Entry-0012

Why n^{n-2}? Eight Perspectives on Cayley's Tree Formula

Area: Graph Theory • **Subarea:** Enumeration of Trees

Difficulty: UG-Upper • **Published:** 2026-11-26 • **Updated:**

Abstract

Cayley's remarkable formula states that there are n^{n-2} trees on n labeled vertices. We present eight perspectives on this theorem, drawing upon bijections, double counting, linear algebra, probability, recursion, symmetry, and generating polynomials. Several of these perspectives lead to more general enumeration theorems for rooted forests and trees with prescribed or restricted vertex degrees. A concluding synthesis reveals connections among the eight proofs, showing how results obtained from one perspective can be recovered from another. The essay illustrates the mathematical discoveries made possible by approaching a single theorem from different perspectives.

Introduction

Let $\mathcal{T}_n$ denote the set of trees whose vertices are labeled $1, 2, \ldots, n$, where $n \geq 2$.

Cayley's formula.

$$\boxed{|\mathcal{T}_n| = n^{n-2}.}$$

The simplicity of this formula is remarkable. We are counting objects whose defining conditions—connectedness and absence of cycles—impose global restrictions on their structure. Yet the answer is the same as the number of words of length $n-2$ over an alphabet of n letters, whose entries can be chosen independently.

Why should this be so?

Prüfer's celebrated proof answers this question by constructing a bijection between labeled trees and such words. Once we have understood that bijection, we have a complete proof of Cayley's formula. But have we exhausted the mathematical ideas behind the formula?

A theorem may admit several proofs, each illuminating something that the others leave in the background. A different proof may reveal a more general theorem, an unexpected connection with another branch of mathematics, or a technique whose usefulness extends far beyond the theorem under consideration. Cayley's formula is an exceptional illustration of this phenomenon.

Eight perspectives

I. **Prüfer — A bijection between trees and words.** Successive leaf deletions encode a tree as a word; a reverse construction recovers the tree.

II. **Pitman — Counting forest histories.** Count the histories of joining components, both forward and backward.

III. **Kirchhoff — Counting trees with determinants.** A determinant distinguishes spanning trees from other selections of edges.

IV. **Joyal — From functions to trees.** Transform the cyclic part of a function into a recoverable path.

V. **Chapuy–Perarnau — A telescoping probability.** Explore a random function in rounds and watch the conditional probabilities telescope.

VI. **Riordan–Rényi — Generalize before recursing.** Enlarge the family of trees to forests so that deletion yields a useful recurrence.

VII. **Cayley–Rényi — A polynomial determined by symmetry.** A specialization, symmetry and low degree determine a polynomial recording the vertex degrees.

VIII. **Child positions and generating polynomials.** Fill ordered child positions and then forget the ordering to explain factorial weights in tree enumeration.

Each perspective can be read independently. Readers are encouraged to pause at the central idea of each proof and explore how far that idea might take them before proceeding to the details. Optional PZoom assistance in the HTML edition invites discovery without being necessary for the main argument.

Reading conventions. Throughout, $[n] = \{1, \ldots, n\}$. The eight perspectives are numbered with Roman numerals; sections, equations and figures within each perspective retain Arabic numbering. Rooted trees have a designated root; when directions are needed, Perspectives II and VIII use arcs away from the root, whereas functional digraphs in Perspectives IV and V point toward their cycles. These are different conventions for the same underlying rooted-tree structure.

Our objective is not simply to prove Cayley's formula eight times. It is to discover what mathematics becomes visible when we look at the same theorem in eight different ways.

I. Prüfer sequences: a bijective proof

Theorem (Cayley–Prüfer). *For $n \geq 2$, the number of labeled trees on $[n] = \{1, \ldots, n\}$ is*

$$\boxed{|\mathcal{T}_n| = n^{n-2}}.$$

The proof constructs *two explicit maps* and verifies both inverse identities. Write

$$\mathcal{T}(V) = \{\text{trees with vertex set } V\}, \qquad \mathcal{W}(V) = V^{|V|-2} \quad (|V| \geq 2).$$

Here V is any finite set of distinct integers, ordered by their usual order; put $\mathcal{T}_n = \mathcal{T}([n])$.

Encoding $f_V : \mathcal{T}(V) \to \mathcal{W}(V)$. Repeatedly delete the *smallest-labeled leaf*, recording its unique neighbor. Stop when two vertices remain. The sequence of $|V| - 2$ recorded labels forms $f_V(T)$.

For a word $P = (a_1, \ldots, a_r)$, write $\operatorname{supp}(P) = \{a_1, \ldots, a_r\}$, with $\operatorname{supp}(()) = \varnothing$.

Decoding $g_V : \mathcal{W}(V) \to \mathcal{T}(V)$. For $P = (a_1, \ldots, a_{m-2})$, set $V_1 = V$, $E_0 = \varnothing$, where $m = |V|$. For $i = 1, \ldots, m-2$, define

$$\boxed{\begin{aligned} P_i &= (a_i, \ldots, a_{m-2}), \\ v_i &= \min(V_i \setminus \operatorname{supp}(P_i)), \\ E_i &= E_{i-1} \cup \{\{v_i, a_i\}\}, \\ V_{i+1} &= V_i \setminus \{v_i\}. \end{aligned}}$$

Join the last two vertices; write $g_V(P)$ for the resulting graph. At every step,

$$\operatorname{supp}(P_i) \subseteq V_i, \qquad |P_i| = |V_i| - 2 \quad \Longrightarrow \quad |V_i \setminus \operatorname{supp}(P_i)| \geq 2.$$

Thus v_i exists and $a_i \in V_i \setminus \{v_i\}$. Each selected vertex is joined to a vertex that survives longer; the final pair is joined. Hence the graph is connected, with $m - 1$ edges, so $g_V(P) \in \mathcal{T}(V)$.

Degree lemma (encoding). For every $T \in \mathcal{T}(V)$, $x \in V$,

$$\boxed{\deg_T(x) = 1 + \operatorname{occ}_x(f_V(T))}, \qquad \operatorname{occ}_x(P) := |\{i : a_i = x\}|. \tag{1}$$

Every deletion of a neighbor of x records x and lowers its current degree by one. The process stops involving x when its degree is one (either x is deleted or it belongs to the final pair). Thus the number of recordings is $\deg_T(x) - 1$.

Degree lemma (decoding). For every $P \in \mathcal{W}(V)$ and $x \in V$,

$$\boxed{\deg_{g_V(P)}(x) = 1 + \operatorname{occ}_x(P).} \tag{2}$$

Indeed, each occurrence of $x = a_i$ creates one edge incident to x. One additional incident edge is created when $x = v_i$, or when x belongs to the final pair. Since a vertex is removed at most once and cannot be used after its removal, these edges are distinct and account for all edges incident to x.

The two algorithms, viewed step by step

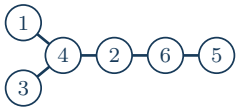
Our example is the tree T on $[6]$ with edge set

$$E(T) = \{14, 34, 42, 26, 65\}, \qquad \text{where } ij := \{i, j\}.$$

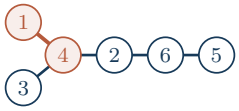
The two diagram sequences are mirror-arranged: read encoding $\longrightarrow$ across each row, decoding $\longleftarrow$ across each row (then continue on the next row). At the upper left and lower left respectively, both display the same completed tree T. Vertex positions are fixed; orange marks the edge being removed or added. Panel numbers correspond to the tables.

Encoding $\longrightarrow$: erase the smallest leaf

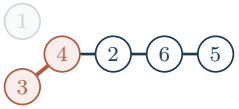
Start: the full tree T



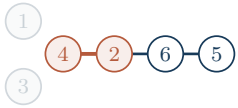
Step 1: delete 1, record 4



Step 2: delete 3, record 4



Step 3: delete 4, record 2



Step 4: delete 2, record 6



Finish: leave 5–6



Step i	Deleted leaf v_i	Recorded a_i	Word recorded so far
1	1	4	(4)
2	3	4	$(4, 4)$
3	4	2	$(4, 4, 2)$
4	2	6	$(4, 4, 2, 6)$

$$\boxed{f_{[6]}(T) = (4, 4, 2, 6).}$$

Decoding $\longleftarrow$: insert the edges in the same order

Step 2: add 34



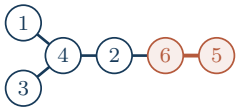
Step 1: add 14



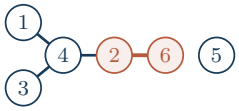
Start: no edges



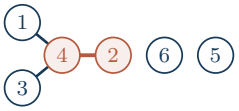
Finish: add 56



Step 4: add 26



Step 3: add 42



Step	Current word (CW)	Labels absent from CW	Chosen v_i	Edge added
1	$(4, 4, 2, 6)$	$1, 3, 5$	1	14
2	$(4, 2, 6)$	$3, 5$	3	34
3	$(2, 6)$	$4, 5$	4	42
4	(6)	$2, 5$	2	26
Final	$()$	$5, 6$	—	56

The two inverse identities

Equations (1) and (2) give the common mechanism:

$$\boxed{\begin{aligned}\mathrm{Leaves}(T) &= V \setminus \mathrm{supp}(f_V(T)),\\ \mathrm{Leaves}(g_V(P)) &= V \setminus \mathrm{supp}(P).\end{aligned}} \tag{3}$$

We prove both inverse identities for finite ordered V by induction on $m = |V|$. For $m = 2$, both maps associate the unique edge and the empty word.

Left inverse: $g_V \circ f_V = \mathrm{id}_{\mathcal{T}(V)}$**.** Let $T \in \mathcal{T}(V)$, $P = f_V(T) = (a_1, P')$. By (3),

$$v = \min \mathrm{Leaves}(T) = \min(V \setminus \mathrm{supp}(P)).$$

Here a_1 is the neighbor of v. Deleting v produces the tree $T - v$ on $V' = V \setminus \{v\}$, and

$$f_V(T) = (a_1, f_{V'}(T - v)).$$

The first decoding step reconstructs $\{v, a_1\}$; the rest is decoding on V'. Thus, using the induction hypothesis,

$$\begin{aligned}E\big(g_V(f_V(T))\big) &= \{\{v, a_1\}\} \cup E\big(g_{V'}(f_{V'}(T - v))\big)\\ &= \{\{v, a_1\}\} \cup E(T - v) = E(T).\end{aligned} \tag{4}$$

Both trees have vertex set V; hence $g_V(f_V(T)) = T$.

Right inverse: $f_V \circ g_V = \mathrm{id}_{\mathcal{W}(V)}$**.** Let $P = (a_1, P') \in \mathcal{W}(V)$, $T = g_V(P)$, and $v = \min(V \setminus \mathrm{supp}(P))$. By (3), the smallest leaf of T is v, and its neighbor is a_1. Consequently, the first encoding step records a_1 and deletes v. The remaining tree is $g_{V'}(P')$, where $V' = V \setminus \{v\}$. By induction,

$$\begin{aligned}f_V(g_V(P)) &= (a_1, f_{V'}(g_{V'}(P')))\\ &= (a_1, P') = P.\end{aligned} \tag{5}$$

Equations (4) and (5) establish

$$\boxed{g_V \circ f_V = \mathrm{id}_{\mathcal{T}(V)}, \qquad f_V \circ g_V = \mathrm{id}_{\mathcal{W}(V)}\,.}$$

Hence $f_{[n]}$ and $g_{[n]}$ are inverse bijections; therefore

$$\boxed{|\mathcal{T}_n| = |[n]^{n-2}| = n^{n-2}.} \qquad \square$$

II. Pitman: grow a rooted forest

Central idea. Rather than count trees directly, count the *histories* by which one rooted forest grows into a rooted tree. Every terminal tree turns out to have the same number of histories.

Double counting. A fundamental method of combinatorial enumeration is to count the same finite set of objects in two different ways and equate the results. Pitman's proof illustrates this method by counting construction histories in two ways: first by successively adding arcs to a given rooted forest, and then by deleting arcs from rooted spanning trees to recover that forest.

Theorem (Pitman's rooted-forest extension identity). Let F_k be a rooted forest on $[n] = \{1, \dots, n\}$ with k components. Let $N(F_k)$ be the number of rooted spanning trees on $[n]$ that extend F_k, preserving its arcs. Then

$$\boxed{N(F_k) = n^{k-1}.}$$

In particular, the number of unrooted labeled trees is n^{n-2} for $n \geq 2$.

Convention. Orient every edge *away from its component root.* Hence a root has indegree 0 and every other vertex has indegree 1. For any rooted forest G, let $A(G)$ be its set of arcs (directed edges), and define

$$\mathcal{R}(F_k) := \{R : \ R \text{ is a rooted spanning tree on } [n],\ A(F_k) \subseteq A(R)\}, \qquad N(F_k) = |\mathcal{R}(F_k)|.$$

The numbers of arcs in R and F_k are $n-1$ and $(n-1)-(k-1) = n-k$, respectively. Write $\mathcal{T}_n$ for the set of unrooted labeled trees on $[n]$.

1. Count construction histories forward

A *history* is a chain

$$F_k \longrightarrow F_{k-1} \longrightarrow \cdots \longrightarrow F_1 = R, \qquad F_j \text{ has } j \text{ rooted components},$$

where F_k is fixed, intermediate F_j may vary with the history, and each arrow inserts one arc as follows:

$$\boxed{\text{Choose } u \in [n], \qquad v = \text{root}(C),\ C \neq C(u), \qquad \text{insert } u \to v.} \tag{1}$$

Here $C(u)$ is the component containing u, and C is any other component. The added edge creates no cycle; v ceases to be a root, while the root of $C(u)$ remains the root of the merged component. Thus every insertion in (1) is valid.

At a stage with j components, there are $j-1$ possible target roots for each of the n choices of u:

$$\boxed{\#\{\text{admissible insertions from } F_j\} = n(j-1).} \tag{2}$$

Therefore, if $\mathcal{H}(F_k)$ is the set of all histories starting from F_k,

$$\boxed{\begin{aligned} |\mathcal{H}(F_k)| &= \prod_{j=2}^{k} n(j-1) \\ &= n^{k-1}(k-1)!. \end{aligned}} \tag{3}$$

2. Count the same histories from their terminal tree

Fix a terminal rooted tree $R \in \mathcal{R}(F_k)$. The arcs added to F_k are

$$B_R := A(R) \setminus A(F_k), \qquad |B_R| = (n-1) - (n-k) = k-1. \tag{4}$$

Key observation. Deleting an arc $u \to v$ from a rooted tree whose edges point away from its root produces *two rooted trees*: the component containing u retains the original root, while the component containing v has root v.

Indeed, each nonroot vertex has indegree 1 (exactly one arc directed into it). Deleting $u \to v$ changes only the indegree of v from 1 to 0; it does not change the directions of the remaining edges.

We may therefore delete the $k-1$ arcs of B_R *in any order*. Every intermediate graph is a rooted forest containing all the arcs of F_k. After deleting every arc outside F_k, the resulting forest is precisely F_k.

Reversing the deletions gives a valid construction history from F_k to R: each reinserted arc $u \to v$ joins two components, with v the root of the component it enters. Conversely, every history ending at R yields one such ordering. Hence

$$\boxed{|\{H \in \mathcal{H}(F_k) : \mathrm{end}(H) = R\}| = (k-1)!.} \tag{5}$$

Counting all construction histories ending at the $N(F_k)$ rooted trees R extending F_k,

$$|\mathcal{H}(F_k)| = \sum_{R \in \mathcal{R}(F_k)} (k-1)! = N(F_k)(k-1)!. \tag{6}$$

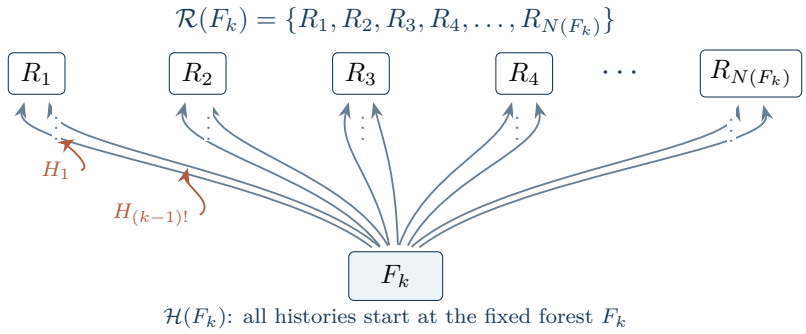


Figure 1: The histories partition into *equal-size bundles* by terminal tree. Each long curved arrow denotes one complete construction history, not one arc of a tree. The three dots between each pair indicate further histories. The shorter orange arrows label the first and last displayed histories to R_1. For each terminal tree there are $(k-1)!$ histories in all.

Combining this with (3),

$$\boxed{N(F_k)(k-1)! = n^{k-1}(k-1)! \quad \Longrightarrow \quad N(F_k) = n^{k-1}.} \qquad \square$$

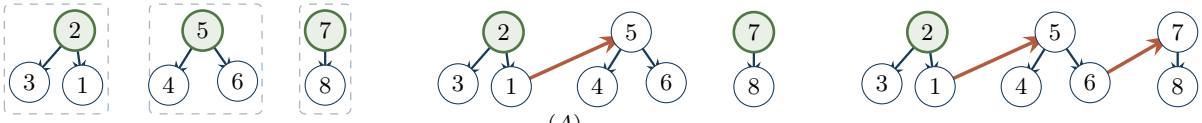
3. The same completed tree, two different histories

Let $n = 8$. Begin with three components rooted at 2, 5, and 7:

$$A(F_3) = \{2 \to 1,\ 2 \to 3,\ 5 \to 4,\ 5 \to 6,\ 7 \to 8\}.$$

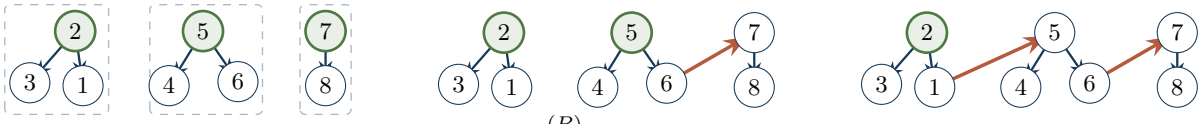
The three rooted components are drawn in separate dashed regions. The additional arcs will be $1 \to 5$ and $6 \to 7$. Two insertion orders are possible. In both rows below, vertex positions remain fixed; orange arrows mark inserted arcs and green circles mark the current component roots.

History A: first $1 \to 5$, then $6 \to 7$



$$F_3 \xrightarrow{1\to5} F_2^{(A)} \xrightarrow{6\to7} R$$

History B: first $6 \to 7$, then $1 \to 5$



$$F_3 \xrightarrow{6\to7} F_2^{(B)} \xrightarrow{1\to5} R$$

For this F_3, forward counting gives $8(3-1) \cdot 8(2-1) = 128$ histories, while each terminal rooted tree has $2! = 2$ histories. Thus $N(F_3) = 128/2 = 8^2 = 64$.

4. Cayley's formula and what survives beyond it

Take F_n to consist of n isolated vertices, each regarded as a one-vertex rooted tree. Every rooted spanning tree extends F_n. Therefore

$$\#\{\text{rooted labeled trees on } [n]\} = n^{n-1}.$$

Each unrooted labeled tree has precisely n choices of root, so

$$\boxed{n\,|\mathcal{T}_n| = n^{n-1} \quad \Longrightarrow \quad |\mathcal{T}_n| = n^{n-2}.}$$

A probabilistic consequence. Start from any fixed F_k and, when j components remain, choose uniformly among the $n(j-1)$ admissible additions. Every history has probability

$$\frac{1}{n^{k-1}(k-1)!}.$$

By (6), every $R \in \mathcal{R}(F_k)$ is reached with the same probability:

$$\boxed{\Pr\{\text{terminal tree} = R \mid F_k\} = \frac{(k-1)!}{n^{k-1}(k-1)!} = \frac{1}{n^{k-1}}.} \tag{7}$$

Thus *the same construction is an exact uniform sampler of rooted-tree extensions.*

Why "coalescent"? At a stage with j components, fix two of sizes a and b. An admissible arc can run from any of the a vertices of the first component to the root of the second, or from any of the b vertices of the second to the root of the first. Thus, under the uniform insertion rule,

$$\boxed{\Pr\{\text{these components merge next} \mid F_j\} = \frac{a+b}{n(j-1)}.} \tag{8}$$

The merger probability is proportional to $a+b$, the characteristic rule of the *additive coalescent.* Thus Pitman's enumeration also describes a random-clustering process.

Sources. J. Pitman, "Coalescent Random Forests," *Journal of Combinatorial Theory, Series A* **85** (1999), 165–193; M. Aigner and G. M. Ziegler, *Proofs from THE BOOK*, 4th ed. (Springer, 2010), *Cayley's formula for the number of trees*, pp. 201–206. This is an independent explanatory presentation of their counting method.

III. Kirchhoff: counting trees with determinants

Central idea. Select $n-1$ edges. Their reduced incidence matrix has determinant 0 if they do not form a tree, and determinant ± 1 if they do. A determinant identity then sums the squares of these numbers, so that *each spanning tree contributes exactly one.*

Kirchhoff's Matrix–Tree Theorem. *Let $G = (V, E)$ be a finite simple undirected graph on $n \geq 2$ vertices. Let L be its Laplacian (diagonal entries $\deg(v)$, and off-diagonal entry -1 for adjacent vertices, 0 otherwise). For any $r \in V$, delete row and column r from L, obtaining $L^{(r)}$. Then*

$$\boxed{\tau(G) = \det L^{(r)},}$$

where $\tau(G)$ is the number of spanning trees of G.

1. From a graph to a determinant

Give each edge an arbitrary orientation. Form the *incidence matrix* D, with one row for each vertex and one column for each edge:

$$D_{ve} = \begin{cases} +1, & e \text{ enters } v, \\ -1, & e \text{ leaves } v, \\ 0, & \text{otherwise.} \end{cases}$$

Remove row r and call the resulting $(n-1) \times |E|$ matrix B. Each column of D contains exactly one $+1$ and one -1. The inner product of row u of B with itself is $\deg_G(u)$, while the inner product of two distinct rows u, v is -1 if they correspond to adjacent vertices, and 0 otherwise. Hence

$$\boxed{BB^{\mathsf{T}} = L^{(r)}.} \tag{1}$$

These products are independent of the arbitrary edge orientations.

2. Why the selected-edge determinants are 0 or ± 1

Fix $S \subseteq E$ with $|S| = n-1$, and write B_S for the square matrix consisting of the corresponding columns of B (in any fixed order).

The determinant test for a tree.

$$\boxed{\det B_S = \begin{cases} \pm 1, & (V, S) \textit{ is a spanning tree,} \\ 0, & (V, S) \textit{ is not a spanning tree.} \end{cases}} \tag{2}$$

A cycle forces zero. An acyclic graph on n vertices with $n-1$ edges is a tree. Hence, if (V, S) is not a tree, it contains a cycle C. Traverse C in either direction and assign $\varepsilon_e = 1$ when the assigned orientation agrees with the traversal, -1 otherwise. The signed incidence columns cancel at each vertex of the cycle:

$$\sum_{e \in C} \varepsilon_e D_e = 0, \qquad \varepsilon_e \in \{-1, 1\}.$$

Since the signed sum is zero in every row, deleting row r preserves the relation. Hence $\sum_{e \in C} \varepsilon_e B_e = 0$. Since $C \subseteq S$, this is a nontrivial linear dependence among the columns of B_S. Therefore $\det B_S = 0$.

A tree gives ± 1. If (V, S) is a tree, choose a leaf $v \neq r$ (a tree with at least two vertices has at least two leaves). Row v of B_S contains just one nonzero entry, $+1$ or -1. Expansion along that row gives $\det B_S = \pm \det B_{S'}$, where $B_{S'}$ is the reduced incidence matrix of the smaller tree obtained by deleting v and its edge, still with row r removed. Continue until one edge remains. Its reduced incidence matrix is (1) or (-1), so $|\det B_S| = 1$. This proves (2).

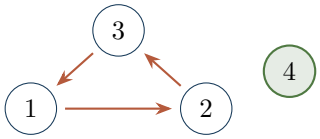
Three edges forming a path



$$B_S = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

$$\det B_S = -1$$

Three edges forming a cycle



$$B_S = \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

$$\det B_S = 0$$

Figure 1. Two selections of three edges on four vertices, with the row of vertex $r = 4$ deleted. The green vertex marks r; orange arrows highlight a cycle. Each column corresponding to a cycle edge is nonzero; it is their signed sum that is zero.

3. Cauchy–Binet, starting from permutations

For a square $m \times m$ matrix $A = (a_{ij})$, the determinant is defined by

$$\det A = \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \prod_{i=1}^{m} a_{i,\sigma(i)}. \tag{3}$$

Here S_m is the set of permutations of $\{1, \ldots, m\}$, and $\operatorname{sgn}(\sigma)$ is $+1$ or -1 according to the parity of σ. Since each summand uses one entry from each row, (3) makes the determinant *multilinear in its rows*. Relabeling permutations after swapping two rows changes every sign: the determinant is *alternating*, so equal rows give determinant 0. The same properties hold for columns. Also, replacing σ by σ^{-1} proves $\det A^{\mathsf{T}} = \det A$.

A glimpse at 2×3. For $B = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$, direct expansion gives

$$\begin{aligned} \det(BB^{\mathsf{T}}) &= (a^2 + b^2 + c^2)(d^2 + e^2 + f^2) - (ad + be + cf)^2 \\ &= (ae - bd)^2 + (af - cd)^2 + (bf - ce)^2. \end{aligned} \tag{4}$$

The three squares are the determinants of the three 2×2 column selections, squared. Why must this happen in every dimension?

Cauchy–Binet, in the form we need. *If B is an $m \times q$ matrix with $q \geq m$ and B_S denotes its $m \times m$ submatrix with columns indexed by the m-element subset S, in increasing order, then*

$$\det(BB^{\mathsf{T}}) = \sum_{\substack{S \subseteq \{1,\ldots,q\} \\ |S| = m}} (\det B_S)^2. \tag{5}$$

Proof. Write b_e for column e of B and b_{ie} for its entry in row i. The ith row of BB^{T} is $\sum_{e=1}^{q} b_{ie}\, b_e^{\mathsf{T}}$. By multilinearity in rows and $\det C^{\mathsf{T}} = \det C$,

$$\begin{aligned} \det(BB^{\mathsf{T}}) &= \sum_{e_1,\dots,e_m=1}^{q} \left(\prod_{i=1}^{m} b_{ie_i}\right) \det \begin{pmatrix} b_{e_1}^{\mathsf{T}} \\ \vdots \\ b_{e_m}^{\mathsf{T}} \end{pmatrix} \\ &= \sum_{e_1,\dots,e_m=1}^{q} \left(\prod_{i=1}^{m} b_{ie_i}\right) \det(b_{e_1}, \dots, b_{e_m}). \end{aligned} \tag{6}$$

If two indices e_i coincide, the last determinant has identical columns and vanishes. Therefore only *distinct* column indices matter.

Fix a set $S = \{s_1 < \cdots < s_m\}$. Every ordering of its columns has the form $(s_{\pi(1)}, \dots, s_{\pi(m)})$, for some $\pi \in S_m$. Rearranging the columns multiplies their determinant by $\operatorname{sgn}(\pi)$. The total contribution of this one set S to (6) is

$$\begin{aligned} &\sum_{\pi \in S_m} \left(\prod_{i=1}^{m} b_{i,s_{\pi(i)}}\right) \det(b_{s_{\pi(1)}}, \dots, b_{s_{\pi(m)}}) \\ &\quad = \det(B_S) \sum_{\pi \in S_m} \operatorname{sgn}(\pi) \prod_{i=1}^{m} b_{i,s_{\pi(i)}} \\ &\quad = \det(B_S) \det(B_S) = (\det B_S)^2, \end{aligned} \tag{7}$$

where the second determinant is *exactly the permutation sum* (3). Summing (7) over all m-element column sets proves (5). □

4. Each spanning tree contributes one

Apply (5) to the reduced incidence matrix B, with $m = n - 1$. By (2),

$$\begin{aligned} \det L^{(r)} = \det(BB^{\mathsf{T}}) &= \sum_{\substack{S \subseteq E \\ |S| = n-1}} (\det B_S)^2 \\ &= \sum_{\substack{S \subseteq E \\ (V,S) \text{ spanning tree}}} 1 = \tau(G). \end{aligned} \tag{8}$$

If $|E| < n - 1$, the graph has no spanning tree and BB^{T} has rank less than $n - 1$, so both sides are 0. Thus the theorem holds without assuming that G is connected.

5. For the complete graph, the determinant collapses

The Laplacian of a graph is $L = \Delta - A$, where Δ is the diagonal matrix of vertex degrees and A is the adjacency matrix.

Every vertex of K_n has degree $n - 1$, and every pair of distinct vertices is adjacent. Hence its Laplacian has diagonal entries $n - 1$ and off-diagonal entries -1. Therefore

$$\boxed{L = nI_n - J_n,}$$

where J_n is the $n \times n$ matrix of all ones. Put $m = n - 1$; then

$$L^{(r)} = nI_m - J_m.$$

Every row of this $m \times m$ matrix sums to $n - m = 1$. Add rows $2, \ldots, m$ to row 1, so its first row becomes $(1, 1, \ldots, 1)$. Subtract column 1 from every other column. Neither operation changes the determinant, and the resulting matrix is

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ -1 & n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & \cdots & n \end{pmatrix}.$$

It is lower triangular, so

$$\boxed{\tau(K_n) = \det L^{(r)} = n^{m-1} = n^{n-2}.} \tag{9}$$

Every labeled tree on $[n]$ is a spanning tree of K_n. Hence $|\mathcal{T}_n| = n^{n-2}$: *Cayley's formula.* □

Perspective. Prüfer constructs a bijection; Pitman counts histories; Kirchhoff turns each candidate edge set into an algebraic 0 or 1 and adds those indicators.

Beyond the proof: squared rectangles. A rectangle dissected into squares gives an electrical network: maximal horizontal segments become vertices, and each small square becomes an edge of unit resistance. Potential differences represent square heights, while currents represent widths. Kirchhoff's equations involve the graph Laplacian, whose reduced determinant counts spanning trees. A *perfect squared square* additionally requires a square outer boundary and pairwise distinct constituent square sizes.

IV. Joyal: from functions to doubly rooted trees

Central idea. Given a function $f : [n] \to [n]$, its *functional digraph* has vertex set $[n]$ and one directed arc $v \to f(v)$ for each vertex v. The vertices on directed cycles form a set on which the function acts as a permutation. Joyal turns this permutation into a path, then uses two distinguished vertices to make the path recoverable. The remaining edges become branches of a tree.

Joyal's bijection. *Let $\mathcal{F}_n$ be the set of functions $f : [n] \to [n]$, and let*

$$\mathcal{D}_n = \{(T, a, b) : T \text{ is a labeled tree on } [n],\ a, b \in [n]\}.$$

The pair (a, b) is ordered*; its entries may coincide. There is a bijection $\mathcal{F}_n \longleftrightarrow \mathcal{D}_n$.*

For a fixed T, there are n^2 ordered choices of (a, b). Consequently,

$$|\mathcal{F}_n| = n^n, \qquad |\mathcal{D}_n| = n^2|\mathcal{T}_n|.$$

Constructing the bijection will prove Cayley's formula.

1. Cycles and the cyclic set

In a functional digraph, every vertex has exactly one outgoing arc, though it may have many incoming arcs. Loops are allowed: $f(v) = v$ is a directed cycle of length one.

Starting from any v, repeatedly follow the outgoing arcs:

$$v,\ f(v),\ f^2(v), \ldots.$$

The process never stops. Since there are only n vertices, at least one vertex repeats among the first $n+1$ terms. The section between its first and second visits is a directed cycle; thereafter the sequence repeats that cycle.

Why exactly one cycle per component of the function's digraph? If $u \to v$ is an arc, the sequences starting at u and at v reach the same cycle: the first is just the second preceded by u. For example,

$$3 \to 2 \to 1 \to 4 \to 2 \to \cdots, \qquad 2 \to 1 \to 4 \to 2 \to \cdots$$

both reach the cycle $2 \to 1 \to 4 \to 2$. This agreement propagates along paths when directions are ignored. Thus each connected component of the underlying undirected graph has exactly one directed cycle, with other vertices feeding toward it.

Let $M \subseteq [n]$ consist of *all vertices on directed cycles*. Since f rotates the vertices of each cycle, $f|_M : M \to M$ is a permutation. Moreover, it is the *unique largest* subset on which f acts as a permutation: if $f|_S : S \to S$ is a permutation on some $S \subseteq [n]$, every vertex of S belongs to a cycle of that finite permutation. Hence

$$S \subseteq M.$$

2. An eight-vertex example

Consider

x	1	2	3	4	5	6	7	8
$f(x)$	4	1	2	2	7	5	5	8

.

Its directed cycles are (4 2 1), (7 5), and (8). The noncycle arcs are $3 \to 2$ and $6 \to 5$.

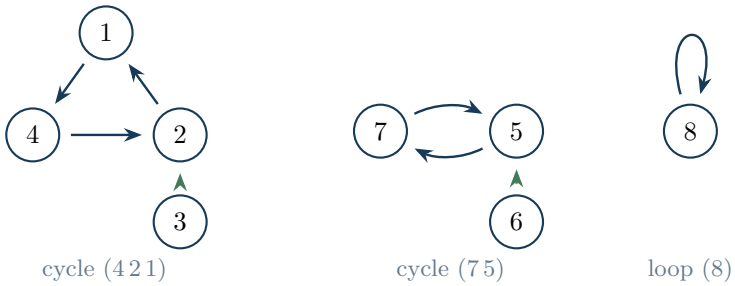


Figure 1. The functional digraph. Each of its three components has one directed cycle; green arcs lead toward cycles.

The cyclic set and its increasing order are

$$M = \{1, 2, 4, 5, 7, 8\}, \qquad m_1 < m_2 < \cdots < m_6 : \quad 1, 2, 4, 5, 7, 8.$$

The corresponding images, in that order, are

$$\boxed{f(m_1), \ldots, f(m_6) = 4, 1, 2, 7, 5, 8.}$$

3. From a function to a doubly rooted tree

For an arbitrary f, list its cyclic vertices as $m_1 < \cdots < m_k$ and set

$$p_i = f(m_i) \qquad (1 \le i \le k).$$

Because $f|_M$ is a permutation, $(p_1, \ldots, p_k)$ lists each vertex of M exactly once. *Remove every cycle arc*, retain the noncycle edges without directions, and join consecutive vertices in the ordered list to make a path

$$p_1 - p_2 - \cdots - p_k.$$

Deleting the cycle arcs leaves a forest of k trees, one rooted at each vertex of M: every other vertex has a unique forward route to its eventual cyclic vertex. Joining these k roots by one path makes the whole graph a tree T.

Why two distinguished vertices? An unmarked tree need not tell us which path was inserted. But *between any two vertices of a tree there is exactly one path.* Mark its endpoints in order,

$$\boxed{a = p_1, \qquad b = p_k.}$$

Then (a, b) specifies both the path and the direction in which to read it. The endpoints may coincide when $k = 1$, and they need not be leaves of T.

For our example, this gives the path $4 - 1 - 2 - 7 - 5 - 8$, with the edges $3 - 2$ and $6 - 5$ attached:

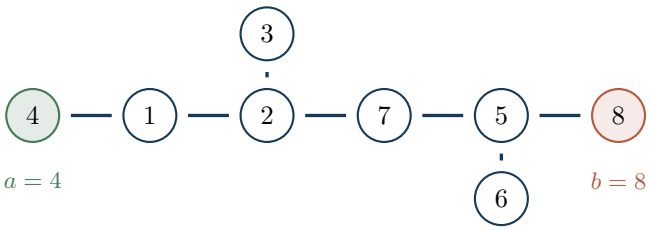


Figure 2. The doubly rooted tree. Green and orange vertices mark the ordered endpoints of the distinguished path.

4. The inverse construction

Start with *any* labeled tree T and choose an ordered pair (a, b) of its vertices, possibly with $a = b$. Follow the unique path from a to b in that direction:

$$p_1 = a,\ p_2, \ldots, p_k = b.$$

Let its vertex set be M, and arrange the same vertices increasingly as $m_1 < \cdots < m_k$. Define

$$\boxed{f(m_i) = p_i \qquad (1 \le i \le k).}$$

This is a permutation of M. For each vertex $v \notin M$, there is exactly one adjacent vertex along the unique route from v to the distinguished path. *Define $f(v)$ to be that neighbor.* Thus all other edges point toward the path. In Figure 2, the path $4, 1, 2, 7, 5, 8$ yields

m_i	1	2	4	5	7	8
$p_i = f(m_i)$	4	1	2	7	5	8

,

and the branches give $f(3) = 2$, $f(6) = 5$. We recover the original eight-vertex function.

5. Why the constructions are inverses

From a function to a doubly rooted tree—and back. Every vertex outside M follows a unique chain of noncycle arcs toward M; these are exactly the edges retained as branches of the tree. The ordered vertices $a = p_1$ and $b = p_k$ recover the inserted path, because a tree has a unique path between any two vertices. Reading this path from a to b recovers $(p_1, \ldots, p_k)$ and hence $f(m_i) = p_i$, where $m_1 < \cdots < m_k$ are its vertices in increasing order. Directing the branches toward the path recovers $f(v)$ for every $v \notin M$.

From a tree to a function—and back. Begin with any labeled tree T and an ordered choice of vertices (a, b). Their unique path from a to b has vertex set M. The reconstructed function satisfies $f(m_i) = p_i$, so it permutes M; hence every vertex of M is cyclic. If $v \notin M$, each application of f moves it one step closer to the path until it reaches M. Thereafter it cannot leave M, since f permutes M. Thus v can never return to itself and is not cyclic. The cyclic set of f is therefore exactly M. Applying the forward construction now recovers the same ordered path, retains all original branches, and marks the same vertices a, b.

Thus the constructions give inverse bijections, and

$$\boxed{n^n = |\mathcal{F}_n| = |\mathcal{D}_n| = n^2|\mathcal{T}_n| \quad \Longrightarrow \quad |\mathcal{T}_n| = n^{n-2}.}$$

The bijection also works for $n = 1$, when $a = b$.

Boundary cases. If f is the identity, then $M = [n]$ and the path is $1 - 2 - \cdots - n$, with $(a, b) = (1, n)$. If $f(v) = r$ for every v, then $M = \{r\}$, T is the star centered at r, and $(a, b) = (r, r)$.

What this perspective adds. Prüfer's bijection encodes a tree by removing leaves. Joyal instead exploits the structure of a function: its cyclic vertices form a set on which it acts as a permutation, and that permutation becomes an ordered path. The unique-path property of trees explains why two distinguished vertices suffice to recover the path; the remaining branches recover all other function values.

The bijection also has a probabilistic consequence. If M is the cyclic set of f and (T, a, b) is its corresponding doubly rooted tree, then the distinguished path has exactly the vertices of M. Hence

$$\boxed{|M| = d_T(a, b) + 1,}$$

where $d_T(a, b)$ is the number of edges on the unique path. Thus the number of cyclic vertices of a uniformly random function has the same distribution as $d_T(a, b) + 1$ for a uniformly random labeled tree T and two independently uniform vertices a, b. This connects random mappings with distances in random trees. Pitman explores probabilistic consequences of such correspondences; Heilman uses a tree–endofunction bijection to obtain concentration inequalities for random-tree statistics.

References and further reading.

1. A. Joyal, "Une théorie combinatoire des séries formelles," *Advances in Mathematics* **42** (1981), 1–82. DOI.
2. M. Aigner and G. M. Ziegler, *Proofs from THE BOOK*, "Cayley's formula for the number of trees," chapter on Cayley's formula.
3. J. Pitman, "Random mappings, forests, and subsets associated with Abel–Cayley–Hurwitz multinomial expansions," *Séminaire Lotharingien de Combinatoire* **46** (2001), Article B46h, 45 pp. Paper.
4. S. Heilman, "Tree/Endofunction Bijections and Concentration Inequalities," *Electronic Journal of Combinatorics* **29**(2) (2022), #P2.33. DOI.

V. Chapuy–Perarnau: a telescoping proof

Central idea. Functions $f : [n] \to [n]$ whose functional digraphs have exactly one vertex belonging to a directed cycle are in bijection with rooted trees on $[n]$. Chapuy and Perarnau compute the probability of this event by exploring a random function in rounds: the conditional probabilities telescope.

1. From functions to rooted trees

Let $[n] = \{1, \dots, n\}$, and choose $f : [n] \to [n]$ uniformly from the n^n functions. Thus the values $f(v)$ are independent and uniform on $[n]$. The *functional digraph* has an arc $v \to f(v)$ for every vertex v; loops are allowed. A vertex is *cyclic* if it belongs to a directed cycle.

If f has exactly one cyclic vertex r, its only cycle is the loop $r \to r$. Every other vertex eventually reaches r. Delete the loop, discard directions, and obtain a tree rooted at r. Conversely, orient each edge of a rooted tree toward its root and add a loop there. Hence

$$\boxed{\{f : \text{exactly one cyclic vertex}\} \longleftrightarrow \{\text{rooted labeled trees}\}.}$$

There are n ways to root each labeled tree, so if $\mathcal{T}_n$ is the set of labeled trees on $[n]$, then

$$\Pr(f \text{ has exactly one cyclic vertex}) = \frac{n\,|\mathcal{T}_n|}{n^n}.$$

We shall prove the following identity of Chapuy and Perarnau.

The probability identity.

$$\boxed{\Pr(f \textit{ has exactly one cyclic vertex}) = \frac{1}{n}.} \tag{1}$$

Equation (1) immediately yields $|\mathcal{T}_n| = n^{n-2}$.

2. Reveal the function in rounds

Initially none of the outgoing arcs is known. An arc is *revealed* when we learn its destination; a vertex is *explored* when its outgoing arc has been revealed. Choose the smallest unexplored vertex, reveal its outgoing arc, then continue from the vertex reached. *End the round* when the newly revealed arc reaches a vertex whose outgoing arc is now already known (possibly the vertex from which that arc just left).

If unexplored vertices remain, start a new round at the smallest such vertex. Continue until every outgoing arc has been revealed. Let K be the number of rounds and

$$0 = T_0 < T_1 < \cdots < T_K = n,$$

where T_i is the total number of explored vertices after round i. The T_i and K are random.

Example ($n = 7$). Consider the function

v	1	2	3	4	5	6	7
$f(v)$	2	2	4	2	6	3	6

.

The successive rounds are:

round	newly revealed arcs	final destination	T_i
1	$1 \to 2,\ 2 \to 2$	2 (loop)	2
2	$3 \to 4,\ 4 \to 2$	2 (earlier round)	4
3	$5 \to 6,\ 6 \to 3$	3 (earlier round)	6
4	$7 \to 6$	6 (earlier round)	7

Table 1. The closing arc in each round has a destination among the vertices visited by that time.

Each round after the first has attached to an earlier one, so this function has just one cyclic vertex, namely 2.

3. Why are all possible closing destinations equally likely?

A distinction worth making. Just before a prospective closing arc of round i is revealed, let x_i be its source and S_i the *set* of vertices visited so far, including x_i. On the event that this arc ends round i, $|S_i| = T_i$. Its destination $f(x_i)$ is a previously unrevealed function value.

Given the history before this reveal, the destination $f(x_i)$ is equally likely to be any vertex of $[n]$. The round ends precisely if this arc points into S_i, an event of probability T_i/n. Therefore, for each $w \in S_i$,

$$\Pr(f(x_i) = w \mid f(x_i) \in S_i,\ \text{history before this reveal}) = \frac{1/n}{T_i/n} = \boxed{\frac{1}{T_i}}. \tag{2}$$

Knowing *that* the round ends does not tell us *which* vertex of S_i the closing arc reaches.

Why may we also condition on all the exploration sizes? Fix the *exploration skeleton*: the start of each round, its successive new vertices, all nonclosing arcs, and the point at which each round ends, but *not the destinations of the closing arcs*. Changing any closing destination to another vertex already visited in that round still ends the round at the same step. It changes neither the explored vertex set nor any later choice of starting vertex. Hence the skeleton, and in particular $(T_1, \ldots, T_K; K)$, carry no information favoring one allowed closing destination over another.

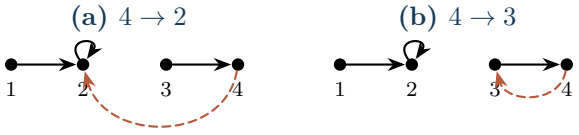


Figure 1. Two completions of the same exploration skeleton. The dashed closing arc changes from $4 \to 2$ to $4 \to 3$. In both, $T_1 = 2$ and $T_2 = 4$, although the number of cycles changes.

The skeleton fixes the continuation arcs and restricts each closing destination separately to S_i; it imposes no relation between different closing destinations. Since these arcs originate at distinct vertices whose

function values were chosen independently, their destinations remain independent and equally likely over their respective sets S_i, even after the skeleton is fixed.

The conditions required for exactly one cyclic vertex depend separately on these closing destinations. They are therefore conditionally independent, so their probabilities may be multiplied.

4. Which closing arcs give exactly one cyclic vertex?

Round 1. Its closing arc necessarily creates a directed cycle. Let x_1 be the source of that arc. The cycle has length one precisely when $f(x_1) = x_1$. Since x_1 is one of the T_1 possible destinations, exactly one choice produces a loop. Hence

$$\Pr(\text{first cycle is a loop} \mid \text{skeleton}) = \boxed{\frac{1}{T_1}}. \tag{3}$$

Round $i \geq 2$. Exactly T_{i-1} of the T_i potential closing destinations lie in *previous* rounds. Choosing one of them attaches the new chain to an existing component and produces no new cycle. Any destination within the current round closes a new cycle. Thus

$$\Pr(\text{no new cycle in round } i \mid \text{skeleton}) = \boxed{\frac{T_{i-1}}{T_i}}. \tag{4}$$

For example, suppose Round 1 reveals $1 \to 2$ and $2 \to 2$, and Round 2 begins with $3 \to 4$. Conditional on Round 2 ending with the outgoing arc of vertex 4, its destination may be any of $1, 2, 3, 4$, each with probability $1/4$.

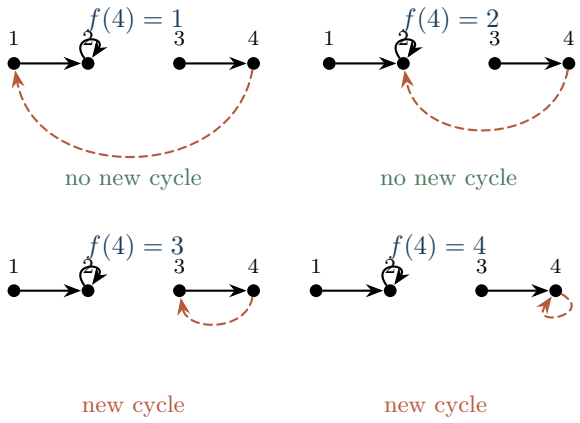


Figure 2. Conditional on Round 2 closing from vertex 4, the four equally likely destinations are $f(4) = 1, 2, 3,$ or 4. The choices $1, 2$ avoid a new cycle, while $3, 4$ create one. Hence the probability of avoiding a new cycle is $T_1/T_2 = 2/4$.

5. The telescoping product

A function has exactly one cyclic vertex if and only if the first round closes in a loop and every later round attaches to an earlier round. Conditional on the exploration skeleton, equations (3)–(4) and the

independence of closing choices give

$$\begin{aligned}\Pr(\text{exactly one cyclic vertex} \mid \text{skeleton}) &= \frac{1}{T_1} \cdot \frac{T_1}{T_2} \cdot \frac{T_2}{T_3} \cdots \frac{T_{K-1}}{T_K} \\ &= \frac{1}{T_K} = \boxed{\frac{1}{n}}.\end{aligned} \tag{5}$$

The probability is $1/n$ for *every* exploration skeleton. Averaging over these skeletons proves (1), and therefore

$$\boxed{|\mathcal{T}_n| = n^{\,n-2}.}$$

What this perspective adds. Joyal reorganizes the cycles of a function into a recoverable path; Chapuy and Perarnau instead reveal the function progressively. The same functional digraph leads to a counting proof through conditional probabilities. Its striking feature is that the randomness of the exploration lengths disappears in the telescoping product.

Reference. G. Chapuy and G. Perarnau, "A Telescopic Proof of Cayley's Formula," *American Mathematical Monthly* **131** (2024), no. 10, 899–902. doi:10.1080/00029890.2024.2396774. Preprint: arXiv:2306.12918.

VI. Riordan–Rényi: a recursion for labeled forests

Central idea. Rather than count labeled trees directly, count forests whose components are rooted trees, with their roots forming a prescribed set. Deleting one of these roots gives a reversible decomposition and a recurrence. The resulting formula contains Cayley's theorem as the case of a single component.

Forest-counting theorem. *Let $A \subseteq [n]$ have k vertices. Let $T_{n,k}$ be the number of forests on $[n]$ consisting of exactly k rooted trees, with A as the set of their roots. Then, for $1 \le k \le n$,*

$$\boxed{T_{n,k} = k\, n^{\,n-k-1}.} \tag{1}$$

Here a root is simply a designated vertex of its component; we do not orient the edges. By relabeling, take $A = \{1, \dots, k\}$. When $k = 1$, a forest counted by $T_{n,1}$ is just a labeled tree, and (1) becomes Cayley's formula.

1. What happens when we delete vertex 1?

Choose a forest F counted by $T_{n,k}$, and suppose $\deg_F(1) = i$. Because no other vertex of A lies in the component containing 1, its i neighbors lie in $[n] \setminus A$.

Deleting vertex 1 breaks its component into i components: two of its neighbors cannot remain connected after the deletion, for otherwise F would have contained a cycle. The other $k - 1$ components are unchanged. Thus the new forest has

$$\boxed{(k-1) + i \text{ components.}} \tag{2}$$

Together, the i former neighbors of 1 and the vertices $2, \dots, k$ lie in pairwise distinct components of the new forest.

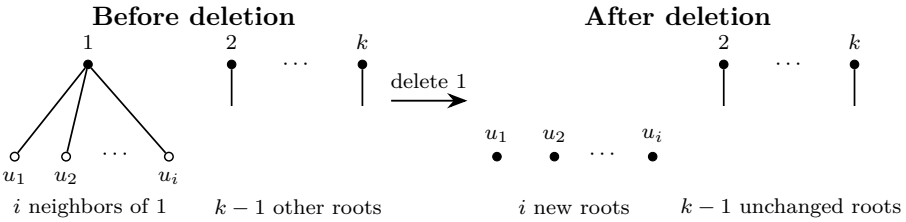


Figure 1. Deleting root 1 of degree i separates its component into i rooted trees, whose roots are its former neighbors $u_1, \dots, u_i$. The other $k-1$ rooted trees remain unchanged; short downward edges indicate unshown portions of those components. Thus k components become $k - 1 + i$.

2. The reversible decomposition and its recurrence

Let $B \subseteq [n] \setminus A$ be the set of the i neighbors of 1. There are $\binom{n-k}{i}$ choices for B. Deleting 1 gives a forest on $[n] \setminus \{1\}$ with $k - 1 + i$ components, each containing exactly one member of

$$A' = \{2, \dots, k\} \cup B.$$

For each fixed B, the number of such forests is $T_{n-1,k-1+i}$: the answer depends only on the size of the designated root set, not its labels.

Why is the construction reversible? Given one of these forests and its specified set B, reinsert vertex 1 and connect it to every vertex of B. Their components are distinct, so no cycle is created. The resulting forest has exactly k components, with root set A; moreover, the neighbors of 1 are precisely B.

Summing over the possible values of i gives the recurrence

$$\boxed{T_{n,k} = \sum_{i=0}^{n-k} \binom{n-k}{i} T_{n-1,k-1+i}.} \tag{3}$$

For $s \geq 1$ we set $T_{s,0} = 0$; this accounts for the impossible case $k = 1$, $i = 0$. Note that $T_{n,n} = 1$, since the only forest with n components is the edgeless forest.

3. Solving the recurrence

We prove (1) by induction on n. The case $n = 1$ is immediate. For $n \geq 2$, the case $k = n$ gives $T_{n,n} = 1 = n\,n^{-1}$.

Now fix $1 \leq k < n$ and assume the formula is known for forests on $n-1$ vertices. Put $m = n - k$, so $m \geq 1$. Substitution into (3) yields

$$T_{n,k} = \sum_{i=0}^{m} \binom{m}{i} (k-1+i)(n-1)^{m-i-1}. \tag{4}$$

For $k = 1$ and $i = 0$ the summand is interpreted as 0, consistently with $T_{n-1,0} = 0$. The useful observation is

$$k - 1 + i = (n-1) - (m-i).$$

Thus (4) becomes a difference of two sums. By the binomial theorem,

$$\sum_{i=0}^{m} \binom{m}{i} (n-1)^{m-i} = n^m.$$

For the other sum, use $(m-i)\binom{m}{i} = m\binom{m-1}{i}$ for $0 \leq i < m$:

$$\begin{aligned} \sum_{i=0}^{m} \binom{m}{i} (m-i)(n-1)^{m-i-1} &= m \sum_{i=0}^{m-1} \binom{m-1}{i} (n-1)^{m-1-i} \\ &= mn^{m-1}. \end{aligned}$$

Consequently,

$$T_{n,k} = n^m - mn^{m-1} = (n-m)n^{m-1} = \boxed{k\,n^{n-k-1}}.$$

This completes the induction.

4. A small check: eight forests

Take $n = 4$, $A = \{1, 2\}$, and $k = 2$. The recurrence sorts the forests by the degree i of vertex 1:

i	0	1	2	Total
$\binom{2}{i} T_{3,1+i}$	3	$2 \cdot 2 = 4$	1	8

Therefore $T_{4,2} = 8$, agreeing with the eight forests illustrated in *Proofs from THE BOOK*. The formula gives the same answer: $2 \cdot 4^{4-2-1} = 8$.

5. What this perspective adds

The key move is to *generalize before recursing*. Deleting a vertex from a tree usually produces several components, so a recurrence involving tree counts alone does not close. The two-parameter family $T_{n,k}$ is exactly broad enough to remain stable under deletion. Once that family is introduced, the recursion is a reversible counting argument and the algebra reduces to the binomial theorem.

There is a revealing contrast with Pitman's proof: Pitman *joins* components while counting insertion histories; the present proof *deletes* a distinguished vertex and counts the resulting forests. Both lead to Cayley's formula, but they expose different mechanisms.

Source note. The recursive argument follows the *Third Proof (Recursion)* in M. Aigner and G. M. Ziegler, *Proofs from THE BOOK*, chapter "Cayley's formula for the number of trees," p. 144. The book attributes the idea essentially to Riordan and Rényi. See J. Riordan, "Forests of labeled trees," *J. Combinatorial Theory* **5** (1968), 90–103, and A. Rényi, "Some remarks on the theory of trees" (1959), as cited there.

VII. Cayley's polynomial identity: Rényi's proof by symmetry

Central idea. Weight each labeled tree by a monomial recording its vertex degrees. A judicious specialization gives an induction. Surprisingly, *symmetry and low degree* allow that one specialization to determine the entire polynomial.

Cayley stated the identity below in 1889, verifying a particular case. Rényi supplied a general inductive proof in work presented in 1969 and published in 1970. We use a concise divisibility version of the algebraic step in his argument.

Cayley's polynomial identity. *For $n \geq 2$, let $\mathcal{T}_n$ be the set of labeled trees on $[n] = \{1, \ldots, n\}$ and put*

$$C_n(x_1, \ldots, x_n) = \sum_{T \in \mathcal{T}_n} \prod_{i=1}^{n} x_i^{\deg_T(i)-1}.$$

Then

$$\boxed{C_n(x_1, \ldots, x_n) = (x_1 + \cdots + x_n)^{n-2}.} \tag{1}$$

Setting every $x_i = 1$ gives $|\mathcal{T}_n| = n^{n-2}$, Cayley's formula. The polynomial identity gives more information: its coefficients also count trees with specified vertex degrees.

1. Why weight a tree by $\prod_i x_i^{\deg_T(i)-1}$?

For every vertex of a tree on $n \geq 2$ vertices, $\deg_T(i) \geq 1$. Thus the exponents are nonnegative. In particular, the factor belonging to a leaf is $x_i^0 = 1$: *setting $x_i = 0$ retains precisely the trees in which i is a leaf.*

Two further observations make these weights especially useful.

Symmetry. Relabeling vertices permutes the trees and the corresponding variables, so C_n is a symmetric polynomial.

Homogeneity. A tree has $n - 1$ edges. The sum of its vertex degrees is $2(n-1)$; hence each monomial has degree

$$\sum_{i=1}^{n} (\deg_T(i) - 1) = 2(n-1) - n = \boxed{n-2}.$$

Consequently, C_n is homogeneous of degree $n - 2$.

2. The combinatorial step: set $x_n = 0$

A term of $C_n(x_1, \ldots, x_{n-1}, 0)$ survives exactly when vertex n is a leaf. Delete that leaf, leaving a tree T' on $[n-1]$. Conversely, any T' on $[n-1]$ and any chosen $j \in [n-1]$ give a unique tree T with leaf n, by adding edge jn.

Attaching the leaf increases the degree of j by one and does not change the other exponents. Therefore

$$\prod_{i=1}^{n} x_i^{\deg_T(i)-1} = x_j \prod_{i=1}^{n-1} x_i^{\deg_{T'}(i)-1}.$$

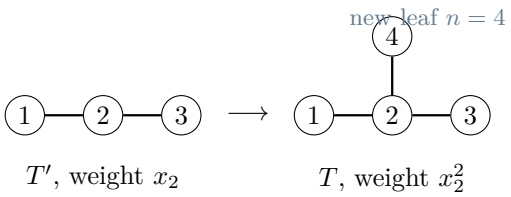


Figure 1. Attaching leaf $n = 4$ at vertex $j = 2$ multiplies the weight by x_2.

Every surviving tree is obtained exactly once from the pair (T', j). Summing first over j and then over T' gives

$$\boxed{C_n(x_1, \ldots, x_{n-1}, 0) = (x_1 + \cdots + x_{n-1}) C_{n-1}(x_1, \ldots, x_{n-1}).} \tag{2}$$

For $n = 2$, (1) is immediate: the unique tree has weight 1. Assuming (1) for $n - 1$, equation (2) becomes

$$\boxed{C_n(x_1, \ldots, x_{n-1}, 0) = (x_1 + \cdots + x_{n-1})^{n-2}.} \tag{3}$$

We have proved the desired identity when $x_n = 0$. Why should that determine a polynomial in *all* n variables?

3. The algebraic step: symmetry and low degree

Symmetry lemma. *If two symmetric polynomials P, Q in n variables have degree less than n and agree when $x_n = 0$, then $P = Q$.*

Proof. Set $R = P - Q$. Regard R as a polynomial in x_n, with coefficients that are polynomials in $x_1, \ldots, x_{n-1}$. Since $R(x_1, \ldots, x_{n-1}, 0) = 0$, its constant term (with respect to x_n) is zero. Thus every term contains a factor x_n, so $x_n \mid R$. By symmetry, the same holds for every variable:

$$x_i \mid R \quad (1 \le i \le n) \quad \Longrightarrow \quad x_1 x_2 \cdots x_n \mid R.$$

But a nonzero polynomial divisible by $x_1 \cdots x_n$ has degree at least n, whereas $\deg R < n$. Hence $R = 0$ and $P = Q$. □

Apply the lemma. Both C_n and $(x_1 + \cdots + x_n)^{n-2}$ are symmetric polynomials of degree $n - 2 < n$. By (3), they agree at $x_n = 0$. The lemma gives identity (1), completing the induction.

Rényi's coefficient viewpoint. A monomial of degree $n - 2$ in n variables must omit at least two variables. By symmetry, the coefficient of any monomial equals the coefficient of a rearranged monomial that omits x_n. Thus the specialization at $x_n = 0$ determines *every* coefficient. This is how Rényi presents the final step; the divisibility lemma expresses the same idea in a compact form.

4. What else does the polynomial identity give?

The coefficient of $x_1^{d_1 - 1} \cdots x_n^{d_n - 1}$ counts trees with $\deg_T(i) = d_i$. Applying the multinomial theorem to (1), if $d_i \ge 1$ and $\sum_i d_i = 2n - 2$, we obtain

$$\boxed{\left|\{T \in \mathcal{T}_n : \deg_T(i) = d_i \text{ for all } i\}\right| = \frac{(n-2)!}{\prod_{i=1}^{n} (d_i - 1)!}.}$$

This is the same degree-counting formula derived from Prüfer sequences, but it has emerged here without encoding any tree.

A bridge to Kirchhoff. Rényi also considered the corresponding weighted polynomial for spanning trees of an arbitrary graph G. Give edge $\{i, j\}$ the weight $x_i x_j$. A spanning tree then has edge-weight product

$$\prod_{\{i,j\} \in E(T)} x_i x_j = \Big(\prod_{i=1}^{n} x_i\Big)\Big(\prod_{i=1}^{n} x_i^{\deg_T(i) - 1}\Big).$$

The weighted Matrix–Tree Theorem therefore links the polynomial for G to a Laplacian cofactor. Our seventh perspective returns us to the linear-algebraic third perspective.

What this perspective adds. Rényi replaces explicit enumeration by *determination*: symmetry, low degree, and a single specialization force the whole polynomial. The motivating combinatorial operation is as elementary as deleting a leaf.

References

[1] A. Cayley, "A theorem on trees," *Quarterly Journal of Pure and Applied Mathematics* **23** (1889), 376–378.

[2] A. Rényi, "On the enumeration of trees," in *Combinatorial Structures and Their Applications*, Proceedings of the Calgary International Conference (1969), Gordon and Breach, New York, 1970, 355–360. Reproduced in *Selected Papers of Alfréd Rényi*, vol. 3, pp. 577–580.

VIII. Child positions and generating polynomials

Central idea. Assign each vertex a prescribed number of ordered child positions, then fill the positions one at a time, always attaching the root of another component. This gives a reversible way to count rooted forests. Forgetting the child orders produces factorial weights, and a generating polynomial then gives Cayley's formula—and counts trees subject to degree restrictions.

1. From degree sequences to a polynomial

For a labeled tree T on $[n] = \{1, \dots, n\}$, where $n \geq 2$, write $d_i = \deg_T(i)$ and $a_i = d_i - 1$. Since T has $n-1$ edges,

$$\sum_{i=1}^{n} a_i = n - 2, \qquad a_i \geq 0.$$

The coefficient $[z^{n-2}](1 + z + \cdots + z^{n-2})^n$ counts the possible assignments of the excess degrees a_i; it does not count the trees realizing each assignment. What coefficients would supply the missing multiplicities? Our answer will be the factorially weighted polynomial

$$\boxed{(n-2)!\,[z^{n-2}] \left(\sum_{j=0}^{n-2} \frac{z^j}{j!} \right)^n .} \tag{1}$$

We shall prove *directly from the structure of trees* why this expression counts them. The key is a lemma about rooted forests.

2. Fill the child positions

Let $1 \leq k \leq n$ and prescribe nonnegative integers $a_1, \dots, a_n$ satisfying

$$\sum_{i=1}^{n} a_i = n - k.$$

Our aim is to construct all rooted forests on $[n]$ satisfying these prescribed numbers of children, and to count them.

More precisely, we consider forests of k rooted trees in which vertex i has exactly a_i children. We impose a linear ordering on the children of each vertex, *independently of the numerical labels they already have.* The k rooted components themselves are not ordered.

For example, if the children of vertex 4 are $2, 5, 8$, we might assign

$$(4,1) \longrightarrow 5, \qquad (4,2) \longrightarrow 2, \qquad (4,3) \longrightarrow 8.$$

This represents the ordering $(5, 2, 8)$. The second coordinate specifies the *child's position, not its vertex label.*

The construction. Begin with n isolated vertices, each regarded as a rooted tree of one vertex. Thus

$$F_0 = \{\{1\}, \{2\}, \dots, \{n\}\}$$

has n components. Give vertex i exactly a_i child positions $(i,1), \dots, (i, a_i)$. **We fill these positions in lexicographic order:** first in increasing order of vertex labels i, then in increasing order of position

numbers at each vertex. This order is fixed *before* any attachments. Thus the vertex receiving the next child is predetermined; we choose only which eligible component root becomes that child.

To fill a position belonging to vertex i, attach to i the root of any component *other than the component containing* i. Each attachment joins two components. After $n-k$ attachments, the resulting forest has exactly k rooted trees with the prescribed numbers of children. The number of components decreases as follows:

$$\boxed{n \longrightarrow n-1 \longrightarrow \cdots \longrightarrow k.}$$

Counting the choices. Since the receiving vertex is predetermined, the only choice is a child. When the current forest has m components, exactly one component contains that vertex; the roots of the other $m-1$ components are eligible. Thus the numbers of choices, in succession, are

$$n-1,\ n-2,\ \ldots,\ k.$$

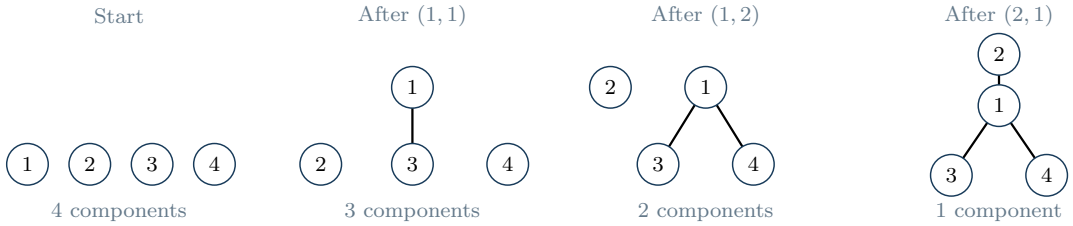


Figure 1. One construction for $(a_1, a_2, a_3, a_4) = (2, 1, 0, 0)$.
The three positions are filled in the displayed order; the available choice counts are $3, 2, 1$.

Child-position counting lemma. *For every* $(a_1, \ldots, a_n)$ *with* $\sum_i a_i = n-k$, *the number of such ordered-child rooted forests is*

$$\boxed{\frac{(n-1)!}{(k-1)!}.} \tag{2}$$

Proof. The vertex receiving each child is predetermined by the fixed order; we count only which eligible root becomes that child. Thus the product $(n-1)(n-2)\cdots k$ counts construction histories. We must show that *every completed ordered-child forest is obtained by exactly one such history.*

Take any completed ordered-child forest F with the prescribed child counts. We shall reconstruct F using our procedure. *Starting afresh with* n *isolated vertices*, insert the edges of F in our predetermined order of child positions. For example, the final tree in Figure 1 is reconstructed by inserting $1 \to 3$, then $1 \to 4$, then $2 \to 1$.

Consider the next child position (i, t) in the fixed order. *Look at the completed forest* F *and identify the vertex* v *occupying this position.* We must now insert the edge $i \to v$, placing v in child position t of i. This attachment is permitted: since the edge $i \to v$ has not yet been inserted, v has no parent in the partially constructed forest and is therefore the root of its component. Furthermore, i and v belong to different components; otherwise, inserting $i \to v$ would create a cycle in F.

Thus each edge of F is an allowed attachment. Because F specifies the occupant of every child position, the fixed order also determines its *unique* construction history. □

Freedom of order. The lexicographic order is chosen for convenience. Any linear ordering of the child positions fixed before the construction begins gives the same enumeration: the proof applies without change.

The factorials now have a meaning. Forget the ordering of children. Each ordinary rooted forest with the specified child counts has exactly $\prod_i a_i!$ orderings of its children. Therefore the number of

unordered-child rooted forests with these counts is

$$\boxed{F_{n,k}(\mathbf{a}) = \frac{(n-1)!}{(k-1)!\prod_{i=1}^{n} a_i!}.} \tag{3}$$

The factorial denominators merely compensate for child orderings introduced to make the construction transparent.

3. The rooted-tree polynomial and Cayley's identity

Take $k = 1$. Let $\mathcal{R}_n$ be the set of all rooted labeled trees on $[n]$, and let $c_i(R)$ be the number of children of vertex i. By (3), for each nonnegative $\mathbf{a}$ with $\sum_i a_i = n-1$, exactly $(n-1)!/\prod_i a_i!$ rooted trees have $c_i(R) = a_i$. Hence, by the multinomial theorem,

$$\boxed{H_n(x_1, \ldots, x_n) := \sum_{R \in \mathcal{R}_n} \prod_i x_i^{c_i(R)} = (x_1 + \cdots + x_n)^{n-1}.} \tag{4}$$

Now start instead with an *unrooted* labeled tree T, and choose its root r. The root has d_r children; every other vertex i has $d_i - 1$ children. Thus

$$\prod_i x_i^{c_i(T,r)} = x_r \prod_i x_i^{d_i(T)-1}.$$

Summing over all choices of r and then over all unrooted trees gives

$$H_n = (x_1 + \cdots + x_n) \sum_{T \in \mathcal{T}_n} \prod_{i=1}^{n} x_i^{d_i(T)-1}.$$

Compare with (4) and cancel the nonzero polynomial $x_1 + \cdots + x_n$. We obtain **Cayley's polynomial identity**:

$$\boxed{\sum_{T \in \mathcal{T}_n} \prod_{i=1}^{n} x_i^{d_i(T)-1} = (x_1 + \cdots + x_n)^{n-2}.} \tag{5}$$

At $x_1 = \cdots = x_n = 1$, this gives $|\mathcal{T}_n| = n^{n-2}$.

4. Return to the generating polynomial: degree restrictions

For a tree T, distinguish an edge e and delete it. The endpoints of e become the roots of two rooted trees. *Every* vertex i then has exactly $a_i = d_i(T) - 1$ children. Conversely, joining the two roots of any such forest recovers a unique pair (T, e): a tree with a distinguished edge. The two components are not ordered, so no additional factor of 2 is involved.

Fix positive integers $d_1, \ldots, d_n$ with $\sum_i d_i = 2(n-1)$. Each tree with these degrees has $n-1$ edges and hence gives rise to $n-1$ different pairs (T, e), or equivalently, to $n-1$ corresponding two-component rooted forests. Therefore,

$$(n-1)\,\#\{T \in \mathcal{T}_n : \deg_T(i) = d_i \text{ for all } i\} = F_{n,2}(d_1 - 1, \ldots, d_n - 1).$$

Applying Formula (3) with $k = 2$ gives

$$\boxed{\#\{T \in \mathcal{T}_n : \deg_T(i) = d_i \text{ for all } i\} = \frac{(n-2)!}{\prod_{i=1}^{n}(d_i - 1)!}.} \tag{6}$$

The denominator supplies the factorial weights promised in (1).

Arbitrary degree restrictions. If vertex i is required to have degree in $D_i \subseteq \{1, \ldots, n-1\}$, its permitted excess degrees are $d-1$ with $d \in D_i$. Summing (6) over admissible degree sequences yields

$$\boxed{N(D_1, \ldots, D_n) = (n-2)!\,[z^{n-2}] \prod_{i=1}^{n} \left(\sum_{d \in D_i} \frac{z^{d-1}}{(d-1)!} \right).} \tag{7}$$

With no restrictions on vertex degrees, each factor may be replaced by e^z without changing the coefficient of z^{n-2}. Thus (7) becomes

$$(n-2)!\,[z^{n-2}]e^{nz} = n^{n-2}.$$

If all degrees are at most 2, we may use $1+z$ for each factor, giving $(n-2)!\binom{n}{2} = n!/2$, the number of labeled paths ($n \geq 2$). If all degrees are at most 3, we may use $1+z+z^2/2$ for each factor, which counts labeled trees of maximum degree at most 3.

5. What this perspective adds

Prescribe, order, fill, count, and forget. Introducing ordered child positions makes the enumeration of ordered-child forests independent of the distribution of children among the vertices. Filling the positions individually in a predetermined order gives a simple product of successive numbers of eligible roots; each completed ordered-child forest has exactly one such construction history. Forgetting the child orders restores the dependence on the prescribed child counts through factorial denominators.

Relation to the literature. The prescribed-child-count formula (3) and the multivariate forest polynomial are classical; see Pitman [1, Theorem 6 and equation (16)]. Pitman [1, Section 2] counts by choosing the entire set of children of each vertex at once, and his coalescent construction [2] joins rooted components one edge at a time. Guo and Guo [3, Corollaries 4.8 and 4.9 in the preprint] use a recursive construction to count labeled plane trees with prescribed child counts, then remove the child orders. Ekhad and Zeilberger [4] state the specialization of (7) to a common set of permitted degrees, invoking Lagrange inversion.

The approach here connects these ideas through the direct filling of prescribed, ordered child positions. It explains the factorial weights of the generating polynomial without requiring Prüfer sequences or Lagrange inversion. We claim *expository clarity*, not historical priority.

A broader counting principle. Sometimes an object becomes easier to count when we give it more structure rather than less. Counting the enriched objects and then forgetting the auxiliary structure can reveal multiplicities that are otherwise less transparent.

References

[1] J. Pitman, "Enumerations of trees and forests related to branching processes and random walks," in *Microsurveys in Discrete Probability*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, **41** (1998), 163–180. 1997 technical report (PDF).

[2] J. Pitman, "Coalescent random forests," *Journal of Combinatorial Theory, Series A*, **85** (1999), 165–193. Online PDF.

[3] S. Guo and V. J. W. Guo, "A recursive algorithm for trees and forests," *Discrete Mathematics*, **340** (2017), 695–703. arXiv:1702.01744.

[4] S. B. Ekhad and D. Zeilberger, "Automated Counting and Statistical Analysis of Labeled Trees with Degree Restrictions," arXiv:2201.07326 (2022).

Concluding synthesis

We began with one enumeration: $|\mathcal{T}_n| = n^{n-2}$. We have arrived at it by eight routes. Looking back, we find that the routes are neither entirely separate nor merely different versions of the same argument. Certain mathematical ideas recur, sometimes in unexpectedly different forms.

1. Bijections, histories and constant-size fibers

Let $f : A \to B$ be a map between finite sets. If every element of B has exactly $m \geq 1$ preimages, then

$$\boxed{|A| = m|B|.}$$

In **Prüfer's proof**, the encoding $\mathcal{T}_n \to [n]^{n-2}$ is a bijection: every fiber has size 1. In **Joyal's proof**, functions on $[n]$ correspond bijectively to triples (T, a, b), where T is a labeled tree and a, b are an ordered pair of vertices, possibly equal. Forgetting the distinguished vertices gives fibers of size n^2, so $n^n = n^2|\mathcal{T}_n|$.

In **Pitman's proof**, each rooted tree extending a fixed k-component rooted forest F_k has precisely $(k-1)!$ histories. Thus

$$n^{k-1}(k-1)! = N(F_k)(k-1)!.$$

Prüfer establishes invertibility; Joyal uses invertibility followed by forgetting marks; Pitman establishes uniform multiplicity. The eighth perspective also introduces order to make a construction reversible, then forgets the order, dividing by the number of orderings for the prescribed child counts.

When a direct bijection is difficult to construct, look for a correspondence having constant-size fibers. The degree classes in the next sections, by contrast, need not have equal sizes.

2. Forests: enlarging the family of objects

Pitman starts with a forest and joins its components. Riordan and Rényi delete a distinguished root, leaving a forest. The child-position construction starts with isolated vertices and stops at a prescribed number of components. Their three enumeration problems impose different restrictions:

Pitman: a fixed rooted forest F_k with k components must be extended, preserving its arcs. The number of rooted extensions is

$$\boxed{N(F_k) = n^{k-1}.}$$

Riordan–Rényi: a particular set of k roots is prescribed. The number of rooted forests is

$$\boxed{T_{n,k} = k\, n^{n-k-1}.}$$

Child positions: the number of children a_i of each vertex is prescribed, with $\sum_i a_i = n - k$. The number of unordered-child rooted forests is

$$\boxed{F_{n,k}(a_1, \ldots, a_n) = \frac{(n-1)!}{(k-1)! \prod_i a_i!}.}$$

The connection between the second and eighth perspectives is particularly revealing. In Pitman's construction, when there are j components, we have

$$\underbrace{n}_{\text{receiving vertex}} \cdot \underbrace{(j-1)}_{\text{eligible child roots}}$$

choices. In the child-position construction the receiving vertex is predetermined by the next position, leaving just $j-1$ choices. This explains the product $(n-1)(n-2)\cdots k$. The two constructions use the same elementary operation but organize the choices differently. What must be divided out is also different: the ordering of inserted arcs in Pitman's proof and the ordering of children in the child-position proof.

3. Two roles for probability—and a bridge between functions and trees

In Pitman's construction, choosing uniformly among admissible arcs makes every history equally probable. Each terminal tree has the same number of histories, so the resulting rooted tree is uniform among the extensions of F_k. At a stage with j components, two particular components of sizes a, b merge next with probability

$$\frac{a+b}{n(j-1)}.$$

This is the additive-coalescent merger rule.

Chapuy and Perarnau start with a uniformly random function. Conditional on the exploration skeleton, the probability of exactly one cyclic vertex is

$$\frac{1}{T_1}\cdot\frac{T_1}{T_2}\cdots\frac{T_{K-1}}{T_K}=\frac{1}{n}.$$

The intermediate quantities depend on the skeleton; the final product does not. In Pitman's proof, the history count is independent of the terminal extension. In Chapuy–Perarnau, the conditional probability is independent of the exploration skeleton. In both, a quantity that appears to depend on a complicated structure turns out to be constant.

There is a more direct bridge from Joyal to Chapuy–Perarnau. In Joyal's bijection, the cyclic set M becomes the vertex set of the path from a to b. Consequently,

$$|M| = d_T(a,b)+1.$$

In particular, $|M|=1$ if and only if $a=b$. For every tree there are n coincident ordered pairs among the n^2 ordered pairs of distinguished vertices. Hence Joyal's bijection gives

$$\Pr\{|M|=1\}=\frac{n|\mathcal{T}_n|}{n^2|\mathcal{T}_n|}=\frac{1}{n}.$$

The restriction of Joyal's bijection to these coincident endpoints is exactly the rooted-tree correspondence used in the fifth perspective: remove the unique loop and forget the remaining directions. Chapuy and Perarnau obtain the same probability without invoking Joyal, by revealing a function progressively. Their proofs are distinct, but the event whose probability telescopes has a particularly simple interpretation in Joyal's representation.

4. Beyond counting: Cayley's polynomial identity

Three perspectives yield the same stronger identity:

$$\boxed{\sum_{T\in\mathcal{T}_n}\prod_{i=1}^{n} x_i^{\deg_T(i)-1} = (x_1+\cdots+x_n)^{n-2}.}$$

Setting all variables equal to 1 gives the original formula. In **Prüfer**, the exponent $\deg_T(i)-1$ equals the number of occurrences of i in the word, so summing weights of words gives the power on the right. In **Rényi's symmetry proof**, setting one variable equal to zero retains just the trees in which the corresponding vertex is a leaf. Leaf deletion yields an induction; symmetry and low degree determine the polynomial. In **child positions**, the rooted-tree polynomial is first obtained by counting ordered children and forgetting their order. Summing over possible roots then supplies the factor $x_1+\cdots+x_n$, which cancels.

All three routes yield, for $d_i \geq 1$ and $\sum_i d_i = 2n-2$,

$$\boxed{\#\{T\in\mathcal{T}_n : \deg_T(i) = d_i\ \forall i\} = \frac{(n-2)!}{\prod_i (d_i-1)!}.}$$

The factorials have two combinatorial interpretations: repeated labels in Prüfer words and ordered children in rooted forests (followed by forgetting the root). Rényi determines the same coefficients without explicitly counting either kind of ordering. A numerical identity becomes more illuminating when recognized as a specialization of a polynomial identity.

5. From polynomials to determinants

Assign weight x_ix_j to an edge $\{i,j\}$. For any spanning tree T,

$$\prod_{\{i,j\}\in E(T)} x_ix_j = \Big(\prod_i x_i\Big)\prod_i x_i^{\deg_T(i)-1}.$$

The weighted Matrix–Tree Theorem follows from the same Cauchy–Binet argument as in Perspective III: with B_r the reduced incidence matrix and W diagonal with the edge weights, expand $\det(B_rWB_r^{\mathsf{T}})$. Each selected edge set contributes its weight product times the square of its reduced incidence determinant, which is 1 for a spanning tree and 0 otherwise.

For the complete graph, therefore, Cayley's polynomial identity gives

$$\boxed{\det L_w^{(r)} = \Big(\prod_i x_i\Big)(x_1+\cdots+x_n)^{n-2},}$$

where $L_w^{(r)} = B_rWB_r^{\mathsf{T}}$. Setting all variables equal to 1 recovers the determinant evaluation in Perspective III. The same combinatorial information is encoded in a polynomial and in a graph matrix.

6. A forest identity hidden in the tree polynomial

Introduce an additional vertex 0 and consider trees on $\{0, 1, \ldots, n\}$. Set $x_0 = t$ and all other variables equal to 1 in Cayley's polynomial identity:

$$\sum_{T \in \mathcal{T}(\{0,1,\ldots,n\})} t^{\deg_T(0)-1} = (t+n)^{n-1}.$$

The coefficient of t^{k-1} counts trees in which vertex 0 has degree k. It is $\binom{n-1}{k-1} n^{n-k}$. Deleting vertex 0 produces a forest of k trees rooted at its former neighbors. Conversely, adjoining vertex 0 and connecting it to every root recovers the tree. By relabeling, each of the $\binom{n}{k}$ possible root sets occurs equally often. The number with a prescribed root set of size k is therefore

$$T_{n,k} = \frac{\binom{n-1}{k-1} n^{n-k}}{\binom{n}{k}} = \boxed{k\, n^{n-k-1}}.$$

This is the Riordan–Rényi forest formula, recovered from the polynomial identity. It is not another recurrence: it is a direct bridge between Perspectives VI and VII.

There is a second bridge, this time from Perspective VIII to Perspective VI. Sum the child-position formula over all vectors with $\sum_i a_i = n - k$. The multinomial theorem gives the number of rooted forests with an arbitrary set of k roots:

$$\sum_{\sum_i a_i = n-k} F_{n,k}(a_1, \ldots, a_n) = \binom{n-1}{k-1} n^{n-k}.$$

Dividing by $\binom{n}{k}$ again gives the prescribed-root formula. Thus the same forest enumeration arises from three different routes: deletion and recursion, adjoining a vertex and extracting a coefficient, and summing prescribed child counts.

7. What changes when we change our perspective?

The eight proofs have revealed techniques that extend far beyond the enumeration of trees: encoding by words, counting histories rather than objects, algebraic indicators, reversible transformations, conditioning and telescoping, enlarging a family before recursing, determination by symmetry and degree, and introducing order before accounting for its multiplicity.

We began with the question:

Why n^{n-2}?

We end with a broader question:

When we know one proof of a theorem, what mathematics might we discover by seeking another?