

# On inequivalent factorizations of a cycle

G Berkolaiko<sup>1</sup>, J M Harrison<sup>1,2</sup>, M Novaes<sup>3</sup>

<sup>1</sup> Department of Mathematics, Texas A&M University,  
College Station, TX 77843-3368, USA

<sup>2</sup> Department of Mathematics, Baylor University, Waco, TX 76798-7328, USA

<sup>3</sup> School of Mathematics, University of Bristol, Bristol BS8 1TW, UK

December 15, 2008

## Abstract

We introduce a bijection between inequivalent minimal factorizations of the  $n$ -cycle  $(1\ 2\ \dots\ n)$  into a product of smaller cycles of given length and trees of a certain structure. A factorization has the type  $\alpha = (\alpha_2, \alpha_3, \dots)$  if it has  $\alpha_j$  factors of length  $j$ . Inequivalent factorizations are defined up to reordering of commuting factors. A factorization is minimal if no factorizations of a type  $\alpha'$  strictly smaller than  $\alpha$  exist.

The introduced bijection allows us to answer such questions as the number of factorizations with a given number of different (commuting) factors that can appear in the first and in the last positions, and the structure of the set of factors that can be arranged into a product evaluating to  $(1\ 2\ \dots\ n)$ . Important consequences of the discovered structure include monotonicity of the constituent factors and uniqueness of an arrangement into a valid factorization: any two minimal factorizations of  $(1\ 2\ \dots\ n)$  consisting of the same factors must be equivalent.

## 1 Introduction

Counting factorizations of a permutation into a product of cycles of specified length is a problem with rich history, dating back at least to Hurwitz [1], and with many important applications, in particular in geometry (see e.g. [2]). Our interest in such problems is driven by applications encountered in physics, namely semiclassical trajectory-based analysis of quantum chaotic transport [3, 4]. The main ingredient of this analysis is the existence of correlations between sets of long trajectories connecting an input channel of the quantum dot to an output channel. The trajectories organize themselves into families, with the elements of a family differing among

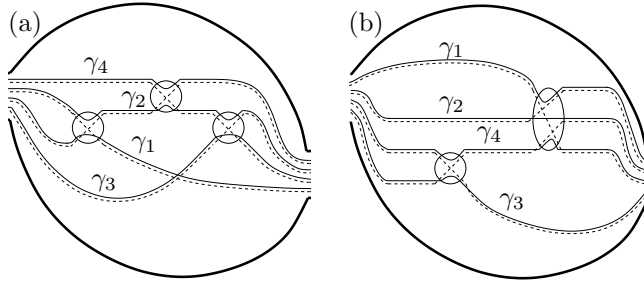


Figure 1: Two schematic examples of correlated sets of classical trajectories through a cavity. One set of trajectories is represented by solid lines, while the other set is drawn in dashed lines. The circles mark the crossing regions, where the trajectories from different sets differ significantly: dotted lines cross while solid lines have avoided crossings.

themselves only by their behavior in small regions (see Fig. 1) in which some of them have crossings while others have anti-crossings. Enumerating possible configurations of crossing regions and their inter-connectivity is a question of combinatorial nature. In a certain special case it was found in [4] to be connected to the *inequivalent minimal* factorizations of the  $n$ -cycle  $(1\ 2\ \dots\ n)$  into a product of smaller cycles.

The base for our results is a simple and highly pictorial bijection between said factorizations and plane trees. This bijection allows us to recover easily some already known results and to answer new questions about the structure of the set of factorizations. To be more specific we need to introduce some notation.

Let  $\sigma_m \cdots \sigma_1$  be a factorization of the cycle  $(1\ 2\ \dots\ n)$  into a product of smaller cycles. By convention, the first entry of a cycle is always its smallest element. We say the factorization is of type  $\alpha$  if among  $\{\sigma_j\}$  there are exactly  $\alpha_2$  2-cycles (transpositions),  $\alpha_3$  3-cycles and so on. Let us define

$$|\alpha| = \sum_{j \geq 2} \alpha_j, \quad \langle \alpha \rangle = \sum_{j \geq 2} (j-1)\alpha_j. \quad (1)$$

The quantity  $\alpha$  satisfies

$$\langle \alpha \rangle \geq n - 1. \quad (2)$$

If the above relation becomes equality, the factorization is called *minimal*. We only consider minimal factorizations.

If two factorizations differ only in the order of commuting factors, they are said to be *equivalent*. An example of two equivalent factorizations is

$$(1\ 2\ 3\ 4) = (3\ 4)(1\ 2)(2\ 4) = (1\ 2)(3\ 4)(2\ 4). \quad (3)$$

From now on we will refer to equivalence classes of factorizations simply as factorizations, unless the distinction is of particular importance. In Theorem 1 we establish a bijection between factorizations of type  $\alpha$  and plane trees with vertex degrees determined by  $\alpha$ . In turn, the trees have been enumerated by Erdélyi and Etherington

[5] (see also Tutte [6] and Stanley [7], theorem 5.3.10). We are interested in counting the factorizations of type  $\alpha$ . This number will be denoted by  $\tilde{H}(\alpha)$ . In Theorem 2 we give an equation for its generating function and its relation to Catalan numbers.

Given an equivalence class of factorizations of the form  $\sigma_m \cdots \sigma_1$ , we refer to the number of different cycles that can appear in the position  $\sigma_m$ , the *number of heads* of the factorization. Similarly, the *number of tails* is the number of cycles that can appear in the position  $\sigma_1$ . For example, the factorization in (3) has 2 heads (transpositions (1 2) and (3 4)) and 1 tail. In Theorem 3 we derive a generating function for the number of inequivalent minimal factorizations with the specified number of heads and tails, denoted by  $\tilde{H}_{\mathbf{h}, \mathbf{t}}(\alpha)$ . The vectors  $\mathbf{h} = (h_2, h_3, \dots)$  and  $\mathbf{t} = (t_2, t_3, \dots)$  characterize the number of heads and tails. Namely,  $h_j$  is the number of  $j$ -cycle heads and  $t_j$  is the number of  $j$ -cycle tails. The quantity  $\tilde{H}_{\mathbf{h}, \mathbf{t}}(\alpha)$  is of importance in applications to quantum chaotic transport [4]. Looking again at Fig. 1, the vector  $\mathbf{h}$  (corresp.  $\mathbf{t}$ ) counts the number of crossings that can happen close to the left (corresp. right) opening of the cavity. For example,  $t_2 = 1$  on the diagram (a) while  $t_2 = 0$  on the diagram (b), since the 3-crossing prevents the 2-crossing from getting close to the right opening.

Finally, we will give a complete characterization (necessary and sufficient conditions) for a set  $\{\sigma_j\}$  of cycles to give rise to a minimal factorization of the  $n$ -cycle  $(1 2 \dots n)$ . This characterization is given in Theorem 4. Here we only mention some of its corollaries. First is a surprising result that a factorization equivalence class is completely determined by the factors. In other words, two factorizations composed of the same factors are equivalent. This can obviously be extended to general permutations but only applies to minimal factorizations: for example the factorizations  $(1 2)(1 2)(1 3) = (1 3)(1 2)(1 2)$  contain the same factors but are not equivalent.

The second corollary gives a simple necessary and sufficient condition for an  $r$ -cycle to appear in a minimal factorization of  $(1 2 \dots n)$ . A cycle  $(p_1 p_2 \dots p_r)$  is called *increasing* if  $p_1 < p_2 < \dots < p_r$ . It turns out that only increasing cycles can appear as factors. For example, this excludes  $(1 3 2)$  from being a part of a minimal factorization of  $(1 2 \dots n)$ . Of course, this restriction is only non-empty for  $r > 2$ .

Some of the results discussed in this paper are already known, although they have been derived by using different methods. Namely, the number of inequivalent factorizations into a product of transpositions (i.e.  $\alpha = (n-1, \mathbf{0})$ ) has been obtained by Eidswick [8] and Longyear [9]. Springer [10] derived a formula for  $\tilde{H}(\alpha)$  using a different bijection to trees of the same type. Irving [11] reproduced the result of Springer using more general machinery involving cacti. The novelty of our approach is in the type (and simplicity!) of the bijection used. Being very visual, our bijection allows us to obtain answers to new questions, namely to count factorizations with specified number of heads and tails and to characterize sets of cycles that can be arranged into a factorization.

Of other related results we would like to mention Hurwitz [1] who suggested a formula for the number of minimal transitive factorizations (counting equivalent factorizations as different) of a general permutation into a product of 2-cycles. A

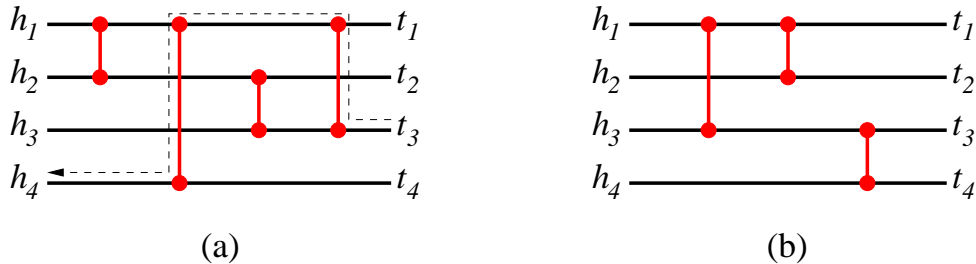


Figure 2: Visualizing a product of transpositions using a “shuttle diagram”. Each term  $(k_1 k_2)$  in the product corresponds to a vertical edge (“shuttle”) connecting lines  $k_1$  and  $k_2$ . The lines are ordered in the same way as the terms in the product. Part (a) depicts the product  $(1\ 2)(1\ 4)(2\ 3)(1\ 3)$  and finding the image of 3 under the resulting permutation (dashed line). Part (b) is the representation of  $(1\ 3)(1\ 2)(3\ 4)$ .

factorization is called *transitive* if the group generated by the factors  $\sigma_1, \dots, \sigma_m$  acts transitively on the set  $1, \dots, n$ . However, Hurwitz gave only a sketch of a proof and his paper was largely unknown to the combinatorialists. For a special case of factorizations of the  $n$ -cycle, the formula was (re-)derived by Dénes [12], with alternative proofs given by Lossers [13], Moszkowski [14], Goulden and Pepper [15]. For general permutations, Strehl [16] reconstructed the original proof of Hurwitz, filling in the gaps, while Goulden and Jackson [17] gave an independent proof. Generalizations of Hurwitz formula to factorizations into more general cycles were considered in Goulden and Jackson [18] and Irving [11]. Finally, *inequivalent* minimal transitive factorizations of a permutation consisting of  $m = 2$  cycles have been counted in Goulden, Jackson and Latour [19] (into transpositions) and in Irving [11] (into general cycles). Generalizations to larger  $m$  appear to be difficult.

## 2 Visualizing a product of cycles

### 2.1 Product of transpositions

A particularly nice way to visualize a product of transpositions was suggested in [20] (see also [21]). A permutation from  $S_n$  is represented as  $n$  labeled horizontal lines with several vertical lines (“shuttles”) connecting some pairs of the horizontal lines, see Fig. 2. The right and left ends of a line  $k$  are labeled with  $t_k$  (for “tail”) and  $h_k$  (for “head”) correspondingly. For every horizontal line, start at the right and trace the line to the left. Whenever an end of a shuttle is encountered, trace this shuttle vertically till its other end and then resume going to the left (towards  $h$ ). Continue in this manner until you reach the left end of one of the horizontal lines. It is clear that the mapping “right ends to left ends” thus described is invertible and therefore one-to-one.

In this construction, a shuttle connecting lines  $k_1$  and  $k_2$  represents the transposition  $(k_1 k_2)$ . The transpositions are ordered in the same way as shuttles: right

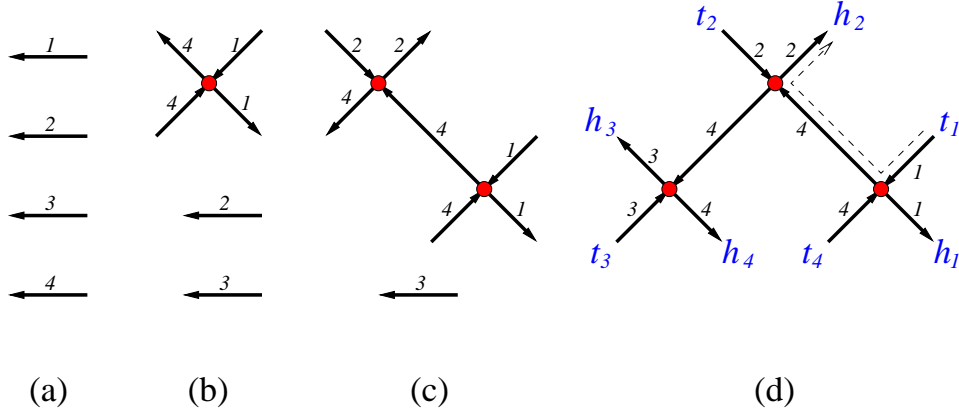


Figure 3: Visualizing a product of transpositions as a directed graph. Depicted are the steps in constructing the graph corresponding to the product  $(34)(24)(14)$ . To read off the image of  $k$  under the resulting permutation we start at  $t_k$  and follow the directions of the edges, choosing the next edge in the counterclockwise order at each vertex, until arriving to  $h_{\pi(k)}$ . The path traced starting with  $t_1$  is illustrated by the dashed line in (d).

to left. If the two neighboring transpositions  $(k_1 k_2)$  and  $(k_3 k_4)$  commute (if and only if all four  $k_j$  are distinct), the corresponding shuttles can be swapped around without affecting the dynamics. We can view the resulting diagram as a graph (with horizontal and vertical edges). If the diagram represents a factorization of an  $n$ -cycle, the graph is connected. By counting vertices and edges, one concludes that if a factorization is *minimal*, the resulting graph is a tree.

Suppose now that the diagram represents a minimal factorization of the cycle  $(12\dots n)$ . In addition to the right-to-left motion described above we define the left-to-right motion as going horizontally, ignoring the shuttles. Then, starting at  $t_1$  and going left we arrive to  $h_2$ . Going right from there we arrive to  $t_2$  and from there, to  $h_3$ . Continuing in this fashion, we obtain a closed walk with several important features. It visits the vertices  $t_1, h_2, t_2, \dots, h_n, t_n, h_1$  in this sequence. It traverses each edge of the graph exactly twice: once in each direction (this follows, for example, from the invertibility of the motion). Since the graph is a tree, we conclude that it goes from one vertex to the next one along the shortest possible route. This walk traversing the entire tree will play an important role in what follows.

Another way to visualize a product of transpositions as a directed plane graph is illustrated on Fig. 3. We start with  $n$  disjoint directed edges labeled 1 to  $n$ . For a product  $\pi = (k_{j-1} k_j) \cdots (k_3 k_4)(k_1 k_2)$ , we start by joining the heads of the edges labeled  $k_1$  and  $k_2$  at a new vertex and add two more *outgoing* edges also labeled  $k_1$  and  $k_2$ . We arrange them around the vertex so that, going counter-clockwise, the outgoing edge  $k_1$  is followed by the incoming  $k_1$ , then by the outgoing  $k_2$  and, finally, by the incoming  $k_2$ . At this and all later stages of the procedure for each  $k = 1, \dots, n$  there is exactly one “free” head of an edge labeled  $k$ , and one free tail of possibly different edge also labeled  $k$ . We now repeat the procedure for the transposition

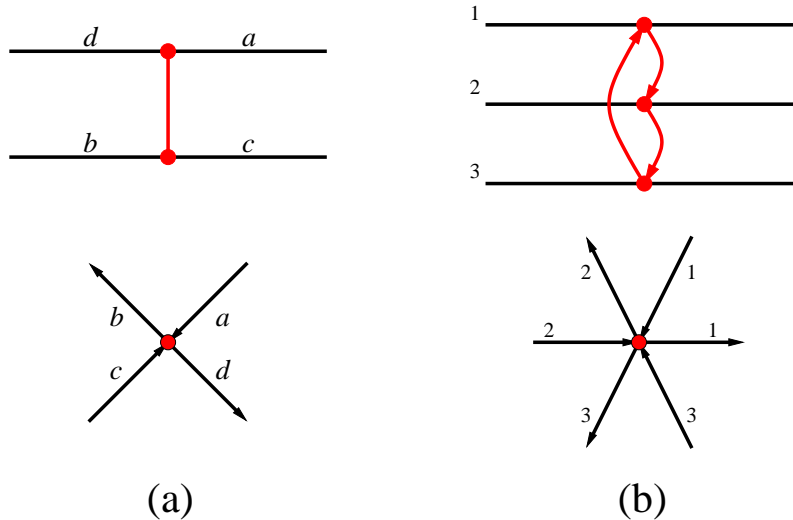


Figure 4: (a) Transforming the diagram representation of a product of transpositions into a directed graph representation. The “shuttle” edge is shrunk and its end-vertices are merged. The edges on the left are labeled in the order they are traversed by the walk  $t_1, h_2, t_2, \dots, h_n, t_n, h_1, t_1$ . (b) Visualization of the cycle  $(123)$  as a “shuttle diagram” and as a directed graph.

$(k_3 k_4)$ , joining free heads of edges marked  $k_3$  and  $k_4$ , adding new outgoing edges to new vertex and ordering the edges in the similar fashion: outgoing  $k_3$ , incoming  $k_3$ , outgoing  $k_4$  and incoming  $k_4$ . On Fig. 3(d) we identified the free heads and tails by labeling them with  $h_j$  and  $t_j$  correspondingly.

The resulting graph is closely related to the diagrams described earlier. Namely, the graph is obtained from the diagram by shrinking the shuttle edges and re-ordering the edges at the newly merged vertices, see Fig. 4(a). Moreover, the ordering of edges has been designed so that, to determine the image of  $k$  under the product permutation  $\pi$ , one would start at  $t_k$  and travel along the direction of the edges, at each vertex taking the next edge in the counterclockwise order, finally arriving to  $h_{\pi(k)}$ . This is illustrated by the dashed line on Fig. 3(d). Starting at  $h_k$  and going *against* the direction of the edges, taking the next counterclockwise edge at each vertex, will get one to  $t_k$ .

Thus, if  $\pi$  is the cycle  $(12\dots n)$ , the corresponding graph is a tree with  $n - 1$  vertices of total degree 4 (henceforth called *internal vertices*),  $2n$  vertices of degree 1 (henceforth called *leaves*) and  $3n - 2$  edges. The walk  $t_1, h_2, t_2, \dots, h_n, t_n, h_1, t_1$ , discussed in the context of diagrams, now circumnavigates the entire tree in the counter-clockwise direction. As before, it traverses each edge exactly once in each direction. The leaves of the tree are thus marked  $h_1, t_1, h_2, t_2, \dots, h_n, t_n$  going counter-clockwise, see Fig. 3.

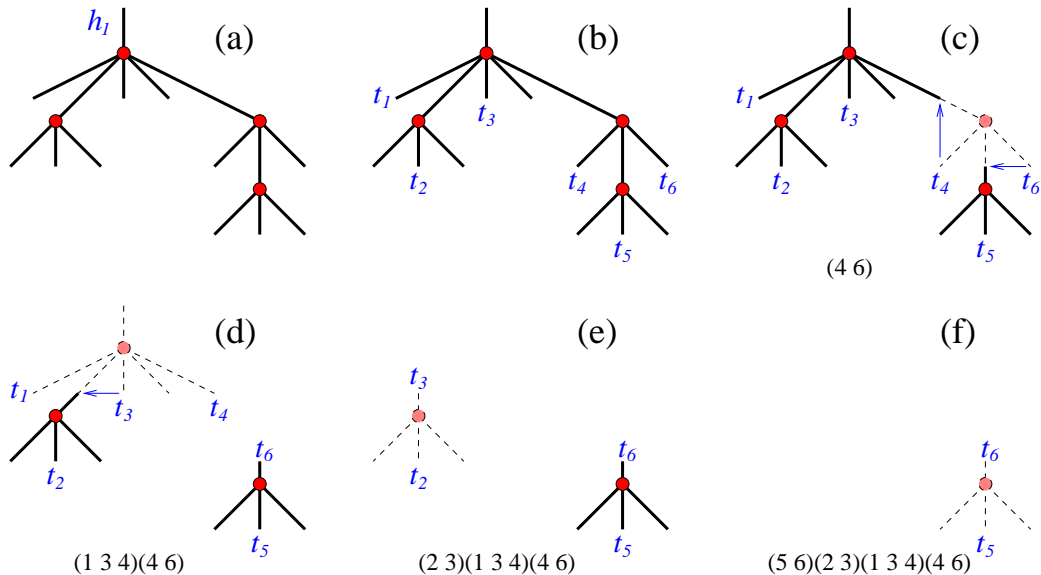


Figure 5: Reconstructing the factorization from an undirected rooted plane tree. Note that at step (d) and (e) one could choose to remove the vertex adjacent to leaves  $t_4$  and  $t_6$  instead.

## 2.2 Product of larger cycles

The generalization of the previous construction from a product of transpositions to a product of general cycles is straightforward. For a  $m$ -cycle  $(k_1 k_2 \dots k_m)$ , the corresponding shuttle is realized as  $m$  directed edges indicating transitions from horizontal line  $k_j$  to horizontal line  $k_{j+1}$ . In the directed graph visualization, the cycle corresponds to a vertex of total degree  $2m$ , with outgoing edge marked  $k_1$  followed by the incoming edge  $k_1$ , then by outgoing edge  $k_2$  and so on. We illustrate this in Fig. 4(b) using the cycle  $(1 2 3)$  as an example.

## 3 Main Results

As described in section 2.1, a factorization (up to equivalence) of the  $n$ -cycle into  $n - 1$  transpositions is naturally represented as a plane tree with  $n - 1$  internal vertices of total degree 4,  $2n$  leaves of degree 1 and  $3n - 2$  edges. If we designate the leaf  $h_1$  as the root of the tree, the labeling of all other leaves and the directions of edges can be reconstructed uniquely. This representation of a factorization as an undirected rooted plane tree turns out to be a bijection.

**Theorem 1.** *Inequivalent minimal factorizations of type  $\alpha$  of the  $n$ -cycle  $(1 2 \dots n)$  are in one-to-one correspondence with undirected rooted plane trees having  $\alpha_j$  vertices of degree  $2j$  and  $2n$  leaves of degree 1.*

*Proof.* The mapping of factorization equivalence classes to trees, described in the second part of section 2.1 and extended to larger factors in section 2.2, is well defined.

Indeed, the construction steps corresponding to commuting factors also commute. We need to show that this mapping is invertible and onto.

The mapping can be inverted by taking the following steps (see Fig. 5 for an example):

1. Label the leaves of the tree with  $h_1, t_1, h_2, \dots, t_n$  starting with the root and going counterclockwise, Fig. 5(b) (to avoid clutter we will omit the  $h$ -labels).
2. For some value of  $j$ , choose a vertex with degree  $2j$  which has  $j$   $t$ -leaves adjacent to it. Such a vertex exists by pigeonhole principle. The indices of the  $t$ -leaves give the next (in the right to left order) factor in the expansion, Fig. 5(c).
3. Remove the vertex. The edges connecting the vertex to leaves are removed entirely. The edges connecting the vertex to other vertices, if any, are cut in half. This creates one or more new leaves and their labels are inherited from the leaves neighboring them in the counterclockwise direction, Fig. 5(c).
4. Repeat from step 2, Fig. 5(d)-(f).

Notice that the number of choices one has when first running step 2 corresponds to the total number of tails.

To verify that the mapping is onto we have to check that the above inversion applied to any tree produces a factorization of the  $n$ -cycle  $(12\dots n)$ . To this end we observe that the deletion-relabeling process coupled with the application of the cycles read at step 2 transports an object initially at  $t_j$  to the leaf  $h_{j+1}$  for all  $j$  (assuming the convention  $n+1 \equiv 1$ ).  $\square$

**Theorem 2.** *The generating function of the number  $\tilde{H}(\boldsymbol{\alpha})$ , defined by*

$$h(\mathbf{x}) = \sum_{\boldsymbol{\alpha}} \tilde{H}(\boldsymbol{\alpha}) x_2^{\alpha_2} x_3^{\alpha_3} \cdots, \quad (4)$$

where the sum over  $\boldsymbol{\alpha}$  is unrestricted, satisfies the recurrence relation

$$h(\mathbf{x}) = 1 + x_2 h^3(\mathbf{x}) + x_3 h^5(\mathbf{x}) + \dots \quad (5)$$

It follows that

$$\sum_{\boldsymbol{\alpha}: \langle \boldsymbol{\alpha} \rangle = n} (-1)^{|\boldsymbol{\alpha}|+n} \tilde{H}(\boldsymbol{\alpha}) = \frac{1}{n+1} \binom{2n}{n} = c_n, \quad (6)$$

where  $c_n$  is the  $n$ -th Catalan number.

The above statement is a simple consequence of the bijection between factorizations and trees and the known results enumerating the trees, see Erdélyi and Etherington [5], Tutte [6] or Stanley [7, Theorem 5.3.10]. We will give a short proof in section 4 to introduce the methods used in the next result.

For factorizations with specified numbers of heads and tails we have

**Theorem 3.** Let  $g(\mathbf{x}, \mathbf{v}, \mathbf{u})$  be the generating function of the number  $\tilde{H}_{\mathbf{h}, \mathbf{t}}(\boldsymbol{\alpha})$  of inequivalent minimal factorizations of the  $n$ -cycle  $(1\ 2\ \dots\ n)$  of type  $\boldsymbol{\alpha}$  with specified number of heads and tails, defined by

$$g(\mathbf{x}, \mathbf{v}, \mathbf{u}) = \sum_{\boldsymbol{\alpha}} \sum_{\mathbf{h}=(0,0,\dots)}^{\boldsymbol{\alpha}} \sum_{\mathbf{t}=(0,0,\dots)}^{\boldsymbol{\alpha}} \tilde{H}_{\mathbf{h}, \mathbf{t}}(\boldsymbol{\alpha}) x_2^{\alpha_2} u_2^{h_2} v_2^{t_2} x_3^{\alpha_3} u_3^{h_3} v_3^{t_3} \dots$$

Then  $g(\mathbf{x}, \mathbf{v}, \mathbf{u})$  can be found as

$$g = f - \sum_{n \geq 2} x_n (1 - u_n) f^n \quad (7)$$

$$= f \hat{f} - \sum_{n \geq 2} x_n (f \hat{f})^n, \quad (8)$$

where  $f$  satisfies the recursion relation

$$f(\mathbf{x}, \mathbf{v}, \mathbf{u}) = 1 + \sum_{n \geq 2} x_n (f^n - 1 + v_n) \hat{f}^{n-1} \quad (9)$$

and  $\hat{f}$  is obtained from  $f$  by exchanging the roles of  $u$  and  $v$ ,

$$\hat{f}(\mathbf{x}, \mathbf{v}, \mathbf{u}) = f(\mathbf{x}, \mathbf{u}, \mathbf{v}).$$

Moving on to the characterization of all possible sets of factors, we remind the reader that a cycle  $(s_1\ s_2\ \dots\ s_{|\sigma|})$  is called *increasing* if  $s_1 < s_2 < \dots < s_{|\sigma|}$ . Here by  $|\sigma|$  we denote the size of the cycle  $\sigma$ .

**Definition 1.** The increasing cycle  $\sigma = (s_1\ s_2\ \dots\ s_{|\sigma|})$  is said to be *higher* than the increasing cycle  $\rho = (r_1\ r_2\ \dots\ r_{|\rho|})$  (denoted  $\sigma \geq \rho$ ) if there is  $j$ ,  $1 \leq j < |\sigma|$ , such that

$$s_j \leq r_1 < \dots < r_{|\rho|} \leq s_{j+1}.$$

The cycle  $\sigma$  is said to be *earlier* than  $\rho$  (denoted  $\sigma \succeq \rho$ ) if

$$s_{|\sigma|} \leq r_1 < \dots < r_{|\rho|}.$$

If any of the relations  $\sigma \geq \rho$ ,  $\sigma \succeq \rho$ ,  $\rho \geq \sigma$  or  $\rho \succeq \sigma$  holds, the cycles are called *comparable*.

**Definition 2.** The cycle  $\rho$  is said to have a *right connector* to the cycle  $\sigma$  if  $\sigma \geq \rho$  and  $r_{|\rho|} = s_{j+1}$ . The cycle  $\rho$  is said to have a *left connector* to the cycle  $\sigma$  if either  $\sigma \geq \rho$  and  $s_j = r_1$  or  $\sigma \succeq \rho$  and  $s_{|\sigma|} = r_1$ .

**Theorem 4.** The cycles  $\sigma_1, \sigma_2, \dots, \sigma_m$  satisfying

$$1 + \sum_{j=1}^m (|\sigma_j| - 1) = n \quad (10)$$

can be arranged into a factorization of the  $n$ -cycle  $(1\ 2\ \dots\ n)$  if and only if all cycles are increasing, pairwise comparable and no cycle has connectors of both types. All factorizations consisting of these cycles are equivalent.

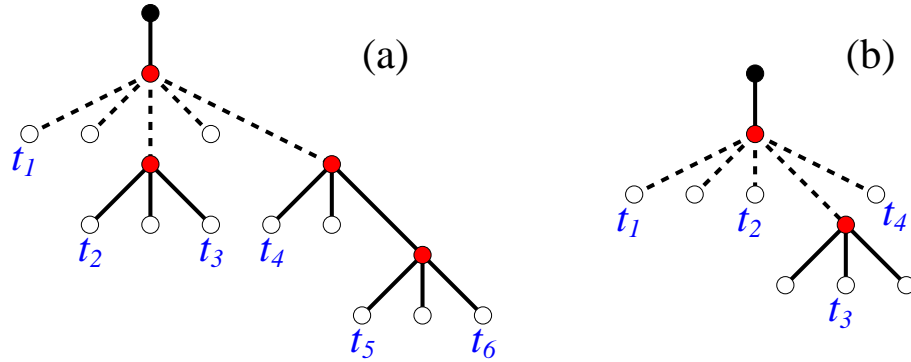


Figure 6: (a) A tree with characteristic  $\alpha = (3, 1, 0)$  separates at the top internal vertex into five subtrees characterized by  $\alpha_1 = \alpha_2 = \alpha_4 = (0)$ ,  $\alpha_3 = (1, 0)$  and  $\alpha_5 = (2, 0)$ . (b) A tree with characteristic  $\alpha = (1, 1, 0)$ .

**Remark 1.** The last condition excludes the following situation: a cycle  $\sigma$  having a left connector to a cycle  $\sigma'$  and a right connector to a cycle  $\sigma''$ . It is possible that a cycle has no connectors at all. From the proof it will become clear that there is only one such cycle in every eligible set  $\{\sigma_1, \sigma_2, \dots, \sigma_m\}$ , namely the cycle corresponding to the top vertex in the tree.

**Example 1.** The cycles  $\{(145), (13), (24)\}$  cannot be arranged into a factorization since  $(13)$  and  $(24)$  are incomparable. The cycles  $\{(145), (23), (34)\}$  cannot be arranged into a factorization since the cycle  $(34)$  has two connectors. The cycles  $\{(145), (12), (23)\}$  satisfy the conditions of Theorem 4 and yield the factorization  $(145)(12)(23)$ .

## 4 Counting factorizations

While Theorem 2 is a simple consequence of Theorem 1 and known counting results for trees ([5], [6] or [7, Theorem 5.3.10]), we provide a brief proof in order to introduce the methods used in the proof of Theorem 3.

*Proof of Theorem 2.* We are going to enumerate the plane trees which have  $\alpha_j$  vertices of degree  $2j$ . The set of all such trees will be denoted by  $\mathcal{T}_\alpha$ , where  $\alpha = (\alpha_2, \alpha_3, \dots)$ .

To derive a recurrence relation for  $|\mathcal{T}_\alpha|$  we break the tree at the top vertex adjacent to the root. The top vertex has degree  $2(d+1)$  for some  $d \geq 1$  and, when splitting the tree, it becomes the root of  $2d+1$  subtrees  $T_1, \dots, T_{2d+1}$ , characterized by vectors  $\alpha_1, \dots, \alpha_{2d+1}$  (some of them possibly empty). Clearly  $\alpha = \sum_{i=1}^{2d+1} \alpha_i + e_d$  where  $e_d$  has 1 in its  $d$ -th component<sup>1</sup> and zero elsewhere, representing the top vertex that was removed. Figure 6(a) shows a tree with characteristic  $(3, 1, \mathbf{0})$ . This tree

<sup>1</sup>we remind the reader that the  $k$ -th component of vector  $\alpha$  is  $\alpha_{k+1}$

splits at the top vertex, degree six ( $d = 2$ ), into five subtrees. The number of all possible trees with the top vertex of degree  $2(d + 1)$  is given by the number of combinations of subtrees,  $\prod_{j=1}^{2d+1} |\mathcal{T}_{\alpha_j}|$ , where  $\sum_{j=1}^{2d+1} \alpha_j = \alpha - \mathbf{e}_d$ . Summing over the possible degrees of the top vertex establishes the recursion relation,

$$\tilde{H}(\alpha) = |\mathcal{T}_\alpha| = \sum_{d \geq 1} \sum_{\alpha_1 \cdots \alpha_{2d+1}} \prod_{j=1}^{2d+1} \tilde{H}(\alpha_j) \delta_{\alpha_1 + \cdots + \alpha_{2d+1}, \alpha - \mathbf{e}_d}. \quad (11)$$

Computing the generating function  $h(\mathbf{x})$ , equation (4), we recover (5).

To relate  $\tilde{H}(\alpha)$  to Catalan numbers (something important in applications, [4]), we take  $x_j = -r^{j-1}$ ,  $j \geq 2$ . These particular values lead to

$$x_2^{\alpha_2} x_3^{\alpha_3} \cdots = (-1)^{|\alpha|} r^{\langle \alpha \rangle}.$$

On the other hand, recursion (5) implies that

$$\tilde{h} = 1 - r\tilde{h}^3 - r^2\tilde{h}^5 - \cdots, \quad \text{where} \quad \tilde{h}(r) = h(-r, -r^2, \dots).$$

The right-hand side is almost a geometric series; we multiply the equation by  $1 - r\tilde{h}^2$  to arrive at

$$r\tilde{h}^2 + \tilde{h} - 1 = 0.$$

This can be solved for  $\tilde{h}$  and results in the well known generating function of  $(-1)^n c_n$ .  $\square$

A recurrence relation for  $\tilde{H}_{\mathbf{h}, \mathbf{t}}(\alpha)$  can be established in a similar manner.

*Proof of Theorem 3.* We recap that we are counting the factorizations with a given number of heads and tails. On a tree, a tail corresponds to a vertex of degree  $2j$  which has  $j$  free  $t$ -labeled edges attached to it. For example, on Fig. 6, there are  $t = 2$  tails. Similarly a head is a degree  $2j$  vertex with  $j$  free  $h$ -labeled edges attached (we omitted  $h$  labels from Fig. 6 and other figures to avoid clutter). Note that the top vertex can be both a tail and a head, although not simultaneously, at least for trees with more than one vertex. The root counts as being  $h$ -labeled and is always free. For example, the tree on Fig. 6(a) has the top vertex as its only head,  $h = 1$ . We also introduce a variable  $h'$  counting all heads excluding the top vertex. We will refer to it as the *reduced head count* and for the tree on Fig. 6(a) it is  $h' = h - 1 = 0$ , while for the tree on Fig. 6(b)  $h' = h = 1$ . We will first derive a recursion counting the trees with a given reduced head count and from there obtain the number of trees with full head count.

The tail and (reduced) head count are further specialized to count the number of heads and tails of a certain degree. Thus, in general,  $\mathbf{h}$ ,  $\mathbf{h}'$  and  $\mathbf{t}$  are infinite vectors with finitely many nonzero components. Let  $\phi$  be a partial generating function with respect to the tail and reduced head count

$$\phi(\alpha, \mathbf{v}, \mathbf{u}) = \sum_{\mathbf{h}'=(0,0,\dots)}^{\alpha} \sum_{\mathbf{t}=(0,0,\dots)}^{\alpha} \tilde{H}_{\mathbf{h}', \mathbf{t}}(\alpha) u_2^{h'_2} v_2^{t_2} u_3^{h'_3} v_3^{t_3} \cdots, \quad \phi(\mathbf{0}, \mathbf{v}, \mathbf{u}) = 1.$$

To establish the recursion relation we again consider breaking the tree into subtrees  $T_1, \dots, T_{2d+1}$  at the top vertex of degree  $2(d+1)$ , numbering the subtrees left to right. As before, the subtrees are characterized by vectors  $\alpha_1, \dots, \alpha_{2d+1}$ . We introduce a special notation for the sum of odd-indexed vectors and for the sum of even-indexed ones,

$$\alpha^o = \sum_{j=0}^d \alpha_{2j+1} \quad \alpha^e = \sum_{j=1}^d \alpha_{2j}. \quad (12)$$

The reduced head count  $\mathbf{h}' = (h_2, h_3, \dots)$  of the full tree can be obtained by summing the appropriate counts for the subtrees, namely

$$\mathbf{h}'(T) = \mathbf{h}'(T_1) + \sum_{j=1}^d (\mathbf{t}(T_{2j}) + \mathbf{h}'(T_{2j+1})). \quad (13)$$

Note that for the even-numbered subtrees, we need to add the number of tails rather than heads. This corresponds to a change in the labeling of the leaves on the subtrees with even index. On subtrees with odd index the first (leftmost) leaf is always  $t$ -labeled, while the first leaf of an even-numbered subtree is  $h$ -labeled, see Fig. 6(b) for an example.

For the tail count of the complete tree, the procedure is analogous, with the addition of the possible contribution of the top vertex. The top vertex is a tail if all the odd subtrees are empty, i.e.  $\alpha_{2j+1} = \mathbf{0}$ ,  $j = 0, \dots, d$ . Figure 6(b) shows a tree where the top vertex is a tail. Therefore,

$$\mathbf{t}(T) = \mathbf{t}(T_1) + \sum_{j=1}^d (\mathbf{h}'(T_{2j}) + \mathbf{t}(T_{2j+1})) + \delta_{\alpha^o, \mathbf{0}} \mathbf{e}_d.$$

Consequently  $\phi(\alpha, \mathbf{v}, \mathbf{u})$  is expressed in terms of functions  $\phi(\alpha_j, \mathbf{v}, \mathbf{u})$  generated by the subtrees,

$$\begin{aligned} \phi(\alpha, \mathbf{v}, \mathbf{u}) = \sum_{d \geq 1} \sum_{\alpha_1 \dots \alpha_{2d+1}} \phi(\alpha_1, \mathbf{v}, \mathbf{u}) \prod_{j=1}^d \phi(\alpha_{2j}, \mathbf{u}, \mathbf{v}) \phi(\alpha_{2j+1}, \mathbf{v}, \mathbf{u}) \\ \times (1 - (1 - v_{d+1}) \delta_{\alpha^o, \mathbf{0}}) \delta_{\alpha^o + \alpha^e, \alpha - \mathbf{e}_d}. \end{aligned} \quad (14)$$

It is important to observe that the functions  $\phi$  with even-indexed vectors  $\alpha_{2j}$  have their arguments  $\mathbf{u}$  and  $\mathbf{v}$  switched around. Calculating the generating function

$$f(\mathbf{x}, \mathbf{u}, \mathbf{v}) = \sum_{\alpha} \phi(\alpha, \mathbf{v}, \mathbf{u}) x_2^{\alpha_2} x_3^{\alpha_3} \dots,$$

we recover recurrence relation (9).

The complete head count can be obtained from the appropriate counts for the subtrees in a slight variation of (13),

$$\mathbf{h}(T) = \mathbf{h}'(T_1) + \sum_{j=1}^d (\mathbf{t}(T_{2j}) + \mathbf{h}'(T_{2j+1})) + \delta_{\alpha^e, \mathbf{0}} \mathbf{e}_d, \quad (15)$$

where  $\alpha^e$  was defined in equation (12).

The partial generating function with respect to the full head count is then

$$\begin{aligned}\psi(\alpha, \mathbf{v}, \mathbf{u}) &= \sum_{\mathbf{h}=(0,0,\dots)}^{\alpha} \sum_{\mathbf{t}=(0,0,\dots)}^{\alpha} \tilde{H}_{\mathbf{h},\mathbf{t}}(\alpha) u_2^{h_2} v_2^{t_2} u_3^{h_3} v_3^{t_3} \cdots \\ &= \sum_{d \geq 1} \sum_{\alpha_1 \cdots \alpha_{2d+1}} \phi(\alpha_1, \mathbf{v}, \mathbf{u}) \prod_{j=1}^d \phi(\alpha_{2j}, \mathbf{u}, \mathbf{v}) \phi(\alpha_{2j+1}, \mathbf{v}, \mathbf{u}) \\ &\quad \times [1 - (1 - v_{d+1})\delta_{\alpha^o, \mathbf{0}} - (1 - u_{d+1})\delta_{\alpha^e, \mathbf{0}}] \delta_{\alpha^o + \alpha^e, \alpha - \mathbf{e}_d}\end{aligned}$$

Opening the square brackets, using the recursion (14) for  $\phi$ , and the fact that  $\alpha^e = \mathbf{0}$  implies  $\phi(\alpha_{2j}, \mathbf{u}, \mathbf{v}) = 1$ , we obtain

$$\psi(\alpha, \mathbf{v}, \mathbf{u}) = \phi(\alpha, \mathbf{v}, \mathbf{u}) - \sum_{d \geq 1} \sum_{\alpha_1 \cdots \alpha_{2d+1}} \prod_{j=0}^d \phi(\alpha_{2j+1}, \mathbf{v}, \mathbf{u}) (1 - u_{d+1}) \delta_{\alpha^o, \alpha - \mathbf{e}_d}.$$

Calculating the full generating function  $g(\mathbf{x}, \mathbf{u}, \mathbf{v})$  we obtain

$$g(\mathbf{x}, \mathbf{v}, \mathbf{u}) = f - \sum_{d \geq 1} x_{d+1} (1 - u_{d+1}) f^{d+1},$$

which is the same as (7) after the substitution  $n = d + 1$ . We now transform this relation to form (8), which confirms that, in contrast to  $f(\mathbf{x}, \mathbf{v}, \mathbf{u})$ , the generating function  $g(\mathbf{x}, \mathbf{v}, \mathbf{u})$  is symmetric with respect to the exchange of  $\mathbf{u}$  and  $\mathbf{v}$ . We exchange  $\mathbf{u}$  and  $\mathbf{v}$  in (9) to obtain a recursion for  $\hat{f}$ ,

$$\hat{f} = 1 + \sum_{n \geq 2} x_n (\hat{f}^n - 1 + u_n) f^{n-1},$$

multiply it by  $f$  and rearrange,

$$f \hat{f} = f - \sum_{n \geq 2} x_n (1 - u_n) f^n + \sum_{n \geq 2} x_n \hat{f}^n f^n,$$

from which (8) immediately follows.  $\square$

## 5 Structure of factors

As seen in Section 2 and Theorem 1, a factorization can be visualized as a tree with vertices representing factors. With this identification in mind, we use the notions of vertex and cycle (factor) interchangeably throughout this section. Consider a rooted tree with the root at the top. Then a *child* of vertex  $v$  is a vertex adjacent to and below  $v$ , the *parent* is the vertex immediately above and the notions of *descendant* and *ancestor* are transitive extensions of the notions of child and parent

correspondingly. For example, a *descendant* of  $v$  can be recursively defined as a child of  $v$  or a child of a descendant of  $v$ .

In the following Lemmas we explore the connection between the descendant-ancestor relations on the tree and the relations introduced in Definition 1.

**Lemma 1.** *Let  $\sigma = (s_1 \dots s_{|\sigma|})$  be a cycle (vertex) in a tree representing a factorization of the  $n$ -cycle  $(12\dots n)$ . Then  $\sigma$  is increasing:  $s_1 < s_2 < \dots < s_{|\sigma|}$ .*

*Proof.* When reading the factorization off the tree as described in the proof of Theorem 1, the labels are assigned initially to the leaves of the tree in the counterclockwise order. The operation of removing a vertex and inheriting the labels preserves this ordering of the labels. If a new connected component is created by the removal operation, its labels are also ordered counterclockwise. Thus, when a vertex  $\sigma$  is removed, the labels of its leaves,  $t_{s_1}, \dots, t_{s_{|\sigma|}}$  satisfy  $s_1 < s_2 < \dots < s_{|\sigma|}$  (provided the starting index  $s_1$  is chosen appropriately). Thus each factor read off the tree is an increasing cycle.  $\square$

**Lemma 2.** *If the vertex  $\sigma$  is an ancestor of the vertex  $\rho$  then  $\sigma \geq \rho$  or  $\sigma \succeq \rho$ . If  $\sigma$  and  $\rho$  belong to different branches of a tree with  $\sigma$  to the left of  $\rho$  then  $\sigma \succeq \rho$ . Thus any two vertices are comparable in the sense of Definition 1.*

*Proof.* First consider the case when  $\sigma = (s_1 \dots s_{|\sigma|})$  is the parent of  $\rho = (r_1 \dots r_{|\rho|})$ . When reading the factorization off the tree as described in the proof of Theorem 1, either  $\sigma$  is processed first or  $\rho$  is. If  $\sigma$  (the parent) is processed first, then  $r_{|\rho|} = s_k$  for some  $2 \leq k \leq |\sigma|$  (see Fig. 5(d)). If the child  $\rho$  is processed first, we have  $s_k = r_1$ ,  $1 \leq k \leq |\sigma|$ . It is easy to see that the two cycles do not share any other entries.

Now consider the origin of the entries of  $\rho$ . They were either assigned initially (for example, labels  $t_1$  and  $t_3$  of the top vertex on Fig. 5), inherited from children (label  $t_4$  of the same vertex) or inherited from the parent (label  $t_6$  of the lower right vertex, Fig. 5(c)). Thus, possibly apart from the parent label, all labels of  $\rho$  come from the set of labels initially assigned to subtree consisting of  $\rho$  and its descendants. These labels are consecutive. If  $\rho$  is processed before its parent  $\sigma$ , then  $s_k = r_1 < r_2 < \dots < r_{|\rho|} < s_{k+1}$ . The last inequality is there only if  $k < |\sigma|$  and is true because otherwise there would be two labels,  $s_k$  and  $s_{k+1}$  that came to  $\sigma$  from the same subtree of  $\rho$  and its descendants. Thus  $\sigma \geq \rho$  if  $k < |\sigma|$  and  $\sigma \succeq \rho$  if  $k = |\sigma|$ , with  $\rho$  having a left connector. On the other hand, if  $\sigma$  is processed first,  $s_{k-1} < r_1 < r_2 < \dots < r_{|\rho|} = s_k$ , i.e.  $\sigma \geq \rho$  and  $\rho$  has a right connector.

We see that the parent and a child are always comparable. By transitivity, an ancestor is always comparable with a descendant, either with a connector or without. For two vertices  $\sigma$  and  $\rho$  belonging to different branches of a tree, suppose  $\sigma$  belongs to a branch to the left of the branch of  $\rho$ . Then the labels on both branches are consecutive and the labels of the left branch are smaller or equal than the labels of the right branch. The equality can only happen if the smallest label of the right branch passed up to its parent and then down to become the largest label of the left branch, see Fig. 7. In any case, we see that  $\sigma \succeq \rho$ .  $\square$

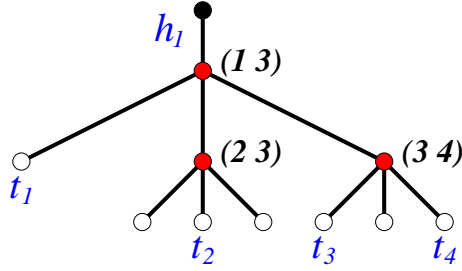


Figure 7: An example of a tree with vertices on different branches,  $(23) \succeq (34)$ , having a connector 3. According to the procedure in the proof of Theorem 1, vertex  $(34)$  gets processed first, whereas its label  $t_3$  moves up to its parent. Then the parent, vertex  $(13)$ , gets processed sending the label  $t_3$  down to its remaining child,  $(23)$ .

**Lemma 3.** *Let the cycle (vertex)  $\sigma$  be the parent of the cycle (vertex)  $\rho$  in a tree representing factorization of the  $n$ -cycle  $(12\dots n)$ . Then **either***

$$\sigma = \min\{\sigma' : \sigma' \succeq \rho\}. \quad (16)$$

**or**

$$\sigma = \max\{\sigma' : \sigma' \succeq \rho, \rho \text{ has a left connector to } \sigma'\}. \quad (17)$$

*Proof.* Most of this lemma has already been obtained in Lemma 2 and its proof. The only vertices  $\sigma'$  that satisfy  $\sigma' \succeq \rho$  are the ancestors of  $\rho$ . Of these, the minimal is the parent of  $\rho$ .

Now suppose  $\sigma \succeq \rho$ . Then  $\rho$  was processed before  $\sigma$  and has a left connector to it. This label is the last one in the code for  $\sigma$  and, therefore, cannot be passed to the parent of  $\sigma$ . Thus  $\rho$  does not have a connector to any of its ancestors, apart from its parent. It can have a connector to a descendant of  $\sigma$ , see Fig. 7, but of these  $\sigma$  is obviously the maximal vertex.  $\square$

*Proof of Theorem 4.* Necessity: all properties of cycles described in Theorem 4 have been proven in Lemma 1, Lemma 2 and Lemma 3, apart from the uniqueness of a connector. A connector can appear through either inheriting a label from the parent (right connector), or sending a label up to the parent (left connector). In both of these possibilities the edge linking the vertex to its parent is broken and no other connector can appear.

Sufficiency: given a set of cycles, we can look for a parent for each cycle using first equation (17) and then equation (16). We need to verify that only one cycle will not have a parent. To do it, we count the number of distinct entries appearing in the cycles. Each cycle  $\sigma$  contributes  $|\sigma|$  new entries minus one if it has a parent: the connector entry will be counted in the parent. The total number of entries is thus

$$\sum_{j=1}^m (|\sigma_j| - 1) + P,$$

where  $P$  is the number of parentless cycles. Since the total number of entries should be  $n$ , equation (10) implies that  $P = 1$ .

To fully reconstruct the tree now we represent each cycle  $\sigma = (s_1 \dots s_{|\sigma|})$  with a vertex with  $2|\sigma|$  edges, labeled  $h_{s_1}, t_{s_1}, \dots, h_{s_{|\sigma|}}, t_{s_{|\sigma|}}$  in counterclockwise order. Let  $\rho$  be a child of  $\sigma$  with a connector  $r$ . If it is a left connector, we join the edge of  $\rho$  marked  $h_r$  to the edge  $t_r$  of  $\sigma$ . If it is a right connector, we join  $t_r$  of  $\rho$  with  $h_r$  of  $\sigma$ . It is easy to see that for each  $\sigma$  there cannot be more than one child  $\rho$  with the same  $r$  as the same-side connector (otherwise one would be the child of the other, not of  $\sigma$ ). Thus the joining above is well-defined. Since we have precisely  $m - 1$  parent-child couples, the process stops when all vertices are connected together into a tree.

Uniqueness: the tree with given vertices is unique because equations (16) and (17) determine the child-parent couples uniquely and ordering of branches is specified by the counterclockwise ordering of the labels.  $\square$

## 6 Conclusions and outlook

The simple pictorial bijection introduced in Theorem 1 has allowed us to perform an in-depth analysis of the set of inequivalent minimal factorizations of the  $n$ -cycle. The next logical step is to apply similar ideas to inequivalent minimal transitive factorizations of a general permutation. Even just counting such factorizations is a hard task, with no published results for  $m > 2$ , where  $m$  is the number of cycles in the cycle representation of the target permutation ( $m = 1$  corresponds to an  $n$ -cycle). Still, our preliminary explorations showed that the ideas of the present manuscript provide if not a complete answer, then at the very least a method for deriving a recursion for the generating function for any finite  $m$ . However, these findings will be reported elsewhere.

## Acknowledgment

The authors acknowledge the discussions they had with P. Lima-Filho and F. Sottile and thank them for making useful suggestions. The authors are extremely grateful to J. Irving for sending us a copy of manuscript [10].

## References

- [1] A. Hurwitz, "Ueber Riemann'sche Flächen mit gegebenen Verzweigungspunkten," *Math. Ann.*, vol. 39, no. 1, pp. 1–60, 1891.

- [2] T. Ekedahl, S. Lando, M. Shapiro, and A. Vainshtein, “Hurwitz numbers and intersections on moduli spaces of curves,” *Invent. Math.*, vol. 146, no. 2, pp. 297–327, 2001.
- [3] S. Müller, S. Heusler, P. Braun, and F. Haake, “Semiclassical approach to chaotic quantum transport,” *New J. Phys.*, vol. 9, no. 1, p. 12, 2007.
- [4] G. Berkolaiko, J. M. Harrison, and M. Novaes, “Full counting statistics of chaotic cavities from classical action correlations,” *J. Phys. A: Math. Theor.*, vol. 41, no. 36, p. 365102 (12pp), 2008.
- [5] A. Erdélyi and I. M. H. Etherington, “Some problems of non-associative combinations. II,” *Edinburgh Math. Notes*, vol. 1941, no. 32, pp. 7–12, 1941.
- [6] W. T. Tutte, “The number of planted plane trees with a given partition,” *Amer. Math. Monthly*, vol. 71, pp. 272–277, 1964.
- [7] R. P. Stanley, *Enumerative combinatorics. Vol. 2*, vol. 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge: Cambridge University Press, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.
- [8] J. A. Eidswick, “Short factorizations of permutations into transpositions,” *Discrete Math.*, vol. 73, no. 3, pp. 239–243, 1989.
- [9] J. Q. Longyear, “A peculiar partition formula,” *Discrete Math.*, vol. 78, no. 1-2, pp. 115–118, 1989.
- [10] C. M. Springer, “Factorizations, trees, and cacti,” in *Eighth International Conference on Formal Power Series and Algebraic Combinatorics, University of Minnesota, June 25-29*, pp. 427–438, 1996.
- [11] J. Irving, “Minimal transitive factorizations of permutations into cycles.” arXiv:math/0610735v1 [math.CO], 2006.
- [12] J. Dénes, “The representation of a permutation as the product of a minimal number of transpositions, and its connection with the theory of graphs,” *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, vol. 4, pp. 63–71, 1959.
- [13] A. J. Schwenk and O. P. Lossers, “Problems and Solutions: Solutions of Elementary Problems: E3058,” *Amer. Math. Monthly*, vol. 93, no. 10, pp. 820–821, 1986.
- [14] P. Moszkowski, “A solution to a problem of Dénes: a bijection between trees and factorizations of cyclic permutations,” *European J. Combin.*, vol. 10, no. 1, pp. 13–16, 1989.
- [15] I. P. Goulden and S. Pepper, “Labelled trees and factorizations of a cycle into transpositions,” *Discrete Math.*, vol. 113, no. 1-3, pp. 263–268, 1993.

- [16] V. Strehl, “Minimal transitive products of transpositions—the reconstruction of a proof of A. Hurwitz,” *Sém. Lothar. Combin.*, vol. 37, pp. Art. S37c, 12 pp. (electronic), 1996.
- [17] I. P. Goulden and D. M. Jackson, “Transitive factorisations into transpositions and holomorphic mappings on the sphere,” *Proc. Amer. Math. Soc.*, vol. 125, no. 1, pp. 51–60, 1997.
- [18] I. P. Goulden and D. M. Jackson, “Transitive factorizations in the symmetric group, and combinatorial aspects of singularity theory,” *European J. Combin.*, vol. 21, no. 8, pp. 1001–1016, 2000.
- [19] I. P. Goulden, D. M. Jackson, and F. G. Latour, “Inequivalent transitive factorizations into transpositions,” *Canad. J. Math.*, vol. 53, no. 4, pp. 758–779, 2001.
- [20] M. Gardner, “Mathematical games,” *Scientific American*, vol. 201, Dec 1959.
- [21] A. Bogomolny, “What, how, and the web: play with braids and knots from interactive mathematics miscellany and puzzles.” <http://www.cut-the-knot.org/SimpleGames/TransExample.shtml>.