Cayley's Tree Formula via Lagrange Inversion

Pitchayut Saengrungkongka

January 15, 2026

This is the notes for 18.A34 lecture on generating functions. The goal is to prove the following theorem.

Theorem 0.1 (Cayley's Tree Formula)**.**

There are exactly n^{n-2} spanning trees on vertices $1, 2, \ldots, n$.

There are many beautiful proofs of this fact. However, what we are going to do here is to try the most straightforward approach to it, and see what it leads to.

§1 Recursion

We let t_n denote the number of spanning trees on n vertices. We casework on

- the degree d of vertex 1.
- the number of vertices $a_1, a_2, \ldots, a_d$ of the subtrees corresponding to neighbors of 1.

Given $a_1, \ldots, a_d$ such that $a_1 + \cdots + a_d = n - 1$, one count the number of trees by:

- partition the $n-1$ vertices into d indistinguishable groups of sizes $a_1, \ldots, a_d$, which gives $\frac{(n-1)!}{d! a_1! \ldots a_d!}$ ways,
- form trees on each groups, giving $t_{a_1} \cdots t_{a_d}$ ways; and
- pick a designated vertex for each tree to connect to 1, giving $a_1 \cdots a_d$ ways.

Taking product of the above three steps and then summing across all $(a_1, \ldots, a_d)$, we find that for all $n \geq 1$,

$$\begin{aligned} t_n &= \sum_{d=0}^{n-1} \sum_{a_1+\cdots+a_d=n-1} \frac{(n-1)!}{d! a_1! \ldots a_d!} (t_{a_1} \cdots t_{a_d})(a_1 \cdots a_d) \\ &= \sum_{d=0}^{\infty} \frac{1}{d!} \sum_{a_1+\cdots+a_d=n-1} \frac{(n-1)!}{(a_1-1)! \cdots (a_d-1)!} t_{a_1} \cdots t_{a_d}. \end{aligned}$$

Note that the equation above also implies $t_1 = 1$ because empty product is 1. Thus, this equation uniquely determines t_n, so theoretically we should be able to forget combinatorics and solve for t_n now.

§2 Generating Functions and Lagrange Inversion Formula

Motivated by the recurrence relation above, we let

$$T(x) = \sum_{n=1}^{\infty} \frac{t_n}{(n-1)!} x^n.$$

We will not dwell into details about convergence issue and only note the following: throughout this talk, all power series will be thought of an analytic function defined over some sufficiently small neighborhood of 0. This is sufficient to justify convergence.

From the recursion above, we have

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{t_n}{(n-1)!} x^n &= \sum_{n=1}^{\infty}\sum_{d=0}^{\infty} \frac{1}{d!} \sum_{a_1+\cdots+a_d=n-1} \frac{t_{a_1}}{(a_1-1)!}\cdots\frac{t_{a_d}}{(a_d-1)!} x^n \\ T(x) &= x\cdot\sum_{d=0}^{\infty}\frac{1}{d!}\sum_{a_1,\dots,a_d}\frac{t_{a_1}}{(a_1-1)!}\cdots\frac{t_{a_d}}{(a_d-1)!}x^{a_1+\cdots+a_d} \\ &= x\left(1+\cdot\sum_{d=0}^{\infty}\frac{1}{d!}T(x)^d\right) \\ &= xe^{T(x)}.\end{aligned}$$

Therefore, we have to find the coefficient of x^n in $T(x)$, where $T(x)$ is the function such that

$$T(x) = xe^{T(x)}.$$

At this point, we get a particular nice description of $T(x)$, but we seem to be hopeless about extracting the coefficients from its power series. However, it turns out that there is a tool to do this easily: **Lagrange Inversion Formula**.

For each integer n and a Laurent series $F(x)$ (which is a power series of the form $\sum_{n=-k}^{\infty} a_n x^n$), we let $[x^n]F(x)$ denote the coefficient in front of x^n of F.

Theorem 2.1 (Lagrange Inversion Formula)**.**

Let $F(x)$ and $G(x)$ be two power series such that $F(G(x)) = x$ and $G(0) = 0$. Then

$$[x^n]F(x) = \frac{1}{n}([x^{-1}]G(x)^{-n})$$

Applying this to $F(x) = T(x)$ and $G(x) = xe^{-x}$, we get that

$$\begin{aligned}\frac{t_n}{(n-1)!} = [x^n]T(x) &= \frac{1}{n}[x^{-1}](xe^{-x})^{-n} \\ &= \frac{1}{n}[x^{n-1}]e^{nx} \\ &= \frac{1}{n}\frac{n^{n-1}}{(n-1)!},\end{aligned}$$

which indeed gives $t_n = n^{n-2}$.

Thus, all that remains is to prove the Lagrange Inversion Formula. The proof that we will present uses complex analysis. It is possible to phrase the argument formally (i.e., working entirely with coefficients) so that it does not involve any complex analysis, but in my opinion, this obscures the motivation behind the proof.

§3 Complex Analysis Crash Course

§3.1 Holomorphic Function

In what follows, let U be an open, simply-connected (i.e., no holes) subset of $\mathbb{C}$.

Definition 3.1.

A function $f : U \to \mathbb{C}$ is said to be **differentiable at point p** if

$$L = \lim_{z \to p} \frac{f(z) - f(p)}{z - p}$$

exists. Equivalently, for all $\varepsilon > 0$, there exists $\delta > 0$ such that if $|z - p| < \delta$, then

$$\left| \frac{f(z) - f(p)}{z - p} - L \right| < \varepsilon.$$

The derivative at p, denoted $f'(p)$ is equal to the limit L.

A function $f : U \to \mathbb{C}$ is **holomorphic** if it is differentiable at every point in U.

Example 3.2.

$f(z) = z^2$, $f(z) = e^z$, $f(z) = \sin z$, and $f(z) = e^{\sin z} + (z^2 + 1)\cos z$ are all holomorphic in the entire $\mathbb{C}$.

$f(z) = \overline{z}$ is not differentiable at any point in $\mathbb{C}$.

Differentiability in $\mathbb{C}$ is a fairly strong condition. Thus, complex analysis works on a narrower class of functions than real analysis, but those are functions that are nicely behaved that we can get striking results out of it. One shocking fact (that we will not prove) is

Fact 3.3.

If $f : U \to \mathbb{C}$ is holomorphic, then

(a) f is infinitely differentiable at every point in U.

(b) for any point $p \in U$, f has a power series at p that converges around an open neighborhood of U.

Enough differentiation, let's do integration. The complex plane is 2-dimensional, so we are allowed to integrate along the curve. I won't formally define it, but I will do an example.

Example 3.4.

Let γ be the curve tracing along the unit circle $|z| = 1$ counterclockwise. Let's compute for each integer n the integral

$$I_n := \oint_\gamma z^n \, dz,$$

where we integrate z^n along γ. In order to do this, we parametrize γ by $z = e^{it}$, where t ranges from 0 to 2π. We have $dz = ie^{it}\,dt$, so the integral becomes

$$I_n = \int_0^{2\pi} e^{int} \cdot ie^{it}\,dt = i\int_0^{2\pi} e^{i(n+1)t}\,dt = \begin{cases} 0 & n \neq -1 \\ 2\pi i & n = -1 \end{cases}.$$

The same also holds even when γ is a circle $|z| = r$ counterclockwise for any $r > 0$.

In particular, if f has a power series around 0: $f(z) = a_0 + a_1 z + a_2 z^2 + \ldots$, then intuitively, $\oint_\gamma f(z)\,dz$ should be 0 because each term in the power series integrates to 0. The following result is true, and is proven before the fact that holomorphic function has a power series.

Theorem 3.5 (Cauchy-Goursat)**.**

Let $f : U \to \mathbb{C}$ be holomorphic, and let γ be a loop inside U. Then

$$\oint_\gamma f(z)\,dz = 0.$$

§3.2 Meromorphic Functions

As you can see above, integrating a holomorphic function isn't much fun. We now consider a quotient of two holomorphic functions, called **meromorphic function**.

Definition 3.6.

A function $f : U \dashrightarrow \mathbb{C}$ is **meromorphic** if it is a quotient of two holomorphic functions $U \to \mathbb{C}$.

The dashed arrow $\dashrightarrow$ is to emphasize that such functions need not be defined everywhere; it's not defined at the points at which the denominator is 0. The **pole** is where the function is undefined.

This time, the integral of a meromorphic function f around a loop need not be 0. However, there is a stunning formula that allows us to determine this integral using only local information around the poles.

First, for any point $p \in U$ and any meromorphic $f : U \dashrightarrow \mathbb{C}$, there is a Laurent series expansion that converges around a neighborhood of p:

$$f(z) = \sum_{n=-m}^{\infty} a_n (z-p)^n.$$

As we have seen when integrating z^n around a contour, the power that matters is z^{-1}. This motivates the following definition.

Definition 3.7.

The **residue** at $z = p$ of f (denoted $\operatorname{Res}_{z=p} f$) is the coefficient a_{-1} in front of $(z-p)^{-1}$.

Example 3.8.

Consider the function $f(z) = \frac{1}{(z-1)(z+2)}$ and $p = 1$, which is a pole. We can expand the Laurent series at $z = 1$ by

$$\frac{1}{(z-1)(z+2)} = \frac{1}{3}(z-1)^{-1} - \frac{1}{9} + \frac{1}{27}(z-1) - \dots,$$

so $\operatorname{Res}_{z=1} f(z) = \frac{1}{3}$.

Now, we can state the big theorem.

Theorem 3.9 (Residue theorem).

For any meromorphic function $f : U \dashrightarrow \mathbb{C}$ and a simple loop γ, we have

$$\oint_\gamma f(z)\,dz = 2\pi i \sum_{\text{poles } p \text{ inside } \gamma} \operatorname{Res}_{z=p} f(p).$$

▶ *Proof Sketch.* Draw new loop γ' that takes a detour at every pole and make a small loop close to them. The integral along γ' is 0 by Cauchy-Goursat. The contribution around each pole is $2\pi i \operatorname{Res}_{z=p} f(p)$ from integrating z^n as in Example 3.4. This gives the result. □

§4 Proof of Lagrange Inversion

▶ *Proof.* Take a very small loop γ around 0. Then we have

$$\begin{aligned}
[x^n]F(x) &= [x^{-1}]\frac{F(x)}{x^{n+1}} \\
&= \operatorname{Res}_{z=0} \frac{F(z)}{z^{n+1}} \\
&= \frac{1}{2\pi i} \oint_\gamma \frac{F(z)}{z^{n+1}}\,dz && \text{(Residue Theorem)} \\
&= \frac{1}{2\pi i} \oint_{\gamma'} \frac{w}{G(w)^{n+1}} G'(w)\,dw && \text{(Substitute } z = G(w)) \\
&= \frac{1}{2\pi i} \left(-\oint_{\gamma'} \frac{-1}{nG(w)^n}\,dw \right) && \text{(Integration by parts)} \\
&= \frac{1}{n}[x^{-1}]G(w)^{-n} && \text{(Residue Theorem)}
\end{aligned}$$

as desired. Here, γ' is image $G(\gamma)$, which is still a small loop around 0 because $F(0) = G(0) = 0$. Notice that we secretly change the contour, but the residue theorem guarantees that the integral remains the same. □