A remark on the enumeration of rooted labeled trees

Alan D. Sokal

Department of Mathematics
University College London
Gower Street
London WC1E 6BT
UNITED KINGDOM
sokal@math.ucl.ac.uk

Department of Physics
New York University
726 Broadway
New York, NY 10003
USA
sokal@nyu.edu

October 31, 2019

Abstract

Two decades ago, Chauve, Dulucq and Guibert showed that the number of rooted trees on the vertex set $[n+1]$ in which exactly k children of the root are lower-numbered than the root is $\binom{n}{k}\, n^{n-k}$. Here I give a simpler proof of this result.

Key Words: Trees, rooted trees, labeled trees, enumeration, Chu–Vandermonde identity, Abel identity.

Mathematics Subject Classification (MSC 2010) codes: 05A15 (Primary); 05A10, 05A19, 05C05 (Secondary).

It is well known that the set $\mathcal{T}_{n+1}$ of rooted trees on the vertex set $[n+1] \stackrel{\text{def}}{=} \{1, \ldots, n+1\}$ has cardinality $(n+1)^n$; and from the binomial theorem we have the obvious identity

$$(n+1)^n \;=\; \sum_{k=0}^{n} \binom{n}{k} n^{n-k} \,. \tag{1}$$

So it is natural to seek a combinatorial explanation of this identity: Can we find a partition of $\mathcal{T}_{n+1}$ into subsets $\mathcal{T}_{n+1,k}$ $(0 \le k \le n)$ such that $|\mathcal{T}_{n+1,k}| = \binom{n}{k} n^{n-k}$?

A solution to this problem was found two decades ago by Chauve, Dulucq and Guibert [4, 5]: they showed that the number of rooted trees on the vertex set $[n+1]$ in which exactly k children of the root are lower-numbered than the root is $\binom{n}{k} n^{n-k}$ [16, A071207]. Their proof was bijective but rather complicated.[1] Here I would like to give a simpler proof.

Let $T(n; i, k, \ell, m)$ be the number of rooted trees on the vertex set $[n+1]$ in which the root is i, the root has k children $< i$ and ℓ children $> i$, and the forest whose roots are the children $< i$ (resp. $> i$) has m (resp. $n - m$) vertices. We can obtain an explicit formula for $T(n; i, k, \ell, m)$ as follows: Given $i \in [n+1]$, we choose the k children $< i$ in $\binom{i-1}{k}$ ways, and the ℓ children $> i$ in $\binom{n+1-i}{\ell}$ ways. Then we choose $m - k$ additional vertices for the first forest from the remaining $n - k - \ell$ vertices, in $\binom{n-k-\ell}{m-k}$ ways. This also fixes the $n - m - \ell$ additional vertices for the second forest. And finally, we recall [23, Proposition 5.3.2] that the number of forests on m total vertices with k $(\le m)$ fixed roots is

$$\phi_{m,k} \;=\; \begin{cases} 1 & \text{if } m = 0 \text{ (and hence } k = 0) \\ k\, m^{m-k-1} & \text{if } m \ge 1 \text{ (and } 0 \le k \le m) \end{cases} \tag{2}$$

[16, A232006]. In the same way, the number of forests on $n - m$ total vertices with ℓ $(\le n - m)$ fixed roots is $\phi_{n-m,\ell}$. It follows that

$$T(n; i, k, \ell, m) \;=\; \binom{i-1}{k} \binom{n+1-i}{\ell} \binom{n-k-\ell}{m-k} \phi_{m,k}\, \phi_{n-m,\ell} \,. \tag{3}$$

This is defined for $n \ge 0$, $1 \le i \le n+1$, $0 \le k \le n$, $0 \le \ell \le n-k$ and $k \le m \le n-\ell$. For $n = 0$ the only term is $T(0; 1, 0, 0, 0) = 1$, so we can assume henceforth that $n \ge 1$.

We now proceed to sum (3) over i and m. Note that i appears only in the first two factors on the right-hand side of (3), while m appears only in the final three factors. So we can perform these two sums separately.

Sum over *i*. We claim that for any integers $n, k, \ell \ge 0$, we have

$$\sum_{i=1}^{n+1} \binom{i-1}{k} \binom{n+1-i}{\ell} \;=\; \binom{n+1}{k+\ell+1} \,. \tag{4}$$

[1] In [4, section 3], the same authors also gave a simple algebraic proof of the special case $k = 0$, based on exponential generating functions and the Lagrange inversion formula.

This identity has a simple combinatorial proof: the right-hand side is the number of ways of choosing $k+\ell+1$ elements from the set $[n+1]$; if we arrange these elements in increasing order and call the $(k+1)$st of them i, then the two binomial coefficients on the left-hand side give the number of ways of choosing the first k elements and the last ℓ elements, respectively. The identity (4) can also be derived algebraically as a corollary of the Chu–Vandermonde identity; we discuss this in Appendix A.1.

From the right-hand side, we see in particular that (4) depends on k and ℓ only via their sum.

Sum over m. We claim that for any integers $n,k,\ell \geq 0$ with $k+\ell \leq n$, we have

$$\sum_{m=k}^{n-\ell} \binom{n-k-\ell}{m-k} \phi_{m,k}\, \phi_{n-m,\ell} \;=\; \phi_{n,k+\ell}\,. \tag{5}$$

This identity too has a simple combinatorial proof: the right-hand side counts the forests on the vertex set $[n]$ with $k+\ell$ fixed roots, while the left-hand side partitions this count according to the number m of vertices that belong to the subforest associated to the first k roots. The identity (5) can also be derived algebraically as a corollary of an Abel identity; we discuss this in Appendix A.2.

From the right-hand side, we see in particular that (5) depends on k and ℓ only via their sum.

Combining the two sums. Combining (3) with (4) and (5), we have for $n \geq 1$

$$\sum_{i=1}^{n+1} T(n;i,k,\ell,m) \;=\; \binom{n+1}{k+\ell+1}\binom{n-k-\ell}{m-k} \phi_{m,k}\, \phi_{n-m,\ell} \tag{6}$$

$$\sum_{m=k}^{n-\ell} T(n;i,k,\ell,m) \;=\; \binom{i-1}{k}\binom{n+1-i}{\ell} (k+\ell)\, n^{n-k-\ell-1} \tag{7}$$

$$\sum_{i=1}^{n+1}\sum_{m=k}^{n-\ell} T(n;i,k,\ell,m) \;=\; \binom{n+1}{k+\ell+1} (k+\ell)\, n^{n-k-\ell-1} \tag{8}$$

The right-hand side of (8) depends on k and ℓ only via their sum; we denote this quantity by $g_n(k+\ell)$, i.e. we define

$$g_n(K) \;\stackrel{\text{def}}{=}\; \binom{n+1}{K+1} K\, n^{n-K-1} \quad \text{for } n \geq 1 \text{ and } 0 \leq K \leq n\,. \tag{9}$$

Sum over ℓ. The final step is to sum (8) over ℓ at fixed k, i.e. to compute

$$G_n(k) \;\stackrel{\text{def}}{=}\; \sum_{\ell=0}^{n-k} g_n(k+\ell) \;=\; \sum_{K=k}^{n} g_n(K)\,. \tag{10}$$

We prove that $G_n(k) = \binom{n}{k} n^{n-k}$, as follows: From (10), $G_n(k)$ manifestly satisfies the backward recurrence

$$G_n(k) \;=\; G_n(k+1) \,+\, \binom{n+1}{k+1} k\, n^{n-k-1} \tag{11}$$

with initial condition $G_n(n) = 1$. A simple calculation shows that $\widehat{G}_n(k) = \binom{n}{k} n^{n-k}$ satisfies the same recurrence and the same initial condition. Hence $G_n(k) = \widehat{G}_n(k)$. QED

Xi Chen (private communication) has found an alternate proof of $G_n(k) = \binom{n}{k} n^{n-k}$ that *derives* it (rather than simply pulling it out of a hat, as the foregoing proof does); this proof is presented in Appendix A.3.

Three final remarks.

1. The special case $k = 0$ of (8) was found by Chauve *et al.* [5, Proposition 2].

2. By summing (7) over ℓ, we can compute the number of rooted trees in $\mathcal{T}_{n+1,k}$ that have a specified element i as the root. This sum is easily performed using the binomial theorem and its derivative, and gives

$$\sum_{\ell=0}^{n+1-i} \sum_{m=k}^{n-\ell} T(n; i, k, \ell, m) \;=\; \binom{i-1}{k} \big[(k+1)(n+1) - i\big]\, n^{i-k-2}\, (n+1)^{n-i} \,. \tag{12}$$

For the special case $k = 0$, this result was obtained bijectively by Chauve *et al.* [5, proof of Proposition 1].

3. We can also compute the number of rooted trees on $n+1$ labeled vertices in which the root has exactly K children: it suffices to sum (8) over $k, \ell \geq 0$ with $k + \ell = K$, yielding

$$(K+1) \binom{n+1}{K+1} K\, n^{n-K-1} \;=\; (n+1) \binom{n}{K} K\, n^{n-K-1} \,. \tag{13}$$

Here $n+1$ counts the number of choices for the root, and the remaining factor $f_{n,k} = \binom{n}{K} K\, n^{n-K-1} = \binom{n}{K} \phi_{n,K}$ counts the number of K-component forests of rooted trees on n labeled vertices. This latter result is essentially equivalent to (2), and is well known.[2]

Appendix: Algebraic proofs

A.1 A corollary of the Chu–Vandermonde identity

The identity (4) is a special case of a slightly more general binomial identity, namely

$$\sum_{j=k-m}^{n-\ell} \binom{m+j}{k} \binom{n-j}{\ell} \;=\; \binom{m+n+1}{k+\ell+1} \,, \tag{A.1}$$

[2] See e.g. [6], [14, pp. 26–27], [7, p. 70], [23, pp. 25–28] or [2]. See also [12,19,22,24] and [1, pp. 235–240] for related information.

valid for integers k, ℓ, m, n with $k, \ell \geq 0$ and $m + n \geq -1$. Although this identity can be found in several places in the literature[3], I have been unable to find any place where it is stated clearly with its optimal conditions of validity. I will therefore give here a detailed derivation, keeping careful track of the conditions of validity for each step.

The binomial coefficients are defined as usual by [11, p. 154]

$$\binom{r}{k} = \begin{cases} \dfrac{r(r-1)\cdots(r-k+1)}{k!} & \text{for integer } k \geq 0 \\ 0 & \text{for integer } k < 0 \\ \text{undefined} & \text{if } k \text{ is not an integer} \end{cases} \qquad \text{(A.2)}$$

Here r can be any element of any commutative ring containing the rationals; in particular, it can be an indeterminate in a ring of polynomials over the rationals. The binomial coefficients satisfy

$$\binom{r}{k} = (-1)^k \binom{-(r-k+1)}{k} \qquad \text{for integer } k \qquad \text{(A.3)}$$

("upper negation") and

$$\binom{n}{k} = \binom{n}{n-k} \qquad \text{for integer } n \geq 0 \text{ and integer } k \qquad \text{(A.4)}$$

("symmetry"). Finally, they satisfy the *Chu–Vandermonde identity*

$$\sum_{j=0}^{N} \binom{x}{j}\binom{y}{N-j} = \binom{x+y}{N} \qquad \text{for integer } N\,, \qquad \text{(A.5)}$$

where x and y can be indeterminates. Applying (A.3) to all three binomial coefficients in the Chu–Vandermonde identity and then replacing $x \to -x$, $y \to -y$, we obtain the *dual Chu–Vandermonde identity*

$$\sum_{j=0}^{N} \binom{x+j-1}{j}\binom{y+N-j-1}{N-j} = \binom{x+y+N-1}{N} \qquad \text{for integer } N\,. \qquad \text{(A.6)}$$

Now suppose that x, y are integers ≥ 1 and that $x+y+N \geq 1$; then we can apply the symmetry (A.4) to the three binomial coefficients in (A.6). Writing $x = k+1$ and $y = \ell + 1$ with integers $k, \ell \geq 0$, we have

$$\sum_{j=0}^{N} \binom{k+j}{k}\binom{N+\ell-j}{\ell} = \binom{k+\ell+N+1}{k+\ell+1}$$
$$\text{for integers } k, \ell, N \text{ with } k, \ell \geq 0 \text{ and } k+\ell+N \geq -1\,. \qquad \text{(A.7)}$$

[3] See e.g. [10, p. 22, eq. (3.3)] and [11, p. 169, eq. (5.26) and pp. 243, 527, Exercise 5.14].

Now change variables $j = j' + m - k$ and $N = m + n - k - \ell$:

$$\sum_{j'=k-m}^{n-\ell} \binom{m+j'}{k} \binom{n-j'}{\ell} \;=\; \binom{m+n+1}{k+\ell+1}$$

$$\text{for integers } k, \ell, m, n \text{ with } k, \ell \geq 0 \text{ and } m+n \geq -1 \;. \tag{A.8}$$

Dropping primes, this is (A.1).

A.2 Abel identity

The identity (5) can also be derived algebraically, as follows: We begin from the well-known Abel identity [20, p. 73]

$$\sum_{M=0}^{N} \binom{N}{M} x(x+M)^{M-1} \, y(y+N-M)^{N-M-1} \;=\; (x+y)\,(x+y+N)^{N-1} \tag{A.9}$$

(see also [18, p. 20, eq. (20)] multiplied by xy).[4] Since all the terms in this identity (even the ones with $M = 0$ and $M = N$) are polynomials in x and y, the variables x and y can be specialized without restriction. (Note, however, that in applying this identity, we must first fix N and M and then specialize x and y.) Setting $N = n-k-\ell$ and changing variables by $M = m - k$ yields

$$\sum_{m=k}^{n-\ell} \binom{n-k-\ell}{m-k} x(x+m-k)^{m-k-1} \, y(y+n-m-\ell)^{n-m-\ell-1}$$

$$=\; (x+y)\,(x+y+n-k-\ell)^{n-k-\ell-1} \;. \tag{A.10}$$

Specializing now to $x = k$ and $y = \ell$, we see that $x(x+m-k)^{m-k-1}\big|_{x=k} = \phi_{m,k}$ even when $m = k = 0$, and likewise $y(y+n-m-\ell)^{n-m-\ell-1}\big|_{y=\ell} = \phi_{n-m,\ell}$ even when $n - m = \ell = 0$. It follows that

$$\sum_{m=k}^{n-\ell} \binom{n-k-\ell}{m-k} \phi_{m,k}\, \phi_{n-m,\ell} \;=\; (k+\ell)\, n^{n-k-\ell-1} \;=\; \phi_{n,k+\ell} \;, \tag{A.11}$$

valid for $n \geq 1$ and $k, \ell \geq 0$ with $k + \ell \leq n$.

We remark, finally, that many Abel identities, including (A.9), can be proven combinatorially: see e.g. [8, 17, 21].

---

[4] The identity (A.9) asserts that the polynomials $P_N(x) = x(x+N)^{N-1}$, which are a specialization of the celebrated *Abel polynomials* $A_n(x; a) = x(x - an)^{n-1}$ [8, 15, 20, 21] to $a = -1$, form a *sequence of binomial type* [9, 15, 20]. See also [13] [3, Section 3.1] for a purely combinatorial approach to sequences of binomial type, employing the theory of species.

A.3 Alternate proof of $G_n(k) = \binom{n}{k} n^{n-k}$ (due to Xi Chen)

We compute the row-generating polynomials $\mathcal{G}_n(x) \stackrel{\rm def}{=} \sum\limits_{k=0}^{n} G_n(k)\, x^k$, as follows:

$$\begin{aligned}
\mathcal{G}_n(x) &= \sum_{k=0}^{n} \sum_{K=k}^{n} \binom{n+1}{K+1} K\, n^{n-K-1}\, x^k && \text{(A.12a)}\\
&= n^{n-1} \sum_{K=0}^{n} \binom{n+1}{K+1} K\, n^{-K} \sum_{k=0}^{K} x^k && \text{(A.12b)}\\
&= n^{n-1} \sum_{K=0}^{n} \binom{n+1}{K+1} K\, n^{-K}\, \frac{1-x^{K+1}}{1-x} && \text{(A.12c)}\\
&= \frac{n^{n-1}}{1-x} \left[\sum_{K=0}^{n} \binom{n+1}{K+1} K\, \frac{1}{n^K} \;-\; x \sum_{K=0}^{n} \binom{n+1}{K+1} K\, \frac{x^K}{n^K} \right] && \text{(A.12d)}\\
&= \frac{n^{n-1}}{1-x} \big[\mathcal{F}_n(1/n) \;-\; x\, \mathcal{F}_n(x/n) \big] && \text{(A.12e)}
\end{aligned}$$

where

$$\mathcal{F}_n(x) \;\stackrel{\rm def}{=}\; \sum_{K=0}^{n} \binom{n+1}{K+1} K\, x^K \;. \tag{A.13}$$

A simple computation, using the derivative of the binomial theorem, shows that

$$\mathcal{F}_n(x) \;=\; (n+1)(x+1)^n \;-\; \frac{(x+1)^{n+1}-1}{x} \;. \tag{A.14}$$

Therefore

$$\mathcal{F}_n(1/n) \;=\; n \qquad \text{and} \qquad x\, \mathcal{F}_n(x/n) \;=\; \frac{1}{n^{n-1}} (x-1)(x+n)^n \;+\; n \;, \tag{A.15}$$

and inserting these into (A.12) gives

$$\mathcal{G}_n(x) \;=\; (x+n)^n \;. \tag{A.16}$$

Taking the coefficient of x^k in $\mathcal{G}_n(x)$, we conclude that $G_n(k) = \binom{n}{k} n^{n-k}$.

Acknowledgments

I wish to thank Xi Chen and Alex Dyachenko for helpful conversations. I am especially grateful to Xi Chen for giving me permission to include her proof of $G_n(k) = \binom{n}{k} n^{n-k}$ in Appendix A.3.

This work was immeasurably facilitated by the On-Line Encyclopedia of Integer Sequences [16]. I warmly thank Neil Sloane for founding this indispensable resource, and the hundreds of volunteers for helping to maintain and expand it.

This research was supported in part by Engineering and Physical Sciences Research Council grant EP/N025636/1.

References

[1] M. Aigner and G.M. Ziegler, *Proofs from The Book*, 6th ed. (Springer-Verlag, Berlin, 2018).

[2] A. Avron and N. Dershowitz, Cayley's formula: A page from The Book, Amer. Math. Monthly **123**, 699–700 (2016).

[3] F. Bergeron, G. Labelle and P. Leroux, *Combinatorial Species and Tree-Like Structures* (Cambridge University Press, Cambridge–New York, 1998).

[4] C. Chauve, S. Dulucq and O. Guibert, Enumeration of some labelled trees, Research Report RR-1226-99, LaBRI, Université Bordeaux I (1999).

[5] C. Chauve, S. Dulucq and O. Guibert, Enumeration of some labelled trees, in *Formal Power Series and Algebraic Combinatorics* (FPSAC'00, Moscow, June 2000), edited by D. Krob, A.A. Mikhalev and A.V. Mikhalev (Springer-Verlag, Berlin, 2000), pp. 146–157.

[6] L.E. Clarke, On Cayley's formula for counting trees, J. London Math. Soc. **33**, 471–474 (1958).

[7] L. Comtet, *Advanced Combinatorics: The Art of Finite and Infinite Expansions* (Reidel, Dordrecht–Boston, 1974). [French original: *Analyse Combinatoire*, tomes I et II, Presses Universitaires de France, Paris, 1970.]

[8] J. Françon, Preuves combinatoires des identités d'Abel, Discrete Math. **8**, 331–343 (1974); corrigendum **10**, 193 (1974).

[9] A.M. Garsia, An exposé of the Mullin–Rota theory of polynomials of binomial type, Linear and Multilinear Algebra **1**, 47–65 (1973).

[10] H.W. Gould, *Combinatorial Identities* (Henry W. Gould, Morgantown, W.Va., 1972).

[11] R.L. Graham, D.E. Knuth and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, 2nd ed. (Addison-Wesley, Reading, Mass., 1994).

[12] S. Guo and V.J.W. Guo, A recursive algorithm for trees and forests, Discrete Math. **340**, 695–703 (2017).

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

[14] J.W. Moon, *Counting Labelled Trees* (Canadian Mathematical Congress, Montreal, 1970).

[15] R. Mullin and G.-C. Rota, On the foundations of combinatorial theory. III. Theory of binomial enumeration, in *Graph Theory and Its Applications*, edited by B. Harris (Academic Press, New York–London, 1970), pp. 167–213.

[16] The On-Line Encyclopedia of Integer Sequences, published electronically at `http://oeis.org`

[17] J. Pitman, Forest volume decompositions and Abel–Cayley–Hurwitz multinomial expansions, J. Combin. Theory A **98**, 175–191 (2002).

[18] J. Riordan, *Combinatorial Identities* (Wiley, New York, 1968). [Reprinted with corrections by Robert E. Krieger Publishing Co., Huntington NY, 1979.]

[19] J. Riordan, Forests of labeled trees, J. Combin. Theory **5**, 90–103 (1968).

[20] S. Roman, *The Umbral Calculus* (Academic Press, New York, 1984).

[21] B.E. Sagan, A note on Abel polynomials and rooted labeled forests, Discrete Math. **44**, 293–298 (1983).

[22] P.W. Shor, A new proof of Cayley's formula for counting labeled trees, J. Combin. Theory A **71**, 154–158 (1995).

[23] R.P. Stanley, *Enumerative Combinatorics*, vol. 2 (Cambridge University Press, Cambridge–New York, 1999).

[24] J. Zeng, A Ramanujan sequence that refines the Cayley formula for trees, Ramanujan J. **3**, 45–54 (1999).