

# COMBINATORICS OF CYCLIC SHIFTS IN PLACTIC, HYPOPLACTIC, SYLVESTER, AND RELATED MONOIDS

ALAN J. CAIN AND ANTÓNIO MALHEIRO

ABSTRACT. The cyclic shift graph of a monoid is the graph whose vertices are elements of the monoid and whose edges link elements that differ by a cyclic shift. For certain monoids connected with combinatorics, such as the plactic monoid (the monoid of Young tableaux) and the sylvester monoid (the monoid of binary search trees), connected components consist of elements that have the same evaluation (that is, contain the same number of each generating symbol). This paper discusses new results on the diameters of connected components of the cyclic shift graphs of the finite-rank analogues of these monoids, showing that the maximum diameter of a connected component is dependent only on the rank. The proof techniques are explained in the case of the sylvester monoid.

## 1. INTRODUCTION

In a monoid  $M$ , two elements  $s$  and  $t$  are related by a cyclic shift, denoted  $s \sim t$ , if and only if there exist  $x, y \in M$  such that  $s = xy$  and  $t = yx$ . In the plactic monoid (the monoid of Young tableaux, denoted  $\text{plac}$ ; see [Lot02, Ch. 5]), elements that have the same evaluation (that is, which contain the same number of each symbol) can be obtained from each other by iterated application of cyclic shifts [LS81, § 4]. Furthermore, in the plactic monoid of rank  $n$  (denoted  $\text{plac}_n$ ), it is known that  $2n - 2$  applications of cyclic shifts are sufficient [CM13, Theorem 17].

To restate these results in a new form, define the *cyclic shift graph*  $K(M)$  of a monoid  $M$  to be the undirected graph with vertex set  $M$  and, for all  $s, t \in M$ , an edge between  $s$  and  $t$  if and only if  $s \sim t$ . Connected components of  $K(M)$  are  $\sim^*$ -classes (where  $\sim^*$  is the reflexive and transitive closure of  $\sim$ ), since they consist of elements that are related by iterated cyclic shifts. Thus the results discussed above say that each connected component of  $K(\text{plac})$  consists of precisely the elements with a given evaluation, and that the diameter of a connected component of  $K(\text{plac}_n)$  is at most  $2n - 2$ . Note that connected components are of unbounded size, despite there being a bound on diameters that is dependent only on the rank.

The plactic monoid is celebrated for its ubiquity, appearing in many diverse contexts (see the discussion and references in [Lot02, Ch. 5]). It

---

The first author was supported by an Investigador FCT fellowship (IF/01622/2013/CP1161/CT0001).

For both authors, this work was partially supported by by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2013 (Centro de Matemática e Aplicações), and the project PTDC/MHC-FIL/2583/2014.

TABLE 1. Monoids and corresponding combinatorial objects.

<i>Monoid</i>	<i>Symbol</i>	<i>Combinatorial object</i>	<i>Citation</i>
Plactic	plac	Young tableau	[Lot02, ch. 5]
Hypoplactic	hypo	Quasi-ribbon tableau	[Nov00]
Stalactic	stal	Stalactic tableau	[HNT07]
Sylvester	sylv	Binary search tree	[HNT05]
Taiga	taig	Binary search tree with multiplicities	[Pri13, § 5]
Baxter	baxt	Pair of twin binary search trees	[Gir12]

TABLE 2. Properties of connected component of the cyclic shift graph for rank- $n$  monoids: whether they are characterized by evaluation, and known values and bounds for their maximum diameters.

<i>Monoid</i>	<i>Char. by evaluation</i>	<i>Maximum diameter</i>			
		<i>Known value</i>	<i>Conjecture</i>	<i>Known bounds</i>	
				<i>Lower</i>	<i>Upper</i>
plac <sub><math>n</math></sub>	Y	?	$n - 1$	$n - 1$	$2n - 3$
hypo <sub><math>n</math></sub>	Y	$n - 1$	—	—	—
stal <sub><math>n</math></sub>	N	$\begin{cases} n - 1 & \text{if } n < 3 \\ n & \text{if } n \geq 3 \end{cases}$	—	—	—
sylv <sub><math>n</math></sub>	Y	?	$n - 1$	$n - 1$	$n$
taig <sub><math>n</math></sub>	Y	?	$n - 1$	$n - 1$	$n$
baxt <sub><math>n</math></sub>	N	?	?	?	?

is, however, just one member of a family of ‘plactic-like’ monoids that are closely connected with combinatorics. These monoids include the hypoplactic monoid (the monoid of quasi-ribbon tableaux) [KT97, Nov00], the sylvester monoid (binary search trees) [HNT05], the taiga monoid (binary search trees with multiplicities) [Pri13], the stalactic monoid (stalactic tableaux) [HNT07, Pri13], and the Baxter monoid (pairs of twin binary search trees) [Gir12]. These monoids, including the plactic monoid, arise in a parallel way. For each monoid, there is a so-called insertion algorithm that allows one to compute a combinatorial object (of the corresponding type) from a word over the infinite ordered alphabet  $\mathcal{A} = \{1 < 2 < 3 < \dots\}$ ; the relation that relates pairs of words that give the same combinatorial object is a congruence (that is, it is compatible with multiplication in the free monoid  $\mathcal{A}^*$ ). The monoid arises by factoring the free monoid  $\mathcal{A}^*$  by this congruence; thus elements of the monoid (equivalence classes of words) are in one-to-one correspondence with the combinatorial objects. Table 1 lists these monoids and their corresponding objects.

Analogous questions arise for the cyclic shift graph of each of these monoids. In a forthcoming paper [CM], the present authors make a comprehensive study of connected components in the cyclic shift graphs of each of these monoids. For several of these monoids, it turns out that each connected

component of its cyclic shift graph consists of precisely the elements with a given evaluation, and that the diameters of connected component in the rank- $n$  case are bounded by a quantity dependent on the rank. (Again, it should be emphasized that there is no bound on the *size* of these connected components.) In each case, the authors either establish the exact value of the maximum diameter or give bounds; Table 2 summarizes the results from [CM]. Also, although these monoids are multihomogeneous (words in  $\mathcal{A}^*$  representing the same element contain the same number of each symbol), the authors also exhibit a rank 4 multihomogeneous monoid for which there is no bound on the diameter of connected components. Thus it seems to be the underlying combinatorial objects that ensure the bound on diameters. This also is of interest because cyclic shifts are a possible generalization of conjugacy from groups to monoids; thus the combinatorial objects are here linked closely to the algebraic structure of the monoid.

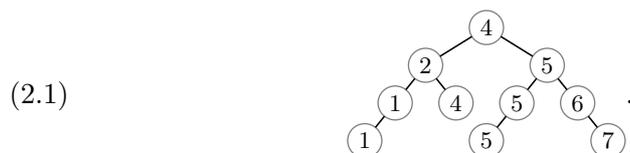
The present paper illustrates these results by focussing on the sylvester monoid (denoted  $\text{sylv}$  or  $\text{sylv}_n$  in the rank- $n$  case). (The authors previously proved that each connected component of  $K(\text{sylv})$  consists of precisely the elements with a given evaluation [CM15, § 3].) Section 2 recalls the definition and necessary facts about the sylvester monoid. Section 3 gives a complete proof that there is a connected component in  $K(\text{sylv}_n)$  with diameter at least  $n - 1$ ; this establishes the lower bound on the maximum diameter shown in Table 2. The complete proof that every connected component of  $K(\text{sylv}_n)$  has diameter at most  $n$ , establishing the upper maximum diameter shown in Table 2, is very long and complicated. Thus Section 4 gives the proof for connected components consisting of elements that contain each symbol from  $\{1, \dots, k\}$  (for some  $k$ ) exactly once; this avoids many of the complexities of the general case.

## 2. BINARY SEARCH TREES AND THE SYLVESTER MONOID

This section gathers the relevant definitions and background on the sylvester monoid; see [HNT05] for further reading.

Recall that  $\mathcal{A}$  denotes the infinite ordered alphabet  $\{1 < 2 < \dots\}$ . Fix a natural number  $n$  and let  $\mathcal{A}_n = \{1 < 2 < \dots < n\}$  be the set of the first  $n$  natural numbers, viewed as a finite ordered alphabet. A word  $u \in \mathcal{A}^*$  is *standard* if it contains each symbol in  $\{1, \dots, |u|\}$  exactly once.

A (*right strict*) *binary search tree* (BST) is a rooted binary tree labelled by symbols from  $\mathcal{A}$ , where the label of each node is greater than or equal to the label of every node in its left subtree, and strictly less than the label of every node in its right subtree. An example of a binary search tree is:



The following algorithm inserts a new symbol into a BST, adding it as a leaf node in the unique place that maintains the property of being a BST.

**Algorithm 2.1.** *Input:* A binary search tree  $T$  and a symbol  $a \in \mathcal{A}_n$ .

If  $T$  is empty, create a node and label it  $a$ . If  $T$  is non-empty, examine the label  $x$  of the root node; if  $a \leq x$ , recursively insert  $a$  into the left subtree of the root node; otherwise recursively insert  $a$  into the right subtree of the root node. Output the resulting tree.

For  $u \in \mathcal{A}^*$ , define  $P_{\text{sylv}}(u)$  to be the right strict binary search tree obtained by starting with the empty tree and inserting the symbols of the word  $u$  one-by-one using Algorithm 2.1, proceeding *right-to-left* through  $u$ . For example,  $P_{\text{sylv}}(5451761524)$  is (2.1). Define the relation  $\equiv_{\text{sylv}}$  by

$$u \equiv_{\text{sylv}} v \iff P_{\text{sylv}}(u) = P_{\text{sylv}}(v),$$

for all  $u, v \in \mathcal{A}^*$ . The relation  $\equiv_{\text{sylv}}$  is a congruence, and the *sylvester monoid*, denoted  $\text{sylv}$ , is the factor monoid  $\mathcal{A}^*/\equiv_{\text{sylv}}$ ; the *sylvester monoid of rank  $n$* , denoted  $\text{sylv}_n$ , is the factor monoid  $\mathcal{A}_n^*/\equiv_{\text{sylv}}$  (with the natural restriction of  $\equiv_{\text{sylv}}$ ). Each element  $[u]_{\equiv_{\text{sylv}}}$  (where  $u \in \mathcal{A}^*$ ) can be identified with the binary search tree  $P_{\text{sylv}}(u)$ . The monoid  $\text{sylv}$  is presented by  $\langle \mathcal{A} \mid \mathcal{R}_{\text{sylv}} \rangle$ , where

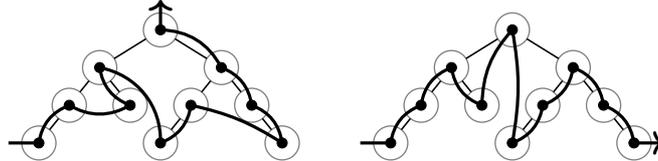
$$\mathcal{R}_{\text{sylv}} = \{ (cavb, acvb) : a \leq b < c, v \in \mathcal{A}^* \};$$

the monoid  $\text{sylv}_n$  is presented by  $\langle \mathcal{A}_n \mid \mathcal{R}_{\text{sylv}} \rangle$ , where the set of defining relations  $\mathcal{R}_{\text{sylv}}$  is naturally restricted to  $\mathcal{A}_n^* \times \mathcal{A}_n^*$ . Notice that  $\text{sylv}$  and  $\text{sylv}_n$  are multihomogeneous.

A *reading* of a binary search tree  $T$  is a word formed from the symbols that appear in the nodes of  $T$ , arranged so that the child nodes appear before parents. A word  $w \in \mathcal{A}^*$  is a reading of  $T$  if and only if  $P_{\text{sylv}}(w) = T$ . The words in  $[u]_{\equiv_{\text{sylv}}}$  are precisely the readings of  $P_{\text{sylv}}(u)$ .

A binary search tree  $T$  with  $k$  nodes is *standard* if it has exactly one node labelled by each symbol in  $\{1, \dots, k\}$ ; clearly  $T$  is standard if and only if all of its readings are standard words.

The *left-to-right postfix traversal*, or simply the *postfix traversal*, of a rooted binary tree  $T$  is the sequence that ‘visits’ every node in the tree as follows: it recursively perform the postfix traversal of the left subtree of the root of  $T$ , then recursively perform the postfix traversal of the right subtree of the root of  $T$ , then visits the root of  $T$ . The *left-to-right infix traversal*, or simply the *infix traversal*, of a rooted binary tree  $T$  is the sequence that ‘visits’ every node in the tree as follows: it recursively performs the infix traversal of the left subtree of the root of  $T$ , then visits the root of  $T$ , then recursively performs the infix traversal of the right subtree of the root of  $T$ . Thus the postfix and infix traversals of any binary tree with the same shape as (2.1) visit nodes as shown on the left and right below:



The following result is immediate from the definition of a binary search tree, but it is used frequently:

**Proposition 2.2.** *For any binary search tree  $T$ , if a node  $x$  is encountered before a node  $y$  in an infix traversal, then  $x \leq y$ .*

In this paper, a *subtree* of a binary search tree will always be a rooted subtree. Let  $T$  be a binary search tree and  $x$  a node of  $T$ . The *complete subtree of  $T$  at  $x$*  is the subtree consisting of  $x$  and every node below  $x$  in  $T$ . The *path of left child nodes in  $T$  from  $x$*  is the path that starts at  $x$  and enters left child nodes until a node with empty left subtree is encountered.

Let  $B$  be a subtree of  $T$ . Then  $B$  is said to be *on* the path of left child nodes from  $x$  if the root of  $B$  is one of the nodes on this path. The *left-minimal* subtree of  $B$  in  $T$  is the complete subtree at the left child of the left-most node in  $B$ ; the *right-maximal* subtree of  $B$  in  $T$  is the complete subtree at the right child of the right-most node in  $B$ .

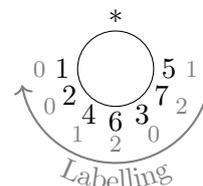
In diagrams of binary search trees, individual nodes are shown as round, while subtrees as shown as triangles. An edge emerging from the top of a triangle is the edge running from the root of that subtree to its parent node. A vertical edge joining a node to its parent indicates that the node may be either a left or right child. An edge emerging from the bottom-left of a triangle is the edge to that subtree's left-minimal subtree; an edge emerging from the bottom-right of a triangle is the edge to that subtree's right-maximal subtree.

### 3. LOWER BOUND ON DIAMETERS

Let  $u \in \mathcal{A}_n^*$  be a standard word. The *cocharge sequence* of  $u$ , denoted  $\text{cochseq}(u)$ , is a sequence (of length  $n$ ) calculated from  $u$  as follows:

- (1) Draw a circle, place a point  $*$  somewhere on its circumference, and, starting from  $*$ , write  $u$  anticlockwise around the circle.
- (2) Label the symbol 1 with 0.
- (3) Iteratively, after labelling some  $i$  with  $k$ , proceed clockwise from  $i$  to  $i + 1$ . If the symbol  $i + 1$  is reached *before*  $*$ , label  $i + 1$  by  $k + 1$ . Otherwise, if the symbol  $i + 1$  is reached *after*  $*$ , label  $i + 1$  by  $k$ .
- (4) The sequence whose  $i$ -th term is the label of  $i$  is  $\text{cochseq}(u)$ .

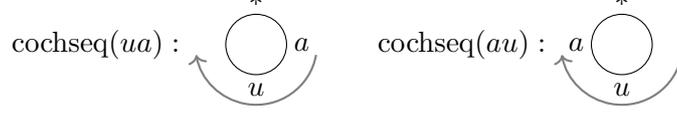
For example, for the word  $u = 1246375$ , the labelling process is shown on the right, and it follows that  $\text{cochseq}(u) = (0, 0, 0, 1, 1, 2, 2)$ . Notice that the first term of a cocharge sequence is always 0, and that each term in the sequence is either the same as its predecessor or greater by 1. Thus the  $i$ -th term in the sequence always lies in the set  $\{0, 1, \dots, i - 1\}$ .



The usual notion of ‘cocharge’ is obtained by summing the cocharge sequence (see [Lot02, § 5.6]).

- Lemma 3.1.**
- (1) Let  $u \in \mathcal{A}_n^*$  and  $a \in \mathcal{A}_n \setminus \{1\}$  be such that  $ua$  is a standard word. Then  $\text{cochseq}(ua)$  is obtained from  $\text{cochseq}(au)$  by adding 1 to the  $a$ -th component.
  - (2) Let  $x, y \in \mathcal{A}_n^*$  be such that  $xy \in \mathcal{A}_n^*$  is a standard word. Then corresponding components of  $\text{cochseq}(xy)$  and  $\text{cochseq}(yx)$  differ by at most 1.

*Proof.* Consider how  $a$  is labelled during the calculation of  $\text{cochseq}(ua)$  and  $\text{cochseq}(au)$ :

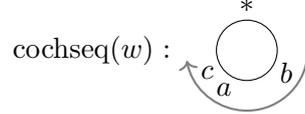


In the calculation of  $\text{cochseq}(ua)$ , the symbol  $a - 1$  receives a label  $k$ , and then  $a$  is reached *after*  $*$  is passed; hence  $a$  also receives the label  $k$ . In the calculation of  $\text{cochseq}(au)$ , the symbols  $1, \dots, a - 1$  receive the same labels as they do in the calculation of  $\text{cochseq}(ua)$ , but after labelling  $a - 1$  by  $k$  the symbol  $a$  is reached *before*  $*$  is passed; hence  $a$  receives the label  $k + 1$ ; after this point, labelling proceeds in the same way. This proves part 1). For part 2), notice that one of  $x$  and  $y$  does not contain the symbol 1; the result is now an immediate consequence of part 1).  $\square$

**Proposition 3.2.** *Let  $u, v \in \mathcal{A}_n^*$  be standard words such that  $u \equiv_{\text{sylv}} v$ . Then  $\text{cochseq}(u) = \text{cochseq}(v)$ .*

*Proof.* It suffices to prove the result when  $w$  and  $w'$  differ by a single application of a defining relation  $(cavb, acvb) \in \mathcal{R}_{\text{sylv}}$  where  $a \leq b < c$ . In this case,  $w = pcavbq$  and  $w' = pacvbq$ , where  $p, q, v \in \mathcal{A}_n^*$  and  $a, b, c \in \mathcal{A}_n$  with  $a \leq b < c$ . Since  $w$  and  $w'$  are standard words,  $a < b$ .

Consider how labels are assigned to the symbols  $a, b$ , and  $c$  when calculating the cocharge sequence of  $w$ :

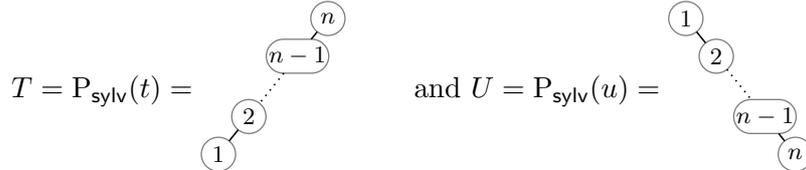


Among these three symbols,  $a$  will receive a label first, then  $b$ , then  $c$ . Thus, after  $a$ , the labelling process will pass  $*$  at least once to visit  $b$  and only then visit  $c$ . Thus if we interchange  $a$  and  $c$ , we do not alter the resulting labelling. Hence  $\text{cochseq}(w) = \text{cochseq}(w')$ .  $\square$

For any standard binary tree  $T$  in  $\text{sylv}_n$ , define  $\text{cochseq}(T)$  to be  $\text{cochseq}(u)$  for any standard word  $u \in \mathcal{A}_n^*$  such that  $T = P_{\text{sylv}}(u)$ . By Proposition 3.2,  $\text{cochseq}(T)$  is well-defined.

**Proposition 3.3.** *The connected component of  $K(\text{sylv}_n)$  consisting of standard elements has diameter at least  $n - 1$ .*

*Proof.* Let  $t = 12 \cdots (n - 1)n$  and  $u = n(n - 1) \cdots 21$ , and let



Since  $T$  and  $U$  have the same evaluation, they are  $\sim^*$ -related by [CM15, § 3], and so in the same connected component of  $K(\text{sylv}_n)$ . Let  $T = T_0, T_1, \dots, T_{m-1}, T_m = U$  be a path in  $K(\text{sylv}_n)$  from  $T$  to  $U$ . Then for  $i = 0, \dots, m - 1$ , we have  $T_i \sim T_{i+1}$ . That is, there are words  $u_i, v_i \in \mathcal{A}_n^*$  such

that  $T_i = P_{\text{sylv}}(u_i v_i)$  and  $T_{i+1} = P_{\text{sylv}}(v_i u_i)$ . By Lemma 3.1(2),  $\text{cochseq}(T_i)$  and  $\text{cochseq}(T_{i+1})$  differ by adding 1 or subtracting 1 from certain components. Hence corresponding components of  $\text{cochseq}(T)$  and  $\text{cochseq}(U)$  differ by at most  $m$ . Since  $\text{cochseq}(T) = (0, 0, \dots, 0, 0)$  and  $\text{cochseq}(U) = (0, 1, \dots, n-2, n-1)$ , it follows that  $m \geq n-1$ . Hence  $T$  and  $U$  are a distance at least  $n-1$  apart in  $K(\text{sylv}_n)$ .  $\square$

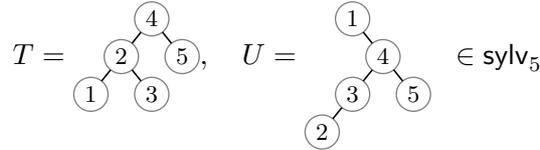
4. UPPER BOUNDS ON DIAMETERS

**Proposition 4.1.** *Any two standard elements of  $\text{sylv}_n$  are a distance at most  $n$  apart in  $K(\text{sylv}_n)$ .*

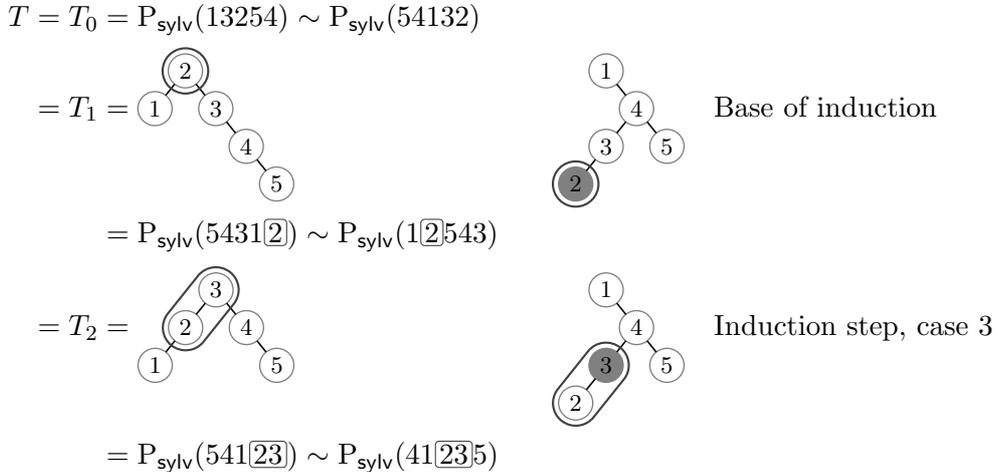
*Proof.* Since  $\text{sylv}_m$  embeds into  $\text{sylv}_n$  for all  $m \leq n$ , and since  $K(\text{sylv}_m)$  is the subgraph of  $K(\text{sylv}_n)$  induced by  $\text{sylv}_m$ , this result follows from Lemma 4.3 below.  $\square$

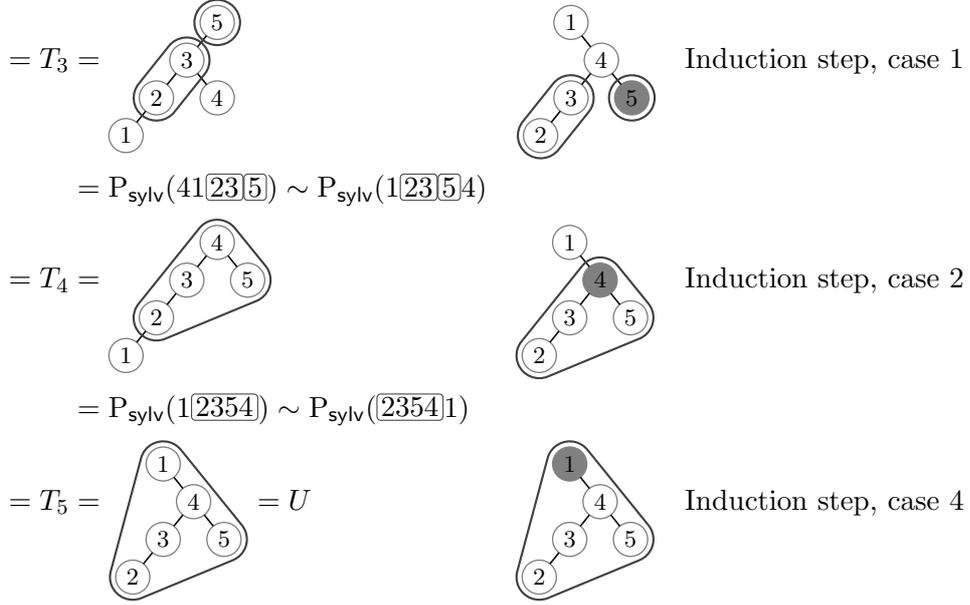
Lemma 4.3 proves that in  $K(\text{sylv}_n)$  there is a path of length at most  $n$  between two standard elements with the same number of nodes. First, however, the strategy used to construct such a path is illustrated in the following example.

**Example 4.2.** Let



The aim is to build a sequence  $T = T_0 \sim T_1 \sim T_2 \sim T_3 \sim T_4 \sim T_5 = U$ . Consider the postfix traversal of  $U$ . The 5 steps in this traversal are shown below on the right, together with the relevant cases in the proof of Lemma 4.3. The parts of  $U$  that have been visited already at each step are outlined. The idea is that the  $h$ -th cyclic shift leads to a tree  $T_h$  where copies of the outlined parts of  $U$  appear on the path of left child nodes from the root of  $T_h$ . Note that cyclic shifts never break up the subwords (outlined) that represent the already-built subtrees. (The difficulty in the general proof is showing that a suitable cyclic shift exists at each step.)





**Lemma 4.3.** *Let  $T, U \in \text{sylv}_n$  be standard and have  $n$  nodes. Then there is a sequence  $T = T_0, T_1, \dots, T_n = U$  with  $T_h \sim T_{h+1}$  for  $h = 0, \dots, n-1$ .*

*Proof.* This proof is only concerned with standard BSTs; thus for brevity nodes are identified with their labels. Notice that each of  $T$  and  $U$  has exactly one node labelled by each symbol in  $\mathcal{A}_n$ .

Consider the left-to-right postfix traversal of  $U$ ; there are exactly  $n$  steps in this traversal. Let  $u_h$  be the node visited at the  $h$ -th step of this traversal.

For  $h = 1, \dots, n$ , let  $U_h = \{u_1, \dots, u_h\}$  and let  $U_h^\top$  be the set of nodes in  $U_h$  that do not lie below any other node in  $U_h$ . Since a later step in a postfix traversal is never below an earlier one, it follows that  $u_h \in U_h^\top$  for all  $h$ . Let  $B_h$  be the subtree of  $U$  consisting of  $u_h$  and every node that is below  $u_h$ .

The aim is to construct inductively the required sequence. Let  $h = 1, \dots, n$  and suppose  $U_h^\top = \{u_{i_1}, \dots, u_{i_k}\}$  (where  $i_1 < \dots < i_k = h$ ). Then the tree  $T_h$  will satisfy the following conditions:

P1 The subtree  $B_{i_k}$  appears at the root of  $T_h$ .

P2 The subtrees  $B_{i_k}, \dots, B_{i_1}$  appear, in that order (but not necessarily consecutively), on the path of left child nodes from the root of  $T_h$ .

(Note that conditions P1 and P2 do not apply to  $T_0$ .)

*Base of induction.* Set  $T_0 = T$ . Take any reading of  $T_0$  and factor it as  $wu_1w'$ . Let  $T_1 = \text{P}_{\text{sylv}}(w'u_1w)$ . Clearly  $T_1$  has root node  $u_1$ . Since  $B_1$  consists only of the node  $u_1$  (since  $u_1$  is the first node in  $U$  visited by the postfix traversal and is thus a leaf node),  $T_1$  satisfies P1. Further,  $T_1$  trivially satisfies P2. Finally, note that  $T_0 \sim T_1$ . (For an illustration, see the definition of  $T_1$  in Example 4.2.)

*Induction step.* The remainder of the sequence of trees is built inductively. Suppose that the tree  $T_h$  satisfies P1 and P2; the aim is to find  $T_{h+1}$  satisfying P1 and P2 with  $T_h \sim T_{h+1}$ . There are four cases, depending on the relative positions of  $u_h$  and  $u_{h+1}$  in  $U$ :



FIGURE 1. Induction step, case 1.

- (1)  $u_h$  is a left child node and  $u_{h+1}$  is in the right subtree of the parent of  $u_h$ ;
- (2)  $u_h$  is the right child of  $u_{h+1}$ , and  $u_{h+1}$  has non-empty left subtree;
- (3)  $u_h$  is the left child of  $u_{h+1}$  (which implies, by the definition of the postfix traversal, that  $u_{h+1}$  has empty right subtree);
- (4)  $u_h$  is the right child of  $u_{h+1}$ , and  $u_{h+1}$  has empty left subtree.

*Case 1.* Suppose that, in  $U$ , the node  $u_h$  is a left child node and  $u_{h+1}$  is in the right subtree of the parent of  $u_h$ . (For an illustration of this case, see the step from  $T_2$  to  $T_3$  in Example 4.2.) Then  $B_{h+1}$  consists only of the node  $u_{h+1}$ , since by the definition of a postfix traversal  $u_{h+1}$  is a leaf node. Furthermore,  $U_{h+1}^\top = U_h^\top \cup \{u_{h+1}\}$ .

By P1,  $B_h$  appears at the root of  $T_h$ . By Proposition 2.2 applied to  $U$ , the symbol  $u_{h+1}$  is greater than every node of  $B_h$ , so  $u_{h+1}$  must be in the right-maximal subtree of  $B_h$  in  $T_h$ .

As shown in Figure 1, let  $\lambda$  be a reading of the left-minimal subtree of  $B_h$ . Let  $\delta$  be a reading of the right-maximal subtree of  $B_h$  outside of the complete subtree at  $u_{h+1}$ . Let  $\alpha$  and  $\beta$  be readings of the left and right subtrees of  $u_{h+1}$ , respectively. Note that the subtrees  $B_i$  for  $u_i \in U_h^\top$  are contained in  $\lambda$ .

Thus  $T_h = \text{P}_{\text{syIv}}(\alpha\beta u_{h+1}\delta\lambda B_h)$ . Let  $T_{h+1} = \text{P}_{\text{syIv}}(\delta\lambda B_h\alpha\beta u_{h+1})$ ; note that  $T_h \sim T_{h+1}$ .

In computing  $T_{h+1}$ , the symbol  $u_{h+1}$  is inserted first and becomes the root node. Since  $B_{h+1}$  consists only of the node  $u_{h+1}$ , the tree  $T_{h+1}$  satisfies P1. Since every symbol in  $B_h$  and  $\lambda$  is strictly less than every symbol in  $\alpha$ ,  $\beta$ , or  $\delta$ , the trees  $B_h$  and  $\lambda$  are re-inserted on the path of left child nodes from the root of  $T_{h+1}$ . Thus all the subtrees  $B_i$  for  $u_i \in U_{h+1}^\top$  are on the path of left child nodes from the root, and so  $T_{h+1}$  satisfies P2.

*Case 2.* Suppose that in  $U$ , the node  $u_h$  is the right child of  $u_{h+1}$ , and  $u_{h+1}$  has non-empty left subtree. (For an illustration of this case, see the step from  $T_3$  to  $T_4$  in Example 4.2.) Let  $u_g$  be the left child of  $u_{h+1}$ . Thus  $B_{h+1}$  consists of  $u_{h+1}$  with left subtree  $B_g$  and right subtree  $B_h$ . By the definition of the postfix traversal,  $u_g$  was visited before the  $h$ -th step, but no node above  $u_g$  has been visited. That is,  $u_g \in U_h^\top$ . Hence  $U_{h+1}^\top = (U_h^\top \setminus \{u_g, u_h\}) \cup \{u_{h+1}\}$ .

By P1,  $B_h$  appears at the root of  $T_h$ ; by P2,  $B_g$  is next subtree  $B_{i_j}$  on the path of left child nodes from the root of  $T_h$  (and is thus in the left-minimal subtree of  $B_h$ ). By Proposition 2.2 applied to  $U$ , the symbol  $u_{h+1}$  is the unique symbol that is greater than every node of  $B_g$  and less than every node of  $B_h$ . Then the node  $u_{h+1}$  may be in one of two places in  $T_h$ , leading to the two sub-cases below. In both cases, as shown in Figure 2, let  $\lambda$  be a

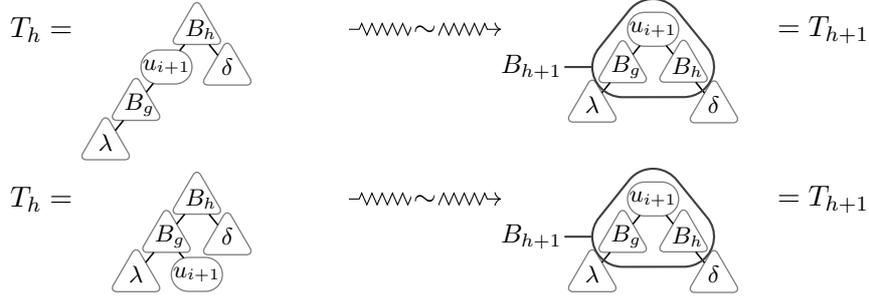


FIGURE 2. Induction step, case 2, two sub-cases.

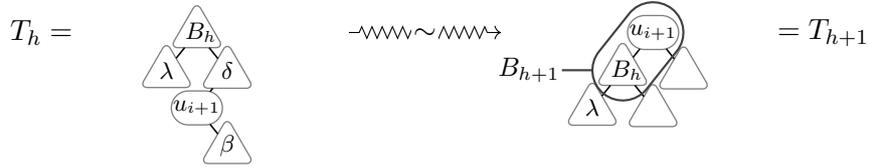


FIGURE 3. Induction step, case 3.

reading of the left subtree of  $B_g$  and let  $\delta$  be a reading of the right-maximal subtree of  $B_h$ ; note that the subtrees  $B_i$  for  $u_i \in U_h^\top \setminus \{u_g, u_h\}$  are contained in  $\lambda$ .

- (1) Suppose  $u_{h+1}$  is the unique node on the path of left child nodes between  $B_g$  and  $B_h$ . In this case, as shown in Figure 2(top),  $T_h = \text{P}_{\text{syIv}}(\lambda B_g u_{h+1} \delta B_h)$ . Let  $T_{h+1} = \text{P}_{\text{syIv}}(\delta B_h \lambda B_g u_{h+1})$ ; note that  $T_h \sim T_{h+1}$ .
- (2) Suppose  $u_{h+1}$  is the unique node in the right-maximal subtree of  $B_g$  and there are no nodes between  $B_g$  and  $B_h$  on the path of left child nodes. In this case, as shown in Figure 2(bottom),  $T_h = \text{P}_{\text{syIv}}(u_{h+1} \lambda B_g \delta B_h)$ . Let  $T_{h+1} = \text{P}_{\text{syIv}}(\lambda B_g \delta B_h u_{h+1})$ ; note that  $T_h \sim T_{h+1}$ .

In computing  $T_{h+1}$ , for both sub-cases, the symbol  $u_{h+1}$  is inserted first and becomes the root node. Every symbol in  $B_g$  and  $\lambda$  is less than  $u_{h+1}$ , so these trees are re-inserted into the left subtree of  $u_{h+1}$ . Every symbol in  $B_h$  and  $\delta$  is greater than  $u_{h+1}$ , so these trees are re-inserted into the right subtree of  $u_{h+1}$ . Since  $B_{h+1}$  consists of  $u_{h+1}$  with  $B_g$  as its left subtree and  $B_h$  as its right subtree, the subtree  $B_{h+1}$  appears at the root and so  $T_{h+1}$  satisfies P1. All the other subtrees  $B_i$  for  $u_i \in U_{h+1}^\top$  are contained in  $\lambda$ , so  $T_{h+1}$  satisfies P2.

*Case 3.* Suppose  $u_h$  is the left child of  $u_{h+1}$ . Then, by the definition of the postfix traversal,  $u_{h+1}$  has empty right subtree in  $U$ , and so  $B_{h+1}$  consists of  $u_{h+1}$  with left subtree  $B_h$  and right subtree empty. (For an illustration of this case, see the step from  $T_1$  to  $T_2$  in Example 4.2.) Proceeding in a similar way to the previous cases, one sees that, as in Figure 3,  $T_h = \text{P}_{\text{syIv}}(\beta u_{h+1} \delta \lambda B_h)$ . Let  $T_{h+1} = \text{P}_{\text{syIv}}(\delta \lambda B_h \beta u_{h+1})$ ; then  $T_h \sim T_{h+1}$  and  $T_{h+1}$  satisfies P1 and P2.

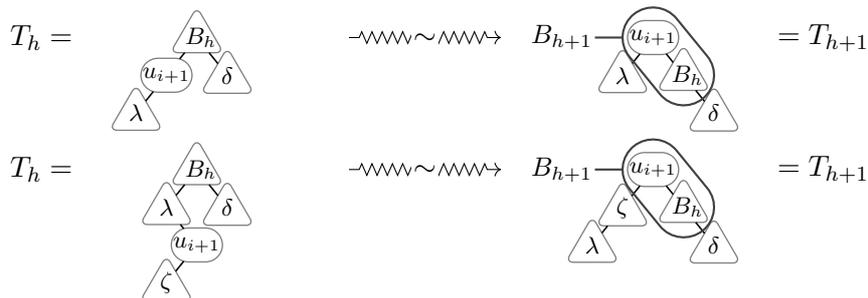


FIGURE 4. Induction step, case 4, two sub-cases.

*Case 4.* Suppose that, in  $U$ , the node  $u_h$  is the right child of  $u_{h+1}$ , and  $u_{h+1}$  has empty left subtree. (For an illustration of this case, see the step from  $T_4$  to  $T_5$  in Example 4.2.) Thus  $B_{h+1}$  consists of the node  $u_{h+1}$  with empty left subtree and right subtree  $U_h$ . Proceeding in a similar way to the previous cases, one sees that there are two sub-cases, as in Figure 4:

- (1)  $T_h = \text{P}_{\text{sylv}}(\lambda u_{h+1} \delta B_h)$ . Let  $T_{h+1} = \text{P}_{\text{sylv}}(\delta B_h \lambda u_{h+1})$ .
- (2)  $T_h = \text{P}_{\text{sylv}}(\zeta u_{h+1} \lambda \delta B_h)$ . Let  $T_{h+1} = \text{P}_{\text{sylv}}(\lambda \delta B_h \zeta u_{h+1})$ .

In both sub-cases,  $T_h \sim T_{h+1}$  and  $T_{h+1}$  satisfies P1 and P2.

*Conclusion.* Thus there is a sequence  $T = T_0, T_1, \dots, T_n = U$  with  $T_h \sim T_{h+1}$  and  $T_{h+1}$  satisfying P1 and P2 for  $h = 0, \dots, h-1$ . In particular,  $T_n$  satisfies P1 and so the subtree  $B_n = U$  appears in  $T_n$ , with its root at the root of  $T_n$ . Hence  $T_n = U$ .  $\square$

## REFERENCES

- [CM] A. J. Cain & A. Malheiro. ‘Combinatorics of cyclic shifts in plactic, hypoplactic, sylvester, and related monoids’. In preparation.
- [CM13] C. Choffrut & R. Mercas. ‘The lexicographic cross-section of the plactic monoid is regular’. In J. Karhumäki, A. Lepistö, & L. Zamboni, eds, *Combinatorics on Words*, no. 8079 in *Lecture Notes in Comput. Sci.*, pp. 83–94. Springer, 2013. DOI: 10.1007/978-3-642-40579-2\_11.
- [CM15] A. J. Cain & A. Malheiro. ‘Deciding conjugacy in sylvester monoids and other homogeneous monoids’. *Internat. J. Algebra Comput.*, 25, no. 5 (2015). DOI: 10.1142/S0218196715500241.
- [Gir12] S. Giraud. ‘Algebraic and combinatorial structures on pairs of twin binary trees’. *J. Algebra*, 360 (2012), pp. 115–157. DOI: 10.1016/j.jalgebra.2012.03.020.
- [HNT05] F. Hivert, J.-C. Novelli, & J.-Y. Thibon. ‘The algebra of binary search trees’. *Theoret. Comput. Sci.*, 339, no. 1 (2005), pp. 129–165. DOI: 10.1016/j.tcs.2005.01.012.
- [HNT07] F. Hivert, J.-C. Novelli, & J.-Y. Thibon. ‘Commutative combinatorial Hopf algebras’. *J. Algebraic Combin.*, 28, no. 1 (2007), pp. 65–95. DOI: 10.1007/s10801-007-0077-0.
- [KT97] D. Krob & J.-Y. Thibon. ‘Noncommutative Symmetric Functions IV: Quantum Linear Groups and Hecke Algebras at  $q = 0$ ’. *J. Algebraic Combin.*, 6, no. 4 (1997), pp. 339–376. DOI: 10.1023/A:1008673127310.
- [Lot02] M. Lothaire. *Algebraic Combinatorics on Words*. No. 90 in *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 2002.
- [LS81] A. Lascoux & M.-P. Schützenberger. ‘Le monoïde plaxique’. In *Noncommutative structures in algebra and geometric combinatorics*, no. 109 in *Quaderni de "La Ricerca Scientifica"*, pp. 129–156, Rome, 1981. CNR. URL: <http://igm.univ-mlv.fr/~berstel/Mps/Travaux/A/1981-1PlaxiqueNaples.pdf>.

- [Nov00] J.-C. Novelli. ‘On the hypoplactic monoid’. *Discrete Mathematics*, 217, no. 1–3 (2000), pp. 315–336. DOI: 10.1016/S0012-365X(99)00270-8.
- [Pri13] J.-B. Priez. ‘A lattice of combinatorial Hopf algebras: Binary trees with multiplicities’. In *Formal Power Series and Algebraic Combinatorics*, Nancy, 2013. The Association. Discrete Mathematics & Theoretical Computer Science. URL: <http://www.dmtcs.org/pfpapers/dmAS0196.pdf>.

CENTRO DE MATEMÁTICA E APLICAÇÕES, FACULDADE DE CIÊNCIAS E TECNOLOGIA,  
UNIVERSIDADE NOVA DE LISBOA, 2829–516 CAPARICA, PORTUGAL  
*E-mail address:* `a.cain@fct.unl.pt`

DEPARTAMENTO DE MATEMÁTICA & CENTRO DE MATEMÁTICA E APLICAÇÕES, FAC-  
ULDADE DE CIÊNCIAS E TECNOLOGIA, UNIVERSIDADE NOVA DE LISBOA, 2829–516 CA-  
PARICA, PORTUGAL  
*E-mail address:* `ajm@fct.unl.pt`