ISOMORPHISM OF WEIGHTED TREES AND STANLEY'S CONJECTURE FOR CATERPILLARS

MARTIN LOEBL AND JEAN-SÉBASTIEN SERENI

Abstract. This paper contributes to a programme initiated by the first author: 'How much information about a graph is revealed in its Potts partition function?'. We show that the W-polynomial distinguishes non-isomorphic weighted trees of a *good* family. The framework developed to do so also allows us to show that the W-polynomial distinguishes non-isomorphic caterpillars. This establishes Stanley's conjecture for caterpillars, an extensively studied problem.

1. Introduction

Consider the following data set $D(T)$ associated with a tree T: for every integer n and every partition P of n, we are given the number of subsets X of edges of T such that P is equal to the multiset formed by the orders of the components of $T - X$. Note that this number is 0 if n is not the number of vertices of T. Note also that if P is composed of t integers, the corresponding subsets X, if any, all have cardinality $t-1$. For instance, one can determine the number of vertices of T by checking, for each positive integer n, whether the trivial partition $\{n\}$ returns a non-zero value (which, necessarily, will be 1). Once the number n of vertices of T is known, the number of leaves of T is precisely the number returned by the partition $\{n-1, 1\}$, which corresponds to the number of edges e such that $T - e$ has one component of order 1. The problem is to know whether this information distinguishes non-isomorphic trees. In other words, if T and T' are two trees such that $D(T) = D(T')$, is it true that necessarily T and T' are isomorphic? That such a reconstruction is always possible was suggested by different authors. We note that there could be non-constructive proofs of the statement. Thus it is a different (harder) problem to be able to effectively recover the tree T from the knowledge of $D(T)$. We explain in subsections 2.1, 2.2 and 2.3 why studying the strength of the information contained in $D(T)$ for an arbitrary tree T helps to understand the strength of the partition function of the Potts model in a magnetic field, for general graphs.

1.1. **State of the Art.** Extensive efforts were dedicated (personal communication with Noble) to proving that $D(T)$ distinguishes non-isomorphic caterpillars — a *caterpillar* is a tree where all edges not incident with a leaf form a path, and a *leaf* is a vertex of degree one. Part of the Ph.D. thesis of Zamora [32] (under the supervision of M. L.) is dedicated to this problem. In addition, Aliste-Prieto and Zamora [2], established the statement restricted to the class of proper caterpillars: a caterpillar is *proper* if every vertex is a leaf or adjacent to a leaf. Prior to that, partial results had been obtained by Martin, Morin and Wagner [18] who had established

Date: April 2, 2017.

2010 *Mathematics Subject Classification.* 05C31, 05C60.

Key words and phrases. W-polynomial, tree, graph reconstruction, graph isomorphism, U-polynomial, Stanley's conjecture, Potts partition function.

This work was done within the scope of the International Associated Laboratory STRUCO.

The authors were partially supported by the Czech Science Foundation under the contract number P202-13-21988S (M. L.) and by P.H.C. Barrande 31231PF of the French M.A.E. (J.-S. S.).

the statement for a subclass of proper caterpillars (where no two non-leaf vertices are adjacent to the same number of leaves) and also to the class of spiders, which is composed of all trees with a unique vertex of degree greater than two. Other related results can be found in the undergraduate thesis by Fougere [11] and the MSc thesis by Morin [19]. Finally, it is reported that Tan checked by computer that Stanley's conjecture [26], which we present in Subsection 2.2, is true for trees with at most 23 vertices (see [18, p. 238]).

1.2. **Main Contribution.**

We solve affirmatively Stanley's conjecture restricted to the class of caterpillars. We also investigate a weighted version of the problem, bearing in mind its connections with graph polynomials, graph colouring and the Potts model. First we summarise the background and motivations.

2. Motivation

In this section we summarise the background (the Noble and Welsh conjecture and the Stanley conjecture) and describe our motivation.

2.1. **The Noble and Welsh Conjecture.**

Motivated by the combinatorial aspects of the relationship between chord diagrams and Vassiliev invariants of knots, Noble and Welsh [22] introduced a polynomial of weighted graphs, the W-polynomial, which includes several specialisations in combinatorics, such as the Tutte polynomial, the matching polynomial (of ordinary graphs) and the polymatroid polynomial of Oxley and Whittle [23]. We need to introduce some terminology to define W.

A *weighted graph* is a graph $G = (V, E)$ together with a function $w\colon V \to \mathbf{Z}^+$. The *weight* of a subset V' of vertices is $w(V') \coloneqq \sum_{v\in V'} w(v)$. If $A \subseteq E$, we let $c_V(A)$ be the number of components of the graph (V, A), where we may omit the subscript when there is no risk of confusion. Further, let $n_1, \ldots, n_{c(A)}$ be the weights of the vertex sets of these components, listed in decreasing order: $n_1 \geqslant \cdots \geqslant n_{c(A)}$. We write $x(A)$ to mean $\prod_{i=1}^{c(A)} x_{n_i}$. Let

$$W_G(z, x_1, x_2, \ldots) \coloneqq \sum_{A\subseteq E} x(A)(z-1)^{|A|-|V|+c(A)}.$$

In particular, W_G depends on z if and only if G contains a cycle [22, Proposition 5.1-1)]. Unlike the Tutte polynomial, the W-polynomial is $\#P$-hard to compute even for trees [22, Theorems 7.3 and 7.12] and for complete graphs [22, Theorems 7.11 and 7. 14].

In the case of unweighted graphs, which corresponds here to the weight function w being identically 1, Noble and Welsh refers to the W-polynomial as to the *U-polynomial.* While computing W is hard for complete graphs, Annan [1] proved that $U_{K_n}(z, x_1, x_2, \ldots)$ can be computed in polynomial time, which is also the case for the Tutte polynomial, for instance. However, U also exhibits differences with the Tutte polynomial: while finding two non-isomorphic graphs with the same Tutte polynomial is easy, the same problem is harder for U. Sarmiento [24] managed to achieve such a construction, but the question remains open for trees: does the U-polynomial distinguishes non-isomorphic trees? That this is the case became known as the *Noble and Welsh conjecture.* This is clearly equivalent to our initial problem: 'Does $D(T)$ distinguish non-isomorphic trees?'

Noble and Welsh also discovered a very interesting specialization of U: they demonstrated the U-polynomial to be equivalent to the *symmetric function generalisation of the bad colouring polynomial*, a function introduced by Stanley [26].

2.2. **The Stanley Conjecture.** To introduce Stanley's isomorphism conjecture let us first define graph colouring. A *colouring* of a graph $G = (V, E)$ is a mapping $s\colon V \to \mathbf{N}^+$. We define $b(s)$ to be the number of *monochromatic edges* in s, that is, the number of edges uv such that $s(u) = s(v)$. The mapping s is a *k-colouring* if $s(V) \subseteq \{1, \dots, k\}$ and s is *proper* if $b(s) = 0$, that is, $s(u) \neq s(v)$ whenever u and v are two adjacent vertices of G. We let $\mathrm{Col}(G; k)$ be the set of proper k-colourings of G and $\mathrm{Col}(G)$ be the set of all proper colourings of G.

In the mid 1990s, Stanley [26] introduced the *symmetric function generalization of the chromatic polynomial*, defined to be

$$X_G(x_1, x_2, \dots) \coloneqq \sum_{s \in \mathrm{Col}(G)} \prod_{v \in V} x_{s(v)}.$$

This is a homogeneous symmetric function in $(x_1, x_2, \dots)$ of degree $|V|$. As is expectable, X_G does not distinguish non-isomorphic graphs: there exist two non-isomorphic graphs on 5 vertices with the same function X. However, Stanley [26] asked whether the polynomial X_G distinguishes non-isomorphic trees. The assertion that it does became known as *Stanley's isomorphism conjecture.*

Further, Stanley [27] later initiated the study of a common generalisation of X and the Tutte polynomial, namely the *symmetric function generalisation of the bad colouring polynomial*, defined for every graph $G = (V, E)$ by

$$X_G(t, x_1, x_2, \dots) \coloneqq \sum_{s\colon V \to \mathbf{N}^+} (1 + t)^{b(s)} \prod_{v \in V} x_{s(v)}.$$

Note that the sum runs over all colourings of G, not only the proper ones. Noble and Welsh [22, Theorem 6.2] proved $X_G(t, x_1, x_2, \dots)$ to be a specialisation of the U-polynomial of G.

The equivalence between Stanley's conjecture and our question was clarified by Thatte. Let us introduce the following notation: given a tree $T = (V, E)$ and two integer vectors $\mathbf{v} = (v_1, \dots, v_r)$ and $\mathbf{e} = (e_1, \dots, e_r)$, let $\theta(T, \mathbf{v}, \mathbf{e})$ be the number of (ordered) partitions $(V_1, \dots, V_r)$ of V such that for each $i \in \{1, \dots, r\}$, the subgraph of T induced by V_i has precisely v_i vertices and e_i edges. Thatte [29, Lemma 4.11] established that θ can be computed from X_T.

2.3. **Loebl's Conjectures.** Loebl [15] introduced the q-chromatic functions. Let $k \in \mathbf{N}$. The *q-chromatic function* of a graph $G = (V, E)$ is

$$M_G(k, q) \coloneqq \sum_{s \in \mathrm{Col}(G;k)} q^{\sum_{v \in V} s(v)}. \tag{2.1}$$

It is known [15] that

$$M_G(k, q) = \sum_{A \subset E} (-1)^{|A|} \prod_{C \in \mathscr{C}(A)} (k)_{q^{|C|}},$$

where $\mathscr{C}(A)$ is the set of components of the spanning subgraph (V, A) and $|C|$ is the number of vertices in the component C. Moreover Loebl also introduced the *q-dichromate*, defined as

$$B_G(x, y, q) \coloneqq \sum_{A \subset E} x^{|A|} \prod_{C \in \mathscr{C}(A)} (y)_{q^{|C|}}.$$

Loebl [15] conjectured the following.

- The q-dichromate is equivalent to the U-polynomial.
- The U-polynomial distinguishes non-isomorphic chordal graphs.

There could be a close link between the latter conjecture and that of Stanley: chordal graphs have a very distinguished tree structure. Indeed, a folklore theorem [4] states that the class of chordal graphs is precisely the class of intersection graphs of subtrees of a tree, that is, for each chordal graph G, there exists a tree T and a mapping f that assigns to each vertex of G a subtree T such that: two vertices u and v of G are adjacent if and only if $f(u) \cap f(v) \neq \varnothing$.

The motivation for Loebl's conjectures is formula (2.2) below, which connects the k-state Potts model partition function and the q-dichromate.

Potts model. We consider a standard model where magnetic materials are represented as lattices: vertices are atoms and weighted edges are nearest-neighbourhood interactions. We assume that each atom has one out of k possible magnetic moments, for a fixed positive integer k. Thus we let $S := \{0, \dots, k-1\}$. Every element of S is called a *spin*. A *state* of a graph $G = (V, E)$ is then an assignment of a single spin to each vertex of G, that is a function $s\colon V \to S$. We assume that all the coupling constants (nearest-neighbourhood interactions) are equal to a constant J. For each state s, the *Potts model energy of the state s* is then $E(P^k)(s) := \sum_{uv\in E} J\delta(s(u), s(v))$ where, as is customary, δ is the Kronecker delta function defined by $\delta(a, b) := 1$ if $a = b$ and $\delta(a, b) := 0$ otherwise. The *k-state Potts model partition function* is then

$$\sum_{s:V\to S} M(s, J)e^{E(P^k)(s)}$$

where $M(s, J)$ is a function describing the magnetic field contribution.

Loebl proved that for each real J,

$$B_G(e^J - 1, k, q) = \sum_{s:V\to S} q^{\sum_{v\in V} s(v)} e^{E(P^k)(s)}. \tag{2.2}$$

This means that the q-dichromate specializes to the k-state Potts model partition function with a certain magnetic field contribution.

Recently a variant of the q-dichromate, $B_{r,G}(x, k, q)$, was proposed by Klazar, Loebl and Moffatt [14]:

$$B_{r,G}(x, k, q) := \sum_{A\subseteq E} x^{|A|} \prod_{C\in\mathscr{C}(A)} \sum_{i=0}^{k-1} r^{|C|q^i}.$$

They established that if $(k, r) \in \mathbf{N}^2$ with $r > 1$ and $x := \mathrm{e}^{\beta J} - 1$, then

$$B_{r,G}(x, k, q) = \sum_{\sigma\colon V\to S} \mathrm{e}^{\beta \sum_{uv\in E(G)} J\delta(\sigma(u),\sigma(v))} r^{\sum_{v\in V} q^{\sigma(v)}}. \tag{2.3}$$

Hence $B_{r,G}(x, k, q)$ is the k-state Potts model partition function with magnetic field contribution $r^{\sum_{v\in V} q^{\sigma(v)}}$. They also proved that $B_{r,G}$ is equivalent to U_G, which can be seen as a first step towards Loebl's programme:

The polynomial U_G is equivalent to the Potts partition function of G with a magnetic field contribution.

A well-known fact is that the isomorphism problem for general graphs is equivalent to the isomorphism problem restricted to chordal graphs: given a graph $G = (V, E)$, consider the chordal graph $G' = (V', E')$ so that $V' := V \cup E$ and $E' = \binom{V}{2} \cup \{\{u, e\}, \{v, e\} : \{u, v\} = e \in E\}$. It clearly holds that G and H are isomorphic if and only if G' and H' are isomorphic. It thus seems particularly interesting to determine whether the U-polynomial does distinguish non-isomorphic chordal graphs, as conjectured by Loebl. If true, we would obtain a surprising conclusion:

The Potts partition function with a magnetic field contribution contains essentially (modulo a simple preprocessing) all the information about the underlying graph.

In that respect, it seems natural to study weighted trees. The tree mentioned in the characterisation of the class of chordal graphs can be chosen to be a *clique-tree*, where the vertices of the tree are the maximal cliques of the graph. Now, if v is a vertex of a weighted tree with weight $w(v)$, one can think of v as a clique of order $w(v)$, thus obtaining an unweighted chordal graph. This is what motivates to work in the (seemingly harder) setting of weighted trees.

2.4. **Main Results.** Two weighted graphs are isomorphic if there is an isomorphism of the graphs that preserves the vertex weights. A purpose of this work is to prove that the W-polynomial distinguishes non-isomorphic *weighted trees* when restricting to collections of weighted trees satisfying some properties made precise later. We call any such collection a *good family.*

Let (T, w) be a weighted tree. We write $V(T)$ and $E(T)$ for the vertex set and the edge set of T, respectively. We define $\mathrm{Ex}(T)$ to be the multi-set composed of all the vertex weights (with multiplicities) of T. If $e \in E(T)$, then $T - e$ is the disjoint union of two trees, which we consider to be weighted and rooted at the endvertex of e that they contain. A rooted weighted tree (S, w_S) is a *shape* of (T, w) if $2 \leqslant |V(S)| \leqslant |V(T)| - 2$ and there exists an edge $e \in E(T)$ such that S is one of the two components of $T - e$; moreover w_S is the restriction of w to the vertex set of S. We consider S rooted at the end-vertex of e. We usually shorten the notation and write S for the shape (S, w_S). In a tree, a vertex of degree one is called a *leaf.*

Definition 2.1. A set $\mathcal{T}$ of weighted trees (T, w) is *good* if it satisfies the following properties.

1. If a vertex of T is adjacent to a leaf, then all its neighbours but possibly one are leaves.
2. If v is a leaf or has a neighbour that is a leaf, then $w(v) = 1$.
3. Let $(T, w), (T', w') \in \mathcal{T}$ and let S be a shape of T and such that $w(S) \leqslant w(T)/2$. Let S' be a shape of T' such that $\mathrm{Ex}(S') = \mathrm{Ex}(S)$. Then S' and S are isomorphic as rooted trees.

Theorem 1. *The W-polynomial distinguishes non-isomorphic weighted trees in any good set.*

Our proof of Theorem 1 is not constructive in the sense that we are not able to reconstruct the weighted tree (T, w) from $W_{(T,w)}$. The difficulty in proving the theorem is that while the main defining property of a good family is about shapes, the W-polynomial does not "see" shapes. However, shapes turn out to be a useful and rather powerful notion: it allowed us to unlock the case of general caterpillars, thereby confirming Stanley's conjecture for the class of (general) caterpillars.

Theorem 2. *Each caterpillar can be reconstructed from its U-polynomial.*

Note that Theorem 2, contrary to Theorem 1, allows for a full reconstruction of the tree.

3. The Structure of the Proofs

We write down a procedure and with its help prove both theorems. The rest of the paper then describes our realisation of the procedure. We fix a good set of weighted trees and, from now on, we say that a weighted tree is *good* if it belongs to this set.

A *j-form* is an isomorphism class of rooted weighted trees with total weight j. Thus a j-form F is a collection of rooted weighted trees and, viewing a shape of a tree T as a rooted weighted tree, a shape can belong to a j-form. Note in particular that two shapes S and S' of a tree belong to a given j-form F if and only if S and S' are isomorphic as rooted trees. We start with an observation.

Observation 3.1. *Let (T, w) be a weighted tree such that every leaf has weight 1. Assume that we know the total weight $w(T)$ of T and that, for each $j \leqslant w(T)/2$ and each j-form F, we know the number of shapes of (T, w) that belong to F. Then we know T.*

Proof. We use an easy but important observation that if two shapes of T have a common vertex then one is contained in the other. We order the shapes of (T, w) of weight at most $w(T)/2$ decreasingly according to their weights. Let m be the maximum weight of such a shape of T and let $S_1, \ldots, S_a$ be the shapes with weight m. Note that we know precisely these a trees. In addition, either the shapes $S_1, \ldots, S_a$ are joined in T to the same vertex, or $a = 2$ and $m =$

$w(T)/2$. In the latter case ($m = w(T)/2$) we know that T consists of the two weighted rooted trees S_1 and S_2 (each of weight m) with an edge between their roots: this ends the proof for this case. Assume that $m < w(T)/2$. We let r be the additional vertex to which we link each of $S_1, \dots, S_a$.

We show by descending induction on $j \in \{2, \dots, m\}$ that we know the subtree of T induced by all shapes of T with weight in $\{j, \dots, \lfloor W(T)/2 \rfloor\}$. The induction has thus been initialized above, so assume that $j \leqslant m-1$. Let $S_1, \dots, S_t$ be the shapes of T with weight in $\{j+1, \dots, \lfloor W(T)/2 \rfloor\}$. Note that we know, in particular, each of these t trees. The shapes of T of weight equal to j, if any, are either shapes of $S_1, \dots, S_t$ or joined to r by an edge from their root. Fix a j-form F. Since we do know the total number of shapes belonging to F and contained in each of $S_1, \dots, S_t$ (because we know precisely those subtrees), we can deduce the number of shapes that belong to F and are attached to r. As this argument applies to all j-forms F, we infer that we know the subtree of T formed by all shapes with weight contained in $\{j, \dots, \lfloor w(T)/2 \rfloor\}$. The reconstruction of T is almost finished: letting w_0 be the total weight of the tree we built so far, it only remains to add $w(T) - w_0$ new leaves, each joined to the vertex r. This concludes the proof. □

Let (T, w) be a weighted tree. Let $\alpha(T) = (\alpha_1, \dots, \alpha_n)$ be the weights of the shapes of T, with $\alpha_1 < \cdots < \alpha_n$. The definition of a shape implies that $\alpha_1 \geqslant 2$. We shall consider both partitions of the integer $w(T)$ and partitions of the tree T. To distinguish between them clearly, partitions of an integer are referred to as *expressions*. For each partition P of T, the weights of the parts of T form an expression of $w(T)$, which we call the *characteristic* of P.

- A *j-expression of an integer m* is a partition of m where one of the parts is equal to $m - j$. In particular, if S is a shape of T with weight α_j, then $(\mathrm{Ex}(S), w(T) - \alpha_j)$ is an α_j-expression of $w(T)$.
- A *j-partition* of T is a partition of T whose characteristic is a j-expression of $w(T)$. In other words, one of the components of the partition has order $w(T) - j$.
- A j-partition $(T_0, \dots, T_k)$ of T with $w(T_0) = w(T) - j$ is *shaped* if there exists an edge e of T such that T_0 is one of the components of $T - e$. Any such edge e is then *associated* to $(T_0, \dots, T_k)$.
- If S is a shape of T with weight α_j and vertex set $V(S) = \{v_1, \dots, v_s\}$, we define $P(S)$ to be $(V(T) \setminus V(S), \{v_1\}, \dots, \{v_s\})$, which is a shaped α_j-partition of T.

For an expression E of a positive integer, we let $\theta(T, w, E)$ be the number of partitions of (T, w) with characteristic E. Note that this number is 0 if E is not an expression of $w(T)$. We notice that, for each expression E, the polynomial $W_{(T,w)}$ determines $\theta(T, w, E)$. We note that among the partitions of T corresponding to a given expression, some are shaped and others are not. If all the vertex weights are equal to one, we abbreviate $\theta(T, w, E)$ as $\theta(T, E)$. The proof of Theorem 1 relies on the following procedure.

Procedure 1.
INPUT: The polynomial $W_{(T,w)}$, an integer $j \in \{\alpha_1 + 1, \dots, w(T)/2\}$, a j-expression E and, for each $j' < j$ and each j'-form F, the number of shapes S of T that are isomorphic to a member of F as weighted *but not rooted* trees.
OUTPUT: The number of shaped j-partitions of T with characteristic E.

Let us see how this procedure allows us to establish Theorem 1.

Proof of Theorem 1. Fix two good weighted trees (T, w) and (T', w') with $W_{(T,w)} = W_{(T',w')}$. By Observation 3.1, (T, w) and (T', w') are isomorphic if $w(T) = w'(T')$ and for each j-form F where $j \leqslant w(T)/2$, the numbers of shapes of T and of T' that belong to F are equal. To establish this, first note that the vector $\alpha(T) = (\alpha_1, \dots, \alpha_n)$ can be computed from $W_{(T,w)}$, since the coordinates correspond to the partitions of T into two subtrees (each with at least two vertices). Thus $\alpha(T') = \alpha(T)$.

We prove by induction on $j \in \{\alpha_1, \dots, \lfloor w(T)/2 \rfloor\}$ that for every j-form F, the numbers of shapes of T and of T' that belong to F are the same. So suppose first that $j = \alpha_1$. Recall that $\alpha_1 \geqslant 2$. Furthermore, a shape S of T or T' belongs to an α_1-form if and only if S is the star on α_1 vertices rooted at its centre. This is because the leaves and their neighbours have weight 1. It follows that the number of shapes of T of weight α_1 can be calculated from $W_{(T,w)}$ and thus this number is the same for (T', w').

Now let $j \in \{\alpha_1 + 1, \dots, \lfloor w(T)/2 \rfloor\}$. We assume that the following statement is true for every $j' \in \{\alpha_1, \dots, j-1\}$ and we establish it for $j' = j$.
'For every j'-form F, the numbers of shapes of T and of T' that belong to F are the same.'

To this end, we set a partial order on the j-forms, which allows us to link tree partitions with j-forms. Given a j-form F, we define $\mathrm{Ex}(F)$ to be $\mathrm{Ex}(f)$ for an arbitrary representative f of F. (This definition is valid, since all representatives of a j-form are isomorphic rooted weighted trees.) A j-form F' is *smaller than* a j-form F if $\mathrm{Ex}(F')$ is a proper refinement of $\mathrm{Ex}(F)$. If $P = (T_0, \dots, T_k)$ is a shaped j-partition of T where $w(T_0) = w(T) - j$, we define $S(P)$ to be the shape of T formed by the union of all parts of T different from T_0, that is, $S(P) \coloneqq \cup_{i=1}^{k} T_i$.

A key observation is that if P is a shaped j-partition of T with characteristic $\mathrm{Ex}(F)$ for some j-form F, then $\mathrm{Ex}(S(P))$ is a refinement of $\mathrm{Ex}(F)$, possibly equal to $\mathrm{Ex}(F)$: actually, there is equality if and only if S belongs to F.

We are now ready to argue the final step of the proof. If F is a j-form, let $n_T(F)$ be the number of shapes of T that belong to F; we use a similar notation for T'. Fix an arbitrary j-form F: our goal is to prove that $n_T(F) = n_{T'}(F)$.

We proceed by induction on the j-form F considered (with respect to the partial order defined above). So we first deal with the case where T has no shape that belongs to a j-form F' such that $\mathrm{Ex}(F')$ is a proper refinement of $\mathrm{Ex}(F)$. We demonstrate the following assertion, which readily implies that $n_T(F) = n_{T'}(F)$ for the case considered.

Assertion 3.2. *The number of shaped j-partitions of T with characteristic* $\mathrm{Ex}(F)$ *is equal to $n_T(F)$.*

To establish Assertion 3.2, suppose first that $n_T(F) = 0$. Then T has no shaped j-partition with characteristic $\mathrm{Ex}(F)$. So assume now that $n_T(F) > 0$. Each shape of T that belongs to F provides exactly one shaped j-partition of T with characteristic $\mathrm{Ex}(F)$. On the other hand, if P is a shaped j-partition of T with characteristic $\mathrm{Ex}(F)$, then $\mathrm{Ex}(S(P))$ is a refinement of $\mathrm{Ex}(F)$, which by our hypothesis on F must be equal to $\mathrm{Ex}(F)$. Hence $S(P)$ gives rise to precisely one shaped j-partition of T with characteristic $\mathrm{Ex}(F)$, namely P. As $\mathrm{Ex}(F) = \mathrm{Ex}(S(P))$ and T is good, it follows that $S(P)$ belongs to F, which ends the proof of Assertion 3.2.

For the general case, we may now assume that $n_T(F') = N_{T'}(F')$ for every j-form F' such that $\mathrm{Ex}(F')$ is a proper refinement of $\mathrm{Ex}(F)$. Observe that for each j-form F' with $F' < F$, each shape of T that belongs to F' gives rise to a certain number of shaped j-partition of T with characteristic $\mathrm{Ex}(F)$, and this number depends only on F'. Thus the number of shaped j-partitions of T with characteristic $(\mathrm{Ex}(F), w(T) - j)$ such that $S(P) \notin F$ depends only on the multi-set $\{n_T(F') : F' < F\}$. As $\{n_T(F') : F' < F\} = \{n_{T'}(F') : F' < F\}$, the conclusion follows. □

As we see next, the notion of a shape and Procedure 1 turn out to be essential tools to study Stanley's conjecture restricted to caterpillars.

4. Caterpillars

In this section, we only consider weights to be 1; since there is then no risk of confusion, we abbreviate $|V(T)|$ as $|T|$ for every tree T. Let T be a caterpillar with at least three vertices. The *spine* of T is the unique path P of T such that every leaf of T is at distance exactly one from a vertex of P.

Before proving Theorem 2, we formalize a simple but crucial observation, which is used repeatedly and implicitly in the proof of Theorem 2.

Observation 4.1. *Every shape of a caterpillar T is rooted at a vertex of the spine of T.*

Given a rooted tree T, a shape T' of T is *strict* if either $T' = T$ or T' does not contain the root of T. If T is a caterpillar, all vertex weights being 1, and E is an expression of $s \leqslant |T|/2+1$ so that no part of E is equal to $|T|-s$, then we define $\theta_s(T,E)$ to be the number of **shaped** s-partitions of T with expression $(|T|-s, E)$. We are now ready to proceed with the proof of Theorem 2.

Proof of Theorem 2. Let S_k be the star on k vertices — thus S_1 is a single vertex. We always consider a star to be rooted at its center, and in the rest of the proof the root of a tree is always excluded from the set of its leaves: this means that the star S_k contains precisely $k-1$ leaves even if k equals 2. If T is a rooted tree then we define $S_k \to T$ to be the tree rooted at the center of S_k and obtained by joining the root of T to that of S_k by an edge. Hence if T is a rooted caterpillar, then $S_k \to T$ is also a rooted caterpillar.

Let T be a caterpillar. We proceed by induction on the number of vertices of T, the theorem being true if $|T| < 4$. We now deal with the inductive step. As before, we note that the vector $\alpha(T) = (\alpha_1, \ldots, \alpha_n)$ can be computed from U_T, since the coordinates correspond to the partitions of T into two subtrees (each with at least two vertices). We prove by induction on $j \in \{\alpha_1, \ldots, \lfloor |T|/2 \rfloor\}$ that for every j-form F, we can deduce from U_T the number of shapes of T that belong to F. Observation 3.1 ensures then that we can reconstruct T. Analogously as in the previous proof the number of shapes of T of size α_1 can be calculated from U_T. This number is one or two since T is a caterpillar.

It follows from Observation 4.1 that for every integer $j \in \{2, \ldots, \lfloor |T|/2 \rfloor\}$, the number of shapes of T with j vertices belongs to $\{0,1,2\}$. We proceed inductively and, at each step of the inductive process, we update our knowledge of the two ends of T, by increasing the size of our knowledge of (at least) one end of T. It is important to note that to know the number of shapes of T that belong to a given j-form F for some $j \geqslant 2$, it is enough to know both ends of T of order j. At any given step, we let R_1 and R_2 be the currently known shapes of the two ends of T. Hence after the first step $R_1 = S_{\alpha_1}$ and $R_2 = \varnothing$ or $R_2 = R_1$, regarding whether $\theta_s(|T|-\alpha_1, \alpha_1)$ equals 1 or 2. (As reported earlier, this number can be deduced from the U-polynomial of T.)

Let $j \in \{\alpha_1+1, \ldots, \lfloor |T|/2 \rfloor\}$. We assume that for each $j' \in \{\alpha_1, \ldots, j-1\}$ and each j'-form F we know the number of shapes of T that belong to F. Let us establish this last statement for $j' = j$. If $j \notin \{\alpha_2, \ldots, \alpha_n\}$, then we know that the sought number is 0, by the definition of $(\alpha_1, \ldots, \alpha_n)$. So we suppose now that $j = \alpha_k$ for some integer $k \in \{2, \ldots, n\}$. We set $m \coloneqq \alpha_k - \alpha_{k-1}$. (Recall that this number can be deduced from the U-polynomial.) Let $\alpha_{k-1} = |R_1| \geqslant |R_2|$, with R_2 possibly empty. Set $p \coloneqq \alpha_k - |R_2|$, let $R_1' \coloneqq S_m \to R_1$ and $R_2' \coloneqq S_p \to R_2$.

We distinguish three cases.

[(1)] Let T have two α_k-shapes.
Then we update both R_1 and R_2, that is, we set $R_1 \coloneqq R_1'$ and $R_2 \coloneqq R_2'$.

[(2)] Let T have exactly one α_k-shape, i.e., either R_1' or R_2'. Moreover let R_1' and R_2' be not isomorphic as unrooted trees.

We recall that $\alpha_k \leqslant |T|/2$. As $|R'_i| < |T|$, we know by induction that $U_{R'_1} \neq U_{R'_2}$. Hence there is an expression E' of $|R'_1| = \alpha_k$ such that $r_1 \coloneqq U_{R'_1}(E') \neq r_2 \coloneqq U_{R'_2}(E')$. Further, we observe that necessarily, $\theta_s(T, E') \in \{r_1, r_2\}$. Therefore, there is a unique $i \in \{1, 2\}$ such that $\theta_s(T, E') = r_i$ and we can determine it by Procedure 1. We set $R_i \coloneqq R'_i$ and leave R_{3-i} unchanged.

Before going further we introduce some terminology. For $i \in \{1, \dots, \ell\}$, let m_i be an integer and E_i an expression of m_i.

- We define $[E_1, \dots, E_\ell]$ to be the expression of $\sum_{i=1}^{\ell} m_i$ equal to the concatenation of $E_1, \dots, E_s$.
- Let s be a non-negative integer and assume that at least two integers among m_1, m_2 and s are positive. We define $d(T, E_1, E_2, s)$ to be the number of partitions of T with characteristic $(E_1, E_2, 1^s, |T| - s - m_1 - m_2)$ such that the parts of E_1 partition an m_1-shape S_1 of T, the parts of E_2 partition an m_2-shape S_2 of T disjoint from S_1, and the s singletons are leaves of T (outside of S_1 and S_2).
- Let E be an expression. We set
$$d(T, E) \coloneqq \sum_{E=[E_1,E_2,1^s]} d(T, E_1, E_2, s).$$
We observe that if E is an expression of $m \leqslant |T|/2 + 1$ so that no part of E is equal to $|T| - m$, then $\theta(T, [E, |T| - m]) = \theta_s(T, E) + d(T, E)$.

[(3)] Let T have exactly one α_k-shape, i.e., either R'_1 or R'_2. Moreover let R'_1 and R'_2 be isomorphic as unrooted trees.
We know, for each α_k-form F, the number of shapes of T that belong to F. If $k = n$ then necessarily T consists of R_1 and R_2 joined by an edge between their roots and we can reconstruct T. Hence set $q \coloneqq \alpha_{k+1} - \alpha_k$.

By Procedure 1, we know for each α_{k+1}-expression E the number of shaped α_{k+1}-partitions of T with characteristic E. If $\alpha_{k+1} > |T|/2$, then T has a unique α_{k+1}-shape and the possible candidates for the α_{k+1}-shape of T are mutually non-isomorphic as unrooted trees. This case can thus be solved similarly as case (2). Hence from now on we assume that $\alpha_{k+1} \leqslant |T|/2$.

There are four candidates for an α_{k+1}-shape of T, namely $S_{1,1} \coloneqq S_q \to S_m \to R_1 = S_q \to R'_1$, $S_{2,1} \coloneqq S_{q+m} \to R_1$, $S_{1,2} \coloneqq S_{q+p} \to R_2$ and $S_{2,2} \coloneqq S_q \to S_p \to R_2 = S_q \to R'_2$. For $i \in \{1, 2\}$, let T_i be the tree obtained from T by replacing one α_{k+1}-shape by $S_{i,1}$ and the other one by $S_{i,2}$. We know that $T = T_1$ or $T = T_2$.

[(3.1)] Let T have two α_{k+1}-shapes.
Let $\mathcal{A}$ be the collection of rooted caterpillars A such that

- A is a single vertex; or
- A is a rooted edge; or
- $|A| \geqslant 3$ and the root of A is either an end-vertex of the spine or a leaf attached to an end-vertex of the spine.

If $A \in \mathcal{A}$ then the *reverse* $\tilde{A}$ of A is defined as follows. If A is a single vertex then $\tilde{A} \coloneqq A$. If A is a rooted edge then $\tilde{A}$ is the same edge rooted at the other end-vertex. If A has at least three vertices and the root is an end-vertex of the spine then $\tilde{A}$ is obtained from A by resetting the root at the other end-vertex of the spine. If A has at least three vertices and the root is a leaf attached to an end-vertex of the spine then $\tilde{A}$ is obtained from A by resetting the root at an arbitrary leaf attached to the other end-vertex of the spine. (We note that such a leaf always exists by the definition of the spine.)

Observation 4.2. *Let $A, B \in \mathcal{A}$ such that A and B are isomorphic as unrooted trees but not isomorphic as rooted trees. Let o, o_1 and o_2 be positive integers.*

(1) The caterpillars $S_o \to A$ and $S_o \to B$ are not isomorphic; and
(2) neither are the caterpillars $S_{o_2} \to S_{o_1} \to A$ and $S_{o_2} \to S_{o_1} \to B$.

Proof. The statements are true if $|A| \leqslant 2$, so we assume that A has at least three vertices — and thus so has B. Given an element $C \in \mathcal{A}$ with $|C| \geqslant 3$, we let r_C be the root of C and we define the degree sequence s_C of C as follows. Let $w_1 \dots w_t$ be the spine of C, where w_1 is closest to r_C. The degree sequence of C is $s_C \coloneqq (\deg(w_1), \dots, \deg(w_t))$. The *reverse* of s_C is then the sequence $(\deg(w_t), \dots, \deg(w_1))$. We observe that two elements C and C' of $\mathcal{A}$ (with at least three vertices) are isomorphic as unrooted trees if and only if $s_C = s_{C'}$ or $s_{C'}$ is the reverse of s_C. Furthermore, C and C' are isomorphic (as rooted trees) if and only if $s_C = s_{C'}$ and $\deg(r_C) = \deg(r_{C'})$ (that is, either both roots have degree one, or both roots have degree greater than one).

Let us make another preliminary remark. If $\deg_A(r_A) = 1 \neq \deg_B(r_B)$, then in each of (1) and (2) the caterpillars obtained from A and from B have spines of different lengths, so they are not isomorphic. We can thus assume that either both of r_A and r_B have degree one, or both have degree greater than one. This implies that $s_A \neq s_B$, as otherwise A and B would be isomorphic as rooted trees. Consequently, s_B is the reverse of s_A. Let us write $s_A = (a_1, \dots, a_t)$.

(1). For convenience, set $A' \coloneqq S_o \to A$ and $B' \coloneqq S_o \to B$. We know that $s_B = (a_t, \dots, a_1) \neq s_A$. Suppose first that $\deg_A(r_A) = 1 = \deg_B(r_B)$. Then $s_{A'} = (o, 2, a_1, \dots, a_t)$ if $o > 1$ while $s_{A'} = (2, a_1, \dots, a_t)$ if $o = 1$. Similarly, $s_{B'} = (o, 2, a_t, \dots, a_1)$ if $o > 1$ while $s_{B'} = (2, a_t, \dots, a_1)$ if $o = 1$. In either case, we see that $s_{A'} \neq s_{B'}$ as $s_A \neq s_B$. So suppose for a contradiction that $s_{B'}$ is the reverse of $s_{A'}$. In the former case, i.e. $o > 1$, this means that $(o, 2, a_1, \dots, a_t) = (a_1, \dots, a_t, 2, o)$. Then $a_{2i+1} = o$ and $a_{2i} = 2$ for $i \in \{0, \dots, \lceil t/2 \rceil - 1\}$. In addition, $a_t = o$ and $a_{t-1} = 2$, showing that t must be odd unless $o = 2$. However, either way this yields that $s_A = s_B$, a contradiction. In the latter case, i.e. $o = 1$, we have $(2, a_1, \dots, a_t) = (a_1, \dots, a_t, 2)$, so $a_i = 2$ for each $i \in \{1, \dots, t\}$ which again contradicts that $s_A \neq s_B$.

It remains to deal with the case where $\deg_A(r_A) \neq 1 \neq \deg_B(r_B)$. If $o > 1$, then $s_{A'} = (o, 1 + a_1, a_2, \dots, a_t)$ and $s_{B'} = (o, 1 + a_t, a_{t-1}, \dots, a_1)$. If $o = 1$, then $s_{A'} = (1 + a_1, a_2, \dots, a_t)$ and $s_{B'} = (1 + a_t, a_{t-1}, \dots, a_1)$. In either case, note that $s_{A'} \neq s_{B'}$ because $s_A \neq s_B$. Further, if $s_{B'}$ is the reverse of $s_{A'}$, then it implies that $o > 1$, $a_t = o = a_1$ and $a_i = o+1$ for $i \in \{2, \dots, t-1\}$, leading to $s_A = s_B$, a contradiction. This ends the proof of (1).

(2). For convenience, set $A' \coloneqq S_{o_2} \to S_{o_1} \to A$ and $B' \coloneqq S_{o_2} \to S_{o_1} \to B$. Assume first that $\deg_A(r_A) = 1 = \deg_B(r_B)$. Then we infer as before that

$$s_{A'} = \begin{cases} (2, 2, a_1, \dots, a_t) & \text{if } o_1 = 1 \text{ and } o_2 = 1, \\ (o_2, 2, 2, a_1, \dots, a_t) & \text{if } o_1 = 1 \text{ and } o_2 > 1, \\ (1 + o_1, 2, a_1, \dots, a_t) & \text{if } o_1 > 1 \text{ and } o_2 = 1, \\ (o_2, 1 + o_1, 2, a_1, \dots, a_t) & \text{if } o_1 > 1 \text{ and } o_2 > 1. \end{cases}$$

and

$$s_{B'} = \begin{cases} (2, 2, a_t, \dots, a_1) & \text{if } o_1 = 1 \text{ and } o_2 = 1, \\ (o_2, 2, 2, a_t, \dots, a_1) & \text{if } o_1 = 1 \text{ and } o_2 > 1, \\ (1 + o_1, 2, a_t, \dots, a_1) & \text{if } o_1 > 1 \text{ and } o_2 = 1, \\ (o_2, 1 + o_1, 2, a_t, \dots, a_1) & \text{if } o_1 > 1 \text{ and } o_2 > 1. \end{cases}$$

We see that in each of the four possible cases $s_{A'} \neq s_{B'}$ as $s_A \neq s_B$. In addition, in none of these fours cases can $s_{B'}$ be the reverse of $s_{A'}$, showing that A' and B' are not isomorphic. For instance, in the second case it would imply that t is 1 modulo 3 and $a_i = o_2$ if i is either equal to 1 modulo 3, or $i = t$, while $a_i = 2$ otherwise; however this would yield that $s_A = s_B$, a contradiction. To check the fourth case, it is useful to consider the value of t modulo 3.

It remains to deal with the case where $\deg_A(r_A) \neq 1 \neq \deg_B(r_B)$. We infer the following expressions.

$$s_{A'} = \begin{cases} (2, 1+a_1, a_2, \dots, a_t) & \text{if } o_1 = 1 \text{ and } o_2 = 1, \\ (o_2, 2, 1+a_1, a_2, \dots, a_t) & \text{if } o_1 = 1 \text{ and } o_2 > 1, \\ (1+o_1, 1+a_1, a_2, \dots, a_t) & \text{if } o_1 > 1 \text{ and } o_2 = 1, \\ (o_2, 1+o_1, 1+a_1, a_2, \dots, a_t) & \text{if } o_1 > 1 \text{ and } o_2 > 1. \end{cases}$$

and

$$s_{B'} = \begin{cases} (2, 1+a_t, a_{t-1}, \dots, a_1) & \text{if } o_1 = 1 \text{ and } o_2 = 1, \\ (o_2, 2, 1+a_t, a_{t-1}, \dots, a_1) & \text{if } o_1 = 1 \text{ and } o_2 > 1, \\ (1+o_1, 1+a_t, a_{t-1}, \dots, a_1) & \text{if } o_1 > 1 \text{ and } o_2 = 1, \\ (o_2, 1+o_1, 1+a_t, a_{t-1}, \dots, a_1) & \text{if } o_1 > 1 \text{ and } o_2 > 1. \end{cases}$$

It follows that in none of the four cases the sequence $s_{B'}$ ie equal to $s_{A'}$ or to the reverse of $s_{A'}$, again relying on the fact that $s_A \neq s_B$. □

Observation 4.3.

(1) All $S_{i,j}$ for $i, j \in \{1, 2\}$ are mutually non-isomorphic (as unrooted trees).

(2) If o is a positive integer, then all $S(o)_{i,j} := S_o \to S_{i,j}$ for $i, j \in \{1, 2\}$ are non-isomorphic (as unrooted trees).

Proof. (1). Comparing the lengths of the spines, the only possible pairs of isomorphic trees are: $S_{1,1}$ with $S_{2,2}$, and $S_{1,2}$ with $S_{2,1}$. However, the fact that R'_1 and R'_2 are isomorphic prevents each of these pairs to consist of isomorphic trees, using Observation 4.2(1) for the former one.

(2). Comparing again the lengths of the spines, the only possible pairs of isomorphic trees are: $S(o)_{1,1}$ with $S(o)_{2,2}$, and $S(o)_{1,2}$ with $S(o)_{2,1}$. The first pair cannot consist of isomorphic trees by Observation 4.2(2), since R'_1 and R'_2 are isomorphic.

Let us assume, for a contradiction, that $S(o)_{1,2}$ and $S(o)_{2,1}$ are isomorphic. Necessarily $o > 1$ since R'_1 and R'_2 are isomorphic. For the same reason, $m = p = o$, and the smallest shape of both R_1 and R_2 is S_o. For $i \in \{1, 2\}$, let $S^*_{i,3-i}$ be obtained from $S_{i,3-i}$ by replacing its smallest shape by S_{m+q}. Then $S^*_{1,2}$ and $S^*_{2,1}$ are isomorphic as unrooted trees, but not as rooted trees. If $S(o)_{1,2}$ and $S(o)_{2,1}$ are isomorphic then $S_{m+q} \to S^*_{1,2}$ and $S_{m+q} \to S^*_{2,1}$ are isomorphic. This however contradicts Observation 4.2(1). □

Observation 4.4. *Let E be an α_{k+1}-expression. There is an identification F between the shaped α_{k+1}-partitions of T_1 with characteristic E and the shaped α_{k+1}-partitions of T_2 with characteristic E so that for each partition P, the class of P containing the root of the (α_{k+1})-shape is identified with the class of $F(P)$ containing the root of the (α_{k+1})-shape.*

Proof. If a shaped partition P partitions the unique α_k-shape of T_1 then in particular P partitions the shape $S_{1,1}$ of T_1. We let $F(P)$ be the corresponding partition of $S_{2,2}$ in T_2. Otherwise P partitions the shape $S_{1,j}$ of T_1, and we let $F(P)$ partition the shape $S_{2,j}$ of T_2. □

[(3.1.1)] Let the numbers of leaves in R'_1 and in R'_2 be different. As R'_1 and R'_2 are isomorphic as unrooted trees, this means that for some $i \in \{1, 2\}$, the root of R'_i has degree 1 while that of R'_{3-i} has degree greater than 1. (Recall that the root of a tree is never considered to be a leaf.)

We recall that

$$\theta(T, |T| - \alpha_k - 1, \alpha_k, 1) = \theta_s(T, \alpha_k, 1) + d(T, \alpha_k, 1).$$

We observe that

$$\theta_s(T, \alpha_k, 1) = \theta_s(T_1, \alpha_k, 1) = \theta_s(T_2, \alpha_k, 1).$$

Indeed, if $\alpha_{k+1} > \alpha_k+1$ then all are zero, otherwise it follows from Observation 4.4. Moreover, we can determine $\theta_s(T_1, \alpha_k, 1)$ from our knowledge of T_1. Summarising we can determine $d(T, \alpha_k, 1)$, the number of leaves of T outside of its unique α_k-shape. We also know the total number of leaves of T and we thus deduce the number of leaves of the unique α_k-shape of T. Since we assume that R'_1 and R'_2 have different numbers of leaves, this allows us to deduce whether $T = T_1$ or $T = T_2$. Hence case (3.1.1) is solved.

[(3.1.2)] Hence let R'_1 and R'_2 have the same number of leaves (as rooted trees). Because R'_1 and R'_2 are isomorphic as unrooted trees, this assumption means that the root of R'_1 has degree 1 if and only if the root of R'_2 has degree 1. As a consequence, we make the following two observations.

Observation 4.5. *The trees T_1 and T_2 have the same number of leaves and $d(T_1, \alpha_k, 1) = d(T_2, \alpha_k, 1)$.*

Observation 4.6. *The number of leaves of $S_{1,1}$ equals the number of leaves of $S_{2,2}$ equals the number of leaves of $S_{1,2}$ minus one equals the number of leaves of $S_{2,1}$ minus one.*

For $\ell \geqslant 2$ and $i \in \{1, 2\}$, let $R_i(\ell)$ be the strict shape of R_i with exactly ℓ vertices, if it exists. Let ℓ_1 be the least integer $\ell \geqslant 2$ such that $R_1(\ell)$ and $R_2(\ell)$ are not isomorphic, or one exists and the other one does not. By the choice of ℓ_1 we know that if $\ell' < \ell_1$, then the rooted trees $R_1(\ell')$ and $R_2(\ell')$, when they exist, are isomorphic (as rooted trees). We note that $\ell_1 \leqslant |R_1|$ since R_1 and R_2 are not isomorphic as rooted trees. It follows that $R_i(\ell_1)$ exists for exactly one index $i \in \{1, 2\}$. Let ℓ_2 be the least integer ℓ such that $\ell = \ell_1$ or T has a unique $(\alpha_k + \ell)$-shape.

Observation 4.7. *Let $\ell \in \{1, \ldots, \ell_2\}$ such that T has either zero or two $(\alpha_k + \ell)$-shapes. Let E be an $(\alpha_k+\ell)$-expression. There is an identification F between the shaped $\alpha_k+\ell$-partitions of T_1 with characteristic E and the shaped $(\alpha_k+\ell)$-partitions of T_2 with characteristic E such that for each partition P, the part of P containing the root of the $(\alpha_k + \ell)$-shape is identified with the part of $F(P)$ containing the root of the $(\alpha_k + \ell)$-shape.*

Proof. We proceed by induction on ℓ, the basic (non-trivial) case being dealt with in Observation 4.4. Assume that $\alpha_k + \ell > \alpha_{k+1}$ and let o be the largest index such that $\alpha_o < \alpha_k + \ell$. So $o \geqslant k+1$. If P is a shaped $(\alpha_k+\ell)$-partition then we let P' be its restriction to the corresponding strict α_o-shape. We define the corresponding partition $F(P)$ to be obtained from $F(P')$ in the same way as P is obtained from P'. We observe that F defined in this way identifies the part of P containing the root of the $(\alpha_k + \ell)$-shape with the class of $F(P)$ containing the root of the $(\alpha_k + \ell)$-shape. □

[(3.1.2.1)] Let there be no integer o such that $\alpha_k+\ell_2 = \alpha_o$. Then $\ell_2 = \ell_1$. Let $i \in \{1, 2\}$ be such that $R_i(\ell_1)$ exists. Then $\theta(T_i, |T|-\alpha_k-\ell_1, \alpha_k, \ell_1) = 0$ and $\theta(T_{3-i}, |T|-\alpha_k-\ell_1, \alpha_k, \ell_1) = 1$. As we know $\theta(T, |T| - \alpha_k - \ell_1, \alpha_k, \ell_1)$, this case is solved.

[(3.1.2.2)] Let $\alpha_k + \ell_2 = \alpha_o$ and let there be two α_o-shapes. Then again $\ell_2 = \ell_1$. Let $i \in \{1, 2\}$ be such that $R_i(\ell_1)$ exists.

Then $\theta(T_i, |T| - \alpha_k - l_1, \alpha_k, \ell_1)$ is equal to the number of shaped $(\alpha_k + l_1)$-partitions of T_i with characteristic (α_k, l_1) and $\theta(T_{3-i}, |T| - \alpha_k - l_1, \alpha_k, \ell_1)$ is equal to 1 plus the number of shaped (α_k+l_1)-partitions of T_{3-i} with characteristic (α_k, l_1). Moreover Observation 4.7 implies that the number of shaped $(\alpha_k + l_1)$-partitions of T_i with characteristic (α_k, l_1) is equal to the number of shaped $(\alpha_k + l_1)$-partitions of T_{3-i} with characteristic (α_k, l_1). Hence **we assume that $\alpha_o = \alpha_k+\ell_2$ and T has a unique α_o-shape.** It follows that $o \geqslant k+2$ since the assumption of case (3.1) is that T has two α_{k+1}-shapes.

[(3.1.2.3)] Let $\alpha_k+\ell_2 = \alpha_o$ and $o \geqslant k+2$. Suppose that either T has two α_{k+2}-shapes, or T has a unique α_{k+2}-shape and then $\alpha_{k+2} - \alpha_{k+1} > 1$. We recall that $\alpha_{k+1} \geqslant |T|/2$.

Let E be an α_{k+1}-expression such that $\theta(S_{1,2},E) =: r_1 \neq r_2 := \theta(S_{2,1},E)$. No part of E is equal to $\alpha_{k+1}-1$ since $S_{1,2}$ and $S_{2,1}$ have the same number of leaves. For $i \in \{1,2\}$,

$$\theta(T_i, |T| - \alpha_{k+1} - 1, E, 1) = d(T_i, E, 1) + \theta_s(T_i, E, 1) + D_i,$$

where

$$D_i := \sum_{[E_1,E_2,s]\in\P} d(T_i, E_1, E_2, 1^s)$$

and

$$\P := \{(E_1, E_2, s) : [E,1] = [E_1, E_2, 1^s] \text{ and the size of each of } E_1 \text{ and } E_2 \text{ is at most } \alpha_k\}.$$

Observation 4.8. *For each $(E_1,E_2,s) \in \P$, we have $d(T_1,E_1,E_2,1^s) = d(T_2,E_1,E_2,1^s)$.*

Proof. If both E_1 and E_2 have size smaller than α_k then the observation follows from the form of T_1 and T_2. Hence let the size of E_1 be α_k. By the considerations before Observation 4.7 we know that the size of E_2 is at most ℓ_2, which is less than α_k. We assume that we are not in case (3.1.1) and we thus know the number of leaves of T_1 to be equal to the number of leaves of T_2 (Observation 4.5). We also know the number of leaves of R'_1 to be equal to the number of leaves of R'_2. Moreover for $\ell < \ell_2$, the rooted trees $R_1(\ell)$ and $R_2(\ell)$ are isomorphic. Hence the observation follows if the size of E_2 is smaller than ℓ_2. We now assume the size of E_2 to be ℓ_2. Then $\alpha_{k+2} \leqslant \alpha_o = \alpha_k + \ell_2 \leqslant \alpha_{k+1}+1$. It follows that $s=0$ and $\alpha_{k+2} = \alpha_{k+1}+1$. By the assumptions of case (3.1.6) we thus deduce that T has two α_{k+2}-shapes. We are thus in case (3.1.5), contrary to our assumption. □

Observation 4.9. *We have $d(T_1,E,1) \neq d(T_2,E,1)$.*

Proof. For $i,j \in \{1,2\}$, let $h_{i,j}$ be the number of leaves of T_i outside of $S_{i,j}$. Let us set $h := h_{1,1}$. By Observations 4.5 and 4.6, we know that $h_{2,2} = h$ and $h_{2,1} = h_{1,2} = h-1$. For $i \in \{1,2\}$ let $\mathcal{T}_{i,0}$ be the set of shaped partitions of $S_{i,i}$ with characteristic E that partition the strict α_k-shape. We recall that the identification F granted by Observation 4.4 satisfies that $P \in \mathcal{T}_{1,0}$ if and only if $F(P) \in \mathcal{T}_{2,0}$. Moreover let $\mathcal{T}_{i,i}$ be the number of shaped partitions of $S_{i,i}$ with characteristic E that do not belong to $\mathcal{T}_{i,0}$, and let $\mathcal{T}_{3-i,i}$ be the number of shaped partitions of $S_{3-i,i}$ of characteristic E. We have

$$d(T_1,E,1) = h\,|\mathcal{T}_{0,1}| + h\,|\mathcal{T}_{1,1}| + (h-1)\,|\mathcal{T}_{1,2}|$$

and

$$d(T_2,E,1) = h\,|\mathcal{T}_{0,2}| + h\,|\mathcal{T}_{2,2}| + (h-1)\,|\mathcal{T}_{2,1}|.$$

Finally, Observation 4.4 implies that $|\mathcal{T}_{0,1}| = |\mathcal{T}_{0,2}|$, $|\mathcal{T}_{1,1}| = |\mathcal{T}_{2,1}| = r_1$ and $|\mathcal{T}_{2,2}| = |\mathcal{T}_{1,2}| = r_2$. □

By the assumptions of (3.1.5) and Observation 4.7 we know that

$$\theta_s(T_1, |T| - \alpha_{k+1} - 1, E, 1) = \theta_s(T_2, |T| - \alpha_{k+1} - 1, E, 1).$$

This along with Observations 4.8 and 4.9 imply that

$$\theta(T_1, |T| - \alpha_{k+1} - 1, E, 1) \neq \theta(T_2, |T| - \alpha_{k+1} - 1, E, 1).$$

Further, we know $\theta(T, |T| - \alpha_{k+1} - 1, E, 1)$. This solves case (3.1.2.3).

[(3.1.2.4)] Finally let $\alpha_k + \ell_2 = \alpha_{k+2}$, T have unique α_{k+2}-shape and $\alpha_{k+2} - \alpha_{k+1} = 1$. The candidates for the α_{k+2}-shape of T are $R_{i,j} = S_1 \to S_{i,j}$ for $i,j \in \{1,2\}$. We first observe, using Observations 4.2 and 4.3, that all $R_{i,j}$ are mutually non-isomorphic. Let R be the unique α_{k+2}-shape of T. The following observation is straightforward.

Observation 4.10. *If $\ell < \ell_1$ then all four ℓ-shapes of $R'_1(\ell)$ and $R'_2(\ell)$ are isomorphic as rooted trees.*

[(3.1.2.4.1)] Let $\ell_1 > \ell_2 = q + 1$. The next observation readily follows.

Observation 4.11. *If $R \in \mathcal{A}_1 \coloneqq \{R_{1,1}, R_{2,2}\}$ then $\theta(T, |T| - \alpha_{k+2}, \alpha_k, \ell_2) = 3$. If $R \in \mathcal{A}_2 \coloneqq \{R_{1,2}, R_{2,1}\}$ then $\theta(T, |T| - \alpha_{k+2}, \alpha_k, \ell_2) = 2$.*

Observation 4.12. *We have $\theta(R_{1,1}, \ell_1, \alpha_{k+2} - \ell_1) \neq \theta(R_{2,2}, \ell_1, \alpha_{k+2} - \ell_1))$ and $\theta(R_{1,2}, \ell_1, \alpha_{k+2} - \ell_1)) \neq \theta(R_{2,1}, \ell_1, \alpha_{k+2} - \ell_1))$.*

Proof. As reported earlier, by the definition of ℓ_1 there is a unique $i \in \{1, 2\}$ such that $R_i(\ell_1)$ exists. We may assume, without loss of generality, that $i = 1$. Since $\ell_1 > \ell_2 = q + 1$, we have

$$\theta(R_{1,1}, \ell_1, \alpha_{k+2} - \ell_1) = \theta(R_{2,1}, \ell_1, \alpha_{k+2} - \ell_1)) \geqslant \theta(R_{2,2}, \ell_1, \alpha_{k+2} - \ell_1)) = \theta(R_{1,2}, \ell_1, \alpha_{k+2} - \ell_1)),$$

$\theta(R_{1,1}, \ell_1, \alpha_{k+2} - \ell_1) \in \{1, 2\}$ and $\theta(R_{2,2}, \ell_1, \alpha_{k+2} - \ell_1) \in \{0, 1\}$. We recall that $\ell_2 = q + 1$. If $\theta(R_{2,2}, \ell_1, \alpha_{k+2} - \ell_1) = 1$ then R'_2 must have a shape with $\ell_1 - \ell_2$ vertices. Then however by Observation 4.10 each of R'_1 and R'_2 has two shapes with $\ell_1 - \ell_2$ vertices and $\theta(R_{1,1}, \ell_1, \alpha_{k+2} - \ell_1) = 2$. □

For $i \in \{1, 2\}$,

$$\theta(T_i, |T| - \alpha_{k+2}, \ell_1, \alpha_{k+2} - \ell_1) = \theta_s(T_i, \ell_1, \alpha_{k+2} - \ell_1) + d(T_i, \ell_1, \alpha_{k+2} - \ell_1).$$

Moreover, $d(T_1, \ell_1, \alpha_{k+2} - \ell_1) = d(T_2, \ell_1, \alpha_{k+2} - \ell_1)$ since both $\alpha_{k+2} - \ell_1$ and ℓ_1 are smaller than α_k. By Observation 4.12 we know that

$$\theta_s(T_1, \ell_1, \alpha_{k+2} - \ell_1) \neq \theta_s(T_2, \ell_1, \alpha_{k+2} - \ell_1).$$

Hence

$$\theta(T_1, |T| - \alpha_{k+2}, \ell_1, \alpha_{k+2} - \ell_1) \neq \theta(T_2, |T| - \alpha_{k+2}, \ell_1, \alpha_{k+2} - \ell_1)$$

and considering $\theta(T, |T| - \alpha_{k+2}, \ell_1, \alpha_{k+2} - \ell_1)$ solves this case.

[(3.1.2.4.2)] Let $\ell_1 = \ell_2$.
We recall that R'_1 and R'_2 have the same number of leaves — case (3.1.1) — and they are isomorphic. It follows that one is the 'reverse' of the other. Hence, if $m < p$, then $m > 1$. There is exactly one $i \in \{1, 2\}$ such that $R_i(\ell_2)$ exists. Moreover, if $m < p$, then $i = 2$. Hence, we may assume without loss of generality that $i = 2$.

If we know that $R = R_{i,j}$, then we know how the trees T_1 and T_2 look up to α_{k+2}-shapes. Hence the following notation makes sense: if $R = R_{i,j}$ then we for each $x \in \{1, 2\}$ we let $T_x(i, j)$ be T_x. We know from the definitions of T_1 and T_2 that $T_x(1, 1) = T_x(2, 1)$ and $T_x(1, 2) = T_x(2, 2)$ for each $x \in \{1, 2\}$.

The following observation is straightforward.

Observation 4.13. $\theta(T_2(2, 1), |T| - \alpha_{k+2}, \alpha_k, \ell_2) = 0$ *and* $\theta(T_1(1, 1), |T| - \alpha_{k+2}, \alpha_k, \ell_2) = 2$.

Hence, if $R = R_{1,1}$ or $R = R_{2,1}$ then we can calculate $\theta(T_1, |T| - \alpha_{k+2}, \alpha_k, \ell_2)$, $\theta(T_2, |T| - \alpha_{k+2}, \alpha_k, \ell_2)$ and $\theta(T, |T| - \alpha_{k+2}, \alpha_k, \ell_2)$. Further,

$$\theta(T_1, |T| - \alpha_{k+2}, \alpha_k, \ell_2) \neq \theta(T_2, |T| - \alpha_{k+2}, \alpha_k, \ell_2).$$

Summarising, the cases where $R = R_{1,1}$ or $R = R_{2,1}$ are solved. Hence from now on we assume that $R = R_{2,2}$ or $R = R_{1,2}$. This means that we know all the s-shapes of the trees T_1 and T_2 for $s \leqslant \alpha_{k+2}$. Moreover, $R_{2,2}$ and $R_{1,2}$ have spines of different length. It follows that there is an α_{k+2}-expression E that distinguishes $R_{2,2}$ and $R_{1,2}$ and has no part of size one. We recall that $\alpha_{k+1} + 1 = \alpha_{k+2}$. Hence the number of non-shaped partitions of T_1 with characteristic $(E, |T| - \alpha_{k+2})$ is equal to the number of non-shaped partitions of T_2 with characteristic $(E, |T| - \alpha_{k+2})$;

hence this number can be determined and is equal to the number of non-shaped partitions of T with characteristic $(E, |T| - \alpha_{k+2})$.

This means that the number of shaped α_{k+2}-partitions of T with characteristic $(E, |T| - \alpha_{k+2})$ can be determined. Thus we can determine whether $T = T_1$ or $T = T_2$ and case (3.1.2.4) is solved.

[(3.2)] Let T have a unique α_{k+1}-shape.
If $q > 1$ then all four candidates $S_{1,1}$, $S_{1,2}$, $S_{2,1}$ and $S_{2,2}$ for the α_{k+1}-shape of T are mutually not isomorphic and we can determine the correct one by considering the shaped α_{k+1}-partitions of T.

Hence suppose that $q = 1$. There are two pairs of isomorphic (as unrooted trees) candidates: $S_{1,1}$ is isomorphic to $S_{1,2}$ and $S_{2,1}$ is isomorphic to $S_{2,2}$. We observe that for each pair, its two elements differ in the number of leaves different from the root. Moreover, $S_{1,1}$ and $S_{2,2}$ are not isomorphic. Considering the shaped α_{k+1}-partitions of T we can determine to which pair the unique α_{k+1}-shape of T belongs. We may assume, without loss of generality, that it belongs to $\{S_{1,1}, S_{1,2}\}$.

We note that $n \neq k+1$ since $q = 1$ and we recall that $j = \alpha_k < \alpha_{k+1} \leqslant |T|/2$. Since we know the isomorphism class of the unique α_{k+1}-shape of T, we can determine the number of shaped α_{k+2}-partitions of T. We have

$$\theta(T, |T| - \alpha_{k+1} - 1, \alpha_{k+1}, 1) = \theta_s(T, \alpha_{k+1}, 1) + d(T, \alpha_{k+1}, 1).$$

The considerations above imply that we can determine $d(T, \alpha_{k+1}, 1)$, which is equal to the number of leaves of T outside of the unique α_{k+1}-shape. Since we know the number of leaves of T, we can also determine the number of leaves of the unique α_{k+1}-shape of T that are different from the root. Hence we can determine whether this shape is $S_{1,1}$ or $S_{1,2}$. This finishes case (3.2) and thus case (3).

This ends our updating process and the inductive step of our induction. Consequently, we established that we know, for each $j \in \{\alpha_1, \dots, |T|/2\}$ and each j-form F, the number of shapes of T that belongs to F. Therefore Observation 3.1 ensures that we know T. This concludes the induction on the size of T and thus the proof of Theorem 2. □

5. Designing Procedure 1

An *α_j-situation* σ is a multi-set $((\sigma_1, w_1), \dots, (\sigma_{t(\sigma)}, w_{t(\sigma)}))$ of disjoint weighted non-rooted trees with $t(\sigma) \geqslant 2$ such that $w_1(\sigma_1) \leqslant \dots \leqslant w_{t(\sigma)}(\sigma_{t(\sigma)})$ and $\sum_{i=1}^{t(\sigma)} w_i(\sigma_i) = \alpha_j$. An α_j-situation σ is said to *occur* in a tree T if there exists a subtree T' of T and $t(\sigma)$ distinct edges $e_1, \dots, e_{t(\sigma)}$ with exactly one end in $V(T')$ such that, for each $i \in \{1, \dots, t(\sigma)\}$, there is an isomorphism preserving the weights *but not necessarily the roots* between σ_i and the component of $T - e_i$ different from T'. Note that if σ occurs in T, then for each $i \in \{1, \dots, t(\sigma)\}$ such that σ_i is not a single vertex the tree T has a shape isomorphic to σ_i.

We proceed in two steps, the first one being an exhaustive listing that depends only on α_j.
Step 1. Explicitly list all α_j-situations for $\alpha_j \leqslant w(T)/2$.
Step 2. For each $\alpha_j \leqslant w(T)/2$ and each α_j-situation σ from Step 1, compute the number $m_T(\sigma)$ of times σ occurs in T.

Before designing Step 2, we show how Steps 1 and 2 accomplish Procedure 1. Suppose that the two steps are completed. Let $E = \{w(T) - \alpha_j, E_1, \dots, E_k\}$ be an α_j-expression of $w(T)$.

For each α_j-situation $\sigma = ((\sigma_1, w_1), \dots, (\sigma_{t(\sigma)}, w_{t(\sigma)}))$, let Ψ_σ be the collection of all surjections from the expression $\{E_1, \dots, E_k\}$ to $\{\sigma_1, \dots, \sigma_{t(\sigma)}\}$. Two elements f and g of Ψ_σ are *equivalent* if the multi-set $f^{-1}(\sigma_i)$ is equal to the multi-set $g^{-1}(\sigma_i)$ for every $i \in \{1, \dots, k\}$. We consider the equivalence classes for this relation on Ψ_σ and we form Ψ'_σ by arbitrarily choosing one representative in each equivalent class. We observe that the number X of non-shaped

α_j-partitions of T with characteristic E is

$$\sum_{\alpha_j\text{-situation } \sigma} m_T(\sigma) \sum_{f\in\Psi'_\sigma} \sum_{i=1}^{t(\sigma)} \theta(\sigma_i, w_i, f^{-1}(\sigma_i)), \tag{5.1}$$

where the multi-set $f^{-1}(\sigma_i)$ is naturally interpreted as an expression. Indeed, a non-shaped partition of T with characteristic E corresponds precisely to the occurrence of some α_j-situation $\sigma = ((\sigma_1, w_1), \ldots, (\sigma_{t(\sigma)}, w_{t(\sigma)}))$ where the trees $\sigma_1 \ldots, \sigma_\ell$ are also partitioned (possibly trivially). Recalling that $\theta(\sigma_i, w_i, E')$ is zero if E' is not an expression of $w_i(\sigma_i)$, the formula (5.1) follows. Notice that (5.1) does allow us to compute X when Step 1 and Step 2 are completed. Consequently, we can compute the number of shaped α_j-partitions of T with characteristic E, which is

$$\theta(T, w, E) - X.$$

This accomplishes Procedure 1.

It remains to design Step 2. We fix an α_j-situation $\sigma = ((\sigma_1, w_1), \ldots, (\sigma_t, w_t))$.

Observation 5.1. *For every pair $(i, j) \in \{1, \ldots, t\}^2$, if T_i and T_j are two shapes of a tree T that are isomorphic to σ_i and σ_j, respectively, then either $T_i \subseteq T_j$ or $T_j \subseteq T_i$ or $T_i \cap T_j = \varnothing$.*

To see this, for $k \in \{i, j\}$ let e_k be the edge of T associated to T_k, that is, T_k is a component of $T - e_k$. Then, either $e_j \in E(T_i)$ or $e_j \in E(T - T_i)$. If $e_j \in E(T - T_i)$, then either $T_i \subseteq T_j$ or $T_j \subseteq T - T_i$, in which case $T_j \cap T_i = \varnothing$. If $e_j \in E(T_i)$, then $T_j \subseteq T_i$: otherwise, $T_j \cap T_i \neq \varnothing$ and $T - T_i \subset T_j$, so that $w(T_i) + w(T_j) > w(T)$. This would contradict the hypothesis that $\sum_{k=1}^{t(\sigma)} w_k(\sigma_k) = \alpha_j$, since $\alpha_j \leqslant w(T)/2$. This concludes the proof of Observation 5.1.

Define Λ to be the set of all t-tuples $(T_1, \ldots, T_t)$ such that for each $i \in \{1, \ldots, t\}$,

- T_i is either a shape of T or a leaf;
- T_i is isomorphic to (σ_i, w_i) as a weighted non-rooted tree; and
- if $j \in \{1, \ldots, t\} \setminus \{i\}$, then T_i is not a subtree of T_j.

Observation 5.2. *The number of times that σ occurs in T is equal to $|\Lambda|$.*

Proof. We prove that the elements of Λ are exactly occurrences of σ in T. By the definition, each occurrence of σ gives rise to an element of Λ.

Conversely, let $(T_1, \ldots, T_t)$ be an element of Λ. Observation 5.1 implies that the shapes T_i are mutually disjoint. For each $k \in \{1, \ldots, t\}$, let e_k be the edge of T associated to the shape T_k, that is, e_k connects the root of T_k to $T - T_k$; and let v_k be the endvertex of e_k that does not belong to T_k. Note that $v_k \notin \cup_{j=1}^{t} T_j$ since no tree T_i is a subtree of another tree T_j and $\alpha_j \leqslant w(T)/2$. Set $T'_0 \coloneqq T$ and $T'_k \coloneqq T'_{k-1} - T_k$ for $k \geqslant 1$.

Observe that each of $T_{k+1}, \ldots, T_t$ is either a leaf or a shape of T'_k. Hence T'_k is connected and contains all the vertices $v_1, \ldots, v_t$. Therefore setting $T' \coloneqq T'_t$ shows that $(T_1, \ldots, T_t)$ occurs in T. □

Our goal is to compute $|\Lambda|$. For a weighted tree (T', w'), define $\Lambda_0(T', w')$ to be the set of all t-tuples $(T_1, \ldots, T_t)$ such that for each $i \in \{1, \ldots, t\}$ it holds that T_i is either a leaf or a shape of T' that is isomorphic to (σ_i, w_i) as a weighted non-rooted tree. Set $\Lambda_0 \coloneqq \Lambda_0(T, w)$. In this notation, the weight shall be omitted when there is no risk of confusion. The advantage of Λ_0 is that its size can be computed. Indeed,

$$|\Lambda_0| = \prod_{i=1}^{t} \sharp((\sigma_i, w_i) \hookrightarrow (T, w)),$$

where $\sharp((\sigma_i, w_i) \hookrightarrow (T, w))$ is the number of leaves or shapes of T that are isomorphic to (σ_i, w_i) as weighted non-rooted trees. This number is given in the input of Procedure 1, since $w_i(\sigma_i) < \alpha_j$.

Next, we compute $|\Lambda|$ using the principle of inclusion and exclusion. Setting $I \coloneqq \{1, \dots, t\}^2 \setminus \{(i,i) : 1 \leqslant i \leqslant t\}$, we have

$$|\Lambda| = |\Lambda_0| - \left| \bigcup_{(i,j)\in I} \Lambda_{(i,j)} \right|,$$

where $\Lambda_{(i,j)}$ is the subset of Λ_0 composed of the elements $(T_1, \dots, T_t)$ with $T_i \subseteq T_j$.

By the principle of inclusion-exclusion, we deduce that the output of Step 2 is equal to

$$|\Lambda_0| - \sum_{\varnothing \neq J \subseteq I} (-1)^{|J|-1} \left| \bigcap_{(i,j)\in J} \Lambda_{(i,j)} \right|.$$

It remains to compute $\left|\bigcap_{(i,j)\in J} \Lambda_{(i,j)}\right|$ for each non-empty subset J of I. We start with an observation, which characterises the sets J for which the considered intersection is not empty.

Observation 5.3. *Let $J \subseteq I$. Then, $\bigcap_{(i,j)\in J} \Lambda_{(i,j)} \neq \varnothing$ if and only if for every $(i,j) \in J$, either σ_i is isomorphic to σ_j, or σ_j has a leaf or a shape that is isomorphic to σ_i as a weighted non-rooted tree.*

From now on, we consider an arbitrary contributing set J. We construct four directed graphs A_0, A_1, A_2 and A_3 that depend on J. Each vertex x of A_k is labeled by a subset $\ell(x)$ of $\{(\sigma_1, w_1), \dots, (\sigma_t, w_t)\}$. These labels will have the following properties.

(1) $(\ell(x))_{x\in V(A_k)}$ is a partition of $\{(\sigma_1, w_1), \dots, (\sigma_t, w_t)\}$.
(2) For each vertex x of A_k, all weighted trees in $\ell(x)$ are isomorphic.
(3) $\left|\cap_{(i,j)\in J}\Lambda_{(i,j)}\right|$ is equal to the number of elements $(T_1, \dots, T_t)$ of Λ_0 such that
   - for each vertex x of A_k, if $(\sigma_i, w_i), (\sigma_j, w_j) \in \ell(x)$ then $T_i = T_j$; and
   - for every arc (x, y) of A_k, if $((\sigma_i, w_i), (\sigma_j, w_j)) \in \ell(x) \times \ell(y)$, then $T_i \subseteq T_j$.

The directed graph A_0 is obtained as follows. We start from the vertex set $\{z_1, \dots, z_t\}$. For each $i \in \{1, \dots, t\}$, the label $\ell(z_i)$ of z_i is set to be $\{(\sigma_i, w_i)\}$. For each $(i,j) \in J$, we add an arc from z_i to z_j. Thus A_0 satisfies properties (1)–(3). Note that A_0 may contain directed cycles, but by Observation 5.3, if C is a directed cycle then all elements in $\cup_{x\in V(C)}\ell(x)$ are isomorphic.

Now, A_1 is obtained from A_0 by the following recursive operation. Let (x, y, z) be a triple of vertices such that (x, y) and (x, z) are arcs, but neither (y, z) nor (z, y) are arcs. Let $(\sigma_y, w_y) \in \ell(y)$ and $(\sigma_z, w_z) \in \ell(z)$. We add the arc (y, z) if $|V(\sigma_y)| \leqslant |V(\sigma_z)|$, and the arc (z, y) if $|V(\sigma_z)| \leqslant |V(\sigma_y)|$. (In particular, if $|V(\sigma_y)| = |V(\sigma_z)|$, then both arcs are added.)

We observe that A_1 satisfies (1)–(3). Since neither the vertices nor the labels were changed, the only thing that we need to show is that if the arc (y, z) was added, then for all tuples $(T_1, \dots, T_t) \in \cap_{(i,j)\in J}\Lambda_{(i,j)}$ and all $((\sigma_i, w_i), (\sigma_j, w_j)) \in \ell(y) \times \ell(z)$, it holds that $T_i \subseteq T_j$. This follows from Observation 5.1: since (y, z) was added, there exists $s \in \{1, \dots, t\}$ such that T_s is contained in both T_i and T_j.

The directed graph A_2 is obtained from A_1 by recursively contracting all directed cycles of A_1. Specifically, for each directed cycle C, all the vertices of C are contracted into a vertex z_C (parallel arcs are removed, but not directed cycles of length 2), and $\ell(z_C) \coloneqq \cup_{x\in V(C)}\ell(x)$. We again observe that A_2 satisfies properties (1)–(3).

Finally, A_3 is obtained from A_2 by recursively deleting transitivity arcs, that is, the arc (y, z) is removed if there exists a directed path of length greater than 1 from y to z. Note that A_2 and A_3 have the same vertex-set, and every arc of A_3 is also an arc in A_2. Again, A_3 readily satisfies properties (1)–(3).

Now, let us prove that each component of A_3 is an arborescence, that is a directed acyclic graph with each out-degree at most one. We only need to show that every vertex of A_3 has outdegree at most 1. Assume that (x,y) and (x,z) are two arcs of A_3. First, note that, in A_2, there is no directed path from y to z or from z to y, for otherwise the arc (x,y) or the arc (x,z) would not belong to A_2, respectively. Therefore, regardless whether y and z arose from contractions of directed cycles in A_1, there exist three vertices x', y' and z' in A_1 such that both (x',y') and (x',z') are arcs but neither (y',z') nor (z',y') is an arc. This contradicts the definition of A_1. Consequently, every vertex of A_3 has outdegree at most 1, as wanted.

We define τ_i to be the ordered $(t+1)$-tuple

$$(\sharp((\sigma_i,w_i)\hookrightarrow(T,w)),\sharp((\sigma_i,w_i)\hookrightarrow(\sigma_1,w_1)),\dots,\sharp((\sigma_i,w_i)\hookrightarrow(\sigma_t,w_t)))$$

We recall that $\tau_1,\dots,\tau_t$ are known from the assumptions of Procedure 1. Step 2 is completed by the following procedure.

Procedure 2.
INPUT: A labeled directed forest A of arborescences and the $(t+1)$-tuples $\tau_1,\dots,\tau_t$.
OUTPUT: For each $H\in\{(T,w),(\sigma_1,w_1),\dots,(\sigma_t,w_t)\}$, the number $\mathcal{P}_3(H,A,\tau(T))$ of elements $(T_1,\dots,T_t)$ of $\Lambda_0(H)$ such that

- for each vertex x of A, if $(\sigma_i,w_i),(\sigma_j,w_j)\in\ell(x)$ then $T_i=T_j$; and
- for every arc (x,y) of A, if $((\sigma_i,w_i),(\sigma_j,w_j))\in\ell(x)\times\ell(y)$, then $T_i\subseteq T_j$.

The output of Procedure 2 can be recursively computed as follows. Let $V_{\max}$ be the set of vertices of A with outdegree 0. For each vertex x of A, let (σ^x,w^x) be a representative of $\ell(x)$.

$$\mathcal{P}_3(H,A,\tau(T))=\prod_{x\in V_{\max}}(\sharp((\sigma^x,w^x)\hookrightarrow H))\cdot\mathcal{P}_3((\sigma^w,w^x),\tilde{A}(w),\tau(T)),$$

where $\tilde{A}(w)$ is obtained from the component of A that contains x by removing x.

By property (3) of the labels, the output $\mathcal{P}_3(T,A_3,\tau(T))$ is equal to $\left|\cap_{(i,j)\in J}\Lambda_{(i,j)}\right|$. This concludes the design of Procedure 1.

References

[1] J. D. Annan, *Complexity of Counting Problems*, PhD thesis, Oxford University (1994).

[2] J. Aliste-Prieto and J. Zamora, *Proper caterpillars are distinguished by their chromatic symmetric function*, Discrete Math. **315** (2014), 158–164.

[3] G. D. Birkhoff, *A determinant formula for the number of ways of coloring a map*, Ann. of Math. (2) **14** (1912/13), no. 1-4, 42–46.

[4] A. Brandstädt, V. B. Le and J. P. Spinrad, *Graph classes: a survey*, SIAM Monographs on Discrete Mathematics and Applications, SIAM, Philadelphia, PA (1999).

[5] J. A. Bondy,*A graph reconstructor's manual*, In *Surveys in Combinatorics*, (Guildford, 1991), London Math. Soc. Lecture Note Ser., Vol. **166** (1991), Cambridge Univ. Press, Cambridge, 221–252.

[6] J. A. Bondy and R. L. Hemminger, *Graph reconstruction—a survey*, J. Graph Theory **1** (1977), no. 3, 227–268.

[7] T. Brylawski, *Intersection theory for graphs*, J. Combin. Theory Ser. B **30** (1981), no. 2, 233–246.

[8] S. Chaudhary and G. Gordon, *Tutte polynomials for trees*, J. Graph Theory **15** (1991), no. 3, 317–331.

[9] R. Conti, P. Contucci, and C. Falcolini, *Polynomial invariants for trees: a statistical mechanics approach*, Discrete Appl. Math. **81** (1998), no. 1-3, 225–237.

[10] J. Ellis-Monaghan, I. Moffatt, *The Tutte-Potts connection in the presence of an external magnetic field*, Adv. in Appl. Math. 47 (2011), no. 4, 772–782.

[11] J. Fougere, *On symmetric chromatic polynomials of trees*, undergraduate thesis, Dartmouth College (2003).

[12] F. Harary, *On the reconstruction of a graph from a collection of subgraphs*, In *Theory of Graphs and its Applications* (Proc. Sympos. Smolenice, 1963), Publ. House Czechoslovak Acad. Sci., Prague (1964), 47–52.

[13] P. J. Kelly, *On Isometric Transformations*, Ph.D. thesis, University of Wisconsin (1942).

[14] M. Klazar, M. Loebl, and I. Moffatt, *The Potts model and chromatic functions of graphs*, Ann. Inst. Henri Poincaré D **1** (2014), no. 1, 47–60.

[15] M. Loebl, *Chromatic polynomial, q-binomial counting and colored Jones function*, Adv. Math. **211** (2007), no. 2, 546–565.
[16] L. Lovász, *A note on the line reconstruction problem*, J. Combin. Theory Ser. B **13** (1972), 309–310.
[17] B. D. McKay, *Computer reconstruction of small graphs*, J. Graph Theory **1**, 281–283.
[18] J. L. Martin, M. Morin, and J. D. Wagner, *On distinguishing trees by their chromatic symmetric functions*, J. Combin. Theory Ser. A **115** (2008), 237–253.
[19] M. Morin, *Caterpillars, ribbons, and the chromatic symmetric function*, MS thesis, University of British Columbia (2005).
[20] V. Müller, *The edge reconstruction hypothesis is true for graphs with more than $n \cdot \log_2 n$ edges*, J. Combin. Theory Ser. B **22** (1977), 281–283.
[21] C. S. J. A. Nash-Williams, *The reconstruction problem*, In *Selected Topics in Graph Theory* (L. W. Beineke and R. J. Wilson, eds.), Academic Press, London (1978), 205–236.
[22] S. D. Noble and D. J. A. Welsh, *A weighted graph polynomial from chromatic invariants of knots*, Ann. Inst. Fourier (Grenoble) **49** (1999), no. 3, 1057–1087, Symposium à la Mémoire de François Jaeger (Grenoble, 1998).
[23] J. Oxley and G. Whittle, *Tutte invariants for 2-polymatroids*, Graph structure theory (Seattle, WA, 1991), 9–19, Contemp. Math. **147** (1993), Amer. Math. Soc., Providence, RI.
[24] I. Sarmiento, *Personnal Communication to S. D. Noble and D. J. A. Welsh* (1999).
[25] I. Sarmiento, *The polychromate and a chord diagram polynomial*, Ann. Comb. **4** (2000), no. 2, 227–236.
[26] R. P. Stanley, *A symmetric function generalization of the chromatic polynomial of a graph*, Adv. Math. **111** (1995), no. 1, 166–194.
[27] R. P. Stanley, *Graph colorings and related symmetric functions: ideas and applications. A description of results, interesting applications, & notable open problems*, Discrete Math. **193** (1998), no. 1-3, 267–286.
[28] R. Statman, *Reductions of the graph reconstruction conjecture*, Discrete Math. **36** (1981), no. 1, 103–107.
[29] B. D. Thatte, *Subgraph posets, partition lattices, graph polynomials and reconstruction*, arXiv:0609574 (2006).
[30] S. M. Ulam, *A Collection of Mathematical Problems*, Interscience Tracts in Pure and Applied Mathematics, Vol. 8 (1960), Interscience, New York–London.
[31] H. Whitney, *A logical expansion in mathematics*, Bull. Amer. Math. Soc. **38** (1932), 572–579.
[32] J. Zamora, *Structural Graph Theory*, Ph.D. Thesis, Charles University in Prague (2013).

Dept. of Applied Mathematics, Charles University, Praha, Czech Republic.
E-mail address: `loebl@kam.mff.cuni.cz`

Centre National de la Recherche Scientifique, (LORIA), Vandœuvre-lès-Nancy, France.
E-mail address: `sereni@kam.mff.cuni.cz`