# VARIOUS PROOFS OF CAYLEY'S FORMULA

JEFF SOOSIAH

## 1. Background

Cayley's formula counts the number of labeled trees on $n$ vertices. Put another way, it counts the number of spanning trees of a complete graph $K_n$. Note that it does not count the number of nonisomorphic trees on $n$ vertices. For comparison, there are 6 nonisomorphic trees on 6 vertices, while there are $6^4 = 1296$ labeled trees on 6 vertices. The formula was first discovered by Borchardt in 1860, and extended by Cayley in 1889. Cayley was also the one to use graph theory terms in his paper. His name was the one associated with the formula since then.

## 2. Theorem, Cayley's Formula (Cayley 1889)

Let $T_n$ denote the number of trees on $n$ labeled vertices. Cayley's formula states:

$$T_n = n^{n-2}$$

## 3. Bijection (Prüfer 1918)

*Proof.* For a tree T, consider its vertex set $N = \{1, 2, ..., n\}$. Note that the number of sequences of length $n - 2$ from $N$ is $n^{n-2}$. The goal is thus to construct a bijection between the set of trees on $n$ labeled vertices and the set of these sequences.

To convert a labeled tree with vertices $\{1, 2, ..., n\}$ into a sequence of length $n - 2$, continue to remove the lowest labeled leaf until two vertices remain. Each time a leaf is removed, add its neighbor to the list.

To convert a sequence $S = (t_1, t_2, ..., t_{n-2})$ into a labeled tree $T$, let $s_1$ the first vertex of $N \backslash S$, and join $s_1$ to $t_1$. Then let $s_2$ be the first vertex of $N \backslash \{s_1\} \backslash S$, and join $s_2$ to $t_2$. Continue until the elements of $S$ have been exhausting, at which point $n - 2$ edges have been added. Join the two vertices of $N \backslash \{s_1, s_2, ..., s_{n-2}\}$ to complete the construction of $T$. □

## 4. Bijection (Joyal 1981)

*Proof.* Consider the set of all labeled trees together with two distinguished vertices: the left end, and the right end, and call this set $\mathcal{T}_n$. Then in a labeled tree, there are $n$ choices for the left end and $n$ choices for the right end, so $|\mathcal{T}_n| = n^2 T_n$, so the goal is to prove $|\mathcal{T}_n| = n^n$. The set $N^N$ of all mappings from $N$ into $N$ has size $n^n$, so a bijection from $N^N$ onto $\mathcal{T}_n$ will suffice.

Let $f : N \to N$ be any map, and represent a graph $\vec{G}_f$ with directed edges that start at $i$ and end at $f(i)$. Because each vertex has one edge emanating from it, each component contains an equal number of edges and vertices, so each contains exactly one directed cycle. Let $M \subseteq N$ be the union of the vertex sets of all cycles in $\vec{G}_f$. Now consider $f|_M = \begin{pmatrix} \text{a} & \text{b} & \dots & \text{z} \\ \text{f(a)} & \text{f(b)} & \dots & \text{f(z)} \end{pmatrix}$ where $a, b, \dots, z$ are ordered naturally. Then let $f(a)$ be the left end and $f(z)$ the right end.

To construct the tree $T$ according to $f$, draw $f(a), \dots, f(z)$ as a path from $f(a)$ to $f(z)$, then fill in the remaining vertices from $\vec{G}_f$ (discarding edge direction).

Given a tree $T$, observe the unique path $P$ from the left end to the right end, which gives the set $M$ and the mapping $f|_M$. Then fill in the remaining correspondences $i \to f(i)$ by the unique paths from $i$ to $P$. □

## 5. Double Counting (Pitman 1999)

*Proof.* Let $\mathcal{F}_{n,k}$ denote the set of all rooted forests on $n$ vertices with $k$ rooted trees. Note that $\mathcal{F}_{n,1}$ is then the set of all rooted trees, and that $|\mathcal{F}_{n,1}| = nT_n$ because every tree has $n$ choices for the root. Then let $F_{n,k} \in \mathcal{F}_{n,k}$ denote a directed graph with such properties. Say a forest $F$ contains a forest $F'$ if $F$ contains $F'$ as a directed graph. If $F$ contains $F'$, $F$ has less components than $F'$ (see the figure below: $F_1$ contains $F_2$; $F_2$ contains $F_3$).

Say $F_1, ..., F_k$ is a refining sequence if $F_i \in F_{n,i}$ and $F_i$ contains $F_{i+1}$ for all $i$ (see the figure below: $F_1, F_2, F_3$ is a refining sequence). Let $F_k \in \mathcal{F}_{n,k}$ be a fixed forest. Then let $N(F_k)$ denote the number of rooted trees containing $F_k$, and $N*(F_k)$ denote the number of refining sequences ending in $F_k$. Count $N*(F_k)$ in two ways: first by starting at an $F_1$ and second by starting at $F_k$.

If an $F_1 \in \mathcal{F}_{n,1}$ contains $F_k$, then delete the $k-1$ edges of $F_1 \backslash F_k$ from $F_1$ in any order to get a refining sequence from $F_1$ to $F_k$. Therefore:

$$N*(F_k) = N(F_k)(k-1)! \tag{1}$$

Starting from $F_k$, to produce an $F_{k-1}$, add a directed edge that starts at any vertex and ends at any of the $k-1$ roots that are not in the same component as the start vertex. This amounts to $n(k-1)$ choices.

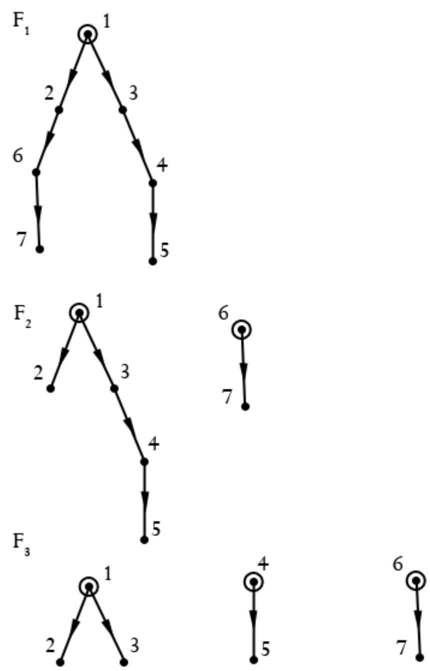
Then going from $F_{k-1}$ to $F_{k-2}$, add a directed edge that starts at any vertex and ends at any of the $k-1$ roots not in the same component as this start vertex, which is $n(k-2)$ choices. Continuing until $F_1$ yields:

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

Equating (1) and (2) provides the solution to $N(F_k)$:

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

Now let $k = n$. Then $F_n$ consists of $n$ vertices and $n$ components: or simply $n$ isolated vertices. Therefore $N(F_n)$ counts the number of all rooted trees (denoted by $\mathcal{F}_{n,1}$), so $|\mathcal{F}_{n,1}| = n^{n-1} \implies T_n = n^{n-2}$. □



## 6. References

(1) Bondy. Graph Theory With Applications. Elsevier Science Publishing Co., Inc. 1976.
(2) Aigner. Proofs from the Book. Springer 2004.
(3) Weisstein, Eric W. Prüfer Code. From MathWorld–A Wolfram Web Resource.