

The number of labeled trees: This proof of Cayley's formula emerged out of conversations I had with my physics colleague at Brown, Dmitri Feldman. In particular, Dmitri supplied the proof of the identity below.

An Identity: Let $0 \leq \ell < n$ be integers. (We care about the case $\ell = n - 2$.) We have

$$\sum_{k=0}^n (-1)^k \binom{n}{k} (n-k)^\ell = \sum_{k=0}^n (-1)^k \binom{n}{k} k^\ell = 0. \quad (1)$$

Proof: Up to sign, the two expressions are equal, so it suffices to establish the second one. By the Binomial Theorem,

$$(1 - e^x)^n = \sum_{k=0}^n (-1)^k \binom{n}{k} e^{kx}. \quad (2)$$

Using the series expansion $1 - e^x = -x + \dots$ we see that the lowest order term on the LHS of Equation 2 is x^n . Hence the (ℓ) th derivative of the LHS, evaluated at 0, is 0. But the (ℓ) th derivative of the RHS, evaluated at 0, is exactly the second sum in Equation 1.

Proof of Cayley's Formula: We show that the number T_n of labeled trees with n vertices is n^{n-2} . This holds for $n = 1, 2$. We pick $n \geq 3$ and suppose by induction that $T_{n-k} = (n-k)^{n-k-2}$ when $k \geq 1$.

How big is the set $S(i_1, \dots, i_k)$ of trees which have labels $i_1 < \dots < i_k$ as leaves? If we omit the edges going to these labels we have T_{n-k} trees. We can return the omitted edges by sticking them on independently: $(n-k)^k$ ways. Hence

$$|S(i_1, \dots, i_k)| = T_{n-k} \times (n-k)^k = (n-k)^{n-2}. \quad (3)$$

By the Inclusion-Exclusion Principle and Equation 3

$$\begin{aligned} T_n &= \sum_i |S(i)| - \sum_{i < j} |S(i, j)| + \sum_{i < j < k} |S(i, j, k)| - \dots = \\ &= \binom{n}{1} (n-1)^{n-2} - \binom{n}{2} (n-2)^{n-2} + \binom{n}{3} (n-3)^{n-2} - \dots \end{aligned} \quad (4)$$

This equals n^{n-2} by Equation 1. We're done.