A refined enumeration of hex trees and related polynomials

Hana Kim[a]*and Richard P. Stanley[b] †

[a] *National Institute for Mathematical Sciences, 70 Yuseong-daero, 1689 beon-gil, Yuseong-gu, Daejeon 305-811, Rep. of Korea*

hkim@nims.re.kr

[b] *Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

rstan@math.mit.edu

Abstract

A hex tree is an ordered tree of which each vertex has updegree 0, 1, or 2, and an edge from a vertex of updegree 1 is either left, median, or right. We present a refined enumeration of symmetric hex trees via a generalized binomial transform. It turns out that the refinement has a natural combinatorial interpretation by means of supertrees. We describe a bijection between symmetric hex trees and a certain class of supertrees. Some algebraic properties of the polynomials obtained in this procedure are also studied.

AMS classifications: Primary:05A15; secondary:12D10

Key words: Hex tree, Supertree, Riordan matrix, zeros of a polynomial, limiting curve.

1 Introduction

An *ordered tree* is a rooted tree where the order of the subtrees at each vertex is significant. The *updegree* of a vertex in an ordered tree is the number of edges incident with the vertex that lead away from the root. A sequence $A = (a_n)_{n\geq 0}$ of numbers defines a class $\mathcal{T}_A$ of ordered trees by assigning the weight a_n to each vertex of updegree n. A weight $a_n = 0$ means that there are no vertices of updegree n. We call the generating function $A(z) = \sum_{n\geq 0} a_n z^n$ and the ordered trees in $\mathcal{T}_A$ a *degree function* and *A-trees*, respectively. As usual, the weight of an A tree is the product of the weights of all edges. If $T(z) = \sum_{n\geq 0} t_n z^n$ where t_n is the sum of the weights of A-trees with n edges, then T satisfies the functional equation

$$T = A \circ (zT). \tag{1}$$

Many interesting classes of ordered trees are defined by degree functions with nonnegative integer coefficients. For instance, the degree functions corresponding to the usual ordered trees, k-ary trees, Motzkin trees, hex trees, Riordan trees, and Schröder trees are $\frac{1}{1-z}$, $1+z^k$,

*The research of the first author was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (2013R1A6A3A03024342).

†The second author was partially supported by NSF grant DMS-1068625.

$1+z+z^2$, $1+3z+z^2$, $\frac{1}{1-z}-z$, and $\frac{1+z}{1-z}$ respectively. The weight a_n is sometimes regarded as the possible number of colors or types for the edges from a vertex of updegree n when $a_n \geq 0$.

A *hex tree* has been introduced by Balaban and Harary [1]. It depicts an edge-rooted polyhex called a *catafusene* [5] in organic chemistry, where a hexagon and an edge shared by two connected hexagons are represented by a vertex and an edge in the tree respectively. Note that hexagons cannot share more than one edge. To be precise, a hex tree is an ordered tree of which each vertex has the updegree 0, 1, or 2, and an edge from a vertex of updegree 1 is either left, median, or right; see Fig. 1.

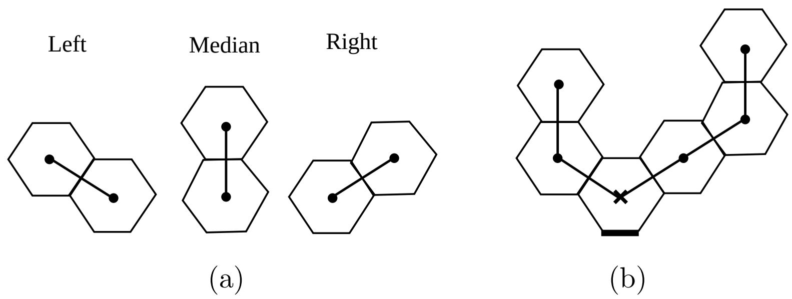


Figure 1. (a) All possible chemical bonds of two benzenoids; and (b) an edge-rooted catafusene with the corresponding hex tree.

Harary and Read [9] derived the generating functions for the numbers of both rooted and unrooted catafusenes. In particular, the generating function $H = \sum_{n\geq 0} h_n z^n$ for the number of hex trees with n edges is given by

$$H = \frac{1-3z-\sqrt{1-6z+5z^2}}{2z^2} = 1+3z+10z^2+36z^3+137z^4+\cdots \quad \text{(A002212)}.$$

The label A002212 refers to that item in [13]. The generating function H can be expressed as $H = \frac{1}{1-z}\cdot\left(C^2 \circ \frac{z}{1-z}\right)$ where $C = \sum_{n\geq 0} C_n z^n = \frac{1-\sqrt{1-4z}}{2z}$ is the generating function for the Catalan numbers. This implies that $(h_n)_{n\geq 0}$ is the binomial transform of the shifted Catalan numbers i.e., $h_n = \sum_{j=0}^{n} C_{j+1}\binom{n}{j}$.

Cyvin et. al. [5, 6] presented explicit formulas for the number of certain unrooted polyhexes belonging to several different symmetries. We say that a hex tree is *symmetric* if it is symmetrical with respect to the vertical line through the root. It is known that the number s_n of symmetric hex trees with even (or odd) number of edges is equal to the binomial transform of the Catalan numbers i.e., $s_n = \sum_{j=0}^{n} C_j\binom{n}{j}$. Equivalently, if $S = \sum_{n\geq 0} s_n z^n$,

$$S = \frac{1}{1-z}\cdot\left(C \circ \frac{z}{1-z}\right) = 1+2z+5z^2+15z^3+51z^4+188z^5+\cdots \quad \text{(A007317)}. \tag{2}$$

In the present paper, we enumerate the symmetric hex trees according to the number of median edges by representing (2) as a matrix-vector equation and taking a power of the Pascal matrix. The theory of the Riordan group simplifies computation in this procedure. It turns out that the numbers obtained from the generalization can be interpreted in terms of supertrees. A bijection between the symmetric hex trees and a certain class of supertrees will be given. Further, we study some algebraic properties of polynomials that arise in the refined enumeration of symmetric hex trees. Specifically, we observe the generating functions, explicit formulas, recurrence relations, and behavior of the zeros of the polynomials. In the last section, we pose some open questions and give several conjectures.

2 The symmetric hex trees and the stem supertrees

The identity (2) can be expressed as the following matrix-vector equation involving the Pascal matrix P:

$$\begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 2 & 1 & & \\ 1 & 3 & 3 & 1 & \\ 1 & 4 & 6 & 4 & 1 \\ & & \cdots & & \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \\ 5 \\ 14 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \\ 15 \\ 51 \\ \vdots \end{bmatrix}.$$

We consider a generalized binomial transform of the Catalan numbers by replacing P with P^x for an indeterminate x:

$$\begin{bmatrix} p_0(x) \\ p_1(x) \\ p_2(x) \\ p_3(x) \\ p_4(x) \\ \vdots \end{bmatrix} := \begin{bmatrix} 1 & & & & \\ x & 1 & & & \\ x^2 & 2x & 1 & & \\ x^3 & 3x^2 & 3x & 1 & \\ x^4 & 4x^3 & 6x^2 & 4x & 1 \\ & & \cdots & & \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \\ 5 \\ 14 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ 1+x \\ 2+2x+x^2 \\ 5+6x+3x^2+x^3 \\ 14+20x+12x^2+4x^3+x^4 \\ \vdots \end{bmatrix}. \quad (3)$$

This gives the polynomials

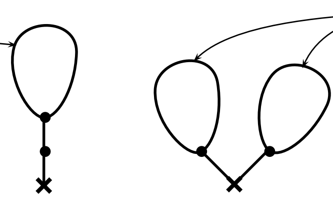
$$p_n(x) = \sum_{j=0}^{n} \frac{1}{j+1}\binom{2j}{j}\binom{n}{j} x^{n-j} \quad (4)$$

of which the specialization when $x = 1$ counts the symmetric hex trees. As we shall see in the next theorem, the coefficient matrix of $(p_n(x))_{n\geq 0}$ gives a refinement of symmetric hex trees. $[x^j]$ denotes an operator that extracts the coefficient of x^j.

Theorem 2.1 *Let $p_n(x)$ be the polynomial defined by (4). Then $[x^j]p_n(x)$ counts the symmetric hex trees with $2n$ edges each of which contains exactly $2j$ median edges.*

Proof. Let $S = S(x, z)$ be the generating function for the symmetric hex trees with $2n$ edges each of which contains $2j$ median edges. Any nonempty symmetric hex tree with $2j$ median edges has one of the following forms:

The first form has the generating function xzS. For the second form, we show that $[z^n x^j](C \circ \frac{z}{1-xz})$ counts the hex trees with $n-1$ edges having j median edges for $n \geq 1$ and $0 \leq j \leq n-1$. A simple computation shows $[z^n x^j](C \circ \frac{z}{1-xz}) = \binom{n-1}{j} C_{n-j}$. Incomplete binary trees are A-trees with $A = 1 + 2z + z^2$, i.e., a vertex of degree 1 has a left or right edge. Solving (1) we get the generating function $C^2 = \frac{1}{z}(C-1)$ for incomplete binary trees.



Hence there are C_{n-j} incomplete binary trees with $n-j-1$ edges. We can obtain a desired hex tree by inserting j median edges at $n-j$ vertices of the incomplete binary tree allowing repetitions in $\binom{n-1}{j}$ ways. Thus there are $\binom{n-1}{j}C_{n-j}$ such hex trees.

Consequently, we obtain

$$S = 1 + xzS + \sum_{n\geq 1}\left(\sum_{j=0}^{n-1}\binom{n-1}{j}C_{n-j}x^j\right)z^n = xzS + \left(C \circ \frac{z}{1-xz}\right),$$

which implies

$$S = \frac{1}{1-xz}\cdot\left(C \circ \frac{z}{1-xz}\right) = \sum_{n\geq 0} p_n(x)z^n. \tag{5}$$

■

We note that the refinement of hex trees (not necessarily symmetric) according to the number of median edges can be found in [9] while that of symmetric hex trees is unknown in authors' knowledge.

On the other hand, the generating function for the polynomials $p_n(x)$ involves a composition of two generating functions as shown in (5). Several combinatorial interpretations for a composition of two functions have been described in [8]. The recursive structure of ordered trees yields a supertree interpretation. A *supertree*[1] is a tree such that an ordered tree grows up from each vertex of a tree τ. We call the tree τ a *base tree*. Thus any class of supertrees is formed by a pair of classes of base trees and growing up trees. A *stem* is a rooted tree each of whose vertices has updegree 1 except the two terminal vertices, i.e., the root and the leaf. The height of a stem is the number of edges in the stem. We define a *stem supertree* to be a supertree with only stems growing up from each vertex of a usual ordered tree. It immediately follows from (5) that $[x^j]p_n(x)$ counts the stem supertrees with n edges in which the sum of the heights of all stems is j; for example, see Fig. 2.

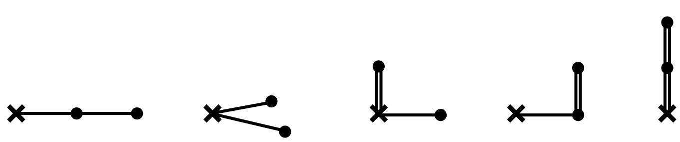


Figure 2. The stem supertrees with 2 edges counted by $p_2(x) = 2 + 2x + x^2$. A base tree consists of the solid edges and the stems consist of the double edges.

[1] We note that this supertree is different from the one [4] arising in Phylogenetics. Supertrees are sometimes called multidimensional trees (for example, see [2]).

Theorem 2.2 *There is a bijection between symmetric hex trees with* $2n$ *edges each of which contains* 2ℓ *median edges and stem supertrees with* n *edges in which the sum of the heights of stems is* ℓ*, where* $n, \ell \geq 0$*.*

Proof. Let $\mathcal{S}(n,\ell)$ and $\mathcal{C}(n,\ell)$ be the sets of symmetric hex trees with $2n$ edges each of which contains 2ℓ median edges, and stem supertrees with n edges in which the sum of the heights of stems is ℓ, respectively. The bijection is a map from $\mathcal{C}(n,\ell)$ to $\mathcal{S}(n,\ell)$ consisting of three steps. Let T be an element of $\mathcal{C}(n,\ell)$ with the base tree τ.

Step 1. Convert to a partially labeled ordered tree: first, label each vertex of the base tree τ by height of the stem growing up from the vertex. Then delete the stem. We then obtain a usual ordered tree T_1 some of whose vertices are labeled, where the total sum of labels is ℓ. See T_1 in Fig. 3.

Step 2. Symmetrization: In order to make the tree symmetric, we make use of a bijection between (complete) binary trees and usual ordered trees. There are a variety of bijections between binary trees and usual ordered trees. Here we define a map ψ to be the inverse of the de Bruijin-Morselt bijection [7] that maps a planted binary tree to a planted ordered tree after some modification. A tree is said to be *planted* if its root has updegree 1. The map ψ is as follows: given an ordered tree T_1, every nonroot vertex v of T_1 becomes a point on the left edge of a pair of edges with common parent p_v. The parent p_v is the left child of the parent of the right sibling (a vertex that has the same parent) of v in T_1. If v is a child of v' then the p_v is the right child of the parent of $p_{v'}$. Let $\widehat{\tau}$ be the resulting binary tree. Then ψ maps the label of a nonroot vertex v in T_1 to the left child of p_v in $\widehat{\tau}$, and the label of the root of T_1 to the root of $\widehat{\tau}$.

In general, $\widehat{\tau}$ is not necessarily symmetric with respect to the vertical line through the root. From bottom to top, compare the children of every vertex at the same level. If the distribution of children is symmetric then do nothing. If it is not symmetric then at each vertex, leave the left edge and move the right edge with all subsequent edges to the symmetric position of the parent of the left edge. All subsequent edges of the left edge remain, and the label of a vertex, if exists, stay with the left edge. Continuing the procedure at each level, from left to right, we end up with a symmetric hex tree s_0 with no median edges; see s_0 in Fig. 3.

Step 3. From a binary tree to a hex tree: the final step is inserting median edges symmetrically according to the labels of vertices of s_0. If the root has a label j then attach the stem of length $2j$ at the root toward the bottom. The end point of the stem becomes the root of a new symmetric hex tree s. If a nonroot vertex v has a label j then attach the stem of length j at v toward the top, and do the same procedure at the vertex in the symmetric position to v.

Clearly, the number of median edges in s is double the number of edges in all stems, which is equal to the sum of the heights of stems in T. Thus s belongs to $\mathcal{S}(n,\ell)$ as shown in Fig. 3.

The inverse map can be similarly obtained by reversing each step. ■

Remark. In the first step of the inverse map, we assign the label ℓ to a vertex v whenever we delete 2ℓ median edges in symmetric position. The vertex v is the lowest vertex in the deleted median edges, and is adjacent to a left edge.

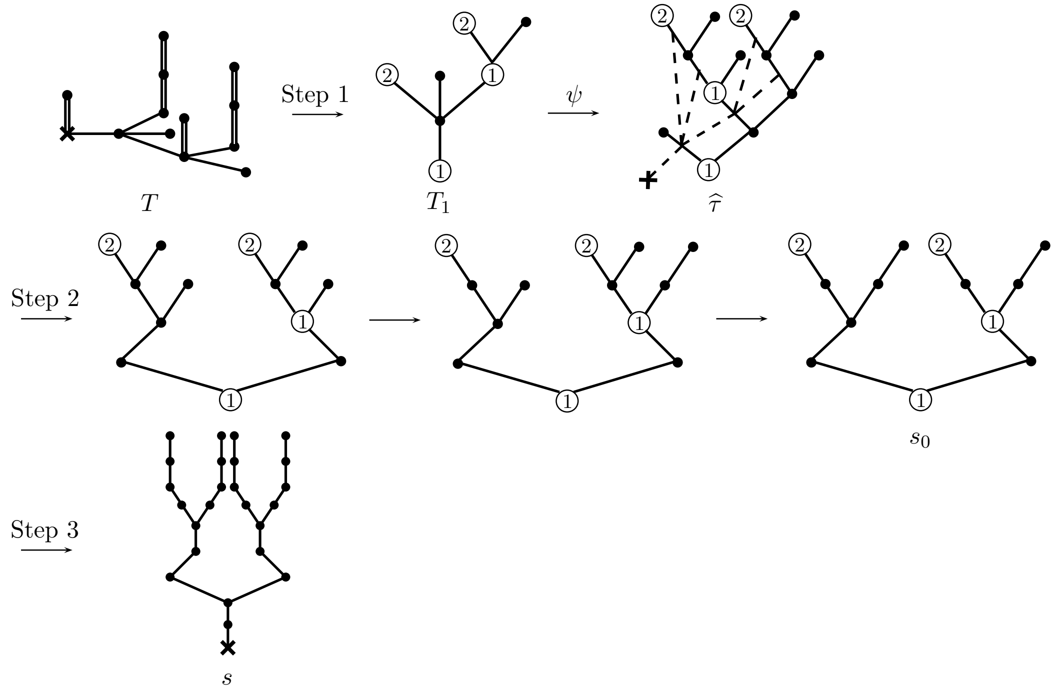


Figure 3. From $T \in \mathcal{C}(12,6)$ to $s \in \mathcal{S}(12,6)$.

3 Properties of the polynomials $p_n(x)$

In this section, we explore some algebraic properties of the polynomials $p_n(x)$ defined by (2). We begin with an extension of the vector $[p_0(x), p_1(x), \ldots]^T$ to a matrix with polynomial entries.

The functional equation (1) connects the class $\mathcal{T}_A$ of trees and a structured matrix defined by two formal power series. Let g and f be elements of the ring of formal power series $\mathbb{C}[[z]]$ over the complex field satisfying $g(0) \neq 0$, $f(0) = 0$, and $f'(0) \neq 0$. A *Riordan matrix* [12] (g, f) is an infinite lower triangular matrix in which the generating function for the kth column is $g \cdot f^k$, $k \geq 0$. The Pascal matrix $P = \left(\frac{1}{1-z}, \frac{z}{1-z}\right)$ is the typical example of a Riordan matrix. If we multiply a column vector $[h_0, h_1, h_2, \ldots]^T$ with the generating function $h = \sum_{n \geq 0} h_n z^n$ by a Riordan matrix (g, f), the generating function for the resulting column vector is $g \cdot (h \circ f)$. This is called the fundamental theorem for Riordan matrices (FTRM, briefly), and enables us to compute the usual matrix multiplication of two Riordan matrices in terms of generating functions as follows:

$$(g, f) * (h, \ell) = (g \cdot (h \circ f), \ell \circ f). \tag{6}$$

The set of all Riordan matrices forms a group under this operation, called the *Riordan group*. The identity is $(1, z)$, and the inverse of (g, f) is $(1/(g \circ \bar{f}), \bar{f})$ where $\bar{f}$ is the compositional inverse of f. It is known [10] that a Riordan matrix $(g, f) = [r_{n,k}]$ is associated to two sequences $A = (a_n)_{n \geq 0}$ and $Z = (z_n)_{n \geq 0}$ with $a_0 \neq 0$ such that $r_{n+1,k+1} = \sum_{j \geq 0} a_j r_{n,k+j}$

and $r_{n+1,0} = \sum_{j\geq 0} z_j r_{n,j}$, or equivalently

$$g = \frac{g(0)}{1 - z(Z \circ f)} \quad \text{and} \quad f = z(A \circ f).$$

Letting $T = f/z$ in the second equation, we have (1). This suggests a combinatorial interpretation for an arbitrary Riordan matrix by means of ordered forests of ordered trees of two kinds. A tree is *planted* if its root has degree 1. The weight of an ordered forest is defined as the product of the weights of all components.

Theorem 3.1 *Let $R = (g, f)$ be a Riordan matrix, and A and B be the analytic solutions of the functional equations $f = z(A \circ f)$ and $g = B \circ (zg)$ respectively. Then the (n, k)-entry of R is the sum of the weights of ordered forests of 1 B-tree and k planted A-trees with n edges, where B-tree precedes all planted A-trees.*

Proof. The series A is the generating function for the A-sequence. Since $g(0) \neq 0$, the functional equation $g = B \circ (zg)$ has the analytic solution B. The series g and f/z are the generating functions for the sum of the weights of B-trees and planted A-trees, respectively. Since the generating function for the kth column of R is $g \cdot f^k$, the result immediately follows from the combinatorial interpretation for a product of two generating functions (see [8, 14]). ■

The interpretation given by Theorem 3.1 looks somewhat complicated, but if $g = 1$ then every entry of $(1, zT)$ is interpreted by just ordered forests of A-trees. In particular, $A = \frac{1}{1-z}$ gives the Riordan matrix $(1, zC)$ where C is the generating function for the Catalan numbers.

We notice that the polynomials $p_n(x)$ form the leftmost column of the Riordan matrix

$$P^x(1, zC)P = \begin{bmatrix} 1 & & & & \\ 1+x & 1 & & & \\ 2+2x+x^2 & 3+2x & 1 & & \\ 5+6x+3x^2+x^3 & 9+9x+3x^2 & 5+3x & 1 & \\ 14+20x+12x^2+4x^3+x^4 & 28+36x+18x^2+4x^3 & 20+20x+6x^2 & 7+4x & 1 \\ & & \cdots & & \end{bmatrix}. \quad (7)$$

What are the other polynomials? Let $p_{n,k}(x)$ be the (n, k)-entry of $P^x(1, zC)P$. For each $k \geq 0$, we obtain a sequence of polynomials defined by the generating function

$$F_k := F_k(x, z) = \sum_{n\geq k} p_{n,k}(x) z^n = \frac{1}{1 - xz} \cdot \left(z^k C^{2k+1} \circ \frac{z}{1 - xz} \right), \quad (8)$$

which is the generating function for the kth column of $P^x(1, zC)P$. It can be easily shown from $C = 1 + zC^2$ that the generating function F_0 for $p_{n,0}(x) = p_n(x)$ satisfies the functional equation

$$(1 - xz)F_0 = 1 + z(1 - xz)F_0^2.$$

For $k \geq 1$, the generating function F_k for $p_{n,k}(x)$ is given by $F_k = z^k(1 - xz)^k F_0^{2k+1}$.

The coefficient matrix of $(p_{n,k}(x))_{n\geq k}$ for each k can be expressed in terms of the exponential Riordan matrix [15]. An *exponential Riordan matrix* $\langle g, f \rangle$ is defined as an infinite lower triangular matrix in which the exponential generating function for the kth column is $\frac{1}{k!}g \cdot f^k$. It can be readily shown that $\langle g, f\rangle = E^{-1}(g, f)E$ where $E = \mathrm{diag}(\frac{1}{0!}, \frac{1}{1!}, \frac{1}{2!}, \ldots)$.

The Pascal matrix P is the only matrix which can be expressed as both Riordan and exponential Riordan matrices:

$$P = \left(\frac{1}{1-z}, \frac{z}{1-z}\right) = \langle e^z, z\rangle .$$

Lemma 3.2 *For $k \geq 0$, $p_{n,k}(x)$ is a polynomial of degree $n-k$ given by*

$$\sum_{n\geq k} p_{n,k}(x)\frac{z^n}{n!} = e^{(2+x)z}\left(I_k(2z) - I_{k+1}(2z)\right)$$

where $I_k(z) = \sum_{j\geq 0} \frac{(z/2)^{2j+k}}{j!(j+k)!}$ is the modified Bessel function of the first kind.

Proof. It suffices to express the ordinary generating function (8) for the kth column of $P^x(1, zC)P$ as an exponential generating function. It follows from $C^{2k+1} = \sum_{n\geq 0} \frac{2k+1}{n+2k+1}\binom{2n+2k}{n} z^n$ that the exponential generating function for $(\frac{2k+1}{n+2k+1}\binom{2n+2k}{n})_{n\geq 0}$ is $e^{2z}\left(I_k(2z) - I_{k+1}(2z)\right)$. By the FTRM, we have

$$\frac{1}{1-xz} \cdot (z^k C^{2k+1} \circ \frac{z}{1-xz}) = \left(\frac{1}{1-xz}, \frac{z}{1-xz}\right) * z^k C^{2k+1}.$$

The right-hand side can be computed in the exponential Riordan group as follows:

$$\langle e^{xz}, z\rangle * e^{2z}\left(I_k(2z) - I_{k+1}(2z)\right) = e^{(2+x)z}\left(I_k(2z) - I_{k+1}(2z)\right).$$

The resulting exponential generating function is the one for the kth column of $P^x(1, zC)P$.■

By Lemma 3.2, the coefficient matrix of the polynomial sequence $(p_{n,k}(x))_{n\geq k}$ is

$$D_k \left\langle z^{-k} e^{2z}\left(I_k(2z) - I_{k+1}(2z)\right), z\right\rangle \tag{9}$$

where D_k is the diagonal matrix with the main diagonal entries $\left(\binom{k}{k}, \binom{k+1}{k}, \binom{k+2}{k}, \ldots\right)$.

It follows from (9) and $I_k(2z) - I_{k+1}(2z) = \sum_{n\geq 0}(-1)^n \binom{n+k}{\lfloor k/2\rfloor}\frac{z^{n+k}}{(n+k)!}$ that

$$p_{n,k}(x) = \sum_{j=0}^{n-k}\left(\sum_{\ell=0}^{n-k-j}(-1)^\ell 2^{n-k-j-\ell}\binom{n-j}{k+\ell}\binom{k+\ell}{\lfloor \ell/2\rfloor}\right)\binom{n}{j}x^j.$$

On the other hand, it follows from (8) that

$$p_{n,k}(x) = \sum_{j=k}^{n}\frac{2k+1}{j+k+1}\binom{2j}{j-k}\binom{n}{j}x^{n-j} = \sum_{j=0}^{n-k}\frac{2k+1}{n+k-j+1}\binom{2n-2j}{n-k-j}\binom{n}{j}x^j.$$

Comparing the coefficient of x^j in $p_{n,k}(x)$, we obtain the following identity after a change of variables:

$$\sum_{\ell=0}^{n-k}(-1)^\ell 2^{n-k-\ell}\binom{n}{k+\ell}\binom{k+\ell}{\lfloor \ell/2\rfloor} = \frac{2k+1}{n+k+1}\binom{2n}{n+k}.$$

Proposition 3.3 *Let $(a_n)_{n\geq 0}$ be the A-sequence of a Riordan matrix $(1, zT)$. Then the A-sequence $(\tilde{a}_n)_{n\geq 0}$ and the Z-sequence $(\tilde{z}_n)_{n\geq 0}$ of $P^x(1, zT)P^y$ are given by*

$$\tilde{a}_n = \sum_{k=1}^{n-1} a_{k+1}\binom{n-2}{k-1}(-y)^{n-k-1} \quad \text{and} \quad \tilde{z}_n = y\sum_{k=1}^{n} a_k\binom{n-1}{k-1}(-y)^{n-k},$$

where $\tilde{a}_0 = a_0$, $\tilde{a}_1 = x + a_0 y + a_1$ and $\tilde{z}_0 = x + a_0 y$.

If $T = C$ i.e., $a_n = 1$ for all $n \geq 0$ then it follows that $\tilde{a}_n = (1-y)^{n-2}$ and $\tilde{z}_n = y(1-y)^{n-1}$. In particular, $y = 1$ gives the A-sequence $(1, x+2, 1, 0, 0, \ldots)$ and Z-sequence $(x+1, 1, 0, 0, \ldots)$. This implies that $p_{n,k}(x)$ satisfy the recurrence relations

$$\begin{aligned} p_{n+1,0}(x) &= (x+1)p_{n,0}(x) + p_{n,1}(x), \\ p_{n+1,k+1}(x) &= p_{n,k}(x) + (x+2)p_{n,k+1}(x) + p_{n,k+2}(x). \end{aligned}$$

We note that the polynomials $p_{n,k}(x)$ in a fixed column do not satisfy a three term recurrence relation. However, using the explicit formula (4) for $p_{n,k}(x)$, we obtain the three term recurrence relation for the polynomials in a fixed row.

Lemma 3.4 *For $n \geq 2$ and $k \geq 0$,*

$$p_{n,k}(x) = \frac{k+1}{n-k}\left(\left(x + 2 - \frac{n+1}{(k+1)(k+2)}\right)p_{n,k+1}(x) + \frac{n+k+3}{k+2}p_{n,k+2}(x)\right).$$

Proof. Making use of (4) with a little computation gives

$$\begin{aligned} &\frac{n-k}{k+1}p_{n,k}(x) - \left(x + 2 - \frac{n+1}{(k+1)(k+2)}\right)p_{n,k+1}(x) \\ = &\sum_{j=0}^{n}\left[\frac{(n+k-j+2)(n-k)(2k+1)}{(n-k-j)(n-k-j-1)(k+1)} - \frac{2j(2k+3)(2n-2j+1)}{(n-k-j)(n-k-j-1)(n+k-j+3)}\right. \\ &\left. - \left(2 - \frac{n+1}{(k+1)(k+2)}\right)\right]\binom{2n-2j}{n-k-j}\binom{n}{j}x^j \\ = &\sum_{j=0}^{n}\frac{n+k+3}{k+2}\cdot\frac{2k+5}{n+k-j+3}\binom{2n-2j}{n-k-j}\binom{n}{j}x^j = \frac{n+k+3}{k+2}p_{n,k+2}(x). \end{aligned}$$

∎

Lemma 3.4 also gives some information on the zeros of polynomials $p_{n,k}(x)$.

Given a polynomial $f(x)$ and $f_1(x) := f'(x)$, the Euclidean algorithm for seeking the greatest common divisor of f and f_1 yields:

$$\begin{aligned} f &= q_1 f_1 - f_2, \\ f_1 &= q_2 f_2 - f_3, \\ &\cdots \\ f_{m-2} &= q_{m-1} f_{m-1} - f_m, \\ f_{m-1} &= q_m f_m. \end{aligned}$$

The sequence $f, f_1, \ldots, f_m$ is called the *Sturm sequence* of f. Sturm's theorem (see, for example, [11]) asserts that if $\omega(x)$ is the number of sign changes in the sequence $f(x), f_1(x), \ldots, f_m(x)$ then the number of distinct roots of f in an interval $(a, b]$ is equal to $\omega(a) - \omega(b)$, where $f(a) \neq 0$, $f(b) \neq 0$ and $a < b$.

By making use of Lemma 3.4 we obtain

Lemma 3.5 *Let $f_{n,k}^{(j)}(x)$ be the Sturm sequence of $p_{n,k}(x)$ i.e., for $j \geq 1$*

$$f_{n,k}^{(j)}(x) = q_{n,k}^{(j+1)}(x) f_{n,k}^{(j+1)}(x) - f_{n,k}^{(j+2)}(x),$$

where $f_{n,k}^{(1)}(x) := p'_{n,k}(x) = (n-k)p_{n-1,k}(x)$ and $q_{n,k}^{(1)} = \frac{1}{n}(x+1)$. Then for $j \geq 2$,

$$f_{n,k}^{(j)}(x) = \begin{cases} (-1)^{\lfloor \frac{j}{2} \rfloor} \frac{1}{n+1} \left(\frac{n+1}{1} \cdot \frac{2}{n-2} \cdot \frac{n+3}{3} \cdot \frac{4}{n-4} \cdot \cdots \cdot \frac{n+j-1}{j-1} \right) p_{n-1,k+j-1}(x) & \text{if } j \text{ is even;} \\ (-1)^{\lfloor \frac{j}{2} \rfloor} n \left(\frac{1}{n-1} \cdot \frac{n+2}{2} \cdot \frac{3}{n-3} \cdot \frac{n+4}{4} \cdot \cdots \cdot \frac{n+j-1}{j-1} \right) p_{n-1,k+j-1}(x) & \text{if } j \text{ is odd,} \end{cases}$$

$$q_{n,k}^{(j)}(x) = \begin{cases} -(n+1)n \left(\frac{1}{n-1} \cdot \frac{n-2}{2} \cdot \cdots \cdot \frac{j-1}{n-j+1} \right) \left(\frac{1}{n+1} \cdot \frac{n+2}{2} \cdot \cdots \cdot \frac{j-1}{n+j-1} \right) \left(x + 2 - \frac{n}{(j-1)j} \right) & \text{if } j \text{ is even;} \\ \frac{1}{(n+1)n} \left(\frac{n-1}{1} \cdot \frac{2}{n-2} \cdot \cdots \cdot \frac{j-1}{n-j+1} \right) \left(\frac{n+1}{1} \cdot \frac{2}{n+2} \cdot \cdots \cdot \frac{j-1}{n+j-1} \right) \left(x + 2 - \frac{n}{(j-1)j} \right) & \text{if } j \text{ is odd.} \end{cases}$$

We note that since the degree of $f_{n,k}^{(j)}(x)$ decreases by 1 at each step, the Sturm sequence of $p_{n,k}(x)$ consists of $n-k+1$ nonzero polynomials.

Theorem 3.6 *The polynomial $p_{n,k}(x)$ has no real zero if $n-k$ is even, and has exactly one real zero x_0 if $n-k$ is odd. Moreover, $-2 < x_0 \leq -1$.*

Proof. Let a be an integer. Since $p_{n,k}(x)$ has positive coefficients, $p_{n,k}(a) > 0$ for $a \geq 0$. If $a = -1$, $[p_{n,k}(-1)] = (R, zM)$ where R and M are the generating functions for the Riordan numbers (A005043) and the Motzkin numbers (A001006) respectively. So $p_{n,k}(-1) > 0$ except $p_{1,0}(-1) = 0$. We now show that for $a \leq -2$,

$$\begin{cases} p_{n,k}(a) > 0 & \text{if } n-k \text{ is even;} \\ p_{n,k}(a) < 0 & \text{if } n-k \text{ is odd.} \end{cases} \tag{10}$$

We proceed by induction on $n-k$. It follows from Lemma 3.4 that

$$p_{n,k}(a) = \frac{k+1}{n-k} \left(\left(a + 2 - \frac{n+1}{(k+1)(k+2)} \right) p_{n,k+1}(a) + \frac{n+k+3}{k+2} p_{n,k+2}(a) \right).$$

If $n-k$ is even (odd, resp.) then by the induction assumption, $p_{n,k+1}(a) < 0$ and $p_{n,k+2}(a) > 0$ ($p_{n,k+1}(a) > 0$ and $p_{n,k+2}(a) < 0$, resp.). Since $a+2 \leq 0$, we obtain $p_{n,k}(a) > 0$ ($p_{n,k}(a) < 0$, resp.).

For a polynomial $f(x)$, define

$$\operatorname{sgn}(f(a)) = \begin{cases} 1 & \text{if } f(a) > 0, \\ -1 & \text{if } f(a) < 0. \end{cases}$$

Let $f_{n,k}^{(j)}(x)$ be the Sturm sequence of $p_{n,k}(x)$ determined by Lemma 3.5, $j = 0, 1, \ldots, n-k$, and let $\omega_{n,k}(a)$ be the number of sign changes in the Sturm sequence when $x = a$. Then

$\mathrm{sgn}(f_{n,k}^{(j)}(a)) = (-1)^{\lfloor j/2 \rfloor}\mathrm{sgn}(p_{n-1,k+j-1}(a))$. Since $p_{n,k}(a) > 0$ for $a \geq -1$, $(\mathrm{sgn}(f_{n,k}^{(j)}(a))) = ((-1)^{\lfloor j/2 \rfloor}) = (1, 1, -1, -1, 1, 1, -1, -1, \ldots)$. Hence $\omega_{n,k}(a) = \lfloor \frac{n-k}{2} \rfloor$. If $a \leq -2$, it follows from (10) that

$$\mathrm{sgn}(f_{n,k}^{(j)}(a)) = \begin{cases} (1, -1, -1, 1, 1, -1, -1, \ldots) & \text{if } \ n-k \ \text{ is even;} \\ (-1, 1, 1, -1, -1, 1, 1, \ldots) & \text{if } \ n-k \ \text{ is odd.} \end{cases}$$

and thus $\omega_{n,k}(a) = \lceil \frac{n-k}{2} \rceil$. Consequently we obtain $\omega_{n,k}(a) - \omega_{n,k}(a+1) \equiv 0$ for $a \neq -2$ and

$$\omega_{n,k}(-2) - \omega_{n,k}(-1) = \begin{cases} 0 & \text{if } \ n-k \ \text{ is even;} \\ 1 & \text{if } \ n-k \ \text{ is odd.} \end{cases}$$

The Sturm's theorem asserts that $p_{n,k}(x)$ has no real zero when $n-k$ is even, and has exactly one real zero in the interval $(-2, -1]$ when $n-k$ is odd. ■

For example, let us consider $p_{5,0}(x) = x^5 + 5x^4 + 20x^3 + 50x^2 + 70x + 42$. It follows from Lemma 3.5 that

$$\begin{aligned}
p_{5,0}(x) &= \frac{1}{5}(x+1)p_{5,0}'(x) - (-p_{4,1}(x)) \\
p_{5,0}'(x) &= -\frac{5}{4}\left(x - \frac{1}{2}\right) \cdot (-p_{4,1}(x)) - \left(-\frac{35}{4}p_{4,2}(x)\right) \\
-p_{4,1}(x) &= \frac{16}{105}\left(x + \frac{7}{6}\right) \cdot \left(-\frac{35}{4}p_{4,2}(x)\right) - \frac{16}{9}p_{4,3}(x) \\
-\frac{35}{4}p_{4,2}(x) &= -\frac{945}{256}\left(x + \frac{19}{12}\right) \cdot \frac{16}{9}p_{4,3}(x) - \frac{945}{64} \\
\frac{16}{9}p_{4,3}(x) &= \frac{4096}{8505}\left(x + \frac{7}{4}\right) \cdot \frac{945}{64}
\end{aligned}$$

Hence the Sturm sequence of $p_{5,0}(x)$ is

$$\left[p_{5,0}(x), 5p_{4,0}(x), -p_{4,1}(x), -\frac{35}{4}p_{4,2}(x), \frac{16}{9}p_{4,3}(x), \frac{945}{64}\right]$$

and the sign changes in the Sturm sequence are

a	$p_{5,0}(a)$	$5p_{4,0}(a)$	$-p_{4,1}(a)$	$-\frac{35}{4}p_{4,2}(a)$	$\frac{16}{9}p_{4,3}(a)$	$\frac{945}{64}$	#(sign changes)
-3	$-$	$+$	$+$	$-$	$-$	$+$	3
-2	$-$	$+$	$+$	$-$	$-$	$+$	3
-1	$+$	$+$	$-$	$-$	$+$	$+$	2
0	$+$	$+$	$-$	$-$	$+$	$+$	2

The sign changes in the Sturm sequence for $a \leq -3$ and for $a \geq 0$ are the same as those for $a = -2$ and for $a = -1$, respectively. Thus $p_{5,0}(x)$ has exactly $3 - 2 = 1$ real zero in the interval $(-2, -1)$. In fact, $p_{5,0}(x)$ has the real zero $x \fallingdotseq -1.47781939$.

Theorem 3.7 *All the zeros of the polynomial $p_{n,k}(x)$ are distinct.*

Proof. Let $R(f, g)$ be the resultant [11] of two polynomials f and g. We make use of the following known facts :

(a) If $f = gq + r$ then $R(f, g) = b^{\deg f - \deg r} R(r, g)$ where b is the leading coefficient of g.

(b) f and g have a common divisor if and only if $R(f, g) = 0$.

Using (a) and the Sturm sequence of $p_{n,k}$, an induction on $n - k$ shows $R(p_{n,k}, p'_{n,k}) \neq 0$. By (b), $p_{n,k}$ and $p'_{n,k}$ do not have a common divisor. Fix k. We now proceed an induction on n to show that all the zeros of $p_{n,k}$ are distinct. Since $(p_{n,k})_{n \geq k}$ is an Appell sequence, $p'_{n,k} = (n - k)p_{n-1,k}$. By the induction assumption, all the zeros of $p_{n-1,k}$ are distinct. Since $p_{n,k}$ and $p_{n-1,k}$ do not have a common zero, all the zeros of $p_{n,k}$ are distinct. ■

More generally, we conjecture that the polynomials $p_{n,k}(x)$ are irreducible over the field of rational numbers.

4 Discussion

We have explored a refinement of symmetric hex trees and a bijection to certain supertrees. The polynomials arising in the refinement are extended to the Riordan matrix $P^x(1, zC)P = [p_{n,k}(x)]$ as in (7). We note that not only $p_{n,0}(x)$ but also the whole entries $p_{n,k}(x)$ can be interpreted as certain ordered forests of hex trees. A supertree is said to be *planted* if no trees grow at the root. A little modification of the bijection introduced in Theorem 2.2 yields a bijection between the planted stem supertrees and the planted hex trees. Specifically, we modify the symmetrization step and the final step as follows: we go through a planted stem supertree from the bottom to the top. After Step 1, the two edges from the root corresponds to a stem from the root, and a pair of edges from the same vertex v correspond to the left edge if v is the left child, and to the right edge if v is the right child. Labeling is the same as that of the bijection in Theorem 2.2. Whenever we meet a vertex with a label j, we insert j median edges from the vertex. This modified bijection maps a planted stem supertree with n edges in which the sum of the heights of stems is ℓ to a planted hex tree with n edges and ℓ median edges. For example, see Fig. 4.

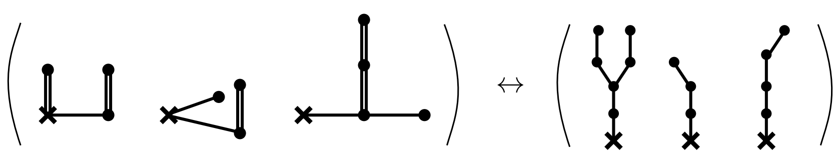


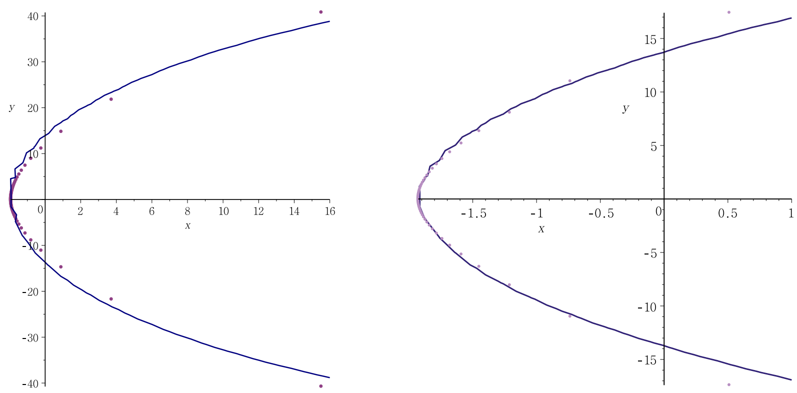
Figure 4. An ordered forest of 1 stem supertree and 2 planted stem supertrees, and the corresponding ordered forest of 1 symmetric hex tree and 2 hex trees.

This yields a combinatorial interpretation for the polynomials $p_{n,k}(x)$ in terms of both certain ordered forests of hex trees and ordered forests of stem supertrees. In particular, as mentioned in the proof of Theorem 3.6, $x = -1$ gives the Pascal conjugation of $(1, zC)$:

$$(R, zM) = \begin{bmatrix} 1 & & & & & & \\ 0 & 1 & & & & & \\ 1 & 1 & 1 & & & & \\ 1 & 3 & 2 & 1 & & & \\ 3 & 6 & 6 & 3 & 1 & & \\ 6 & 15 & 15 & 10 & 4 & 1 & \\ 15 & 36 & 40 & 29 & 15 & 5 & 1 \\ & & & \cdots & & & \end{bmatrix}$$

where $R = \frac{1+z-\sqrt{1-2z-3z^2}}{2z(1+z)}$ and $M = \frac{1-z-\sqrt{1-2z-3z^2}}{2z}$ are the generating functions for the Riordan numbers and the Motzkin numbers, respectively.

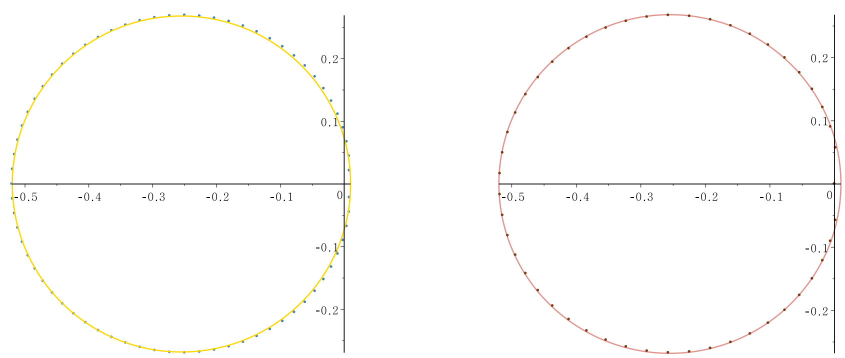
We end this paper with some conjectures concerning the zeros of polynomials $p_{n,k}$. It turns out that they reveal an interesting behavior besides some properties we have observed.



(a) (b)

Figure 5. The zeros of the polynomial (a) $p_{71,0}(x)$, (b) $p_{50,2}(x)$ and the conjectured limiting curves.

Fig. 5 suggests that the zeros of $p_{n,k}(x)$ accumulate to a curve that resembles a quadratic one. In order to guess the equations for the curves, let us consider the polynomials $r_{n,k}(x) := x^{n-k}p_{n,k}(x^{-1})$ which have the same coefficients with $p_{n,k}(x)$ but in reverse order.



(a) (b)

Figure 6. The zeros of the polynomial (a) $r_{71,0}(x)$, (b) $r_{50,2}(x)$, and the conjectured limiting curves.

Based on Fig. 6, one may guess that the zeros of $r_{n,k}(x)$ tend to accumulate to a circle or an ellipse. The centroid of the zeros of $r_{n,k}(x)$ is

$$-\frac{1}{n-k}\frac{[x^{n-k-1}]r_{n,k}(x)}{[x^{n-k}]r_{n,k}(x)} = -\frac{1}{n-k}\frac{[x^1]p_{n,k}(x)}{[x^0]p_{n,k}(x)} = -\frac{1}{n-k}\cdot\frac{\frac{2k+1}{n+k}\binom{2n-2}{n-k-1}\binom{n}{n-1}}{\frac{2k+1}{n+k+1}\binom{2n}{n-k}\binom{n}{n}} = -\frac{n+k+1}{4n-2} \to -\frac{1}{4}$$

as $n \to \infty$, for a fixed k.

Conjecture 1. For a fixed k, the zeros of $r_{n,k}(x)$ accumulate to the ellipse

$$\frac{\left(x+\frac{1}{4}+\frac{1}{4n}\right)^2}{\left(\frac{1}{4}+\frac{1}{n}\right)^2} + \frac{y^2}{\left(\frac{1}{4}+\frac{1}{n-1}\right)^2} = 1.$$

The ellipses in Conjecture 1 converge to the circle centered at $(-\frac{1}{4}, 0)$ with radius $\frac{1}{4}$. The reflection of the ellipse with respect to the unit circle centered at the origin gives the following conjecture on the zero behavior of $p_{n,k}(x)$.

Conjecture 2. For a fixed k, the zeros of $p_{n,k}(x)$ accumulate to the curve

$$\frac{\left(\frac{x}{x^2+y^2} + \frac{1}{4} + \frac{1}{4n}\right)^2}{\left(\frac{1}{4} + \frac{1}{n}\right)^2} + \frac{\left(\frac{y}{x^2+y^2}\right)^2}{\left(\frac{1}{4} + \frac{1}{n-1}\right)^2} = 1.$$

Remark. It seems that not only the zeros of $p_{n,k}(x)$ for fixed n and k tend to accumulate to some curve, but also the zeros of all $p_{n,k}(x)$ for a fixed k tend to form certain fountain-shaped curves, see Fig. 7. They also exhibit a complex version of interlacing [3].

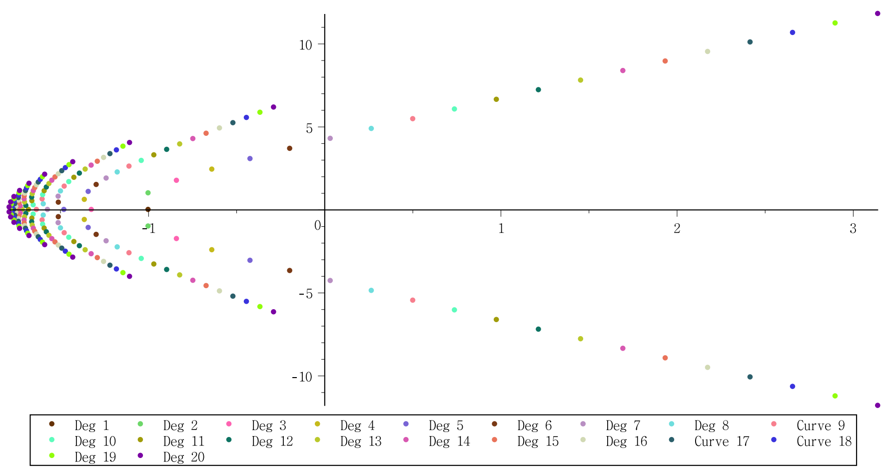


Figure 7. The zeros of polynomials $p_{n,0}(x)$ for $n = 1, \ldots, 20$. Each color depicts the zeros of a different polynomial.

References

[1] A. T. Balaban and F. Harary, Chemical graphs-V : Enumeration and proposed nomenclature of benzenoid cata-condensed polycyclic aromatic hydrocarbons, Tetrahedron 24 (1968), 2505-2516.

[2] W. A. Baldwin and G. O. Strawn, Multidimensional trees, Theor. Comput. Sci. 84 (1991), 293-311.

[3] C. M. Bender, S. Boettcher and V. M. Savage, Conjecture on the interlacing zeros in complex Sturm-Liouville problems, J. Math. Phys. 41 (2000), 6381-6387.

[4] O. R. P. Bininda-Emonds, The evolution of supertrees, Trends. Ecol. Evol. 19 (2004), 315-322.

[5] S. J. Cyvin, B. N. Cyvin and J. Brunvoll, Isomer enumeration of catafusenes, $C_{4h+2}H_{2h+4}$ Benzenoid and Heliceinc Hydrocarbons, MATCH-Commun. Math. Co. 26 (1991), 63-86.

[6] S. J. Cyvin, B. N. Cyvin, J. Brunvoll, Z. Fuji, G. Xiaofeng and R. Toŏić, Graph-theoretical studies on fluoranthenoids and fluorenoids: enumeration of some catacondensed systems, J. Mol. Struct. 285 (1993), 179-185.

[7] N. G. De Bruijin and B. J. M. Morselt, A note on plane trees, J. Comb. Theory 2 (1967), 27-34.

[8] P. Flajolet and R. Sedgewick, Analytic Combinatorics, Cambridge University Press, Cambridge, 2009.

[9] F. Harary and R. C. Read, The enumeration of tree-like polyhexes, Proc. Edinburgh Math. Soc. 17 (1970), 1-13.

[10] D. Merlini, D. G. Rogers, R. Sprugnoli and M. C. Verri, On some alternative characterizations of Riordan arrays, Canad. J. Math. 49 (1997), 301-320.

[11] V. V. Prasolov, Polynomials, Springer, Heidelberg, 2004.

[12] L. W. Shapiro, S. Getu, W. Woan and L. C. Woodson, The Riordan group, Discrete Appl. Math. 34 (1991), 229-239.

[13] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences, http://www.research.att.com/~njas/sequences.

[14] R. P. Stanley, Enumerative Combinatorics, vol. 2, Cambridge University Press, Cambridge, 1999.

[15] W. Wang and T. Wang, Generalized Riordan arrays, Discrete Math. 308 (2008), 6466-6500.