widehat{y} = \overline{y_1} y_2 \overline{y_3}$ . We then obtain a term  $u_{\mathbf{g},i}$  (3).

**Fourth step** In this step, we admit the substitution of any occurrence, in the factors of  $u_{\mathbf{g},i}$  (3) which were obtained from the transformations on the words  $x_j$ , of a factor  $y \in A^+$  of length  $k$  by  $\overline{y}$ . We say that a term obtained from  $u_{\mathbf{g},i}$  (3) using these substitutions is *irreducible* when it is not possible to make more substitutions (that is, when there are no more occurrences of factors  $y \in A^+$  of length  $k$  on the factors which resulted from substitutions on the words  $x_j$ ). We choose an irreducible term and denote it by  $u_{\mathbf{g},i}$  (4). This concludes the process of transformation of the word  $u_{\mathbf{g},i}$ .

We let  $\widehat{u_{\mathbf{g},i}} = u_{\mathbf{g},i}$  (4). This concludes the process of transformation of the factors of  $u_{\mathbf{g}\delta}$  in factorization (14). We then let

$$\widehat{u_{\mathbf{g}\delta}} = \widehat{l_{\mathbf{g}}} \widehat{u_{\mathbf{g},0}} \widehat{c_{\mathbf{g},1}} \cdots \widehat{c_{\mathbf{g},s_{\mathbf{g}}}} \widehat{u_{\mathbf{g},s_{\mathbf{g}}}} \widehat{r_{\mathbf{g}}}. \quad (22)$$

We can finally define  $\delta'$  to be the labelling of  $\Gamma$  by (rank 1  $\kappa$ -terms of)  $\Omega_A^\kappa \mathbf{S}$  such that  $\mathbf{g}\delta' = \widehat{u_{\mathbf{g}\delta}}$  for any  $\mathbf{g} \in \Gamma$ . Notice that, by (15) to (20),  $\mathbf{g}\delta'$  was obtained in such a way that conditions c2) and c3) stated in Section 6.1 are satisfied. Moreover, the remarks made after the definition of the reduction rule R), in the end of that section, clearly imply that  $\mathbf{g}\delta'\psi = u_{\mathbf{g}\delta}\psi$ . Therefore, by Lemma 3.3, we deduce that  $\mathbf{g}\delta'\psi = \mathbf{g}\delta\psi$  and, so, that  $\delta'\psi = \gamma$ . To establish condition c1), it remains to show that  $\delta'p_{\mathbf{LSI}}$  is consistent. This will be done in the next section where we also show that condition c4) holds.

## 6.6 Proof of the consistency of $\delta'p_{\mathbf{LSI}}$

To show that  $\delta'p_{\mathbf{LSI}}$  is consistent, we have to prove that  $\mathbf{e}\alpha\delta'p_{\mathbf{LSI}} \cdot \mathbf{e}\delta'p_{\mathbf{LSI}} = \mathbf{e}\omega\delta'p_{\mathbf{LSI}}$  for every edge  $\mathbf{e}$ . Let  $\mathbf{e}$  be an arbitrary edge. Then, we have to prove that

$$(\mathbf{e}\alpha\delta' \cdot \mathbf{e}\delta')p_{\mathbf{LSI}} = \mathbf{e}\omega\delta'p_{\mathbf{LSI}}. \quad (23)$$

By Proposition 5.1, this equality holds if and only if  $(\mathbf{e}\alpha\delta' \cdot \mathbf{e}\delta')p_{\mathbf{LI}} = \mathbf{e}\omega\delta'p_{\mathbf{LI}}$  and  $\mathbf{e}\alpha\delta' \cdot \mathbf{e}\delta'$  and  $\mathbf{e}\omega\delta'$  have the same bi-infinite factors. From (22) and (21), we have

$$\begin{aligned} \mathbf{e}\alpha\delta' \cdot \mathbf{e}\delta' &= \widehat{u_{\mathbf{e}\alpha\delta}} \cdot \widehat{u_{\mathbf{e}\delta}} \\ &= \widehat{l_{\mathbf{e}\alpha}} \widehat{u_{\mathbf{e}\alpha,0}} \widehat{c_{\mathbf{e}\alpha,1}} \cdots \widehat{c_{\mathbf{e}\alpha,s_{\mathbf{e}\alpha}}} \widehat{u_{\mathbf{e}\alpha,s_{\mathbf{e}\alpha}}} \widehat{r_{\mathbf{e}\alpha}} \cdot \widehat{l_{\mathbf{e}}} \widehat{u_{\mathbf{e},0}} \widehat{c_{\mathbf{e},1}} \cdots \widehat{c_{\mathbf{e},s_{\mathbf{e}}}} \widehat{u_{\mathbf{e},s_{\mathbf{e}}}} \widehat{r_{\mathbf{e}}} \\ &= \widehat{l_{\mathbf{e}\alpha}} \widehat{u_{\mathbf{e}\alpha,0}} \widehat{c_{\mathbf{e}\alpha,1}} \cdots \widehat{c_{\mathbf{e}\alpha,s_{\mathbf{e}\alpha}}} \widehat{u_{\mathbf{e}\alpha,s_{\mathbf{e}\alpha}}} \widehat{c_{\mathbf{e}}} \widehat{u_{\mathbf{e},0}} \widehat{c_{\mathbf{e},1}} \cdots \widehat{c_{\mathbf{e},s_{\mathbf{e}}}} \widehat{u_{\mathbf{e},s_{\mathbf{e}}}} \widehat{r_{\mathbf{e}}} \end{aligned}$$

and

$$e\omega\delta' = \widehat{u_{e\omega\delta}} = \widehat{l_{e\omega}}\widehat{u_{e\omega,0}}\widehat{c_{e\omega,1}} \cdots \widehat{c_{e\omega,s_{e\omega}}}\widehat{u_{e\omega,s_{e\omega}}}\widehat{r_{e\omega}}.$$

Since  $e\alpha$  and  $e\omega$  belong to the same connected component of  $\Gamma$ , i.e.,  $e\alpha \in e\omega\theta$ , (9) implies that  $l_{e\omega} = l_{e\alpha}$ , so that  $\widehat{l_{e\omega}} = \widehat{l_{e\alpha}}$ . Moreover, by (18),  $\widehat{r_e} = \widehat{r_{e\omega}}$ . Therefore  $(e\alpha\delta' \cdot e\delta')p_{\mathbf{LI}} = e\omega\delta'p_{\mathbf{LI}}$ .

Let us now show that  $e\alpha\delta' \cdot e\delta'$  and  $e\omega\delta'$  have the same bi-infinite factors. That is, let us show that  $e\alpha\delta' \cdot e\delta'$  and  $e\omega\delta'$  have the same factors of the form  $x^\omega$  and  $x^\omega y z^\omega$ , with  $x, z \in A^+$  and  $y \in A^*$ . Since each of  $e\alpha\delta' \cdot e\delta'$  and  $e\omega\delta'$  has at least two occurrences of factors of the form  $x^\omega$ , it clearly suffices to show that they have the same factors of the form  $x^\omega y z^\omega$ . Moreover, since each occurrence of a factor  $x^\omega$  appears in an occurrence of a factor of the form  $\bar{u}$ , it suffices to show that  $e\alpha\delta' \cdot e\delta'$  and  $e\omega\delta'$  have the same factors of the form

$$\bar{x}y\bar{z}, \quad \text{with } x, z \in A^+ \text{ and } y \in A^*. \quad (24)$$

Notice first that, since  $(e\alpha\delta \cdot e\delta)q_{\mathbf{LSI}} = e\omega\delta q_{\mathbf{LSI}}$  by consistency of  $\delta q_{\mathbf{LSI}}$  and since  $\equiv_n$  is a congruence on  $\overline{\Omega}_A \mathbf{S}$ , we deduce from Lemma 3.3 that  $u_{e\alpha\delta} \cdot u_{e\delta} \equiv_n u_{e\omega\delta}$ . Let us now notice the following observations.

**Lemma 6.1** *The following statements hold.*

- a)  $\{c_{e\alpha,1}, \dots, c_{e\alpha,s_{e\alpha}}, c_e, c_{e,1}, \dots, c_{e,s_e}\} = \{c_{e\omega,1}, \dots, c_{e\omega,s_{e\omega}}\}$ ;
- b)  $u_{e\alpha\delta} \cdot u_{e\delta}$  and  $u_{e\omega\delta}$  have the same factors of length  $3k - 1$  with relatively-free occurrences (we say that an occurrence of length  $3k - 1$  is relatively-free in  $u_{\mathbf{g}\delta}$  if it occurs in some  $u_{\mathbf{g},i}$ , with  $0 \leq i \leq s_{\mathbf{g}}$ , and it is free in  $u_{\mathbf{g},i}$ );
- c)  $u_{e\alpha\delta} \cdot u_{e\delta}$  and  $u_{e\omega\delta}$  have the same relatively-marked factors (we say that a factor of  $u_{\mathbf{g}\delta}$  is relatively-marked if it is a marked factor in some  $u_{\mathbf{g},i}$  with  $0 \leq i \leq s_{\mathbf{g}}$ ).

**Proof.** Let  $i \in \{1, \dots, s_{e\omega}\}$ . Then  $c_{e\omega,i} = c_f$  for some edge  $f$ , and  $c_f$  is a factor of  $u_{e\omega\delta}$ . So, since  $n \geq |c_f|$ , we deduce from Lemma 3.3 and condition (d) on page 18, that  $\mathbf{w}_f$  is a bi-infinite factor of  $e\omega\delta$ . Hence, by consistency of  $\delta q_{\mathbf{LSI}}$ , we deduce that  $\mathbf{w}_f$  is a bi-infinite factor of  $e\alpha\delta \cdot e\delta$ . Now, Proposition 3.2 implies that  $\mathbf{w}_f$  is a factor of  $e\alpha\delta$ , or a factor of  $e\delta$ , or  $\mathbf{w}_f$  is a shift of the bi-infinite word  $\mathbf{w}_e = e\alpha\delta q_{\mathbf{D}} \cdot e\delta q_{\mathbf{K}}$ . Therefore, we conclude respectively that  $c_f$  is an allowed center of  $u_{e\alpha\delta}$  (if  $c_f$  is non-periodic, then  $\overleftarrow{c_f}$  is a factor of  $e\alpha\delta$  and  $|\overleftarrow{c_f}| \leq L < n$  whence  $\overleftarrow{c_f}$  is also a factor of  $u_{e\alpha\delta}$ ), or that  $c_f$  is an allowed center of  $u_{e\delta}$  (analogously), or, by condition (a) on page 18, that  $c_f = c_e$ . This proves that  $c_{e\omega,i} \in \{c_{e\alpha,1}, \dots, c_{e\alpha,s_{e\alpha}}, c_e, c_{e,1}, \dots, c_{e,s_e}\}$ . Inversely, one deduces analogously that all centers  $c_{e\alpha,i}$ ,  $c_e$  and  $c_{e,j}$  are allowed centers of  $u_{e\omega\delta}$ , since the corresponding bi-infinite words  $\mathbf{w}_f$  can be shown to be factors of  $e\omega\delta$ . This concludes the proof of statement a).

Let now  $x$  be a factor of length  $3k - 1$  with a relatively-free occurrence in  $u_{e\omega\delta}$ . Then  $x$  has a free occurrence  $o$  in  $u_{e\omega,i}$ , for some  $i \in \{1, \dots, s_{e\omega}\}$ . Therefore, by definition of free occurrence, there exists a  $k''$ -neighborhood  $v$  of  $o$  in  $u_{e\omega,i}$  such that  $v$  is  $k'$ -abundant. In particular,  $|v| \leq k''$ . Suppose that  $yzv$  is an extension of  $v$  in  $u_{e\omega\delta}$  with  $|y|, |z| = L$ . Then,  $|yzv| \leq 2L + k'' < n$ . Therefore  $yzv$  has an occurrence in  $u_{e\alpha\delta} \cdot u_{e\delta}$ , since  $u_{e\omega\delta}$  and  $u_{e\alpha\delta} \cdot u_{e\delta}$  have the same factors of length  $n$ . Then the corresponding occurrence of  $v$  is

disjoint from the allowed centers of  $u_{e\alpha\delta} \cdot u_{e\delta}$ . Indeed, if this was not the case, since the periodic centers and the extended-centers of non-periodic centers have length at most  $L$ , then the occurrence of  $v$  in  $u_{e\omega\delta}$  would not be disjoint from the allowed centers of  $u_{e\omega\delta}$ . So, in this case,  $x$  has a free occurrence in  $u_{e\alpha\delta} \cdot u_{e\delta}$ . Suppose now that there is not an extension  $yvz$  of  $v$  in  $u_{e\omega\delta}$  with  $|y|, |z| = L$ . That is, if  $yvz$  is an extension of  $v$  in  $u_{e\omega\delta}$ , then  $|y| < L$  or  $|z| < L$ . In the first case,  $v$  occurs in the prefix of  $u_{e\omega\delta}$  of length  $L + k''$ . Now, as  $u_{e\omega\delta}$  and  $u_{e\alpha\delta}$  have the same prefix of length  $n > 2L + k''$ , one can easily verify that the factorizations of  $u_{e\omega\delta}$  and  $u_{e\alpha\delta}$  for the centers of  $\Gamma$  coincide on their common prefix of length  $L + k''$ . Whence  $v$  occurs in  $u_{e\alpha\delta}$  disjoint from  $l_{e\alpha}$  and from the centers  $c_{e\alpha,i}$ , so that  $x$  has a free occurrence in  $u_{e\alpha\delta}$  and so also in  $u_{e\alpha\delta} \cdot u_{e\delta}$ . The second case is symmetric. This concludes the proof that every factor of length  $3k - 1$  that has a relatively-free occurrence in  $u_{e\omega\delta}$  has also a relatively-free occurrence in  $u_{e\alpha\delta} \cdot u_{e\delta}$ . The reverse could be proved symmetrically, whence statement *b*) is established.

The proof of *c*) can be made analogously. We just notice that  $u_{e\omega\delta}$  and  $u_{e\alpha\delta} \cdot u_{e\delta}$  have the same factors of length  $n$  and that the marked factors are of length  $< k''$  by Lemma 4.2.  $\square$

We are now able to show that  $e\alpha\delta' \cdot e\delta'$  and  $e\omega\delta'$  have the same factors of the form  $\bar{x}y\bar{z}$ . Let  $\bar{x}y\bar{z}$  be an occurrence in  $e\alpha\delta' \cdot e\delta'$ . The following cases may arise:

- $\bar{x}y\bar{z} = \widehat{c}$  for some center  $c \in \{c_{e\alpha,1}, \dots, c_{e\alpha,s_{e\alpha}}, c_e, c_{e,1}, \dots, c_{e,s_e}\}$ . By Lemma 6.1,  $c \in \{c_{e\omega,1}, \dots, c_{e\omega,s_{e\omega}}\}$ . Therefore,  $\bar{x}y\bar{z}$  is also a factor of  $e\omega\delta'$ .
- $\bar{x}y\bar{z}$  is an occurrence in some  $\widehat{u_{g,i}}$  with  $g \in \{e\alpha, e\}$ . Then, by definition of  $\widehat{u_{g,i}}$  in section 6.5, either the occurrence  $xyz \in A^+$  is a marked factor in  $u_{g,i}$ , or it is not contained in a marked factor (since, if  $xyz$  is not a marked factor, then  $\bar{x}y\bar{z}$  was obtained in the third or the fourth steps of the transformations of the word  $u_{g,i}$ ). If  $xyz$  is a marked factor  $w$  in  $u_{g,i}$ , then by Lemma 6.1  $xyz$  is a marked factor in some  $u_{e\omega,j}$ . Therefore, by the first step in section 6.5,  $\bar{x}y\bar{z} = \widehat{w}$  is a factor of  $\widehat{u_{e\omega,j}}$  and so also of  $e\omega\delta'$ .

Suppose now that the occurrence  $xyz$  in  $u_{g,i}$  is not contained in a marked factor. Then  $|y| < k$ , so that  $2k \leq |xyz| < 3k$ , and, by the third step in section 6.5, there exists a free occurrence of  $xyz$  in  $u_{g,i}$ . Hence, by Lemma 6.1,  $xyz$  has a free occurrence in some  $u_{e\omega,j}$  and so, by the third step,  $\bar{x}y\bar{z}$  is a factor of  $\widehat{u_{e\omega,j}}$  and so also of  $e\omega\delta'$ .

- $\bar{x}y\bar{z}$  is an occurrence not disjoint both with  $\widehat{l_{e\alpha}}$  and  $\widehat{u_{e\alpha,0}}$ . In this case,  $\widehat{l_{e\alpha}} = p\bar{x}$  for some  $p \in A^*$  by (19) and (20). Hence, since  $l_{e\alpha} = l_{e\omega}$ , we also have  $\widehat{l_{e\omega}} = p\bar{x}$ . On the other hand, let  $u_{e\alpha,0} = x_0w_1x_1w_2 \cdots w_r x_r$  and  $u_{e\omega,0} = x'_0w'_1x'_1w'_2 \cdots w'_s x'_s$  be, respectively, the marked factorizations of  $u_{e\alpha,0}$  and  $u_{e\omega,0}$ .

If  $|u_{e\alpha,0}| < 3k - 1$ , then  $u_{e\alpha,0} = u_{e\omega,0}$  and  $c_{e\alpha,1} = c_{e\omega,1}$  since  $u_{e\alpha\delta}$  and  $u_{e\omega\delta}$  have the same prefix of length  $n > 3L$ . Therefore, by the first step of the transformation described in section 6.5,

$$\widehat{u_{e\alpha,0}} = w_1 = u_{e\alpha,0} = u_{e\omega,0} = w'_1 = \widehat{u_{e\omega,0}}.$$

so that,  $y = w_1 = w'_1$  and  $\widehat{c_{e\alpha,1}} = \widehat{c_{e\omega,1}} = \bar{z}s$  for some  $s \in \bar{A}^+$ . Therefore,  $\bar{x}y\bar{z}$  is a factor of  $e\omega\delta'$ .

Suppose now that  $|u_{e\alpha,0}| \geq 3k - 1$ . Then, by the second step of the transformation described in section 6.5,  $y\bar{z}$  is either  $x_0\bar{w}_{1,1}$  if  $|x_0| < k$ , or  $\bar{x}_{0,1}$  if  $|x_0| \geq k$ . Since  $u_{e\alpha\delta}$  and  $u_{e\omega\delta}$  have the same prefix of length  $n$ , if  $|x_0| < k$ , then  $x'_0 = x_0$  and  $w_1 = w'_1$ . In this case  $y\bar{z} = x'_0\bar{w}'_{1,1} = x_0\bar{w}_{1,1}$ . If  $|x_0| \geq k$ , then  $x'_{0,1} = x_{0,1}$  and so  $y\bar{z} = \bar{x}'_{0,1} = \bar{x}_{0,1}$ . Therefore, we deduce in both cases that  $\bar{x}y\bar{z}$  is a factor of  $\widehat{l_{e\omega}u_{e\omega,0}}$ , whence it is a factor of  $e\omega\delta'$ .

- $\bar{x}y\bar{z}$  is an occurrence not disjoint both with  $\widehat{u_{e,s_e}}$  and  $\widehat{r_e}$ . This case is symmetric of the previous one.
- for some  $1 \leq i \leq s_g$ ,  $\bar{x}y\bar{z}$  is an occurrence not disjoint both with  $\widehat{u_{g,i-1}}$  and  $\widehat{c_{g,i}}$ , or with  $\widehat{c_{g,i}}$  and  $\widehat{u_{g,i}}$  (where  $g \in \{e\alpha, e\}$ ), or with  $\widehat{u_{e\alpha,s_{e\alpha}}}$  and  $\widehat{c_e}$ , or with  $\widehat{c_e}$  and  $\widehat{u_{e,0}}$ . These cases are similar to the two last ones.

Therefore, we have proved in all cases that  $\bar{x}y\bar{z}$  is a factor of  $e\omega\delta'$ . Analogously, one can show that every factor  $\bar{x}y\bar{z}$  of  $e\omega\delta'$  is also a factor of  $e\alpha\delta' \cdot e\delta'$ . This concludes the proof of the consistency of  $\delta'p_{\mathbf{LSI}}$  and establishes condition c1).

The proof of condition c4) can be made analogously. In fact, if  $v_1$  and  $v_2$  are vertices such that  $v_1\delta q_{\mathbf{LSI}} = v_2\delta q_{\mathbf{LSI}}$ , then, in particular, the words  $u_{v_1\delta}$  and  $u_{v_2\delta}$  are  $\equiv_{n-}$  equivalent by Proposition 3.2 and Lemma 3.3. Therefore, we can assume that  $r_{v_1} = r_{v_2}$  so that we can also assume that  $\widehat{l_{v_1}} = \widehat{l_{v_2}}$  and  $\widehat{r_{v_1}} = \widehat{r_{v_2}}$ . Moreover, one can show that the words  $u_{v_1\delta}$  and  $u_{v_2\delta}$  have the same allowed centers, the same factors of length  $3k - 1$  with relatively-free occurrences and the same relatively-marked factors. The proof that  $v_1\delta'p_{\mathbf{LSI}} = v_2\delta'p_{\mathbf{LSI}}$  is now analogous to the one above.

This proves the following theorem.

**Theorem 6.2** *The pseudovariety **LSI** is  $\kappa$ -tame.*

## References

- [1] D. Albert, R. Baldinger and J. Rhodes, *The identity problem for finite semigroups (the undecidability of)*, J. Symbolic Logic **57** (1992), 179-192.
- [2] J. Almeida, *Finite semigroups and universal algebra*, World Scientific, Singapore, 1994. Revised translation of the original in Portuguese, 1992.
- [3] J. Almeida, *Hyperdecidable pseudovarieties and the calculation of semidirect products*, Int. J. Algebra and Computation **9** (1999), 241-261.
- [4] J. Almeida, *Some key problems on finite semigroups*, Semigroup Forum, to appear.
- [5] J. Almeida, *Finite semigroups: An Introduction to a unified theory of pseudovarieties*, in Semigroups, Algorithms, Automata and Languages, ed. G. Gomes, World Scientific, 2002, pp. 3-64.

- [6] J. Almeida and M. Delgado, *Tameness of the pseudovariety of abelian groups*, Tech. Report CMUP 2001-24, Univ. Porto, 2001.
- [7] J. Almeida and B. Steinberg, *On the decidability of iterated semidirect products and applications to complexity*, Proc. London Math. Soc. **80** (2000), 50-74.
- [8] J. Almeida and P. Weil, *Relatively free profinite monoids: an introduction and examples*, in Semigroups, Formal Languages and Groups, ed. J. Fountain, Kluwer, 1995, pp. 73-117.
- [9] J. Almeida and M. Zeitoun, *Tameness of some locally trivial pseudovarieties*, Communications in Algebra, to appear.
- [10] C. Ash, *Inevitable graphs: a proof of the type II conjecture and some related decision procedures*, Int. J. Algebra and Computation **1** (1991), 127-146.
- [11] J. Brozowski and I. Simon, *Characterization of locally testable events*, Discrete Math. **4** (1973), 243-271.
- [12] J. C. Costa, *Free profinite locally idempotent and locally commutative semigroups*, Journal of Pure and Applied Algebra **163** (2001), 19-47.
- [13] J. C. Costa, *Tameness of joins involving some locally trivial pseudovarieties*, in preparation.
- [14] R. McNaughton, *Algebraic decision procedures for local testability*, Math. Systems Theory **8** (1974), 60-76.
- [15] J.-E. Pin, *Variétés de langages formels*, Masson, Paris, 1984. English translation: *Varieties of formal languages*, Plenum, New York and North Oxford, London, 1986.
- [16] J. Reiterman, *The Birkhoff theorem for finite algebras*, Algebra Universalis **14** (1982), 1-10.
- [17] J. Rhodes, *Undecidability, automata and pseudovarieties of semigroups*, Int. J. Algebra and Computation **9** (1999), 455-473.