



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/79310/>

Version: Accepted Version

---

**Article:**

Marsh, RJ and Schroll, S (2013) A circular order on edge-coloured trees and RNA m-diagrams. *Advances in Applied Mathematics*, 54 (1). 11 - 26. ISSN: 0196-8858

<https://doi.org/10.1016/j.aam.2013.09.009>

---

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# A CIRCULAR ORDER ON EDGE-COLOURED TREES AND RNA $m$ -DIAGRAMS

ROBERT J. MARSH AND SIBYLLE SCHROLL

ABSTRACT. We study a circular order on labelled,  $m$ -edge-coloured trees with  $k$  vertices, and show that the set of such trees with a fixed circular order is in bijection with the set of RNA  $m$ -diagrams of degree  $k$ , combinatorial objects which can be regarded as RNA secondary structures of a certain kind. We enumerate these sets and show that the set of trees with a fixed circular order can be characterized as an equivalence class for the transitive closure of an operation which, in the case  $m = 3$ , arises as an induction in the context of interval exchange transformations.

## 1. INTRODUCTION

Interval exchange transformations are, roughly speaking, a generalization of a rotation of a circle. More precisely, given a partition of the unit interval into smaller segments, an interval exchange transformation is a map from the unit interval to itself where the segments are rearranged according to a choice of permutation. Recently there has been particular interest in the class of interval exchange transformations induced by the permutation  $(k \ k - 1 \ \cdots \ 1)$ . In [2] a combinatorial interpretation in terms of an induction process on labelled trees with edges coloured with 3 possible colours, called trees of relations, has proven a fruitful tool leading to new results in the understanding of the associated languages. In this paper we extend this combinatorics by defining a generalization of the induction process to the case of labelled trees with edges coloured with  $m$  possible colours and study its properties.

Given positive integers  $k$  and  $m$ , an  $m$ -edge-coloured tree with  $k$  vertices is a tree whose edges are labelled with  $m$  possible colours in such a way that no vertex is incident with two edges of the same colour. It is said to be *labelled* if its vertices are labelled with  $\{1, 2, \dots, k\}$ . Such trees were shown to possess a circular order in the case  $m = 3$  in [2] corresponding to the permutation of the interval exchange transformation. By giving a new proof of this fact, we show that the circular order can be generalized to arbitrary  $m$  and we study its relation to the generalized induction process we define.

Firstly, we show that the set  $\mathcal{T}_{k,m}$  of labelled  $m$ -edge-coloured trees with  $k$  vertices and a fixed circular order is in bijection with a collection of combinatorial objects consisting of collections of noncrossing arcs in a disk with coloured marked points on its boundary, satisfying a certain matching rule. Such diagrams can be regarded as generalized RNA secondary structures of a certain kind, in the sense of [13], drawn in the Nussinov circle representation [10], so we refer to them as *RNA  $m$ -diagrams of degree  $k$* . Such a bijection was given for the case  $m = 3$  in [2].

---

*Date:* 4 September 2013.

*Key words and phrases.* tree, edge-coloured, labelled triangulation, Fuss-Catalan number, enumeration, tree of relations, RNA secondary structure, polygon,  $m$ -angulation, induction, interval exchange transformation, RNA  $m$ -diagram.

*2010 Mathematics Subject Classification:* Primary: 05C05, 05A15; Secondary: 37B10.

This work was supported by the Engineering and Physical Sciences Research Council [grant number EP/G007497/1], the Mathematical Sciences Research Institute (MSRI) in Berkeley, California, and by the Leverhulme Trust in the form of an Early Career Fellowship for Sibylle Schroll.

We compute the cardinality of  $\mathcal{T}_{k,m}$  (and thus also the number of RNA  $m$ -diagrams of degree  $k$ ) using the well-known correspondence between trees and  $m$ -angulations of a polygon which induces a bijection with a collection of appropriately labelled  $m$ -angulations

We prove the existence of a transformation of labelled  $m$ -edge-coloured trees preserving the circular order. We call this transformation *generalized induction* as it generalizes the induction in the case  $m = 3$ , studied combinatorially in [2]. In [4] it was shown that this induction (for  $m = 3$ ) gives rise to languages naturally generalizing Sturmian languages and arising from interval exchange transformations. We also give an example showing that the most straightforward generalization of this induction to the case  $m > 3$  does not in general preserve the circular order.

We go on to show that the equivalence classes of the transitive closure of the generalized induction are characterized by the circular order. Thus the set of equivalence classes is in bijection with the set of  $k$ -cycles in the symmetric group of degree  $k$ .

Finally, motivated by snake-triangulations and  $m$ -snakes in cluster theory, we give an interpretation of the generalized induction process in terms of labelled  $m$ -angulations of polygons.

This article is structured as follows. In Section 2, we introduce the combinatorial objects considered in the paper. In Section 3 we give the bijections mentioned above and compute the cardinality of the set of labelled  $m$ -edge-coloured trees with a fixed circular order. In Section 4 we show that the circular order characterizes the equivalence classes of the transitive closure of the generalized induction and in section 5 we show that induction can be given by a composition of flips of diagonals in an  $m$ -angulation.

## 2. SOME COMBINATORIAL OBJECTS

We first introduce the main combinatorial objects we will be considering: edge-coloured trees, RNA  $m$ -diagrams, and  $m$ -angulations.

Given positive integers  $k, m$ , we consider  $m$ -edge-coloured trees on  $k$  vertices, i.e. trees whose edges are coloured with one of the  $m$  symbols  $S_1, S_2, \dots, S_m$  in such a way that no two edges incident with the same vertex have the same colour. We say that such a tree is *labelled* if its vertices are labelled with  $\{1, 2, \dots, k\}$ . As usual, if the connectedness assumption is not satisfied, we refer to the corresponding objects as forests. See Figure 3(b) for an example of a (rooted)  $m$ -edge-coloured tree and (c) for an example of a labelled  $m$ -edge-coloured tree.

Given an  $m$ -edge coloured tree  $\mathcal{S}$  with  $k$  vertices, each symbol  $S_r$  determines a map (with the same name) from the set of vertices of  $\mathcal{S}$  to itself. A vertex is fixed by  $S_r$  unless it is incident with an edge coloured  $S_r$ , in which case it is sent to the vertex at the other end of that edge. Let  $\sigma_{\mathcal{S}}$  be the composition  $S_m S_{m-1} \cdots S_1$ : this is a permutation of the vertices of  $\mathcal{S}$ . We use the same definition for a labelled  $m$ -edge coloured tree,  $G$ , obtaining a permutation  $\sigma_G$  in the symmetric group of degree  $k$ . We refer to  $\sigma_{\mathcal{S}}$  (respectively,  $\sigma_G$ ) as the *circular order* of  $\mathcal{S}$  (respectively,  $G$ ). In the 3-symbol case these maps were considered in [2]. The circular order is always a  $k$ -cycle:

**Lemma 2.1.** *Let  $G$  be a labelled  $m$ -edge-coloured tree with  $k$  vertices. Then the circular order  $\sigma_G$  of  $G$  is a  $k$ -cycle.*

*Proof.* This is clearly true if  $k = 1$  or  $2$ . Suppose it holds for smaller values of  $k$ . Let  $v$  be a vertex of  $G$  which is not a leaf. Suppose that  $v$  is incident with edges  $e_1, e_2, \dots, e_d$  in  $G$ , coloured with  $S_{r_1}, S_{r_2}, \dots, S_{r_d}$  respectively, where  $r_1 < r_2 < \dots < r_d$ . Let the end-points of these edges (other than  $v$ ) be  $v_1, v_2, \dots, v_d$ . Removing  $v$  from  $G$  leaves  $d$  subtrees  $G'_1, G'_2, \dots, G'_d$  incident with  $v_1, v_2, \dots, v_d$  respectively. Let  $G_i$  be the subtree  $G'_i$  with  $v$  and  $e_i$  reattached to  $v_i$ .

By the inductive hypothesis the  $\sigma_{G_i}$  are all cycles. Hence, repeatedly applying  $\sigma_{G_1}$  to  $v$  cycles through the vertices of  $G_1$ . Since  $\sigma_G = \sigma_{G_1}$  on all vertices of  $G_1$  except  $w = \sigma_{G_1}^{-1}(v)$ , repeatedly applying  $\sigma_G$  also cycles through all the vertices of  $G_1$ . Since  $r_2$  is minimal such that  $S_{r_2}$  is a symbol colouring an edge incident with  $v$  with  $r_2 > r_1$ ,  $\sigma_G(w)$  will lie in  $G'_2$ . In fact  $\sigma_G(w) = \sigma_{G_2}(v)$ , since in  $G_2$ ,  $v$  is not incident with any edge with symbol other than  $S_{r_2}$ . Repeatedly applying  $\sigma_G$  then cycles through the vertices of  $G'_2$  before coming back to  $\sigma_{G'_2}^{-1}(v)$ . Repeating this argument, we see that repeatedly applying  $\sigma_G$  to  $v$  first cycles through  $G'_1$ , then through  $G'_2, G'_3, \dots, G'_d$  in order before eventually returning to  $v$ .  $\square$

Next, we define an *RNA  $m$ -diagram of degree  $k$*  as follows. We take  $k$  vertices  $1, 2, \dots, k$ , numbered clockwise around a circle. Each vertex contains the symbols  $S_1, S_2, \dots, S_m$  written in clockwise order. A collection of arcs connect equal symbols at different vertices. A symbol at a vertex can be incident to at most one arc. Such a diagram is *noncrossing* if it can be drawn in such a way that there are no crossings between the arcs. We say that it is *connected* provided there is a path between any two vertices (where moving between symbols at a vertex is allowed). See Figure 3(a) for an example of a connected noncrossing RNA 4-diagram of degree 10.

Diagrams of this kind (but with different rules) have been considered in [1] in the context of Fuss-Catalan algebras. Such diagrams can be regarded as a certain kind of abstract RNA secondary structures in the sense of [13] (see also [8, §5]), drawn in the Nussinov circle representation [10].

Finally, let  $P$  be an  $(m - 2)k + 2$ -sided regular polygon. An  $m$ -angulation of  $P$  is a collection of diagonals dividing  $P$  into  $k$   $m$ -gons. We say that an  $m$ -angulation of  $P$  is *diagonal-coloured* if each diagonal in  $P$  is coloured with a symbol from the set  $\{S_1, S_2, \dots, S_m\}$  in such a way that the colours on the sides of each  $m$ -sided polygon in the  $m$ -angulation are  $S_1, S_2, \dots, S_m$  in clockwise order. We say that it is *rooted* if there is a distinguished  $m$ -sided polygon in the  $m$ -angulation. We say that it is  *$m$ -gon-labelled* if the  $m$ -gons are labelled  $1, 2, \dots, k$ .

### 3. BIJECTIONS

Our main aim in this section is to compute the number of connected noncrossing RNA  $m$ -diagrams of degree  $k$  and to give a bijective proof that this is the same as the number of labelled  $m$ -edge-coloured trees with  $k$  vertices and circular order  $(k \ k - 1 \ \dots \ 1)$ .

**Theorem 3.1.** *There is a bijection between the following sets:*

- (I) *The set of degree  $k$  noncrossing RNA  $m$ -diagrams.*
- (II) *The set of labelled  $m$ -edge-coloured forests  $G$  with  $k$  vertices such that, on writing  $G = \sqcup_i G_i$  as a union of connected components, we have the following:*
  - (a) *If  $i \neq j$  and  $a_1, a_2 \in G_i, b_1, b_2 \in G_j$ , we cannot have that  $a_1 > b_1 > a_2 > b_2$ .*
  - (b) *If  $a \in G_i$  for some  $i$ , then  $\sigma_G(a)$  is the maximal vertex of  $G$  less than  $a$  lying in  $G_i$  (or, if no such vertex exists, it is the largest vertex of  $G$  lying in  $G_i$ ).*

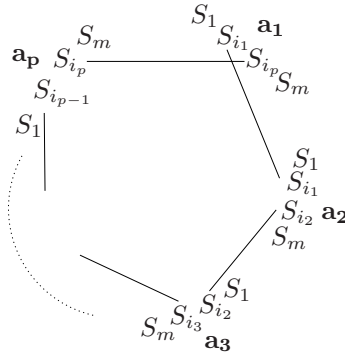
*Proof.* Let  $\Sigma$  be a noncrossing RNA  $m$ -diagram of degree  $k$  on  $m$  symbols as in (I).

Let  $G$  be the labelled  $m$ -edge-coloured graph with  $k$  vertices whose edges given by the arcs of  $\Sigma$ . That is, there is an edge between vertices  $i$  and  $j$  of  $G$  coloured with  $S_k$  if and only if there is an arc in  $\Sigma$  between the instances of the symbol  $S_k$  in vertices  $i$  and  $j$  in  $\Sigma$ .

**Claim:**  $G$  is a forest.

We prove the claim. Suppose, for a contradiction, there is a cycle

$$a_1 \xrightarrow{S_{i_1}} a_2 \xrightarrow{S_{i_2}} \dots \quad a_{p-1} \xrightarrow{S_{i_{p-1}}} a_p \xrightarrow{S_{i_p}} a_1$$

FIGURE 1. A cycle in  $G$  leads to a crossing

in  $G$ , and thus a corresponding cycle in  $\Sigma$ . Without loss of generality, we may assume that  $i_1 < i_2$ . For vertices  $a, b$ , we denote by  $(a, b)$  the set of vertices  $c$  of  $\Sigma$  lying strictly clockwise of  $a$  and strictly anticlockwise of  $b$ .

Then  $a_3 \in (a_2, a_1)$ , since the arc in  $\Sigma$  corresponding to the edge in  $G$  between  $a_2$  and  $a_3$  cannot cross the arc in  $\Sigma$  corresponding to the edge in  $G$  between  $a_1$  and  $a_2$ .

By assumption on  $\Sigma$ ,  $i_2 \neq i_3$ . If  $i_2 > i_3$ , there can be no path in  $\Sigma$  from the symbol  $S_{i_3}$  in vertex  $a_3$  of  $\Sigma$  back to vertex  $a_1$  of  $\Sigma$  without crossings, a contradiction, hence  $i_2 < i_3$ .

Repeating this argument, we see that, moving clockwise on  $\Sigma$  from  $a_1$  we meet vertices  $a_2, a_3, \dots, a_p$  in order before returning to  $a_1$ , and that  $i_1 < i_2 < \dots < i_p$ . But then the arc between  $a_p$  and  $a_1$  (on symbol  $S_{i_p}$  crosses the arc between  $a_1$  and  $a_2$  (on symbol  $S_{i_1}$ ), since  $i_1 < i_p$ ; see Figure 1. Hence  $G$  has no cycles, and must be a tree. The claim is shown.

We next prove that (a) holds. Suppose that  $i \neq j$ ,  $a_1, a_2 \in G_i$ ,  $b_1, b_2 \in G_j$ , and  $a_1 > b_1 > a_2 > b_2$ . Then  $a_1, b_1, a_2, b_2$  follow each other anticlockwise around the circle. Then, since  $a_1, a_2$  are in the same connected component of  $G$ , there is a path in  $\Sigma$  from  $a_1$  to  $a_2$ , and similarly from  $b_1$  to  $b_2$ . The arrangement of  $a_1, a_2, b_1$  and  $b_2$  implies that these two paths cross, a contradiction. Hence no such arrangement can occur, and (a) is shown.

We next prove that (b) holds. It is enough to prove the following claim:

**Claim:** Let  $a \in G_i$ . Then  $\sigma_G(a)$  is the next vertex of  $G_i$  (considered as a vertex of  $\Sigma$ ) anticlockwise on the circle from  $a$ .

To prove the claim, we note that, by the definition of  $\sigma_G$ ,  $\sigma_G(a)$  is connected to  $a$  by a path in  $G$ :

$$a = a_1 \xrightarrow{S_{i_1}} a_2 \xrightarrow{S_{i_2}} \dots \quad \dots \xrightarrow{S_{i_{p-1}}} a_p = \sigma_G(a),$$

where  $i_1 < i_2 < \dots < i_{p-1}$ . Furthermore,  $a_1$  is not incident with any arc with symbol  $S_r$  for  $r < i_1$ ,  $a_p$  is not incident with any arc with symbol  $S_r$  for  $r > i_{p-1}$ , and, for  $2 \leq j \leq p-1$ ,  $a_j$  is not incident with any symbol  $S_r$  for  $i_{j-1} < r < i_j$ .

The existence of the above path implies that  $\sigma_G(a)$  lies in  $G_i$ . Since  $i_1 < i_2 < \dots < i_p$  and there are no crossings, the path must go clockwise around the circle; see Figure 2. The conditions above and the fact there are no crossings imply that no vertex in  $(\sigma_G(a), a)$  has an arc with a vertex in  $(a, \sigma_G(a))$  or with  $a$  or  $\sigma_G(a)$ , so these vertices are connected only amongst themselves. It follows that they do not lie in  $G_i$  and we see that (II)(b) holds. Thus  $G$  is a labelled  $m$ -edge-coloured forest satisfying (II).

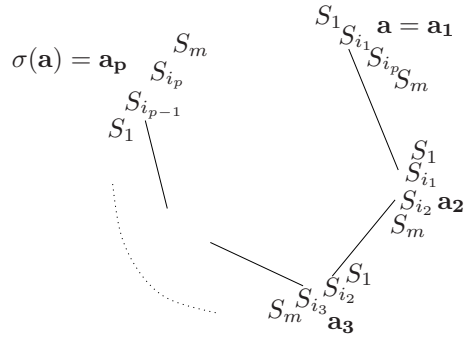


FIGURE 2. The part of  $\Sigma$  between  $a$  and  $\sigma_G(a)$ .

Conversely, suppose that we have a labelled  $m$ -edge-coloured forest with  $k$  vertices satisfying (II). Let  $\Sigma$  be the RNA  $m$ -diagram of degree  $k$  with an arc joining  $S_r$  in vertex  $i$  with  $S_r$  in vertex  $j$  if and only if there is an edge in  $G$  between vertices  $i$  and  $j$  coloured with symbol  $S_r$ . We must check that  $\Sigma$  can be drawn with no crossing arcs, i.e. that it is noncrossing.

We do this by induction on the number of vertices. Suppose first that  $G$  has more than one connected component, i.e. that  $G$  is not connected. By induction, each component  $G_i$  corresponds to a noncrossing RNA  $m$ -diagram (on the vertices of  $G_i$ ).

Suppose that we had  $a_1 > a_2 \in G_i$  and  $b_1 > b_2 \in G_j$  for two distinct components  $G_i$  and  $G_j$ , with arcs between  $a_1$  and  $a_2$  and  $b_1$  and  $b_2$  which cross in  $\Sigma$ . Then, going around the circle anticlockwise, starting at vertex  $k$ , we must encounter  $a_1, b_1, a_2, b_2$  in order, or  $b_1, a_1, b_2, a_2$  in order. Swapping  $G_i$  and  $G_j$  if necessary, we can assume we are in the first case. But then  $a_1 > b_1 > a_2 > b_2$ , contradicting (II)(a). Hence  $\Sigma$  is noncrossing.

So we are reduced to the case in which  $G$  has exactly one connected component, i.e.  $G$  is connected. Suppose that vertex  $k$  is incident with edges  $e_1, e_2, \dots, e_d$  in  $G$ , coloured with symbols  $S_{r_1}, S_{r_2}, \dots, S_{r_d}$  where  $r_1 < r_2 < \dots < r_d$ . Let the endpoints of these edges (other than  $k$ ) be  $v_1, v_2, \dots, v_d$ . Removing vertex  $k$  from  $G$  leaves precisely  $d$  trees  $T_1, T_2, \dots, T_d$  containing vertices  $v_1, v_2, \dots, v_d$  respectively. By (b), we know that  $\sigma_G = S_m S_{m-1} \dots S_1$  induces the permutation  $(k \ k-1 \ \dots \ 1)$  on the vertices of  $G$ .

We apply  $\sigma_G$  to vertex  $k$ , and then repeatedly apply  $\sigma_G$ . By its definition, each application of  $\sigma_G$  corresponds to following a certain path through  $G$ , i.e. passing along the edges corresponding to the symbols in the sequence  $S_1, S_2, \dots, S_m$  in that order, when such incident edges exist. Since the edge  $e_1$  has symbol  $S_{r_1}$ , and no edge incident with  $k$  has smaller symbol, it follows that, after the first application of  $\sigma_G$ , we obtain vertex  $k_1 := k - 1$  in  $T_1$ .

Since  $\sigma_G$  is a  $k$ -cycle, after repeated application of  $\sigma_G$ , we must leave  $T_1$ . Suppose that  $k_2$  is the number of the first vertex reached outside  $T_1$ . Since  $r_2$  is the minimum number of a symbol adjacent to  $k$  greater than  $r_1$ ,  $k_2$  will lie in  $T_2$ . Repeating this argument, we will obtain  $k > k_1 > k_2 > \dots > k_d \geq 1$  such that vertices  $k_{i+1} + 1, \dots, k_i$  lie in tree  $T_i$  for  $i = 1, 2, \dots, d-1$ . At the final step, the first vertex reached on leaving  $T_d$  must be  $k$ . Since  $\sigma_G$  is a  $k$ -cycle, all vertices must have been visited.

Let  $k_{d+1} = 0$ . It follows from the above that tree  $T_i$  contains precisely vertices  $k_{i+1} + 1, \dots, k_i$  for each  $i$ . Thus, the numbering of the vertices of  $G$  is first the vertices of  $T_d$  in some order, then the vertices of  $T_{d-1}$  in some order, then the vertices of  $T_{d-2}$ , and so on, ending with the vertices of  $T_1$  and then finally  $k$ . Each  $T_i$  will correspond (by the inductive hypothesis) to a noncrossing RNA  $m$ -diagram on its vertices. Thus the vertices  $v_1, v_2, \dots, v_d$  in  $G$  will be numbered in decreasing order. The arcs in  $\Sigma$

from  $k$  to these vertices are numbered by symbols  $S_{r_1}, S_{r_2}, \dots, S_{r_d}$ , respectively, with  $r_1 < r_2 < \dots < r_d$ . It follows these arcs do not cross each other or any of the other arcs in  $\Sigma$ . See Figure 3(a) for an example, where  $v_1 = 9$ ,  $r_1 = 1$ , and  $v_2 = 8$ ,  $r_2 = 4$ . Hence,  $\Sigma$  is noncrossing and thus an object in (I) as required.

It is clear that the two maps we have constructed are inverse to each other, so the theorem is proved.  $\square$

The following lemma follows easily from the definitions.

**Lemma 3.2.** *A noncrossing RNA  $m$ -diagram is connected if and only if the corresponding labelled  $m$ -edge-coloured forest is connected, i.e. is a tree.*

**Remark 3.3.** (1) Since a forest on  $k$  vertices is connected if and only if it has exactly  $k - 1$  edges, a noncrossing RNA  $m$ -diagram with  $k$  vertices is connected if and only if it has  $k - 1$  arcs.

(2) In the connected case, the circular order of  $G$  is just the permutation  $(k \ k - 1 \ \dots \ 1)$  and we have a bijection between the following sets:

The set of connected noncrossing RNA  $m$ -diagrams of degree  $k$ .

$\downarrow$

The set of labelled  $m$ -edge-coloured trees with  $k$  vertices and circular order  $(k \ k - 1 \ \dots \ 1)$ .

(3) The labelling on the vertices for a tree in the latter set in (2) is determined by a distinguished vertex, that labelled  $k$ , say, since  $\sigma_G$  then determines the labels on all the other vertices.

By Remark 3.3 and Lemma 2.1 we have:

**Corollary 3.4.** *The bijection in Theorem 3.1 induces a bijection between the following sets:*

- (a) *The set of connected noncrossing RNA  $m$ -diagrams of degree  $k$ .*
- (b) *The set of rooted  $m$ -edge-coloured trees with  $k$  vertices.*
- (c) *The set of labelled  $m$ -edge-coloured trees with  $k$  vertices and circular order  $(k \ k - 1 \ \dots \ 1)$ .*

See Figure 3(a)-(c) for an illustration of this bijection. The following is well-known; see, for example, [12, Sect. 6.2] (adding the labelling is straightforward).

**Theorem 3.5.** *There is a bijection between the following sets:*

- (a) *The set of  $m$ -edge-coloured trees with  $k$  vertices.*
- (b) *The set of diagonal-coloured  $m$ -angulations of an  $(m - 2)k + 2$ -sided regular polygon up to rotation.*

**Corollary 3.6.** *There is a bijection between the following sets:*

- (a) *The set of labelled  $m$ -edge-coloured trees with  $k$  vertices and circular order  $(k \ k - 1 \ \dots \ 1)$ .*
- (b) *The set of rooted diagonal-coloured  $m$ -angulations of an  $(m - 2)k + 2$ -sided regular polygon up to rotation.*

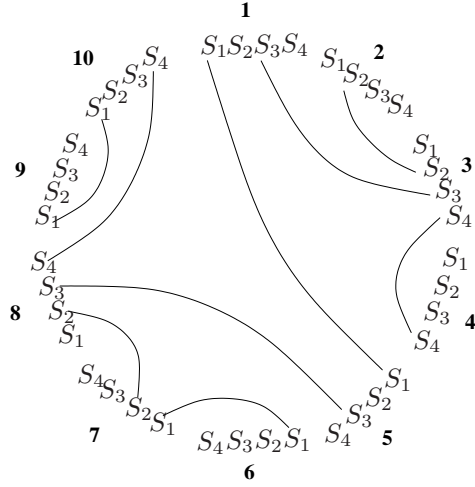
*Proof.* This follows from Theorem 3.5 and Corollary 3.4.  $\square$

**Corollary 3.7.** *The cardinality of each of the following sets:*

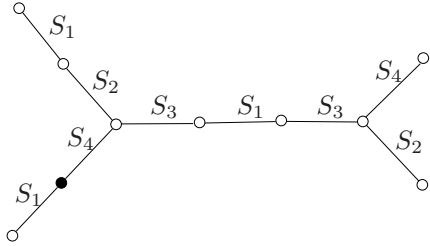
- (a) *The set of labelled  $m$ -edge-coloured trees with  $k$  vertices and circular order  $(k \ k - 1 \ \dots \ 1)$ ;*
- (b) *The set of connected noncrossing RNA  $m$ -diagrams of degree  $k$ ;*

*is given by*

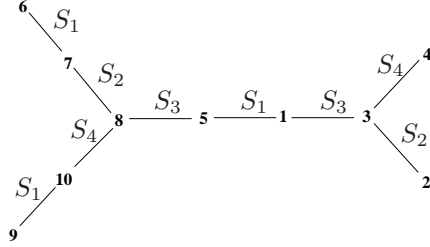
$$T_{k,m} = \frac{m}{(m-2)k+2} \binom{(m-1)k}{k-1}.$$



(a) A connected noncrossing RNA 4-diagram of degree 10.



(b) A rooted 4-edge-coloured tree with 10 vertices.



(c) A labelled 4-edge-coloured tree with 10 vertices and circular order (10 9 ... 1).

FIGURE 3. Objects corresponding to each other under the bijections in Corollary 3.4.

*Proof.* Both sets have the same cardinality by Corollary 3.4 and, by Corollary 3.6, they have the same cardinality as the set of rooted diagonal-coloured  $m$ -angulations of an  $(m - 2)k + 2$ -sided regular polygon up to rotation. The number  $S_{k,m}$  of such  $m$ -angulations without a root, with no labelling of diagonals and ignoring rotational equivalence is well-known (see e.g. [7]). Let  $C_k^m$  be the  $k$ th Fuss-Catalan number of degree  $m$ :

$$C_k^m = \frac{1}{k} \binom{mk}{k-1} = \frac{1}{(m-1)k+1} \binom{mk}{k}.$$

Then

$$S_{k,m} = C_k^{m-1} = \frac{1}{(m-2)k+1} \binom{(m-1)k}{k}.$$

Since there are  $k$   $m$ -sided polygons in an  $m$ -angulation, there are  $k$  possibilities for the root. There are  $m$  possibilities for a labelling since once one diagonal is coloured, all other diagonals in the  $m$ -angulation have determined colours using the rule that each  $m$ -gon must have its edges coloured  $S_1, S_2, \dots, S_m$  clockwise around the boundary. Each orbit of diagonal-coloured rooted  $m$ -angulations under the action of the rotation group of the polygon contains  $(m - 2)k + 2$  elements (the number of sides of  $P$ ). Hence, we have:

$$T_{k,m} = \frac{kmS_{k,m}}{(m-2)k+2}$$

and the result follows.  $\square$

It is interesting to note that by [9, §3], the sequence  $T_{1,m}, T_{2,m}, \dots$  is the  $m$ -fold convolution of the sequence  $S_{0,m}, S_{1,m}, \dots$

**Example 3.8.** For  $m = 3, 4, 5, 6$ , the first few values of  $T_{k,m}$  are given in the following table:

$k$	0	1	2	3	4	5	6
$T_{k,3}$	1	1	3	9	28	90	297
$T_{k,4}$	1	1	4	18	88	455	2448
$T_{k,5}$	1	1	5	30	200	1425	10626
$T_{k,6}$	1	1	6	45	380	3450	32886

The cases  $m = 3, 4$  are sequences A071724 and A006229, respectively, in [11]; the cases  $m = 5, 6$  do not appear. The case  $m = 3$  appears in [2, Prop. 7.5].

**Corollary 3.9.** *The total number of labelled  $m$ -edge-coloured trees with  $k$  vertices is:*

$$U_{k,m} = \frac{m((m-1)k)!}{((m-2)k+2)!}$$

*Proof.* By Lemma 2.1, the circular order of any labelled  $m$ -edge-coloured tree with  $k$  vertices is a  $k$ -cycle. By Corollary 3.7 the number of such trees with a given circular order is  $T_{k,m}$ . Since any  $k$ -cycle can arise, we have  $U_{k,m} = \frac{1}{(k-1)!} T_{k,m}$ , giving the result.  $\square$

Note that this result is already known [3], [12, 5.28, p124].

#### 4. GENERALIZED INDUCTION

In this section we give the definition of a generalized induction on labelled  $m$ -edge-coloured trees with  $k$  vertices. Generalized induction generates new labelled  $m$ -edge-coloured trees with the same number of vertices starting with a given such tree, and the transitive closure is an equivalence relation. We show that the circular order is an invariant, giving rise to a classification of the equivalence classes by  $k$ -cycles in the symmetric group of degree  $k$ .

Given a labelled  $m$ -edge-coloured tree  $G$  and integers  $i, j \in \{1, \dots, m\}$  we define a maximal  $S_i - S_j$  chain  $B$  in  $G$  to be a (linear) subtree of  $G$  whose edges are only coloured  $S_i$  and  $S_j$  such that no other edges incident to  $B$  are coloured by  $S_i$  or  $S_j$ .

**Definition 4.1.** Let  $G$  be a labelled  $m$ -edge-coloured tree with  $k$  vertices. Fix  $i, j \in \{1, \dots, m\}$  with  $i < j$ . Let  $B$  be a maximal  $S_i - S_j$  chain in  $G$ . Define  $R_{i,j}^B(G)$  to be the labelled  $m$ -edge-coloured tree with  $k$  vertices obtained from  $G$  by

- first removing all subtrees in the complement of the maximal chain  $B$
- interchanging the vertices of each edge of  $B$  coloured by  $S_j$
- interchanging the symbols  $S_i$  and  $S_j$  on the whole maximal chain  $B$
- reattaching the previously removed subtrees to  $B$  at the vertices with the same label they were removed from.

Similarly, define  $L_{i,j}^B(G)$ , where in the second bullet point in the above definition we interchange the vertices of each edge coloured by  $S_i$  rather than those coloured by  $S_j$ . We also set  $R_i^B := R_{i,i+1}^B$  and  $L_i^B := L_{i,i+1}^B$  and we will write  $R_i$  and  $L_i$  if  $B$  is clear from the context.

**Remark 4.2.** (1) Induction can also be defined on  $m$ -edge-coloured trees: For an  $m$ -edge-coloured tree with  $k$  vertices, choose an arbitrary vertex-labelling, apply induction, and then remove the vertex labelling. It is clear that this is independent of the vertex-labelling chosen.

- (2) The inductions  $R_{i,j}^B$  and  $L_{i,j}^B$  are mutually inverse maps.

**Lemma 4.3.** *Let  $i, j \in \{1, \dots, m\}$  with  $i < j$  and  $B$  be a maximal  $S_i - S_j$ -chain with no incident edges coloured by  $S_{i+1}, \dots, S_{j-1}$ . Then we have*

$$\begin{aligned} R_{i,j}^B(G) &= L_i L_{i+1} \cdots L_{j-2} R_{j-1} R_{j-2} R_{j-3} \cdots R_i \\ &= R_{j-1} R_{j-2} \cdots R_{i+1} R_i L_{i+1} L_{i+2} \cdots L_{j-1}; \\ L_{i,j}^B(G) &= L_i L_{i+1} \cdots L_{j-2} L_{j-1} R_{j-2} R_{j-3} \cdots R_i \\ &= R_{j-1} R_{j-2} \cdots R_{i+1} L_i L_{i+1} L_{i+2} \cdots L_{j-1}, \end{aligned}$$

where, in each case, the inductions of form  $R_p$  and  $L_p$  are applied to all maximal chains contained in  $B$ . In particular, the induction  $R_{i,j}^B(G)$  can be written as a product of inductions of the form  $R_p$  or  $L_p$  for  $p = i, i+1, \dots, j-1$ .

*Proof.* Suppose first that  $B$  has the following form:

$$a_1 \xrightarrow{S_i} a_2 \xrightarrow{S_j} a_3 \xrightarrow{S_i} a_4 \quad \cdots \quad a_{r-1} \xrightarrow{S_i} a_r .$$

Then, in  $R_{i,j}^B(G)$ ,  $B'$  becomes the chain

$$B' = a_1 \xrightarrow{S_j} a_3 \xrightarrow{S_i} a_2 \xrightarrow{S_j} a_5 \xrightarrow{S_i} a_4 \quad \cdots \quad a_{r-2} \xrightarrow{S_j} a_r .$$

On the other hand, applying  $R_{j-2} R_{j-3} \cdots R_{i+1} R_i$  (with each induction applying to all the maximal chains of the appropriate type contained in  $B$ ), we obtain the maximal  $S_{j-1} - S_j$ -chain

$$a_1 \xrightarrow{S_{j-1}} a_2 \xrightarrow{S_j} a_3 \xrightarrow{S_{j-1}} a_4 \quad \cdots \quad a_{r-1} \xrightarrow{S_{j-1}} a_r .$$

Next, apply  $R_{j-1}^B$  to get the maximal  $S_{j-1} - S_j$  chain:

$$a_1 \xrightarrow{S_j} a_3 \xrightarrow{S_{j-1}} a_2 \xrightarrow{S_j} a_5 \quad \cdots \quad a_{r-2} \xrightarrow{S_j} a_r .$$

It is then clear that subsequently applying  $L_i L_{i+1} \cdots L_{j-2}$  (in each case applying the induction to all the maximal chains of appropriate type contained in the full subgraph on the vertices  $a_1, a_2, \dots, a_r$ ) gives the chain  $B'$ .

The proof works in a similar way for the other configurations of maximal  $S_i - S_j$ -chains, i.e. for the above case with  $S_i$  and  $S_j$  switched and for the chains of the form

$$a_1 \xrightarrow{S_i} a_2 \xrightarrow{S_j} a_3 \xrightarrow{S_i} a_4 \quad \cdots \quad a_{r-1} \xrightarrow{S_j} a_r$$

and

$$a_1 \xrightarrow{S_j} a_2 \xrightarrow{S_i} a_3 \xrightarrow{S_j} a_4 \quad \cdots \quad a_{r-1} \xrightarrow{S_j} a_r .$$

The other identities are proved similarly.  $\square$

**Definition 4.4.** Call two labelled  $m$ -edge-coloured trees with  $k$  vertices  $G, G'$  *induction equivalent* if there is a sequence of inductions taking  $G$  to  $G'$ , either of the form  $L_i$ , for  $1 \leq i \leq m-1$  or of the form  $R_i$ , for  $1 \leq i \leq m-1$ . Clearly this is a reflexive relation. It is symmetric since  $L_i$  and  $R_i$  are inverse maps and it is easy to see that it is transitive. Hence it is an equivalence relation. We write  $\Gamma(G)$  for the equivalence class containing  $G$ .

We next show that two labelled  $m$ -edge-coloured trees with  $k$  vertices are induction equivalent if and only if they have the same circular order. Firstly, Lemmas 4.5 and 4.6 show that the circular order is invariant under an induction of the form  $R_i$  or  $L_i$ . Key to showing the converse is Proposition 4.8, which shows that every labelled  $m$ -edge coloured tree is induction equivalent to one whose labels are only  $S_1$  and  $S_m$ , which is necessarily a line. This, together with a calculation of the order of induction (Lemma 4.9) allows a counting argument for proving the converse; see Theorem 4.11.



*Proof.* We first show that there exists a labelled  $m$ -edge-coloured tree  $G_2$  with  $k$  vertices that is induction equivalent to  $G$ , none of whose edges is coloured with  $S_2$ , by removing the symbols  $S_2$  one by one. Firstly, remove all edges coloured with symbols  $S_4, \dots, S_m$  and call the resulting labelled 3-edge-coloured tree  $\tilde{G}$ . By [2, Prop. 5.2] there exists a sequence of inductions of the form  $R_1$  and  $L_2$  taking  $\tilde{G}$  to a labelled 3-edge-coloured tree  $\tilde{G}_2$  with no edge coloured  $S_2$ . Let  $G_2$  be the labelled  $m$ -edge-coloured tree  $\tilde{G}_2$  with the detached edges reattached (to the vertices with the same label). Since none of the detached edges are coloured with  $S_1, S_2$ , or  $S_3$  this sequence of inductions also takes  $G$  to  $G_2$  by identifying maximal chains in  $\tilde{G}$  with corresponding maximal chains in  $G$ .

Suppose we have shown that  $G$  is induction equivalent to  $G_{l-1}$ , where  $G_{l-1}$  has no edges coloured  $S_2, \dots, S_{l-1}$ .

Then detach all edges coloured  $S_{l+2}, \dots, S_m$  from  $G_{l-1}$ . Call the resulting labelled  $m$ -edge-coloured tree  $\tilde{G}$ . By [2, Prop. 5.2] there is a sequence of inductions of the form  $R_{1,l}$  and  $L_l = L_{l,l+1}$  taking  $\tilde{G}$  to  $\tilde{G}_l$ , where  $\tilde{G}_l$  has no edges coloured  $S_2, \dots, S_l$ . Reattach the detached edges and call the resulting labelled  $m$ -edge-coloured tree  $G_l$ . Since none of the reattached edges are coloured by  $S_1, S_l$ , or  $S_{l+1}$ , this sequence of inductions takes  $G_{l-1}$  to  $G_l$  by identifying the maximal chains in  $\tilde{G}$  with the maximal chains in  $G_{l-1}$ . By Lemma 4.3, each application of  $R_{1,l}$  can be written as a composition of inductions  $R_p, L_p, p = 1, 2, \dots, l-1$ , since  $G_{l-1}$  has no edges coloured  $S_2, \dots, S_{l-1}$ .

Note that none of the symbols  $S_2, \dots, S_l$  appears in  $G_l$ . Hence, by induction on  $l$ , we can construct  $G_{m-1} = G'$  with no symbols  $S_2, \dots, S_{m-1}$  and a sequence of inductions, each of the form  $R_p$  or  $L_p$ , with  $1 \leq p \leq m-1$ , taking  $G$  to  $G'$ .  $\square$

**Lemma 4.9.** (a) *Let  $k$  be odd, let  $i \neq j$  and let  $G$  be a labelled  $m$ -edge-coloured tree with  $k$  vertices whose edges are coloured with  $S_i$  and  $S_j$  only. Let  $\mathcal{S}$  be the underlying unlabelled tree of  $G$ . Then  $R_{i,j}$  and  $L_{i,j}$  have order  $k$  on  $G$ , producing  $k$  distinct trees with shape  $\mathcal{S}$ .*

(b) *Let  $k$  be even,  $i \neq j$  and let  $G$  be a labelled  $m$ -edge-coloured tree with  $k$  vertices and whose edges are coloured with  $S_i$  and  $S_j$  only. Let  $\mathcal{S}$  be the underlying unlabelled tree of  $G$ . Then  $R_{i,j}$  and  $L_{i,j}$  have order  $k$  on  $G$ , producing  $k/2$  distinct labelled trees whose underlying unlabelled tree is  $\mathcal{S}$  in each case and  $k/2$  distinct labelled trees whose underlying unlabelled tree is  $R_{i,j}(\mathcal{S})$ .*

*Proof.* (a) Since  $\mathcal{S}$  contains only the symbols  $S_i$  and  $S_j$ , it is a line. Suppose the line is drawn horizontally and suppose the leftmost edge has colour  $S_j$  and vertices  $a_1$  and  $a_2$  from left to right. Then in  $(R_{i,j}^G)^d(G)$ , if it is drawn with orientation given by  $\cdot \xrightarrow{S_i} a_1 \xrightarrow{S_j} \cdot$  (where one of the edges may not exist), the vertex  $a_1$  is the  $d^{\text{th}}$  vertex from the left. It is clear that all the induced trees  $(R_{i,j}^G)^d(G)$  have the same underlying unlabelled tree (those for  $d$  odd should be read from right to left).

(b) The proof is similar to the one in (a), except that for  $d$  odd, the underlying unlabelled tree of  $(R_{i,j}^G)^d(G)$  is  $R_{i,j}(\mathcal{S})$ .  $\square$

**Remark 4.10.** Since in the context of Lemma 4.9 the underlying unlabelled tree of  $G$  is a line, we can replace  $R_{i,j}$  or  $L_{i,j}$  in Lemma 4.9 by one of the products in Lemma 4.3.

We finally obtain:

**Theorem 4.11.** *Let  $G$  and  $G'$  be labelled  $m$ -edge-coloured trees with  $k$  vertices. Then  $G'$  is in  $\Gamma(G)$  if and only if  $\sigma_G = \sigma_{G'}$ .*

*Proof.* Suppose that  $G'$  is in  $\Gamma(G)$ . We have, by Lemma 4.6, that  $\sigma_G = \sigma_{G'}$ . Conversely, suppose that  $\sigma_G = \sigma_{G'}$ . By Proposition 4.8 there exist labelled  $m$ -edge-coloured trees with  $k$  vertices,  $G_*$ , in  $\Gamma(G)$  and  $G'_*$  in  $\Gamma(G')$  where the underlying unlabelled trees of  $G_*$  and  $G'_*$  contain symbols  $S_1$  and  $S_m$  only.

Suppose that  $k$  is odd and that  $G_*$  and  $G'_*$  are as follows:

$$\begin{aligned} G_* &: a_1 \xrightarrow{S_1} a_2 \xrightarrow{S_m} a_3 \xrightarrow{S_1} a_4 \quad \cdots \quad a_{k-1} \xrightarrow{S_m} a_k \\ G'_* &: a'_1 \xrightarrow{S_1} a'_2 \xrightarrow{S_m} a'_3 \xrightarrow{S_1} a'_4 \quad \cdots \quad a'_{k-1} \xrightarrow{S_m} a'_k. \end{aligned}$$

Then

$$\sigma_{G_*} = (a_1 a_3 a_5 \cdots a_k a_{k-1} a_{k-4} \cdots a_2) = (a'_1 a'_3 a'_5 \cdots a'_k a'_{k-1} a'_{k-4} \cdots a'_2) = \sigma_{G'_*}.$$

Because the underlying unlabelled tree of  $G'_*$  is a line,  $G'_*$  is determined by  $a'_1$  and its circular order. Thus, given  $G_*$  there are at most  $k$  possibilities for  $G'_*$  such that  $\sigma_{G_*} = \sigma_{G'_*}$  holds.

By Lemma 4.9 and Lemma 4.6, repeatedly applying either  $R_{i,j}$ -induction or  $L_{i,j}$ -induction to  $G_*$  gives  $k$  distinct labelled  $m$ -edge-coloured trees  $H$  with  $k$  vertices satisfying  $\sigma_{G_*} = \sigma_H$ . Hence any  $G'_*$  such that  $\sigma_{G_*} = \sigma_{G'_*}$  must be one of these trees  $H$ . Therefore, using Remark 4.10,  $G'_*$  is in  $\Gamma(G_*)$  and thus  $G'$  is in  $\Gamma(G)$ .

Suppose that  $k$  is even. In this case, there are at most  $k/2$  distinct possibilities for  $G'_*$  because of the symmetry of the underlying unlabelled tree. The result then follows from Lemma 4.9, Remark 4.10, and Lemma 4.6 as above.  $\square$

**Lemma 4.12.** *Let  $G$  be a labelled  $m$ -edge-coloured tree with  $k$  vertices containing a maximal  $S_i$ - $S_j$  chain  $B$  with  $l$  vertices and no incident edges coloured  $S_i$ ,  $i < l < j$ . Then  $R_{i,j}^B$  has order  $l$  on  $G$ . In particular,  $R_i$  and  $L_i$  have finite order on any maximal chain in  $G$ .*

*Proof.* This follows easily from Lemma 4.9.  $\square$

Let  $G$  be a labelled 3-edge-coloured tree with  $k$  vertices. Let  $\Gamma'(G)$  be the smallest set of labelled 3-edge-coloured trees with  $k$  vertices closed under  $R_1$  and  $L_2$  (as defined in [2, §4], identifying the symbols  $+$ ,  $=$ ,  $-$  with  $S_1$ ,  $S_2$  and  $S_3$  respectively). Then:

**Corollary 4.13.** (a) *Two labelled  $m$ -edge-coloured trees with  $k$  vertices are induction equivalent if and only if there is a sequence of inductions only of the form  $R_p$  (respectively, only of the form  $L_p$ ) taking one to the other.*  
(b) *If  $G$  is a 3-edge-coloured tree with  $k$  vertices, then  $\Gamma(G) = \Gamma'(G)$ .*

*Proof.* Part (a) follows from Lemma 4.12, using the fact that  $L_p$  (on a given maximal chain) is the inverse of  $R_p$ . For part (b), note that (on a given maximal chain),  $L_1$  is the inverse of  $R_1$  and  $R_1$  has finite order. Similarly,  $R_2$  is the inverse of  $L_2$  and  $L_2$  has finite order.  $\square$

Finally in this section, we note the interesting fact that every induction equivalence class of labelled  $m$ -edge coloured trees contains a labelling of any given unlabelled  $m$ -edge coloured tree. We first have:

**Lemma 4.14.** *Any two (unlabelled)  $m$ -edge coloured trees are induction equivalent.*

*Proof.* Let  $\mathcal{S}$  and  $\mathcal{S}'$  be arbitrary  $m$ -edge-coloured trees with  $k$  vertices. By Lemma 4.8,  $\mathcal{S}$  is induction equivalent to a labelled  $m$ -edge-coloured tree with  $k$  vertices containing only the symbols  $S_1$  and  $S_m$ ; similarly for  $\mathcal{S}'$ . If  $k$  is odd there is only one such tree:

$$\bullet \xrightarrow{S_m} \bullet \xrightarrow{S_1} \bullet \xrightarrow{S_m} \bullet \quad \cdots \quad \bullet \xrightarrow{S_1} \bullet,$$

and it follows that  $\mathcal{S}$  and  $\mathcal{S}'$  are induction equivalent. If  $k$  is even, there are two such trees:

$$\begin{aligned} S_m &: \bullet \xrightarrow{S_m} \bullet \xrightarrow{S_1} \bullet \xrightarrow{S_m} \bullet \quad \cdots \quad \bullet \xrightarrow{S_m} \bullet \\ S_1 &: \bullet \xrightarrow{S_1} \bullet \xrightarrow{S_m} \bullet \xrightarrow{S_1} \bullet \quad \cdots \quad \bullet \xrightarrow{S_1} \bullet. \end{aligned}$$

Then  $R_{1,m}^{\mathcal{S}_m}(\mathcal{S}_m) = \mathcal{S}_1$ , so  $\mathcal{S}_m$  is induction equivalent to  $\mathcal{S}_1$  by Lemma 4.3. It follows that  $\mathcal{S}$  and  $\mathcal{S}'$  are induction equivalent in this case also.  $\square$

**Corollary 4.15.** *Let  $G$  be a labelled  $m$ -edge-coloured tree with  $k$  vertices. Then every possible  $m$ -edge-coloured tree with  $k$  vertices appears as the underlying unlabelled  $m$ -edge-coloured tree of a labelled  $m$ -edge-coloured tree in  $\Gamma(G)$ .*

*Proof.* Given a labelled  $m$ -edge-coloured tree with  $k$  vertices,  $G$ , whose underlying unlabelled tree is  $\mathcal{S}$ , and an arbitrary  $m$ -edge-coloured tree with  $k$  vertices,  $\mathcal{S}'$ , Lemma 4.14 shows that  $\mathcal{S}$  and  $\mathcal{S}'$  are induction equivalent. It follows that  $G$  and a vertex-labelling of  $\mathcal{S}'$  are induction equivalent, and the result follows.  $\square$

## 5. INDUCTION ON $m$ -ANGULATIONS OF POLYGONS

By Theorem 3.5, the set of labelled  $m$ -edge-coloured trees with  $k$  vertices is in bijection with the set of  $m$ -gon-labelled diagonal-coloured  $m$ -angulations of a polygon  $P_n$  with  $n = (m-2)k + 2$  sides, up to rotation, where  $m$ -gon-labelled means that the  $m$ -gons are labelled with  $1, 2, \dots, k$ . The colours satisfy a *boundary rule*, i.e. those bounding each  $m$ -gon in the  $m$ -angulations must be the symbols  $S_1, \dots, S_m$  in clockwise order.

Our aim in this section is to rewrite induction in the language of  $(m-2)$ -clusters. The set of  $m$ -angulations of  $P_n$  containing  $k$   $m$ -gons is in bijection with the set of  $(m-2)$ -clusters of type  $A_{k-1}$  by Fomin-Reading [5]. Mutation corresponds to rotating a diagonal one step anticlockwise in the subpolygon obtained when the diagonal is removed.

We shall see that induction has a description as a composition of such mutations and appropriate recolourings of diagonals and vertex relabellings. An easy induction argument based on cutting an  $m$ -angulation along one of its sides shows that:

**Lemma 5.1.** *Every  $m$ -angulation  $\mathcal{M}$  of a polygon has at least two  $m$ -gons with  $m-1$  boundary edges or is an  $m$ -angulation of an  $m$ -gon.  $\square$*

**Remark 5.2.** If all the diagonals in an  $m$ -angulation are incident with one vertex, then by moving them one after another, we get the same  $m$ -angulation rotated through  $2\pi/n$ .

We label the vertices of  $P_n$   $1, 2, \dots, n$  clockwise around the boundary, and use the notation  $[i, j]$  to denote a diagonal in the polygon connecting vertex  $i$  with vertex  $j$ .

**Lemma 5.3.** *Let  $\mathcal{M}$  be an  $m$ -angulation of  $P_n$  with  $k$   $m$ -gons. Then there is an explicit sequence of mutations taking  $\mathcal{M}$  to its rotation through  $2\pi/n$  anticlockwise.*

*Proof.* We induct on the number  $k$  of  $m$ -gons in the  $m$ -angulation. The result is trivial if  $k = 1$ , so consider the general case. By Lemma 5.1, there is an  $m$ -gon  $M$  with a unique internal edge  $e$ , joining vertices  $[i, i + (m-1)]$  for some  $i$ . Let  $R$  be the subpolygon of  $P_n$  given by the union of the  $m$ -gons incident with  $i$ . Applying Remark 5.2 to the induced  $m$ -angulation of  $R$  to rotate it one step anticlockwise, we obtain a new  $m$ -angulation of  $R$ , and hence of  $P_n$ , containing an  $m$ -gon  $M'$  with vertices  $i-1, i, i+1, \dots, i+(m-1)-1$ . Let  $R'$  be the subpolygon of  $R$  with  $M'$  removed. Applying Remark 5.2 repeatedly we may rotate the induced  $m$ -angulation of  $R'$  one step clockwise.

We obtain a new  $m$ -angulation of  $P_n$ . By induction we may rotate the subpolygon with  $M'$  removed anticlockwise by an explicit sequence. It is easy to check that the total effect of the above is to rotate the original  $m$ -angulation of  $P_n$  one step anticlockwise.  $\square$

**Definition 5.4.** Let  $\mathcal{M}$  be a diagonal-coloured  $m$ -angulation of  $P_n$ . If there is at least one internal diagonal in  $\mathcal{M}$  and all of the internal diagonals of  $\mathcal{M}$  are coloured only with  $S_i$  or  $S_{i+1}$  for fixed  $i$ , we call  $\mathcal{M}$  a *snake  $m$ -angulation*. A subpolygon of  $P_n$  with this property is called a *snake subpolygon*. Note that in any snake subpolygon the internal diagonals must be of the form  $[i_1, i_2]$ ,  $[i_2, i_3]$ , and so on.

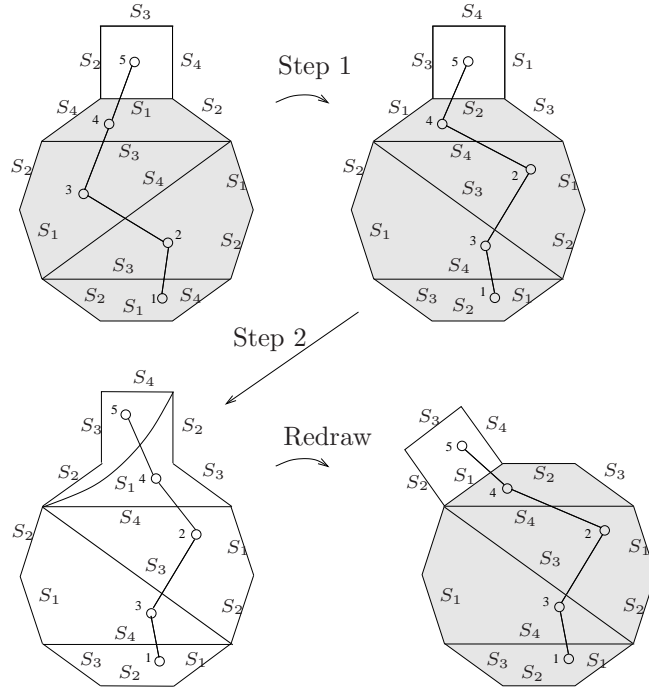


FIGURE 5. Mutations giving rise to  $R_3$  induction (see Proposition 5.5).

In a cluster algebra context, snake triangulations appeared in [6, Sect. 3.5] and for general  $m$  in [5, Sect. 5.1] as  $m$ -snakes.

Let  $\mathcal{M}$  be an  $m$ -gon-labelled diagonal-coloured  $m$ -angulation of  $P_n$  and let  $1 \leq i \leq m$ . Choose a maximal snake subpolygon  $\mathcal{B}$  of  $\mathcal{M}$  with internal diagonals coloured  $S_i$  or  $S_{i+1}$ . Let  $\mathcal{R}_i^{\mathcal{B}}(\mathcal{M})$  be defined using the following procedure.

**Step 1:** Let  $M_1, M_2$  be the two  $m$ -gons in  $\mathcal{B}$  with  $m-1$  boundary edges in  $\mathcal{B}$ , and let  $e = e_1$  be the internal edge of  $M_1$  in  $\mathcal{B}$ . Let  $e_2, \dots, e_N$  be the other internal edges in  $\mathcal{B}$ , numbered so that  $e_{i-1}$  and  $e_i$  are incident for all  $i$ . If  $e_1$  has colour  $S_i$  (respectively,  $S_{i+1}$ ), mutate edges  $e_2, e_4, \dots$  (respectively,  $e_1, e_3, \dots$ ), recolouring the new diagonals with  $S_i$  and using the boundary rule to recolour the rest of the polygon. The labelling of an  $m$ -gon is given by the number of the  $m$ -gon before mutation whose intersection with the boundary of  $\mathcal{B}$  was the same.

**Step 2:** Let  $C_1, C_2$  be the union of the connected components of the complement of  $\mathcal{B}$  in  $P_n$  incident with  $M_1, M_2$  respectively, and let  $D_i = C_i \cup M_i$ . Applying Lemma 5.3 to  $D_1, D_2$  we get a sequence of mutations rotating the induced  $m$ -angulation of each  $D_i$  anticlockwise one step around its boundary. We recolour the diagonals according to the boundary rule. The label of the image of an  $m$ -gon under this rotation is the same.

Note that Step 2 has the same effect as detaching each component of  $C_i$  in the original  $m$ -angulation from an edge of  $M$  coloured  $S_j$ ,  $j \neq i, i+1$ , and reattaching it after Step 1 to a boundary edge of  $M$  with the same symbol  $S_j$  (now one step anticlockwise around the boundary of  $M$ ), keeping the original colouring of the diagonals of  $C_i$  (i.e. from before Step 1). Such a boundary edge of  $M$  always exists, since  $j \neq i, i+1$ .

It is easy to see that the above procedure does not depend on the initial choice of  $M_1$ . Note also that the procedure commutes with any rotational symmetry of  $P_n$  and so gives a well defined induction on a diagonal-coloured  $m$ -angulation of  $P_n$  up to rotation.

Comparing the definitions of  $R_i^{\mathcal{B}}$  and  $\mathcal{R}_i^{\mathcal{B}}$  we see that:

**Proposition 5.5.** *Suppose that  $1 \leq i \leq m$  and let  $G$  be a labelled  $m$ -edge-coloured tree with  $k$  vertices containing a maximal  $S_i - S_{i+1}$  chain  $B$ . Let  $\mathcal{M}$  be the corresponding  $m$ -gon-labelled diagonal-coloured  $m$ -angulation of  $P_n$  up to rotation with maximal snake subpolygon  $\mathcal{B}$  corresponding to  $B$ . Then  $R_i^{\mathcal{B}}(G)$  corresponds to  $R_i^{\mathcal{B}}(\mathcal{M})$ .  $\square$*

For an example of Proposition 5.5, see Figure 5, with the snake subpolygon shaded.

**Acknowledgements:** We would like to thank Luca Q. Zamboni for some helpful conversations. RJM would like to thank Sibylle Schroll and the Department of Mathematics at the University of Leicester for their kind hospitality.

#### REFERENCES

- [1] D. Bisch, V. Jones, Algebras associated to intermediate subfactors, *Invent. Math.* 128 (1997) 89–157.
- [2] J. Cassaigne, S. Ferenczi, L. Q. Zamboni, Combinatorial trees arising in the study of interval exchange transformations, *European J. Combin.* 32 (2011) 1428–1444.
- [3] M. Cho, D. Kim, S. Seo, H. Shin, Colored Prüfer codes for  $k$ -edge colored trees, *Electron. J. Combin.* 11 (2004), 7 pp.
- [4] S. Ferenczi, L. Q. Zamboni, Structure of  $k$ -interval exchange transformations: induction, trajectories, and distance theorems, *J. Anal. Math.* 112 (2010) 289–328.
- [5] S. Fomin, N. Reading, Generalized cluster complexes and Coxeter combinatorics, *Int. Math. Res. Not. IMRN* 44 (2005) 2709–2757.
- [6] S. Fomin, A. Zelevinsky,  $Y$ -systems and generalized associahedra, *Ann. of Math. (2)* 158 (2003) 977–1018.
- [7] P. Hilton, J. Pederson, Catalan numbers, their generalization, and their uses, *Math. Intelligencer* 13 (1991) 64–75.
- [8] I. L. Hofacker, P. Schuster, P. F. Stadler, Combinatorics of RNA secondary structures, *Discrete Appl. Math.* 88 (1998) 207–237.
- [9] V. E. Hoggatt, Jr., M. Bicknell, Catalan and related sequences arising from inverses of Pascal’s triangle matrices, *Fibonacci Quart.* 14 (1976) 395–405.
- [10] R. Nussinov, G. Pieczenik, J. R. Griggs, D. J. Kleitman, Algorithms for loop matchings, *SIAM J. Appl. Math.* 35 (1978) 68–82.
- [11] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences, 2010, [www.oeis.org](http://www.oeis.org).
- [12] R. P. Stanley, *Enumerative combinatorics*, Vol. 2, Cambridge University Press, Cambridge, 1999.
- [13] M. S. Waterman, Secondary structure of single-stranded nucleic acids, in: *Studies in foundations and combinatorics*, Academic Press, New York-London, 1978, pp. 167–212.

SCHOOL OF MATHEMATICS, UNIVERSITY OF LEEDS, LEEDS LS2 9JT, UNITED KINGDOM  
*E-mail address:* `marsh@maths.leeds.ac.uk`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF LEICESTER, UNIVERSITY ROAD, LEICESTER LE1 7RH, UNITED KINGDOM  
*E-mail address:* `ss489@le.ac.uk`