

A Branch Statistic for Trees: Interpreting Coefficients of the Characteristic Polynomial of Braid Deformations

Priyavrat Deshpande[†] and Krishna Menon[‡]

[†]Department of Mathematics, Chennai Mathematical Institute, India
 Email: pdeshpande@cmi.ac.in

[‡]Department of Mathematics, Chennai Mathematical Institute, India
 Email: krishnamenon@cmi.ac.in

Received: April 10, 2022, **Accepted:** October 3, 2022, **Published:** October 7, 2022
 The authors: Released under the CC BY-ND license (International 4.0)

ABSTRACT: A hyperplane arrangement in \mathbb{R}^n is a finite collection of affine hyperplanes. The regions are the connected components of the complement of these hyperplanes. By the theorem of Zaslavsky, the number of regions of a hyperplane arrangement is the sum of the absolute values of the coefficients of its characteristic polynomial. Arrangements that contain hyperplanes parallel to subspaces whose defining equations are $x_i - x_j = 0$ form an important class called the deformations of the braid arrangement. In recent work, Bernardi showed that regions of certain deformations are in one-to-one correspondence with certain labeled trees. In this article, we define a statistic on these trees such that the distribution is given by the coefficients of the characteristic polynomial. In particular, our statistic applies to well-studied families like extended Catalan, Shi, Linal, and semiorder.

Keywords: Braid arrangement; Combinatorial statistic; Hyperplane arrangement; Labeled tree
2020 Mathematics Subject Classification: 52C35; 05C30

1. Introduction

A *hyperplane arrangement* \mathcal{A} is a finite collection of affine hyperplanes (i.e., codimension 1 subspaces and their translates) in \mathbb{R}^n . A *region* of \mathcal{A} is a connected component of $\mathbb{R}^n \setminus \bigcup \mathcal{A}$. The number of regions of \mathcal{A} is denoted by $r(\mathcal{A})$. The poset of non-empty intersections of hyperplanes in an arrangement \mathcal{A} ordered by reverse inclusion is called its *intersection poset* denoted by $L(\mathcal{A})$. The ambient space of the arrangement (i.e., \mathbb{R}^n) is an element of the intersection poset; considered as the intersection of none of the hyperplanes. The *characteristic polynomial* of \mathcal{A} is defined as

$$\chi_{\mathcal{A}}(t) := \sum_{x \in L(\mathcal{A})} \mu(\hat{0}, x) t^{\dim(x)}$$

where μ is the Möbius function of the intersection poset and $\hat{0}$ corresponds to \mathbb{R}^n . Using the fact that every interval of the intersection poset of an arrangement is a geometric lattice, we have

$$\chi_{\mathcal{A}}(t) = \sum_{i=0}^n (-1)^{n-i} c_i t^i \tag{1}$$

where c_i is a non-negative integer for all $0 \leq i \leq n$ [10, Corollary 3.4]. The characteristic polynomial is a fundamental combinatorial and topological invariant of the arrangement and plays a significant role throughout the theory of hyperplane arrangements.

In this article, our focus is on the enumerative aspects of (rational) arrangements in \mathbb{R}^n . In that direction, we have the following seminal result by Zaslavsky.

Theorem 1.1 ([12]). *Let \mathcal{A} be an arrangement in \mathbb{R}^n . Then the number of regions of \mathcal{A} is given by*

$$r(\mathcal{A}) = (-1)^n \chi_{\mathcal{A}}(-1) = \sum_{i=0}^n c_i.$$

When the regions of an arrangement are in bijection with a certain combinatorially defined set, one could ask if there is a corresponding ‘statistic’ on the set whose distribution is given by the c_i ’s. For example, the regions of the braid arrangement in \mathbb{R}^n (whose hyperplanes are given by the equations $x_i - x_j = 0$ for $1 \leq i < j \leq n$) correspond to the $n!$ permutations of $[n]$. The characteristic polynomial of this arrangement is $t(t-1)\cdots(t-n+1)$ [10, Corollary 2.2]. Hence, c_i ’s are the unsigned Stirling numbers of the first kind. Consequently, the distribution of the statistic ‘number of cycles’ on the set of permutations is given by the coefficients of the characteristic polynomial.

In this paper, we consider arrangements where each hyperplane is of the form $x_i - x_j = s$ for some $s \in \mathbb{Z}$. Such arrangements are called *deformations of the braid arrangement*. Recently, Bernardi [3] obtained a method to count the regions of any deformation of the braid arrangement using certain objects called *boxed trees*. For certain special deformations, which he calls *transitive*, he also obtained an explicit bijection between the regions of the arrangement and a certain set of trees. Our main aim is to obtain a statistic on such trees whose distribution is given by the coefficients of the characteristic polynomial of the corresponding arrangement.

For any finite set of integers S , we associate a deformation of the braid arrangement $\mathcal{A}_S(n)$ in \mathbb{R}^n with hyperplanes

$$\{x_i - x_j = k \mid k \in S, 1 \leq i < j \leq n\}.$$

Important examples of such arrangements are the Catalan, Shi, Linial, and semiorder arrangements. These correspond to $S = \{-1, 0, 1\}$, $\{0, 1\}$, $\{1\}$, and $\{-1, 1\}$ respectively. For any $m \geq 1$, the extended Catalan arrangement, or m -Catalan arrangement, in \mathbb{R}^n is $\mathcal{A}_S(n)$ where $S = \{-m, \dots, m\}$. Similarly, the extended Shi, Linial, and semiorder arrangements correspond to $S = \{-m+1, \dots, m\}$, $\{-m+1, \dots, m\} \setminus \{0\}$, and $\{-m, \dots, m\} \setminus \{0\}$ respectively.

If the set S satisfies certain conditions (see Definition 2.4), then the arrangements $\mathcal{A}_S(n)$ are called transitive. The extended Catalan, Shi, Linial, and semiorder arrangements are all transitive. We note here that Bernardi [3] considers a larger class of arrangements to be transitive, but we only focus on arrangements of the form described above.

From [3, Theorem 3.8], we know that if S is transitive, then the regions of $\mathcal{A}_S(n)$ are in bijection with a certain set of trees $\mathcal{T}_S(n)$ (see Definition 2.3). For example, when $S = \{0, 1\}$ which corresponds to the Shi arrangement, $\mathcal{T}_{\{0,1\}}(n)$ is the set of labeled binary trees with n nodes where any right node has a label smaller than its parent.

Example 1.1. A tree in $\mathcal{T}_{\{0,1\}}(4)$ is shown in Figure 1.

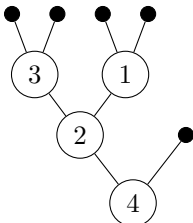


Figure 1: A tree in $\mathcal{T}_{\{0,1\}}(4)$

For such a tree, we define the *trunk* to be the path from the root to the leftmost leaf. Using the nodes on the trunk, we obtain a sequence of numbers. A node in this sequence that is greater than all the nodes after it is called a *branch node*. For the tree in Figure 1, the sequence on the trunk is 4, 2, 3 and the branch nodes are 4 and 3. These definitions can be generalized to trees in $\mathcal{T}_S(n)$ for other sets S .

The main theorem of this article is:

Theorem 1.2. For a transitive set S , the absolute value of the coefficient of t^j in $\chi_{\mathcal{A}_S(n)}(t)$ is the number of trees in $\mathcal{T}_S(n)$ with j branch nodes.

The article begins with a short account of Bernardi’s work [3] in Section 2. In Section 3 the branch statistic is introduced and the main theorem is proved. In Section 4 we derive some properties of the coefficients of the characteristic polynomial and in particular study the extended Catalan arrangements.

An extended abstract of this article [5] will appear in a proceedings volume of Séminaire Lotharingien Combinatoire for FPSAC 2022.

2. Preliminaries

A *tree* is a graph with no cycles. A *rooted tree* is a tree with a distinguished vertex called the root. We will draw rooted trees with their root at the bottom. Children of a vertex v in a rooted tree are those vertices w that are

adjacent to v and such that the unique path from the root to w passes through v . Similarly, we can define the parent of a vertex v to be the vertex w for which v is the child of w . Any non-root vertex has a unique parent. All the vertices that have at least one child are called *nodes* and those that do not are called *leaves*.

A *rooted plane tree* is a rooted tree with a specified ordering for the children of each node. When drawing a rooted plane tree, the children of any node will be ordered from left to right. The *left siblings* of a vertex v are the vertices that are also children of the parent of v but are to the left of v . We denote the number of left siblings of v as $\text{lsib}(v)$.

Definition 2.1. An $(m + 1)$ -ary tree is a rooted plane tree where each node has exactly $(m + 1)$ children. We will denote by $\mathcal{T}^{(m)}(n)$ the set of all $(m + 1)$ -ary trees with n nodes labeled with distinct elements from $[n]$.

For trees in $\mathcal{T}^{(m)}(n)$, we will denote the node having label $i \in [n]$ by just i .

Definition 2.2. If a node i in a tree $T \in \mathcal{T}^{(m)}(n)$ has at least one child that is a node, the *cadet* of i is the rightmost such child, which we denote by $\text{cadet}(i)$.

Example 2.1. Figure 1 shows an element of $\mathcal{T}^{(1)}(4)$ where

- 4 is the root,
- $\text{lsib}(2) = 0, \text{lsib}(3) = 0, \text{lsib}(1) = 1,$
- $\text{cadet}(4) = 2,$ and $\text{cadet}(2) = 1.$

Definition 2.3. For any finite set of integers S with $m = \max\{|s| \mid s \in S\}$, define $\mathcal{T}_S(n)$ to be the set of trees in $\mathcal{T}^{(m)}(n)$, such that if $\text{cadet}(i) = j$:

- $\text{lsib}(j) \notin S \cup \{0\} \Rightarrow i < j.$
- $-\text{lsib}(j) \notin S \Rightarrow i > j.$

Recall that for any finite set of integers S , we defined the arrangement $\mathcal{A}_S(n)$ as the deformation of the braid arrangement in \mathbb{R}^n with hyperplanes

$$\{x_i - x_j = k \mid k \in S, 1 \leq i < j \leq n\}.$$

Though Bernardi [3] derived results for more general deformations, we will only be focused on these.

Definition 2.4. A finite set of integers S is said to be *transitive* if for any $s, t \notin S$,

- $st > 0 \Rightarrow s + t \notin S.$
- $s > 0$ and $t \leq 0 \Rightarrow s - t \notin S$ and $t - s \notin S.$

Example 2.2. As mentioned in Section 1, for any $m \geq 1$, the sets $\{-m, \dots, m\}, \{-m+1, \dots, m\}, \{-m, \dots, m\} \setminus \{0\}$, and $\{-m+1, \dots, m\} \setminus \{0\}$ are all transitive.

We can now state the result for arrangements $\mathcal{A}_S(n)$ where S is transitive.

Theorem 2.1. [3, Theorem 3.8] For any transitive set of integers S , the regions of the arrangement $\mathcal{A}_S(n)$ are in bijection with the trees in $\mathcal{T}_S(n)$.

Before looking at the characteristic polynomials of such arrangements, we recall a few results from [9]. Suppose that $c : \mathbb{N} \rightarrow \mathbb{N}$ is a function and for each $n, j \in \mathbb{N}$, we define

$$c_j(n) = \sum_{\{B_1, \dots, B_j\} \in \Pi_n} c(|B_1|) \cdots c(|B_j|)$$

where Π_n is the set of partitions of $[n]$. Define for each $n \in \mathbb{N}$,

$$h(n) = \sum_{j=0}^n c_j(n).$$

From [9, Example 5.2.2], we know that in such a situation,

$$\sum_{n, j \geq 0} c_j(n) t^j \frac{x^n}{n!} = \left(\sum_{n \geq 0} h(n) \frac{x^n}{n!} \right)^t.$$

Informally, we consider $h(n)$ to be the number of “structures” that can be placed on an n -set where each structure can be uniquely broken up into a disjoint union of “connected sub-structures”. Here $c(n)$ denotes the number of connected structures on an n -set and $c_j(n)$ denotes the number of structures on an n -set with exactly j connected sub-structures.

We now consider the characteristic polynomials of arrangements of the form $\mathcal{A}_S(n)$. For a fixed set S , the sequence of arrangements $(\mathcal{A}_S(1), \mathcal{A}_S(2), \dots)$ forms what is called an *exponential sequence of arrangements* (ESA).

Definition 2.5. [10, Definition 5.14] A sequence of arrangements $(\mathcal{A}_1, \mathcal{A}_2, \dots)$ is called an ESA if

- \mathcal{A}_n is a braid deformation in \mathbb{R}^n .
- For any k -subset I of $[n]$, the arrangement

$$\mathcal{A}_n^I = \{H \in \mathcal{A}_n \mid H \text{ is of the form } x_i - x_j = s \text{ for some } i, j \in I\}$$

satisfies $L(\mathcal{A}_n^I) \cong L(\mathcal{A}_k)$ (isomorphic as posets).

The result on ESAs that we will need is the following.

Theorem 2.2. [10, Theorem 5.17] If $(\mathcal{A}_1, \mathcal{A}_2, \dots)$ is an ESA, then

$$\sum_{n \geq 0} \chi_{\mathcal{A}_n}(t) \frac{x^n}{n!} = \left(\sum_{n \geq 0} (-1)^n r(\mathcal{A}_n) \frac{x^n}{n!} \right)^{-t}.$$

Remark 2.1. We note that this is also a special case of [3, Theorem 5.2].

Using this result, the form of a characteristic polynomial given in (1), and the above discussion on connected structures, we note that interpreting the coefficients of the polynomial $\chi_{\mathcal{A}_S(n)}(t)$ is equivalent to defining a notion of “connected structures” for trees in $\mathcal{T}_S(n)$. We do this in the next section.

3. A branch statistic

A *label set* is a finite set of positive integers. For any label set V , we define $\mathcal{T}^{(m)}(V)$ to be the set of $(m+1)$ -ary trees with $|V|$ nodes labeled distinctly using V . Note that $\mathcal{T}^{(m)}([n]) = \mathcal{T}^{(m)}(n)$.

We now describe the method we use to break up a tree in $\mathcal{T}^{(m)}(V)$ into “connected sub-structures”, which we call *branches*.

Definition 3.1. The *trunk* of a tree in $\mathcal{T}^{(m)}(V)$ is the path from the root to the leftmost leaf. The nodes on the trunk of the tree break up the tree into sub-trees, which we call *twigs* (see Figure 2).

Let the nodes on the trunk of a tree be v_1, v_2, \dots, v_k , where v_1 is the root and v_{i+1} is the leftmost child of v_i for any $i \in [k-1]$. If $v_i = \max\{v_1, \dots, v_k\}$, then the first branch of the tree consists of the twigs corresponding to the nodes v_1, \dots, v_i . If $v_j = \max\{v_{i+1}, \dots, v_k\}$, then the second branch of the tree consists of the twigs corresponding to the nodes v_{i+1}, \dots, v_j . Continuing this way, we break up the tree into branches.

Note that the number of branches of the tree is just the number of right-to-left maxima of the sequence v_1, v_2, \dots, v_k of nodes on the trunk, *i.e.*, the number of v_i such that $v_i > v_j$ for all $j > i$. We will call such v_i the *branch nodes* of the trunk.

Example 3.1. The tree in Figure 2 has 3 twigs and 2 branches. The first branch consists of just the first twig since 6 is the largest node in the trunk. The second branch consists of the second and third twigs since 5 is larger than 4. Here 6 and 5 are the branch nodes.

We use the notation $\mathcal{T}_j^{(m)}(V)$ to denote the trees in $\mathcal{T}^{(m)}(V)$ having j branches. To prove that this is indeed a break-up of trees into connected sub-structures, we have to prove that

$$|\mathcal{T}_j^{(m)}(V)| = \sum_{\{B_1, \dots, B_j\} \in \Pi_V} |\mathcal{T}_1^{(m)}(B_1)| \cdots |\mathcal{T}_1^{(m)}(B_j)|. \tag{2}$$

Hence, “connected” trees are those with exactly one branch, *i.e.*, trees where the last node of the trunk is the one with the largest label.

The connected components associated to a given tree are the branches of the tree.

Example 3.2. The connected components associated with the tree in Figure 2 are given in Figure 3.

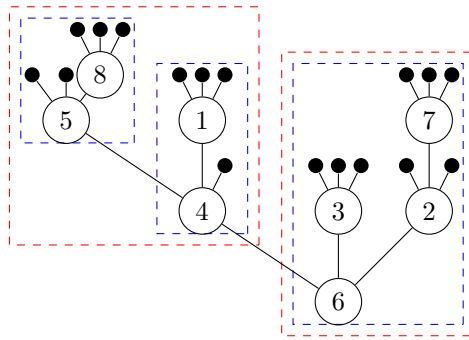


Figure 2: A labeled 3-ary tree with twigs and branches specified.

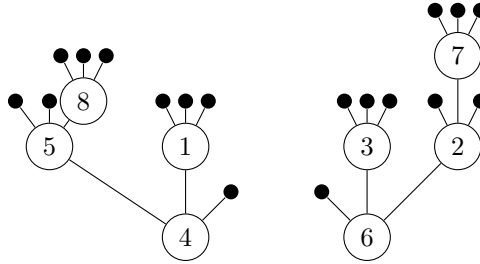


Figure 3: Connected components of the tree in Figure 2.

A collection of connected trees (with disjoint label sets) can be put together in exactly one way to form a tree for which they form the branches. This is done as follows: Find the largest label among those on the trunks of the connected trees. The connected tree T_1 with this label is made the first branch of the tree we are building. Again, find the largest label among those on the trunks of the remaining connected trees. The connected tree T_2 with this label is made the second branch of the tree we are building by gluing it to T_1 . This is done by deleting the leftmost leaf of T_1 and fixing the root of T_2 in its position. This process is repeated until all the connected trees are glued together.

Example 3.3. *The tree associated with the collection of connected trees in Figure 4 is given in Figure 5.*

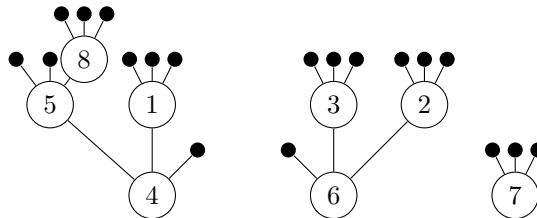


Figure 4: A collection of connected trees.

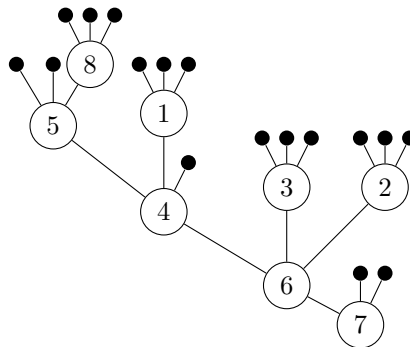


Figure 5: The tree associated with the collection of connected trees in Figure 4.

Recall that for a finite set of integers S with $m = \max\{|s| : s \in S\}$, the set $\mathcal{T}_S(V)$, for some label set V , is the set of trees in $\mathcal{T}^{(m)}(V)$ such that if $\text{cadet}(u) = v$, then

- if $\text{lsib}(v) \notin S \cup \{0\}$, we must have $u < v$, and
- if $-\text{lsib}(v) \notin S$, we must have $u > v$.

We call this set of conditions ‘‘Condition S ’’.

We set $\mathcal{T}_S := \bigcup_V \mathcal{T}_S(V)$ where the union is overall label sets V . We now show that

1. the connected components of any tree in \mathcal{T}_S are also in \mathcal{T}_S , and
2. trees that are built using connected trees in \mathcal{T}_S are also in \mathcal{T}_S .

We first note that statement 1 follows since the condition for a tree to be in \mathcal{T}_S is a local condition. This is also because, if T' is a connected component of the tree T , the cadet of any node in T' (if it exists) is the same as its cadet when considered as a node of T .

To prove statement 2, we only have to check that Condition S is satisfied for the branch nodes of a tree built using connected trees in \mathcal{T}_S . If a branch node does not have a cadet, Condition S is trivially satisfied. If a branch node u has a cadet v , we consider two cases:

- If the cadet is not the first child, then Condition S is satisfied since it is satisfied by the connected components of the tree.
- If the cadet is the first child, then we must have $u > v$ since u is a branch node. This makes sure that Condition S is satisfied since we have $\text{lsib}(v) = 0$ and hence $\text{lsib}(v) \in S \cup \{0\}$.

From the preceding, we get an equation analogous to (2) for the trees \mathcal{T}_S . Hence, from the discussion in Section 2, we get the following result.

Theorem 3.1. *For a transitive set of integers S , the absolute value of the coefficient of t^j in $\chi_{\mathcal{A}_S(n)}(t)$ is the number of trees in $\mathcal{T}_S(n)$ with j branches.*

Example 3.4. *When $S = \{0\}$, we obtain the braid arrangement. Here, $\mathcal{T}_{\{0\}}(n)$ corresponds to permutations of $[n]$ and Theorem 3.1 states that the absolute value of the coefficient of t^j in $\chi_{\mathcal{A}_{\{0\}}(n)}(t)$ is the number of permutations of $[n]$ with j right-to-left maxima. By [8, Corollary 1.3.11], this agrees with the observation in Section 1 that the coefficients are the Stirling numbers of the first kind.*

Example 3.5. *The Shi arrangement \mathcal{S}_n in \mathbb{R}^n is the deformation $\mathcal{A}_{\{0,1\}}(n)$. The trees in $\mathcal{T}_{\{0,1\}}(n)$, called Shi trees, are those labeled binary trees where any right node has a label less than that of its parent. The Shi trees for $n = 3$ are given in Figures 6 and 7. Counting the branches in these trees, we get $\chi_{\mathcal{S}_3}(t) = t^3 + 6t^2 + 9t$, which agrees with the known formula for the characteristic polynomial (for example, see [2, Theorem 3.3]).*

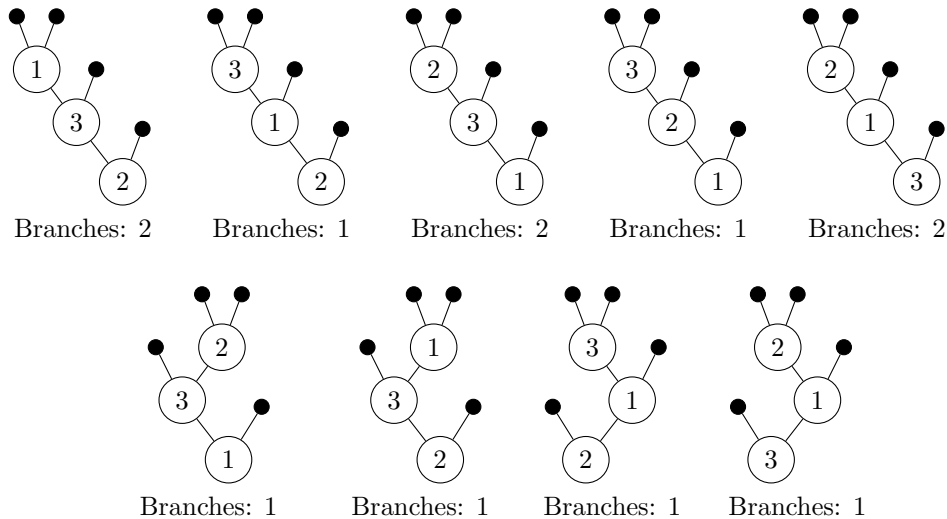


Figure 6: Shi trees for $n = 3$ that are not Linial.

Remark 3.1. *The Shi trees $\mathcal{T}_{\{0,1\}}(n)$ are in bijection with Cayley trees on $n + 1$ vertices. Using a decomposition of Cayley trees, one can show that the coefficient of t^j in $\chi_{\mathcal{S}_n}(t)$ is the number of such Cayley trees where the vertex $n + 1$ has degree j .*

Example 3.6. The Linal arrangement \mathcal{L}_n in \mathbb{R}^n is the deformation $\mathcal{A}_{\{1\}}(n)$. The trees in $\mathcal{T}_{\{1\}}(n)$, called Linal trees, are those Shi trees that also satisfy the property that any left node whose sibling is a leaf has a smaller label than that of its parent. The Linal trees for $n = 3$ are given in Figure 7. Counting the branches in these trees, we get $\chi_{\mathcal{L}_3}(t) = t^3 + 3t^2 + 3t$, which agrees with the known formula for the characteristic polynomial (for example, see [2, Theorem 4.2]).

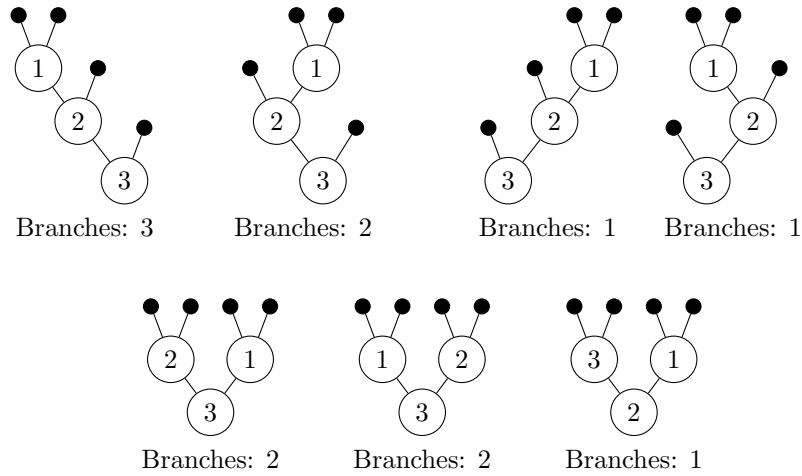


Figure 7: Linal trees for $n = 3$.

4. Properties of coefficients

In this section, we use the combinatorial interpretation of the coefficients to obtain some of their properties, some of which are not obvious. For any transitive set S , we use $C(S, n, j)$ to denote the absolute value of the coefficient of t^j in $\chi_{\mathcal{A}_S(n)}(t)$.

Proposition 4.1. For any transitive sets $S' \subseteq S$, we have

$$C(S', n, j) \leq C(S, n, j).$$

Proof. As stated in [3, Remark 3.10], we have $\mathcal{T}'_S \subseteq \mathcal{T}_S$. When $\max\{|s'| \mid s' \in S'\} < \max\{|s| \mid s \in S\}$, this inclusion can be obtained via adding the required number of leaves to each node of the trees in \mathcal{T}'_S . This inclusion preserves the number of branches and hence gives us the required result. \square

Proposition 4.2. For any transitive set S , we have

$$C(S, n, j) \leq C(S, n + 1, j + 1).$$

Proof. Given a tree $T \in \mathcal{T}_S(n)$, we construct one in $\mathcal{T}_S(n + 1)$ that has root $n + 1$ with T attached as the leftmost child and all other children being leaves. This proves the required result. \square

Proposition 4.3. For any transitive set S , we have

$$C(S, n, j) \leq C(S, n + 1, j).$$

Proof. The result follows by increasing the label of each node and replacing the leftmost leaf with a node labeled 1 with all leaf children. \square

We can derive some more properties for a particular class of transitive sets. This follows by using a different break-up of trees into connected components from the one presented in Section 3.

Proposition 4.4. If S is a transitive set such that $0 \in S$ and there exists some $k \geq 1$ such that $k, -k \in S$, we have

$$C(S, n, 1) \geq \sum_{j=2}^n C(S, n, j).$$

Proof. We have to show that there are more connected trees in $\mathcal{T}_S(n)$ than disconnected trees. To prove this, we use *all* the nodes, not just the ones on the trunk, to break up the tree into *compartments*.

Let the nodes on the trunk of a tree be v_1, v_2, \dots, v_k , where v_1 is the root and v_{i+1} is the leftmost child of v_i for any $i \in [k - 1]$. If the twig corresponding to v_i is the one that contains the label n , then the first compartment of the tree consists of the twigs corresponding to the nodes v_1, \dots, v_i . If the twig corresponding to v_j is the one that has the node with maximum label among the twigs corresponding to v_{i+1}, \dots, v_k , then the second compartment of the tree consists of the twigs corresponding to the nodes v_{i+1}, \dots, v_j . Continuing this way, we break up the tree into compartments. For example, the tree in Figure 1 has 2 compartments and that in Figure 5 has 1 compartment.

It can be checked that this is a valid break-up of trees into connected components since there is no condition relating a node to its leftmost child since $0 \in S$. Hence, $C(S, n, j)$ is the number of trees in $\mathcal{T}_S(n)$ with j compartments. In this situation, disconnected trees are those trees where n is not in the last twig of the tree and connected trees are those where it is. To prove the result, we have to show that the number of disconnected trees is less than the number of connected trees.

Let T be a disconnected tree. First, let us suppose that the label n is not in the first twig of T . Let i be the node on the trunk whose corresponding twig has the node n and j be the parent of i . Fix some $k \geq 1$ such that $k, -k \in S$ which exists by our hypothesis. Let T_j be the subtree of T whose root is the $(k + 1)^{th}$ child of j . We construct a connected tree T' by breaking off the subtree of T whose root is i and attaching it to T_j by replacing the leftmost leaf of T_j with the node i . It can be checked that $k, -k \in S$ implies that T' will indeed be in $\mathcal{T}_S(n)$.

Example 4.1. If T is the disconnected tree in Figure 8 and $k = 1$, then T' is the connected tree in Figure 9.

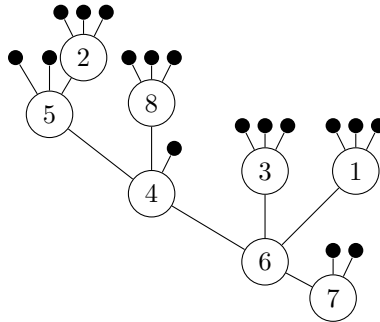


Figure 8: Disconnected tree with the largest label not in the first twig.

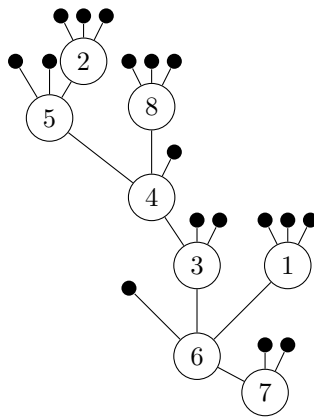


Figure 9: Connected tree associated to the tree in Figure 8.

To obtain T back from T' , we first need the following definition from [3]. We define $\text{drift}(u) = 0$ when u is the root of a tree and for any other vertex v with parent w , define $\text{drift}(v) = \text{lsib}(v) + \text{drift}(w)$.

To obtain T back from T' , we note that i is the vertex with drift k that is furthest from the root in the unique path from the root to the node n . Now T is obtained by breaking off the subtree of T' with root i and replacing the leftmost leaf of T' with this subtree.

Next, suppose that the root of T is i and the label n is in the first twig. Let j be the last node on the trunk, which is necessarily different from i since T is disconnected. Just as before, let T_j be the subtree of T whose root is the $(k + 1)^{th}$ of j and construct T' by removing the first twig of T and using it to replace the leftmost leaf of T_j .

Example 4.2. If T is the disconnected tree on the left in Figure 10 and $k = 1$, then T' is the connected tree on the right.

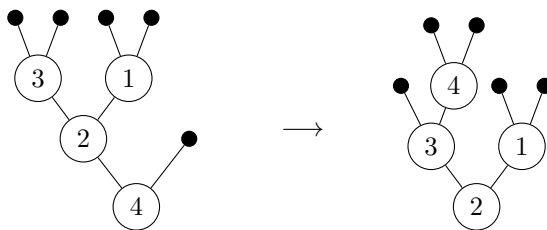


Figure 10: Disconnected tree with the largest label in the first twig and associated connected tree.

We can determine the root i of T from T' just as in the previous case, *i.e.*, i is the vertex with drift k that is furthest from the root in the unique path from the root to the node n . Now T is obtained from T' by breaking off the subtree with root i and making it the first twig.

Note that this is an injection from the set of disconnected trees to the set of connected trees. This is because we can obtain T from T' in each case above and there is no overlap of the connected trees in the two cases. This is because, in the first case, the leftmost child of the node i is always a node, whereas it is always a leaf in the second case. This proves the required result. \square

Proposition 4.5. If S is a transitive set such that $0 \in S$ and there exists some $k \geq 1$ such that $k, -k \in S$, we have

$$C(S, n, j) \geq C(S, n, j + 1).$$

Proof. Using the “connected structure” property of branches and considering the size of the branch containing the label n , we get for $j \geq 2$,

$$C(S, n, j) = \sum_{k=1}^{n-1} \binom{n-1}{k-1} C(S, k, 1) C(S, n-k, j-1).$$

By induction on j , this shows that it is enough to prove that $C(S, n, 1) \geq C(S, n, 2)$ for all $n \geq 1$. This follows from Proposition 4.4. \square

4.1 Extended Catalan arrangement

We now focus on the case when $S = \{-m, -m + 1, \dots, m - 1, m\}$ for some $m \geq 1$. The corresponding arrangement $\mathcal{A}_S(n)$ is called the m -Catalan arrangement in \mathbb{R}^n . We let $C(m, n, j)$ denote the absolute value of the coefficient of t^j in $\chi_{\mathcal{A}_S(n)}(t)$. Here $\mathcal{T}_S(n) = \mathcal{T}^{(m)}(n)$ and hence, from Theorem 3.1, $C(m, n, j)$ is the number of $(m + 1)$ -ary trees with n nodes and j branches. We now compute expressions for $C(m, n, j)$ using this combinatorial interpretation.

Proposition 4.6. We have

$$C(m, n, j) = \sum_{k=j}^n T_m(n, k) \binom{n}{k} c(k, j) (n - k)!$$

where

- $c(k, j)$ is the number of permutations of $[k]$ with j right-to-left maxima (unsigned Stirling number of first kind), and
- $T_m(n, k)$ is the number of unlabeled $(m + 1)$ -ary trees with n nodes, k of which are on its trunk, given by

$$\frac{mk}{(m + 1)n - k} \binom{(m + 1)n - k}{n - k}.$$

Proof. Specifying a tree in $\mathcal{T}^{(m)}(n)$ with j branches is done by

1. choosing an unlabeled $(m + 1)$ -ary tree with n nodes and $k \geq j$ nodes on its trunk,
2. choosing the k labels from $[n]$ for the trunk,
3. labeling the trunk so that it has j right-to-left maxima, and

4. using the rest of the labels for the non-trunk nodes.

The formula for $T_m(n, k)$ follows from the bijection between trees and Dyck paths given in [3, Section 8] and the discussion about returns in Dyck paths in [4, Page 4]. One can check that trunk nodes correspond to returns in the Dyck path under this bijection. \square

Proposition 4.7. *We also have for any $m, n, j \geq 1$,*

$$C(m, n, j) = \sum_{k=j}^n (-1)^{k-j} B_m(n, k) c(k, j)$$

where

$$B_m(n, k) = \frac{(n-1)!}{(k-1)!} \binom{(m+1)n}{n-k}.$$

Proof. It can be checked, for example using [9, Theorem 5.3.10], that $B_m(n, k)$ is the number of ways to partition $[n]$ into k blocks and associate to each block B a tree in $\mathcal{T}^{(m)}(B)$. Using the fact that branches give the trees an exponential structure, another way to partition $[n]$ into k blocks and associate to each block B a tree in $\mathcal{T}^{(m)}(B)$ is to

1. select a collection of $i \geq k$ connected trees whose label sets partition $[n]$ (which is the same as choosing a tree in $\mathcal{T}_i^{(m)}(n)$),
2. partition these trees into k sets, and
3. combine the connected trees in each of these k sets to form k trees.

This shows

$$B_m(n, k) = \sum_{i=k}^n C(m, n, i) S(i, k)$$

where $S(i, k)$ is the number of ways to partition the set $[i]$ into k blocks, *i.e.*, they are Stirling numbers of the second kind. The result now follows by Möbius inversion, since by [8, Proposition 1.9.1], we have

$$\sum_{k \geq j} (-1)^{k-j} S(i, k) c(k, j) = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j. \end{cases}$$

\square

Remark 4.1. *The triangle of numbers $C(1, n, j)$ is listed in the OEIS [6] as A038455. For $m \geq 2$, the triangle $C(m, n, j)$ does not seem to be listed.*

We now derive some properties of $C(m, n, j)$. First, we list properties that are consequences of the general properties we have already seen.

Proposition 4.8. *For any $m, n, j \geq 1$, we have:*

1. $C(m, n, j) \leq C(m+1, n, j)$
2. $C(m, n, j) \leq C(m, n+1, j)$
3. $C(m, n, 1) \geq \sum_{k=2}^n C(m, n, k)$
4. $C(m, n, j) \geq C(m, n, j+1)$

We now list some properties that are specific to the case of the extended Catalan arrangement.

Proposition 4.9. *For any $m, n \geq 1$, we have*

$$C^{(m)}(n) := \sum_{j=1}^n C(m, n, j) = \frac{n!}{mn+1} \binom{(m+1)n}{n}.$$

Proof. This follows from the known formula for $|\mathcal{T}^{(m)}(n)|$ (see [9, Section 5.3]). \square

Proposition 4.10. For any $m, n \geq 1$, we have

$$C^{(m)}(n) \leq C(m + 1, n, 1).$$

Proof. Adding a leaf as the first child of each node in a labeled $(m + 1)$ -ary tree gives the required result. \square

Proposition 4.11. For any $m, n \geq 1$, we have

$$C^{(m)}(n) \leq C(m, n + 1, 1).$$

Proof. The result follows by replacing the leftmost leaf of a tree in $\mathcal{T}^{(m)}(n)$ with a node labeled $n + 1$ with all children being leaves. \square

There are several combinatorial objects that correspond to the regions of the extended Catalan arrangement (especially in the case $m = 1$, see [11]). One such is the generalized Dyck paths. We now describe a corresponding statistic for these Dyck paths.

A labeled m -Dyck path on $[n]$ is a sequence of $(m + 1)n$ terms where

1. n terms are ‘ $+m$ ’,
2. mn terms are ‘ -1 ’,
3. the sum of any prefix of the sequence is non-negative, and
4. each $+m$ term is given a distinct label from $[n]$.

A labeled m -Dyck path on $[n]$ can be drawn in \mathbb{R}^2 in a natural way. Start the path at $(0, 0)$, read the labeled m -Dyck path and for each term move by $(1, m)$ if it is $+m$ and by $(1, -1)$ if it is -1 . Also, label each $+m$ step with its corresponding label in $[n]$.

A Dyck path breaks up into *primitive* parts based on when it touches the x -axis. If a labeled Dyck path has k primitive parts, then we break the path into *compartments* as follows. If the number n is in the i_1^{th} primitive part, then the primitive parts up to the i_1^{th} form the first compartment. Let j be the largest number in $[n] \setminus A$ where A is the set of numbers in the first compartment. If j is in the i_2^{th} primitive part then the primitive parts after the i_1^{th} up to the i_2^{th} form the second compartment. Continuing this way, we break up a labeled Dyck path into compartments.

Example 4.3. The labeled 1-Dyck path on $[7]$ given in Figure 11 has 3 primitive parts and 2 compartments.

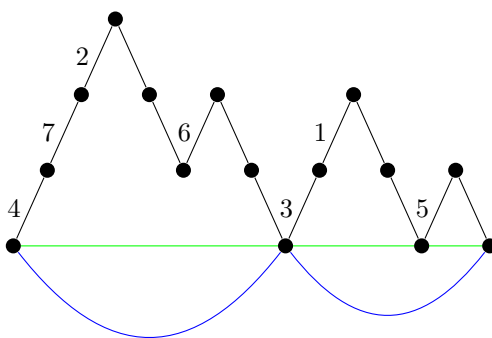


Figure 11: A labeled 1-Dyck path with compartments specified.

It can be checked that this is a valid break-up of Dyck paths into connected structures. Thus have the following.

Theorem 4.1. The number of labeled m -Dyck paths on $[n]$ with j compartments is $C(m, n, j)$.

We say that a labeled Dyck path has j right-to-left maxima if the string of labels before its first down step has j right-to-left maxima. For example, the string of labels before the first down step in the Dyck path in Figure 11 is 4, 7, 2. Hence, it has 2 right-to-left maxima.

Using the bijection between labeled trees and labeled Dyck paths given in [3] and the result in Section 3, we get another statistic on labeled Dyck paths with the same distribution.

Theorem 4.2. The number of labeled m -Dyck paths on $[n]$ with j right-to-left maxima is $C(m, n, j)$.

5. Concluding remarks

We note that a combinatorial interpretation for the coefficients of the characteristic polynomial of the Linnell arrangement is already given in [7, Corollary 4.2]. This is in terms of alternating trees.

For various deformations of the braid arrangement, expressions for the characteristic polynomials are known (for example, see [1, 2]). Hence, for transitive sets S , these can be used to extract coefficients and hence give formulas for the number of trees in \mathcal{T}_S according to the number of nodes and branches.

Acknowledgements

The authors are partially supported by a grant from the Infosys Foundation. We thank the anonymous referee for helpful suggestions.

References

- [1] F. Ardila, *Computing the Tutte polynomial of a hyperplane arrangement*, Pacific J. Math. 230 (2007), 1–26.
- [2] C. A. Athanasiadis, *Characteristic polynomials of subspace arrangements and finite fields*, Adv. Math. 122 (1996), 193–233.
- [3] O. Bernardi, *Deformations of the braid arrangement and trees*, Adv. Math. 335 (2018), 466–518.
- [4] N. T. Cameron and J. McLeod, *Returns and hills on generalized Dyck paths*, J. Integer Seq. 19 (2016), 16.6.1.
- [5] P. Deshpande and K. Menon, *A statistic for regions of braid deformations*, Formal Power Series and Algebraic Combinatorics, 2022.
- [6] N. J. A. Sloane et al., *The On-Line Encyclopedia of Integer Sequences*, 2019. Available at <https://oeis.org>.
- [7] R. P. Stanley, *Hyperplane arrangements, interval orders, and trees*, Proc. Natl. Acad. Sci. USA 93 (1996), 2620–2625.
- [8] R. P. Stanley, *Enumerative Combinatorics, Volume I*, Cambridge University Press, Cambridge, UK, 1997.
- [9] R. P. Stanley, *Enumerative Combinatorics, Volume II*, Cambridge University Press, Cambridge, UK, 2001.
- [10] R. P. Stanley, *An introduction to hyperplane arrangements*, Geometric Combinatorics (2007), Volume 13 IAS/Park City Math. Ser., 389–496, Amer. Math. Soc., Providence, RI.
- [11] R. P. Stanley, *Catalan Numbers*, Cambridge University Press, Cambridge, UK, 2015.
- [12] T. Zaslavsky, *Facing up to arrangements: Face-count formulas for partitions of space by hyperplanes*, Mem. Am. Math. Soc. 154 (1975), Providence, RI: American Mathematical Society (AMS).