ICECA

International Conference
Enumerative Combinatorics and Applications
University of Haifa – Virtual – September 4-6, 2023

A branch statistic for trees

Krishna Menon

Chennai Mathematical Institute

krishnamenon@cmi.ac.in

Abstract: A *hyperplane arrangement* in $\mathbb{R}^n$ is a finite collection of affine hyperplanes. The *regions* of an arrangement are the connected components of the space obtained when its hyperplanes are deleted from $\mathbb{R}^n$. By a theorem of Zaslavsky, the number of regions of an arrangement is the sum of the absolute values of the coefficients of its characteristic polynomial. Arrangements consisting of hyperplanes parallel to those whose defining equations are $x_i - x_j = 0$ form an important class called *deformations of the braid arrangement*. In a recent work, Bernardi showed that regions of certain deformations are in one-to-one correspondence with certain labeled trees. 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, Linial and semiorder. This is based on joint work with Priyavrat Deshpande.

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 $\mathrm{L}(\mathcal{A})$. 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

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

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 $0 \leq i \leq n$ [8, 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.

We have the following seminal result by Zaslavsky for obtaining the number of regions of an arrangement from its characteristic polynomial.

Theorem 1.1. [10] *Let $\mathcal{A}$ be an arrangement in $\mathbb{R}^n$. Then the number of regions of $\mathcal{A}$ is given by*

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

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 $t(t-1)\cdots(t-n+1)$ [8, 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.

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 describe a statistic on such trees, which we call the branch statistic, whose distribution is given by the coefficients of the characteristic polynomial of the corresponding arrangement.

We begin with a short account of Bernardi's work [3] in Section 2. In Section 3 we describe the branch statistic. In Section 4 we exhibit some properties of the coefficients of the characteristic polynomial that can be proved using this combinatorial interpretation. Details can be found in [4].

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 use i to denote the node having label $i \in [n]$.

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

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

Definition 2.4. *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* $\operatorname{cadet}(i) = j$:

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

Example 2.5. $\mathcal{T}_{\{0,1\}}(n)$ *is the set of labeled binary trees with* n *nodes where any right node has a label smaller than its parent. A tree in* $\mathcal{T}_{\{0,1\}}(4)$ *is shown in Figure 1.*

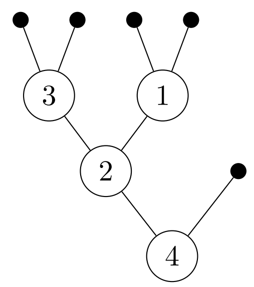

FIGURE 1. A tree in $\mathcal{T}_{\{0,1\}}(4)$

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

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

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

Definition 2.6. *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* $t \le 0 \Rightarrow s - t \notin S$ *and* $t - s \notin S$.

Example 2.7. *For any* $m \ge 1$, *the sets* $\{-m, \dots, m\}$, $\{-m+1, \dots, m\}$, $\{-m, \dots, m\} \setminus \{0\}$, *and* $\{-m+1, \dots, m\} \setminus \{0\}$ *are all transitive. The arrangements* $\mathcal{A}_S(n)$ *corresponding to these sets when* $m = 1$ *are called the Catalan, Shi, semiorder, and Linial arrangements respectively.*

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

Theorem 2.8. [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* $\mathcal{T}_S(n)$.

3. A BRANCH STATISTIC

We first break up a tree into *twigs* and group twigs together to form *branches*.

Definition 3.1. *The* trunk *of a tree in* $\mathcal{T}^{(m)}(n)$ *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 *(blue boxes in 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 group twigs to form 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 3.2. *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.*

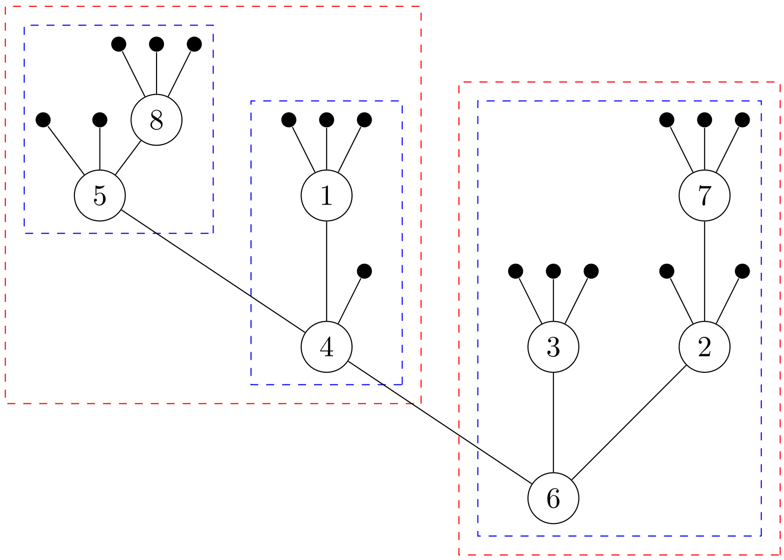

FIGURE 2. A labeled 3-ary tree with twigs and branches specified.

Theorem 3.3. *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.*

The main idea behind the proof is that the sequence of arrangements $(\mathcal{A}_S(n))_{n\geq 0}$ forms an exponential sequence of arrangements [8, Definition 5.14] and that branches give trees an exponential structure [7, Example 5.2.2].

Example 3.4. *When $S = \{0\}$, we obtain the braid arrangement. Here, $\mathcal{T}_{\{0\}}(n)$ corresponds to permutations of $[n]$ and Theorem 3.3 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* [6, 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 Linial arrangement $\mathcal{L}_n$ in $\mathbb{R}^n$ is the deformation $\mathcal{A}_{\{1\}}(n)$. The trees in $\mathcal{T}_{\{1\}}(n)$, called Linial trees, are those binary trees where any node has a larger label than its cadet. The Linial trees for $n = 3$ are given in Figure 3. 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]*).*

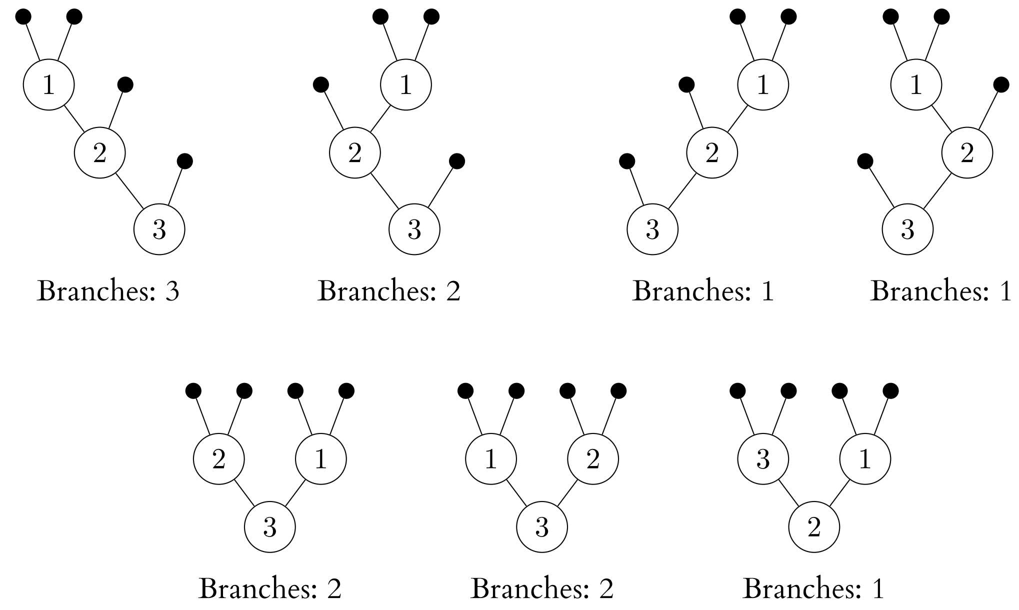

FIGURE 3. Linial trees for $n = 3$.

4. PROPERTIES OF COEFFICIENTS

In this section, we present some properties of the coefficients that are consequences of the combinatorial interpretation presented in the previous section. 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 set S and $n, j \geq 1$, we have the following:*

- $C(S', n, j) \leq C(S, n, j)$ *for any transitive set $S' \subseteq S$.*
- $C(S, n, j) \leq C(S, n+1, j+1)$.
- $C(S, n, j) \leq C(S, n+1, j)$.

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

Proposition 4.2. *Suppose S is a transitive set such that $0 \in S$ and there exists some $s \geq 1$ such that $s, -s \in S$. For any $j \geq 1$, we have*

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

In particular, we have $C(S, n, 1) \geq C(S, n, 2) \geq \cdots \geq C(S, n, n-1) \geq C(S, n, n)$.

4.1. Extended Catalan arrangement.

We now focus on the case when $S = \{-m, -m+1, \ldots, m-1, m\}$ 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 3.3, $C(m, n, j)$ is the number of $(m+1)$-ary trees with n nodes and j branches. We obtain the following expression for $C(m, n, j)$ using this combinatorial interpretation.

Proposition 4.3. *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 $c(k,j)$ is the number of permutations of $[k]$ with j right-to-left maxima (unsigned Stirling number of the first kind) and

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

The above expression follows since one can show, for example using [7, 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 $\mathcal{T}^{(m)}(|B|)$.

We now state some properties of $C(m,n,j)$ that can be easily proved using this combinatorial interpretation. We omit those that are consequences of the general properties we have already seen.

Proposition 4.4. *For any $m, n \geq 1$, we have the following:*

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

There are several combinatorial objects that correspond to the regions of the extended Catalan arrangement (especially in the case $m=1$, see [9]). 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

- n terms are '$+m$',
- mn terms are '-1',
- the sum of any prefix of the sequence is non-negative, and
- 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 the 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 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 4.5. *The labeled* 1*-Dyck path on* [7] *given in Figure 4 has* 3 *primitive parts and* 2 *compartments.*

Using the same methods as for trees, we have the following.

Theorem 4.6. *The number of labeled m-Dyck paths on $[n]$ with j compartments is $C(m,n,j)$.*

5. Concluding remarks

We note that a combinatorial interpretation for the coefficients of the characteristic polynomial of the Linial arrangement is already given in [5, 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 $\mathcal{T}_S(n)$ according to number of branches.

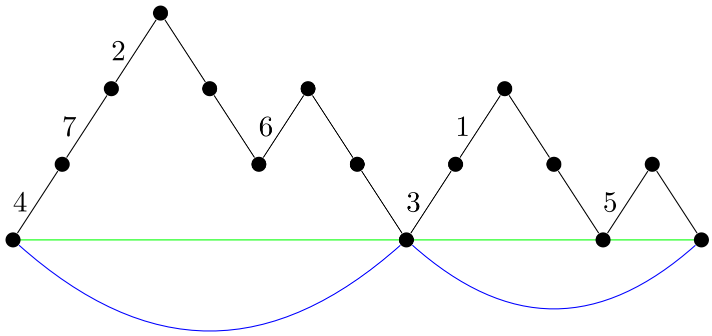

FIGURE 4. A labeled 1-Dyck path with compartments specified.

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] P. Deshpande and K. Menon, *A branch statistic for trees: interpreting coefficients of the characteristic polynomial of braid deformations*, Enumer. Combin. Appl. 3:1 (2023), Article S2R5.

[5] R. P. Stanely, *Hyperplane arrangements, interval orders, and trees*, Proc. Natl. Acad. Sci. USA 93 (1996), 2620–2625.

[6] R. P. Stanley, *Enumerative Combinatorics, Volume I*, Cambridge University Press, Cambridge, UK, 1997.

[7] R. P. Stanley, *Enumerative Combinatorics, Volume II*, Cambridge University Press, Cambridge, UK, 2001.

[8] 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.

[9] R. P. Stanley, *Catalan Numbers*, Cambridge University Press, Cambridge, UK, 2015.

[10] 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).