Enumeration of trees by inversions

Ira M. Gessel [1]
Department of Mathematics, Brandeis University, Waltham, MA 02254-9110, U.S.A.

Bruce E. Sagan [2]
Department of Mathematics, Michigan State University, East Lansing, MI 48824-1027, U.S.A.

Yeong-Nan Yeh [3]
Institute of Mathematics, Academia Sinica, Nankang, Taipei, Taiwan 11529, R.O.C.

May 26, 1999

Key Words: bijection, generating function, hypergeometric series, ordered tree, plane tree, cyclic tree

AMS subject classification (1985): Primary 05C05; Secondary 05A15.

[1]Supported in part by NSF grant DMS 8902666

[2]Supported in part by NSF grant DMS 8805574 and the National Science Council, Taiwan, R.O.C.

[3]Supported in part by Grant NSC-82-0208-M001-042 from the National Science Council, Taiwan, R.O.C.

Proposed running head:

Trees and inversions

Send proofs to:

From August 1993 to August 1994:

Bruce E. Sagan
Matematiska Institutionen
Kungl. Tekniska Högskolan
Fiskartorpsv. 15A
S-100 44 Stockholm
SWEDEN
FAX: 8-723-1788

After August 1994:

Bruce E. Sagan
Department of Mathematics
Michigan State University
East Lansing, MI 48824-1027
U.S.A.
FAX: 517-336-1562

Abstract

Mallows and Riordan [21] first defined the inversion polynomial, $J_n(q)$, for trees with n vertices and found its generating function. In the present work, we define inversion polynomials for ordered, plane and cyclic trees and find their values at $q = 0, \pm 1$. Our techniques involve the use of generating functions (including Lagrange inversion), hypergeometric series and binomial coefficient identities, induction and bijections. We also derive asymptotic formulae for those results for which we do not have a closed form.

1 Introduction

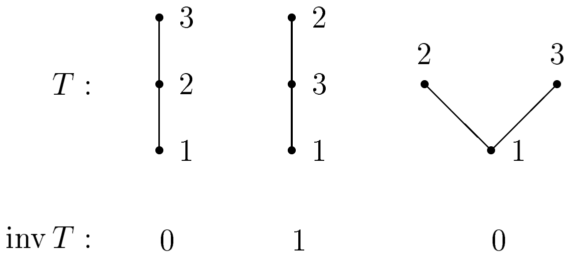
In this paper, a *tree*, T, will mean a rooted tree labeled with the integers $1, 2, \ldots, n$ which has the label 1 at the root. An *inversion in* T is a pair of vertices labeled i, j such that $i > j$ and i is on the unique path from 1 to j in T. Let

$$\operatorname{inv} T = \text{ number of inversions in } T$$

Now we can define the *tree inversion polynomials* by

$$J_n(q) := \sum_{|T|=n} q^{\operatorname{inv} T}$$

where $|\cdot|$ denotes cardinality and $|T|$ is the cardinality of T's vertex set. To illustrate, for n=3 we have listed all three trees and their inversions in Figure 1.

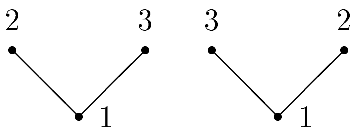


Trees and inversions for $n = 3$

Figure 1

Thus $J_3(q) = q^0 + q^1 + q^0 = 2 + q$. Mallows and Riordan [21] were the first to define the inversion polynomial and find its exponential generating function. We will investigate analogs of the inversion polynomial for three other types of trees.

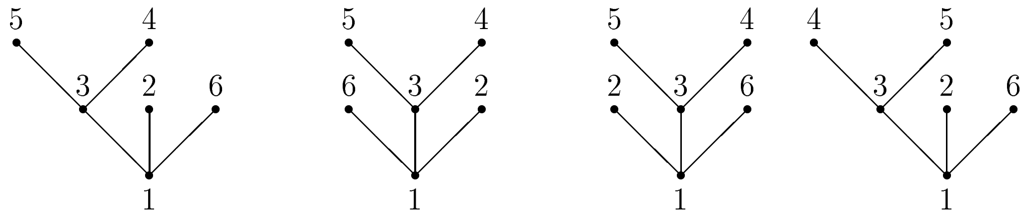
An *ordered tree* is one where the children of each vertex have been ordered left to right. For example, the two trees of Figure 2 are the same as ordinary trees but different as ordered trees.



Two ordered trees

Figure 2

A *plane tree* is an equivalence class of ordered trees, where two trees T and T' are considered identical if T can be transformed into T' by a motion in the plane such that edges don't cross. Since all of our trees are rooted, this is equivalent to the condition that the subtrees of the root of T' are a circular rearrangement of the subtrees of the root of T. Among the trees in Figure 3, only the first two are equal when considered as plane trees.



Four trees

Figure 3

Finally, define a *cyclic tree* to be an equivalence class of ordered trees where T and T' are considered identical if T' can be obtained from T by circularly rearranging the subtrees of every vertex. As cyclic trees, the first, second and fourth trees in Figure 3 are equal, while the third one is still in a different equivalence class. Ordered and plane trees are well known in the literature, but the cyclic variety have only been considered twice before [8, 19].

The *ordered inversion polynomial* is

$$J_n^o(q) := \sum_{\substack{|T|=n \\ T \text{ ordered}}} q^{\operatorname{inv} T}$$

The *plane inversion polynomial*, $J_n^p(q)$, and *cyclic inversion polynomial*, $J_n^c(q)$, are defined similarly. We will calculate the values of these polynomials at $q = 0, \pm 1$. However, to state these results we need a little more notation.

If n is a non-negative integer, then its *double factorial* is

$$n!! := n(n-2)(n-4)\cdots$$

where the last factor of the product is 2 or 1 depending upon whether n is even or odd. Although such factorials can always be expressed in terms of ordinary factorials and powers of 2, this notation will make our results more transparent. (It is also why we have chosen this notation over other, equally conventional, notations.) The *triple factorial* is defined by

$$n!!! := n(n-3)(n-6)\cdots$$

and similarly for higher factorials. Given another non-negative integer, k, we will also need to have *falling factorials*

$$\langle n\rangle_k := n(n-1)(n-2)\cdots(n-k+1)$$

and *rising factorials*

$$(n)_k := n(n+1)(n+2)\cdots(n+k-1)$$

Other useful combinatorial functions include the *binomial coefficients*, $\binom{n}{k}$, which count k-element subsets of an n-element set, and *(signless) Stirling numbers of the first kind*, $c(n,k)$, which count permutations of n elements that decompose into k disjoint cycles. Finally, we need the *exponential integral function* defined by

$$E(t) := \int_t^\infty \frac{du}{ue^u} \tag{1}$$

Also given any function $f(t)$ expanded as a Maclaurin series, we let

$$\mathbf{C}_{t^n} f(t) := \text{ coefficient of } t^n \text{ in } f(t)$$

We are now in a position to state our main result.

Theorem 1.1 *The values of the three polynomials $J_n^o(q), J_n^p(q), J_n^c(q)$ evaluated at $q = 0, \pm 1$ are given in the following table.*

	$q=0$	$q=1$	$q=-1$
$J_n^o(q)$	$(2n-3)!!$	$\frac{1}{n}(4n-6)!!!!$	$(3m-1)!!!(3m-2)!!!$ *if* $n=2m+1$ $(3m-2)!!!(3m-4)!!!$ *if* $n=2m$
$J_n^p(q)$	$(2n-4)!!$	$\frac{1}{2(n-1)}(4n-6)!!!!$	$(n-2)! \sum_{k\le(n-2)/2} \frac{1}{3^k}\binom{n+k-2}{k}$
$J_n^c(q)$	$\mathbf{C}_{t^n/n!}(1-e^{1-E^{-1}(E(1)-\frac{t}{e})})$	$\frac{1}{n}\sum_{k\ge 0} c(n-1,k)\langle n\rangle_k$	?

In the spirit of Moon's book on counting labeled trees [23], we will give as many different proofs as possible of these results. Our techniques involve the use of generating functions (including Lagrange inversion), hypergeometric series and binomial coefficient identities, induction and bijections. These proofs will be found in Sections 2 through 5. We conclude with sections on asymptotic results and open problems.

Our three specializations all have combinatorial interpretations. When $q = 0$ we are counting trees with no inversions [4, 20] which are sometimes called *increasing*. The case $q = 1$ counts all trees of a given type. Finally, setting $q = -1$ gives the net number of trees with an even number of inversions over the number with an odd number. Other interpretations at $q = -1$ can be given (see [24] and Section 7).

The only formulas in the table that we have found in the literature are the ones for $J_n^o(0)$, $J_n^o(1)$ and $J_n^c(1)$. The first independently appeared in the thesis of William Chen [7] where he uses the term "plane tree" to refer to what we have called an ordered tree. He gives a proof using generating functions which is identical to ours, and a bijective demonstration (motivated by a comment of Jay Goldman) which is different. He also outlines an inductive proof of Stanley which is the same as the one we reproduce in Section 4. The second specialization was derived by Rodrigues [27] when considering a problem related to Catalan's [5]. (Or see [10, pp. 21–27].) In particular, he showed that the number of ways to parenthesize a string of n elements which are not in a fixed order is $(4n-6)!!!!$. Thus

$$nJ_n^o(1) = (4n-6)!!!! = n!C_{n-1}$$

where $C_n = \frac{1}{n+1}\binom{2n}{n}$ is the nth Catalan number. Rodrigues' demonstration is combinatorial, and can be converted into the corresponding bijection in Section 5 with a little bit of work. A different bijective proof is given by Chen [7, 8]. The third specialization is also in [8] as well as following from Corollary 1 of Labelle's article [18].

2 Generating functions

We can use generating functions to derive all of the results in Theorem 1.1. Define the *q-analog of n* to be

$$[n] := 1 + q + q^2 + \cdots + q^{n-1}$$

for any non-negative integer n. Also let

$$F_q^o(t) = \sum_{n\geq 0} J_n^o(q)\frac{t^n}{n!}, \quad G_q^o(t) = \sum_{n\geq 1} [n]J_n^o(q)\frac{t^n}{n!}.$$

and similarly for the plane and cyclic cases. It will be seen in the next proposition that $G_q^o(t)$ is the generating function for ordered trees by inversions where any label can be at the root (and similarly for $G_q^p(t), G_q^c(t)$). Finally, let D_t denote the derivative operator with respect to t.

Proposition 2.1 *These generating functions satisfy the following differential equations*

1. $D_t F_q^o(t) = \frac{1}{1-G_q^o(t)}$
2. $D_t F_q^p(t) = 1 + \ln \frac{1}{1-G_q^o(t)}$
3. $D_t F_q^c(t) = 1 + \ln \frac{1}{1-G_q^c(t)}$

Proof. We shall prove the first of these identities, the others being similar. From the theory of exponential series [6, 12, 16, 20, 29] it follows that the derivative counts ordered trees with $n+1$ vertices (as the coefficient of $t^n/n!$) by inversions. Removing the root of such a tree leaves an ordered forest of rooted trees. The generating function for such trees is $G_q^o(t)$ where the extra factor of $[n]$ in each term is needed to account for inversions from the root (which may no longer be the smallest label). Substituting this series into $1/(1-t)$ counts ordered lists of these trees, which is the desired result. ■

$\underline{J_n^o(0)}$. Substituting $q = 0$ in part 1 of Proposition 2.1 and noting that $[n]_{q=0} = 1$ yields

$$D_t F_0^o(t) = \frac{1}{1 - F_0^o(t)}$$

Separating variables and integrating gives

$$F_0^o - \frac{1}{2}\left(F_0^o\right)^2 = t$$

which can be solved as

$$F_0^o = 1 - \sqrt{1-2t}$$

(This generating function is also given in [20].) So, by the binomial theorem,

$$\begin{aligned} J_n^o(0) &= \mathbf{C}_{t^n/n!}(1 - \sqrt{1-2t}) \\ &= -n!\binom{1/2}{n}(-2)^n \\ &= (2n-3)!! \end{aligned}$$

as desired. ■

$\underline{J_n^p(0)}$. Putting $q = 0$ in the second part of Proposition 2.1 we get

$$D_t F_0^p(t) = 1 + \ln \frac{1}{1 - F_0^o(t)} = 1 - \ln \sqrt{1 - 2t}$$

Taking the coefficient of $t^{n-2}/(n-2)!$ in $D_t^2 F_0^p(t)$ finishes the computation. ■

$\underline{J_n^c(0)}$. We are indebted to Herbert Wilf for this proof and the demonstration of the corresponding asymptotic result in Section 6. Using the usual proposition and substitution we obtain

$$D_t F_0^c(t) = 1 + \ln \frac{1}{1 - F_0^c(t)} \tag{2}$$

Substituting

$$1 - F_0^c(t) = \exp(1 - H(t)) \tag{3}$$

allows us to separate variables and get

$$\frac{H'}{He^H} = \frac{1}{e}$$

Now we can integrate using the exponential integral function (1)

$$-E(H(t)) \Big|_0^t = \frac{t}{e} \Big|_0^t$$

or

$$-E(H(t)) + E(H(0)) = \frac{t}{e}$$

But from (3)

$$H(0) = 1 + \ln \frac{1}{1 - F_0^c(0)} = 1$$

so

$$H(t) = E^{-1}(E(1) - t/e) \tag{4}$$

Substitution of this expression back into equation (3) and solving for F_0^c completes the proof. ■

$\underline{J_n^o(1)}$. Start the same way as in the proof of the formula for $J_n^o(0)$ except letting $q = 1$. Multiplying the resulting equation by t yields

$$G_1^o(t) = \frac{t}{1 - G_1^o(t)}$$

Next, solve for F_1^o

$$G_1^o(t) = \frac{1 - \sqrt{1 - 4t}}{2} \tag{5}$$

and extract the proper coefficient. ∎

$\underline{J_n^p(1).}$ Using the usual method and equation (5), we get

$$\begin{aligned}\sum_{n\geq 0} J_{n+1}^p(1)\frac{t^n}{n!} &= 1+\ln\frac{1}{1-G_1^o(t)} \\ &= 1+\ln 2-\ln(1+\sqrt{1-4t})\end{aligned}$$

Taking the derivative and multiplying by t yields

$$\sum_{n\geq 0} nJ_{n+1}^p(1)\frac{t^n}{n!} = \frac{(1-4t)^{-1/2}-1}{2}$$

The coefficient of $t^{n-1}/(n-1)!$ gives the desired formula. ∎

$\underline{J_n^c(1).}$ Now our functional equation is

$$G_1^c(t) = t\left(1+\ln\frac{1}{1-G_1^c(t)}\right)$$

This can be solved by Lagrange inversion. (See [29, pp. 138–141].) If

$$H(u) = 1+\ln\frac{1}{1-u}$$

then

$$\begin{aligned}nJ_n^c(1) &= \mathbf{C}_{t^n/n!}G_1^c(t) \\ &= n!\cdot\frac{1}{n}\mathbf{C}_{u^{n-1}}H(u)^n \\ &= (n-1)!\mathbf{C}_{u^{n-1}}\left(1+\ln\frac{1}{1-u}\right)^n \\ &= \mathbf{C}_{u^{n-1}/(n-1)!}\sum_k\binom{n}{k}\left(\ln\frac{1}{1-u}\right)^k \\ &= \sum_k\binom{n}{k}k!c(n-1,k) \\ &= \sum_k\langle n\rangle_k c(n-1,k) \quad \blacksquare\end{aligned}$$

$\underline{J_n^o(-1).}$ Let

$$E(t) = \sum_{n \text{ even}} J_n^o(-1)\frac{t^n}{n!}$$

and

$$O(t) = \sum_{n \text{ odd}} J_n^o(-1) \frac{t^n}{n!}$$

The substitution $q = -1$ in part 3 of Proposition 2.1 yields

$$D_t F_{-1}^o(t) = E'(t) + O'(t) = \frac{1}{1 - O(t)}$$

Clearing the fraction gives

$$E' - OE' + O' - OO' = 1$$

Equating even and odd powers of t, we get

$$O' - OE' = 1 \tag{6}$$

and

$$E' - OO' = 0 \tag{7}$$

Substituting OO' for E' in equation (6) yields

$$O' - O^2 O' = 1$$

which can be integrated to obtain

$$O - O^3/3 = t \tag{8}$$

Also, integration of equation (7) gives

$$E = O^2/2 \tag{9}$$

Equation (8) can be solved by a more general version of Lagrange inversion that will also give the solution to (9). Writing (8) as

$$O(t) = \frac{t}{1 - O(t)^2/3}$$

we apply Lagrange in the form

$$\mathbf{C}_{t^n} f(O(t)) = \frac{1}{n} \mathbf{C}_{u^{n-1}} (Df(u))(1 - u^2/3)^{-n} \tag{10}$$

Taking the coefficients for $f(u) = u$ and $f(u) = u^2/2$ gives, respectively,

$$O(t) = \sum_{m \geq 0} \frac{(3m)!}{3^m m!} \frac{t^{2m+1}}{(2m+1)!}$$

and

$$E(t) = O(t)^2/2 = \sum_{m\geq 1} \frac{(3m-2)!}{3^{m-1}(m-1)!} \frac{t^{2m}}{(2m)!}$$

These are equivalent to the formulae in our table. ■

$\underline{J^p_n(-1).}$ In the manner to which we have become accustomed

$$D_t F^p_{-1}(t) = 1 + \ln \frac{1}{1 - O(t)}$$

where $O(t)$ is as defined in the previous proof. Applying (10) with the function $f(u) = 1 + \ln(1-u)^{-1}$ yields

$$\begin{aligned} J^p_{n+1}(-1) &= n! \mathbf{C}_{t^n} f(O(t)) \\ &= (n-1)! \mathbf{C}_{u^{n-1}} \frac{1}{(1-u)(1-u^2/3)^n} \end{aligned}$$

Coefficient extraction completes our demonstration of the final formula in Theorem 1.1. ■

3 Hypergeometric series & binomial coefficients

We first derive recurrence relations satisfied by the ordered and plane inversion polynomials. They can all be derived by algebraic manipulations from Proposition 2.1. However, we will give a combinatorial proof.

Proposition 3.1 *The inversion enumerators satisfy*

1. $J^o_{n+1}(q) = \sum_{k\geq 1} \binom{n}{k} [k] J^o_k(q) J^o_{n-k+1}(q)$
2. $J^p_{n+1}(q) = \sum_{k\geq 1} \binom{n-1}{k-1} [k] J^o_k(q) J^o_{n-k+1}(q)$

Proof. Any ordered tree on $n+1$ vertices is composed of its leftmost principle subtree, T_1, together with the subtree T_2 consisting of the root with the rest of the principle subtrees attached. If T_1 has k vertices, then there are $\binom{n}{k}$ ways to choose them (since the root of T_2 must be labeled 1). Also T_1 and T_2 contribute $[k]J^o_k(q)$ and $J^o_{n-k+1}(q)$ inversions, respectively, to the total so we are done with the first formula. The proof of the second equation is similar except that one selects the principle subtree containing the label 2, which accounts for the different binomial coefficient. ■

If $a_1, a_2, \ldots, a_p$ and $b_1, b_2, \ldots, b_q$ are constants, then we can form the *hypergeometric series*

$$ {}_pF_q\left[\begin{array}{cccc|c} a_1, & a_2, & \ldots, & a_p & x \\ & b_1, & \ldots, & b_q & \end{array}\right] := \sum_{k\geq 0} \frac{(a_1)_k(a_2)_k\cdots(a_p)_k}{(b_1)_k\cdots(b_q)_k}\frac{x^k}{k!} $$

For information about such functions, see the books of Bailey [1] or Slater [28]. In order to convert our results into hypergeometric format, we will need to express all our our binomial coefficients and multiple factorials in terms of rising factorials of the form $(c)_k$ where c is constant with respect to k. For example,

$$ \begin{aligned} \binom{n}{k} &= \frac{(-1)^k(-n)_k}{k!} \\ (2k)!! &= 2^k(1)_k \\ (2k-1)!! &= 2^k(\tfrac{1}{2})_k \\ (2n-2k)!! &= \frac{(2n)!!}{(-2)^k(-n)_k} \\ (2n-2k-1)!! &= \frac{(2n-1)!!}{(-2)^k(-n+\frac{1}{2})_k} \end{aligned} $$

$\underline{J_n^o(0).}$ Substituting $q=0$ in Proposition 3.1 gives

$$ J_{n+1}^o(0) = \sum_{k\geq 1}\binom{n}{k} J_k^o(0)J_{n-k+1}^o(0) $$

By induction on n, we can write

$$ J_{n+1}^o(0) = \sum_{k\geq 1}\binom{n}{k}(2k-3)!!(2n-2k-1)!! $$

Converting to hypergeometric format, we get

$$ \begin{aligned} J_{n+1}^o(0) &= \sum_{k\geq 1}\frac{(-1)^k(-n)_k}{k!}(-2^k(-\tfrac{1}{2})_k)\frac{(2n-1)!!}{(-2)^k(-n+\frac{1}{2})_k} \\ &= (2n-1)!!\left(1 - {}_2F_1\left[\begin{array}{cc|c} -n, & -\frac{1}{2} & 1 \\ & -n+\frac{1}{2} & \end{array}\right]\right) \end{aligned} $$

This is a special case of the Chu-Vandermonde Theorem [28, p. 28].

Theorem 3.2 (Chu-Vandermonde) *If n is a positive integer, then*

$$ {}_2F_1\left[\begin{matrix} -n, & a \\ & -n+b \end{matrix}\middle| 1\right] = \frac{(1-b+a)_n}{(1-b)_n} \blacksquare $$

In our case

$$ {}_2F_1\left[\begin{matrix} -n, & -\frac{1}{2} \\ & -n+\frac{1}{2} \end{matrix}\middle| 1\right] = \frac{(0)_n}{(\frac{1}{2})_n} = 0 \tag{11} $$

and the formula $J^o_{n+1}(0) = (2n-1)!!$ follows. $\blacksquare$

$\underline{J^p_n(0),\ J^o_n(1),\ J^p_n(1).}$ These three proofs are all very similar to the one for $J^o_n(0)$. (In the plane cases, we use the formulae derived for ordered trees, instead of induction when substituting into the summation.) The analogs of equation (11) for $J^p_n(0)$, $J^o_n(1)$ and $J^p_n(1)$ are, respectively,

$$ \begin{aligned} {}_2F_1\left[\begin{matrix} -n+1, & \frac{1}{2} \\ & -n+\frac{3}{2} \end{matrix}\middle| 1\right] &= \frac{(2n-2)!!}{(2n-3)!!} \\ {}_2F_1\left[\begin{matrix} -n-1, & -\frac{1}{2} \\ & -n+\frac{1}{2} \end{matrix}\middle| 1\right] &= 0 \\ {}_2F_1\left[\begin{matrix} -n, & \frac{1}{2} \\ & -n+\frac{3}{2} \end{matrix}\middle| 1\right] &= 0 \end{aligned} $$

All of these follow easily from the Chu-Vandermonde Theorem.

$\underline{J^o_n(-1).}$ When $q=-1$ we have $[k] = 0, 1$ depending on whether k is even or odd respectively. Thus this substitution gives

$$ \begin{aligned} J^o_{n+1}(-1) &= \sum_{\substack{k\geq 1 \\ k \text{ odd}}} \binom{n}{k} J^o_k(-1) J^o_{n-k+1}(-1) \\ &= \sum_{l\geq 0} \binom{n}{2l+1} J^o_{2l+1}(-1) J^o_{n-2l}(-1) \end{aligned} $$

To apply induction, we will have to break into two cases since the correct value for $J^o_{n-2l}(-1)$ depends on the parity of n. Also, this time it will be convenient to write our results in terms of binomial coefficients. If $n = 2m+1$ then

$$ \begin{aligned} J^o_{2m+2}(-1) &= \sum_{l\geq 0} \frac{(2m+1)!}{(2l+1)!(2m-2l)!} \frac{(3l)!}{3^l l!} \frac{(3m-3l)!}{3^{m-l}(m-l)!} \\ &= \frac{(2m+1)!}{3^m} \sum_{l\geq 0} \frac{1}{3l+1} \binom{3l+1}{l} \binom{3m-3l}{m-l} \end{aligned} $$

Similarly, if $n = 2m$ then

$$J^o_{2m+2}(-1) = \frac{(2m)!}{3^m} \sum_{l\geq 0} \frac{1}{3l+1}\binom{3l+1}{l}\binom{3m-3l+1}{m-l}$$

To evaluate these sums and the ones that appear in the following proof for $J^p_n(-1)$ we will need a couple of convolutions that can be derived from the generalized binomial and exponential series [15, pp. 200–204]. (See also Gould [14, Formula 3.144].)

Theorem 3.3 *If n is a positive integer and a, b, c are constants, then*

$$\sum_{l\geq 0} \frac{a}{a+cl}\binom{a+cl}{l}\binom{b+c(n-l)}{n-l} = \binom{a+b+cn}{n} \tag{12}$$

and

$$\sum_{l\geq 0} \binom{a+cl}{l}\binom{b+c(n-l)}{n-l} = \sum_k c^k \binom{a+b+cn-k}{n-k} \ \blacksquare \tag{13}$$

Plugging $a = 1, b = 0, c = 3$ into (12) yields the n odd case while $a = 1, b = 1, c = 3$ yields n even.

$\underline{J^p_n(-1).}$ Using both the proof and results of the preceding case we can obtain

$$J^p_{2m+2}(-1) = \frac{(2m)!}{3^m} \sum_{l\geq 0} \binom{3l}{l}\binom{3m-3l}{m-l} \tag{14}$$

and

$$J^p_{2m+1}(-1) = \frac{(2m-1)!}{3^{m-1}} \sum_{l\geq 0} \binom{3l}{l}\binom{3m-3l-2}{m-l-1} \tag{15}$$

To show that these are the same as the sums for $J^p_n(-1)$ in Theorem 1.1, use (13) with $a = 0, b = 0, c = 3, n = m$ (n even) or $a = 0, b = 1, c = 3, n = m-1$ (n odd).

It has been shown that these sums do not have a closed form. Doron Zeilberger has an algorithm for producing recurrences satisfied by hypergeometric sums [31]. First Zeilberger, with the computational help of Shalosh B. Ekhad, provided us with relations satisfied by the sums. Then Marko Petkovšek reduced the order of one of the recursions and applied his algorithm for determining all hypergeometric solutions to linear difference equations with polynomial coefficients [25]. Since these sums were not among them, they have no closed form. For completeness, we state the recurrence relations here; the reader should consult Petkovšek's article for details of the proof. In what follows, E denotes the shift operator with respect to m, i.e., $E(f(m)) = f(m+1)$.

Proposition 3.4 *Let*

$$f(m) = \sum_{l \geq 0} \binom{3l}{l} \binom{3m-3l}{m-l}$$

and

$$F(m) = \sum_{l \geq 0} \binom{3l}{l} \binom{3m-3l-2}{m-l-1}$$

Then $f(m)$ satisfies the second order recurrence

$$[2(m+2)(2m+3)E - 3(3m+2)(3m+4)][4E-27]f(m) = 0$$

and $F(m)$ satisfies the second order recurrence

$$[2(m+1)(2m+3)E - 3(3m+1)(3m+2)][4E-27]F(m) = 0 \blacksquare$$

4 Inductive proofs

For some of the formulae which are in the form of a product, we can give inductive proofs. We will sometimes abuse notation and write $v = m$ when the vertex v is the one with label m. Remember that $v = 1$ is always the root.

$\underline{J_n^o(0).}$ It suffices to show that

$$J_{n+1}^o(0) = (2n-1)J_n^o(0)$$

Take any ordered tree, T, on $n+1$ vertices with no inversions. Then $n+1$ must be a leaf. Delete this leaf to obtain a tree, T', on n vertices. This defines a map

$$T \overset{d}{\longmapsto} T' \tag{16}$$

To finish the proof we need only demonstrate that each such T' is the image of $2n-1$ trees, T.

Consider how many ways one can add back a leaf labeled $n+1$ to T'. Given a vertex $v \in T'$,

$$\text{number of ways to attach } n+1 \text{ to } v = \begin{cases} \deg v & \text{if } v \neq 1 \\ \deg v + 1 & \text{if } v = 1 \end{cases}$$

where $\deg v$ is the degree of v. Thus the total number of ways to add back $n+1$ is

$$1 + \sum_{v \in T'} \deg v = 1 + 2(n-1) = 2n - 1 \tag{17}$$

since the sum of the degrees in any graph is twice the number of edges. ∎

$\underline{J_n^p(0).}$ Now we must show

$$J_{n+1}^p(0) = (2n-2)J_n^p(0)$$

If T is a plane tree, we can insist that it always be written in *standard form* with the children of $v = 1$ ordered so that the smallest child is on the right. We now follow the same argument as for $J_n^o(0)$. The only change is that there are only $\deg v$ ways to add $n+1$ at the root, so the sum becomes $2n-2$ as desired. ∎

$\underline{J_n^o(1).}$ We want $nJ_n^o(1) = (4n-6)!!!!$. Note that

$$o_n := nJ_n^o(1) \tag{18}$$

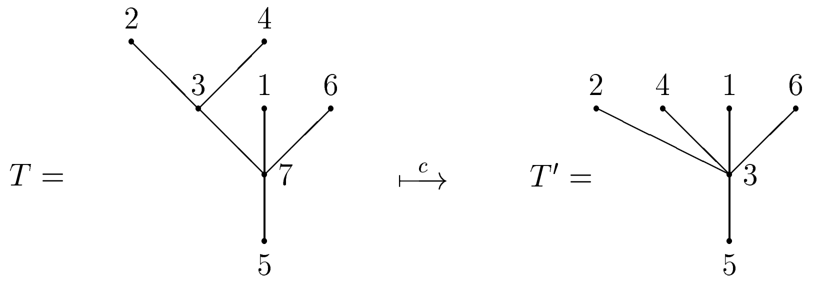
is just the number of labeled ordered trees with any label at the root. So we must show

$$o_{n+1} = (4n-2)o_n$$

Let T be an ordered labeled tree on $n+1$ vertices rooted at an arbitrary vertex, and consider vertex $n+1$. Obtain a tree T' on n vertices, denoted

$$T \overset{c}{\longmapsto} T' \tag{19}$$

as follows. If $n+1$ is a leaf, just delete it. Otherwise, let v be the leftmost child of $n+1$. Then T' is formed by contracting the edge from $n+1$ to v and labeling the node obtained by the identification with the same label as v. Note that the subtrees of v in T become the leftmost subtrees of v in T' in their same relative order, while the subtrees of $n+1$ in T become the rightmost subtrees of v in T' in their same relative order. Figure 4 contains an example.



Contraction

Figure 4

Now given T', how many trees T map to it? We can add $n+1$ as a leaf in any of $2n-1$ ways, according to equation (17). If $n+1$ is to be added back as an internal vertex, then it can replace any $v \in T'$, pushing v off as a new left child. Also consider the subtrees of v in T' which we will call $T_1, T_2, \ldots, T_k$, listed in left-right order. Then in T, v has subtrees $T_1, \ldots, T_i$, and $n+1$ has subtrees corresponding to v and then $T_{i+1}, \ldots, T_k$ for some i with $0 \leq i \leq k$. Since

$$k = \begin{cases} \deg v - 1 & \text{if } v \text{ is not the root} \\ \deg v & \text{if } v \text{ is the root} \end{cases}$$

We get $2n-1$ possible T in this way. Thus the total number of preimages of T' is $(2n-1)+(2n-1) = 4n-2$. ∎

5 Bijections

We can turn the inductive proofs of the previous section into bijective ones as follows.

$\underline{J_n^o(0).}$ It is enough to construct a bijection

$$T \overset{f}{\longmapsto} a_2 a_3 \ldots a_n$$

where T ranges over all ordered trees with $T = n$ and no inversions, and $a_2 a_3 \ldots a_n$ is a sequence of integers with $1 \leq a_i \leq 2i-3$ for all i. Given T, we perform depth-first search, always following the leftmost available edge. Now record the label of each node each time it is visited to form a sequence

$$l = l_1 l_2 \ldots l_{2n-1} \tag{20}$$

For example, the first tree in Figure 3 would yield the sequence

$$l_1 l_2 \ldots l_{11} = 1\ 3\ 5\ 3\ 4\ 3\ 1\ 2\ 1\ 6\ 1 \tag{21}$$

Now define the i-th element of the sequence $a_2 \ldots a_n$ by finding

$$i_f := \text{ number of visits prior to the first visit to } i$$

and computing

$$a_i = i_f - 2 \cdot |\{l_j \ : \ j \leq i_f \text{ and } l_j > i\}| \tag{22}$$

To illustrate, let's compute a_2 for the sequence (21). There are 7 visits prior to the first (indeed, only) visit to 2, and there are 3 numbers among those prior visits larger than 2 (namely the three, four and five), so we have $a_2 = 7 - 2(3) = 1$. Similar computations yield, for this example,

$$T \stackrel{f}{\longmapsto} 1\ 1\ 2\ 2\ 9$$

To show that the function f defined in the previous paragraph is a well-defined bijection, it suffices to prove the following lemma. In it, the map $T \stackrel{d}{\mapsto} T'$ is the map defined in (16).

Lemma 5.1 *Fix an ordered tree T' with $|T'| = n$ and no inversions, and suppose*

$$T' \stackrel{f}{\longmapsto} a_2 a_3 \ldots a_n$$

Let T be any of the $2n-1$ ordered trees on $n+1$ vertices such that $T \stackrel{d}{\mapsto} T'$. Then

$$T \stackrel{f}{\longmapsto} b_2 b_3 \ldots b_{n+1}$$

where $b_i = a_i$ for $i \leq n$ and $1 \leq b_{n+1} \leq 2n-1$. Furthermore, the $2n-1$ choices for T each yield a different b_{n+1}.

Proof. Suppose that $n+1$ is attached to some vertex labeled v in passing from T' to T. Then the effect on the sequence (20) is to insert the pair $n+1\ v$ after some occurrence of v. Now the computation of b_i for $i \leq n$ falls into two cases. If the first visit to i was before the insertion, then clearly $b_i = a_i$. If the visit was after, then the insertion increases the value of i_f by two. But this also increases the cardinality of the set in equation (22) by one, so the net change is $2 - 2(1) = 0$ again.

Now we claim that b_{n+1} takes on the values from 1 to $2n-1$ each exactly once. But adding $n+1$ to T' in all possible ways corresponds to inserting a pair starting with $n+1$ after every possible element of the sequence (20) for T. Thus i_{n+1} takes on these values and the set in equation (22) is empty. The claim and the lemma are now proved. ■

$\underline{J_n^p(0).}$ We can construct a map

$$T \stackrel{f}{\longmapsto} a_3 \ldots a_n$$

where T is a plane tree in standard form and a_i is defined exactly as before. It is easy to verify that the standardness of T implies that we get exactly those sequences with $1 \leq a_i \leq 2i-4$ for all i. This completes the proof. ■

$\underline{J_n^o(1).}$ Now we want a bijection

$$T \stackrel{g}{\longmapsto} a_2 \ldots a_n$$

where T ranges over all ordered trees with any number at the root and $1 \leq a_i \leq 4i-6$. Such a map can be constructed along the same lines as the one for $J_n^o(0)$, but it is more complicated and so not as interesting. We omit the details.

$\underline{J_n^p(1).}$ We will provide a bijection to show that

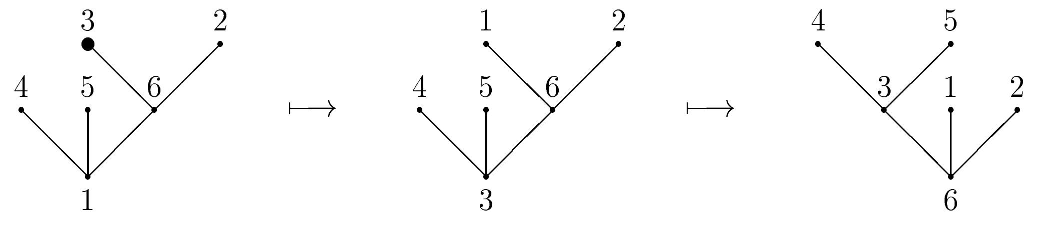
$$2(n-1)J_n^p(1) = o_n$$

where o_n is as defined by (18). This proof is actually a translation of a generating function argument.

The factor $(n-1)J_n^p(1)$ counts plane trees where some vertex other than the root has been marked. Such trees are clearly equinumerous with ordered trees having a marked vertex in the leftmost subtree of the root. Thus $2(n-1)J_n^p(1)$ counts the disjoint union of two sets of ordered trees: one with each leftmost subtree of the root containing a distinguished vertex and the other with each rightmost subtree of the root containing a marked node. If the root of T has only one subtree, then T appears twice, once with the subtree being leftmost and once with it being rightmost. (This is what we mean by the union being disjoint.)

To show that these are in bijection with ordered trees having any label at the root, first consider T with a marked node labeled l in the leftmost subtree. Exchange l with the 1 at the root of T to form a new tree T'. Note that the mark can be removed in T' since the 1 indicates which vertex had been distinguished. This is clearly a bijection whose image is all trees counted by o_n having vertex 1 in the leftmost subtree of the root.

If T has the mark in the rightmost subtree of the root, then start out as before, exchanging the labels l and 1 of the marked vertex and the root, respectively. Now continue by taking the edge connecting l to its rightmost child r and rotating this edge through 90° clockwise. This has the effect of making r the new root of the tree and turning l into its leftmost child. As an illustration, see Figure 5. It is not hard to see that this map gives the rest of the trees counted by o_n.



The bijection with $l = 3$ and $r = 6$

Figure 5

It is possible to turn the above proof into an inductive demonstration by combining it with the mapping $T \overset{c}{\mapsto} T'$. However, there does not seem to be an easy way to describe the result except as a composition of functions. ■

6 Asymptotic results

$J_n^c(0)$. From the formula for F_0^c that we derived in Section 2, we see that the growth of the coefficients is controlled by the singularities of $x = E^{-1}(u)$, the inverse of the exponential integral function. In particular, we see that $x = E^{-1}(u) \to \infty$ as $u \to 0+$. It will be easier to work with $D_t F_0^c$, so combining equations (2), (3) and (4)

$$\begin{aligned} D_t F_0^c(t) &= 1 + \ln \frac{1}{1 - F_0^c(t)} \\ &= E^{-1}(E(1) - t/e) \end{aligned}$$

Thus we want to see what happens as $t \to eE(1)-$.

Integration by parts in the definition of the exponential integral, (1), gives

$$E(x) = \frac{1}{xe^x}(1 + o(1)) \quad \text{as} \quad x \to \infty$$

so

$$x = E^{-1}(u) \sim -\ln u \quad \text{as} \quad u \to 0+$$

and

$$\begin{aligned} E^{-1}(E(1) - t/e) &\sim -\ln(E(1) - t/e) \\ &\sim -\ln(1 - t/eE(1)) \quad \text{as} \quad t \to eE(1)- \end{aligned}$$

But exercise 2, page 147, of De Bruijn's book [9] states the following.

Lemma 6.1 *Suppose*

$$\sum_{k \geq 0} a_k z^k \sim -\ln(1 - z) \quad \textit{as} \quad z \to 1-$$

where

$$a_k > -C \frac{\ln k}{k}$$

for some positive constant C and $k \geq 1$. Then

$$\sum_{k=0}^{n} a_k \sim \ln n \quad \textit{as} \quad n \to \infty \ ■$$

In our case, we have $z = t/eE(1)$ and are approximating the coefficients of

$$\begin{aligned}\sum_{k\geq 0} J^c_{k+1}(0)\frac{t^k}{k!} &= D_t F^c_0(t) \\ &= \sum_{k\geq 0} a_k \frac{t^k}{e^k E(1)^k}\end{aligned}$$

We have already verified the first condition of the lemma and the second is obvious since $a_k \geq 0$ for all k. Thus

$$J^c_{k+1}(0) = \frac{a_k k!}{e^k E(1)^k} \quad \text{where} \quad \sum_{k=0}^{n} a_k \sim \ln n$$

Using the fact that $E(1) = 0.21938\ldots$ and Stirling's approximation, we can also write

$$J^c_{k+1}(0) \sim (ck)^k a_k \sqrt{2\pi k} \quad \text{where} \quad c \approx 0.6169 \quad \text{and} \quad \sum_{k=0}^{n} a_k \sim \ln n. \blacksquare$$

We wish to thank Ed Bender for his help in obtaining the next two results.

$\underline{J^c_n(1).}$ We need the following theorem of Meir and Moon [22].

Theorem 6.2 *Suppose*

$$H(u) = 1 + h_1 u + h_2 u^2 + \cdots$$

is a regular function of u when $|u| < R$ and let

$$G(t) = t + g_2 t^2 + g_3 t^3 + \cdots$$

be the solution of $G(t) = tH(G(t))$ in the neighborhood of $t = 0$. If

1. *$h_i > 0$ for all i, and*
2. *$\nu H'(\nu) = H(\nu)$ for some $0 < \nu < R$*

then $G(t)$ is regular in the disk

$$|t| \leq \rho = \nu/H(\nu)$$

except at $t = \rho$. Furthermore

$$g_n \sim c\rho^{-n} n^{-3/2}$$

where

$$c = \left(\frac{H(\nu)}{2\pi H''(\nu)}\right)^{1/2}. \blacksquare$$

In our case

$$G(t) = G_1^c(t) = \sum_{n\geq 1} \frac{J_n^c(1)}{(n-1)!} t^n$$

and

$$H(u) = 1 + \ln \frac{1}{1-u}.$$

Solving

$$\frac{\nu}{1-\nu} = 1 + \ln \frac{1}{1-\nu}$$

for $0 < \nu < 1$ using the algebra package Maple yields

$$\nu \approx 0.6822$$

Finally, applying Theorem 6.2 together with Stirling's approximation yields

$$\begin{aligned} J_n^c(1) &\sim (n-1)! c \rho^{-n} n^{-3/2} \\ &\sim a b^n n^{n-2} \end{aligned}$$

where

$$a = \left(\frac{H(\nu)}{H''(\nu)}\right)^{1/2} \approx 0.4656 \quad \text{and} \quad b = \frac{H(\nu)}{e\nu} \approx 1.1574. \blacksquare$$

$\underline{J_n^p(-1)}$. We will use a technique which the reader can find outlined in Bender [2, pp. 487-488] or Bender and Williamson [3, p. 386]. Consider

$$J_n^p(-1) = (n-2)! \sum_{k\leq(n-2)/2} \frac{1}{3^k} \binom{n+k-2}{k}.$$

It does no harm to assume that n is even, since the gamma function can be used to continuously define the binomial coefficients for all positive arguments. Now we reverse the order of summation so that the series is weakly decreasing with its maximum term first, i.e., we replace k by $\frac{n-2}{2} - k$ to obtain

$$J_n^p(-1) = \frac{(n-2)!}{3^{(n-2)/2}} \sum_{k\leq(n-2)/2} 3^k \binom{3\left(\frac{n-2}{2}\right) - k}{\frac{n-2}{2} - k}. \tag{23}$$

Let t_k denote the kth term of this sequence. Then

$$\frac{1 - \frac{t_{k+1}}{t_k}}{k} = \frac{2}{3\left(\frac{n-2}{2}\right) - k} \sim \frac{4}{3n}$$

for $n \to \infty$ whenever $k = o(n)$. Since this last fraction is independent of k, we can call it r_n and conclude that the sum in equation (23) is asymptotic to $t_0\sqrt{\pi/2r_n}$. Thus

$$\begin{aligned} J_n^p(-1) &\sim \frac{(n-2)!}{3^{(n-2)/2}}\binom{3\left(\frac{n-2}{2}\right)}{\frac{n-2}{2}}\sqrt{\frac{3\pi n}{8}} \\ &= \frac{1}{3^{(n-2)/2}}3^{3(n-2)/2}\left(\frac{n-2}{2e}\right)^{n-2}\sqrt{\frac{3\pi n}{8}} \end{aligned}$$

by Stirling's approximation. Using the fact that $(n-2)^{n-2} \sim e^{-2}n^{n-2}$ we finally obtain

$$J_n^p(-1) \sim \frac{1}{3}\sqrt{2\pi n}\left(\frac{3}{2e}\right)^n n^{n-2}. \blacksquare$$

7 Open problems

The main result of the paper of Mallows and Riordan [21] is the determination of the generating function for $J_n(q)$ itself. Specifically, they show that

$$\exp\left(\sum_{n\ge1}(q-1)^n J_n(q)t^n/n!\right) = \sum_{n\ge0} q^{\binom{n}{2}}t^n/n!$$

Unfortunately, we have not been able to do the same thing for the other cases. The problem is that, at a crucial step in their proof, they use the fact that $e^{x+y} = e^x e^y$. Since this property is not shared by the other functions in Proposition 2.1, one cannot mimic their demonstration.

From the preceding generating function, one can easily see that

$$J_n(2) = \text{ number of connected graphs on } n \text{ vertices.}$$

It also follows from the work of Robinson [26] that if $N = \deg J_n(q) = \binom{n-1}{2}$ then

$$2^N J_n(\tfrac{1}{2}) = \text{ number of initially connected acyclic digraphs}$$

where a labeled digraph is *initially connected* if there is a directed path from the vertex labeled 1 to any other vertex. Gessel and Wang [13] have given combinatorial proofs of these facts. It is unclear what combinatorial significance, if any, can be attached to the values of our other polynomials for $q = 2$ or $q = \frac{1}{2}$.

Eğecioğlu and Remmel [11] have considered statistics on trees that are related to the major index, maj, of a permutation. For permutations, inv and maj are equidistributed, but there does not seem to be any relation between the two for trees.

William Chen [private communication] has provided answers to a number of the questions that we left open. He has found a formula for $J_n^c(-1)$ using a function similar to the exponential integral. He has also combinatorially explained the product formula for $J_n^o(-1)$ using involutions, bijections and depth first search.

Acknowledgement. The authors would like to thank the referees for comments that improved the exposition of this paper.

References

[1] W. N. Bailey, “Generalized Hypergeometric Series,” Cambridge University Press, Cambridge, 1935.

[2] E. A. Bender, Asymptotic methods in Analysis, *SIAM Review*, **16** (1974), 485–515

[3] E. A. Bender and S. G. Williamson, “Foundations of Applied Combinatorics,” Addison-Wesley, Reading, MA, 1991.

[4] F. Bergeron, P. Flajolet and B. Salvy, Varieties of increasing trees, in “CAAP92, Proceedings,” Lecture Notes in Computer Science, Vol. 581, Springer-Verlag, New York, NY, 24–48.

[5] E. Catalan, Note sur une Équation aux différences finies, *J. de Math.*, **3** (1838), 508–516.

[6] B. Chen and Y.-N. Yeh, Some explanations of Dobinski’s formula, *Studies in Applied Math.*, to appear.

[7] W. Y. C. Chen, “On the Combinatorics of Plethysm,” Ph. D. thesis, M.I.T., Cambridge, MA, 1991.

[8] W. Y. C. Chen, A general bijective algorithm for trees, *Proc. Nat. Acad. Sci. U.S.A.* **87** (1990), 9635–9639.

[9] N. G. De Bruijn, “Asymptotic Methods in Analysis,” Dover Pub., New York, NY, 1981.

[10] H. Dörrie, “100 Great Problems of Elementary Mathematics,” Dover Pub., New York, NY, 1965.

[11] Ö. Eğecioğlu and J. Remmel, Bijections for Cayley trees, spanning trees, and their q-analogues, *J. Combin. Theory Ser. A* **42** (1986), 15–30.

[12] D. Foata, "La Série Génératrice Exponentielle dans les Problémes d'Enumération," Séminaire de Mathématique Supérieurs, No. 54, Presses de l'Université de Montréal, Montréal, 1974

[13] I. M. Gessel and D.-L. Wang, Depth-first search as a combinatorial correspondence, *J. Combin. Theory Ser. A*, **26** (1979), 308–313.

[14] H. W. Gould, "Combinatorial Identities," revised edition, Morgantown Printing and Binding Co., Morgantown, W. Va., 1972.

[15] R. L. Graham, D. E. Knuth and O. Patashnik, "Concrete Mathematics: A Foundation for Computer Science," Addison-Wesley, Reading, MA, 1989.

[16] A. Joyal, Une théorie combinatoire des séries formelles, *Adv. in Math.*, **42** (1981), 1–82.

[17] G. Kreweras, Une famille de polynômes ayant plusieurs propriétés énumératives,*Periodica Math. Hungarica* **11** (1980), 309–320.

[18] G. Labelle, Une nouvelle démonstration combinatoire des formules d'inversion de Lagrange, *Adv. in Math.* **42** (1981), 217–247.

[19] G. Labelle, Counting asymmetric enriched trees, *J. Symbolic Computation* **14** (1992), 211–242.

[20] P. Leroux and G. X. Viennot, Combinatorial resolution of systems of differential equations, I. Ordinary differential equations, in "Combinatoire Enumérative, Montréal, 1985, Proceedings," Lecture Notes in Math., Vol. 1234, Springer-Verlag, New York, NY, 1986, 210–245.

[21] C. L. Mallows and J. Riordan, The inversion enumerator for labeled trees, *Bull. Amer. Math. Soc.*, **74** (1968), 92–94.

[22] A. Meir and J. W. Moon, On the altitude of nodes in random trees, *Canad. J. Math.*, **5** (1978), 997–1015.

[23] J. W. Moon, "Counting Labeled Trees," Canad. Math. Monographs, Vol. 1, Canad. Math. Congress, Montréal, 1970.

[24] J.-J. Pansiot, Nombres d'Euler et inversions dans les arbes, *European J. Combin.* **3** (1982), 259–262.

[25] M. Petkovšek, Hypergeometric solutions of linear recurrences with polynomial coefficients, *J. Symbolic Computation*, to appear.

[26] R. W. Robinson, Counting labeled acyclic digraphs, in "New Directions in the Theory of Graphs," F. Harary, ed., Academic Press, New York, NY, 1973.

[27] M. O. Rodrigues, Sur le nombre de manières d'effectuer un produit de n facteurs, *J. de Math.*, **3** (1838), 549.

[28] L. J. Slater, "Generalized Hypergeometric Functions," Cambridge University Press, Cambridge, 1966.

[29] H. S. Wilf, "Generatingfunctionology," Academic Press, Boston, MA, 1990.

[30] Y.-N. Yeh, The calculus of virtual species and K-species, in "Combinatoire Enumérative, Montréal, 1985, Proceedings," G. Labelle and P. Leroux eds., Lecture Notes in Math., Vol. 1234, Springer-Verlag, New York, NY, 1986, 210–245.

[31] D. Zeilberger, The method of creative telescoping, *J. Symbolic Computation*, **11** (1991), 195-204.