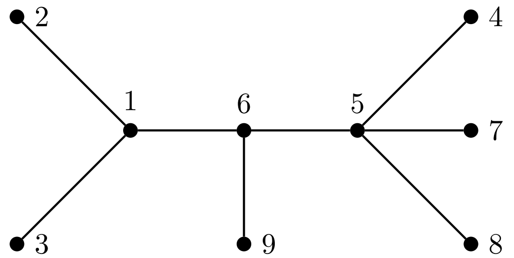
Math 374
Cayley's Formula

Definition. Let S be an n-element subset of $\mathbb{Z}$, and define $\mathcal{T}_S$ to be the set of all trees of order n whose vertices are labeled with distinct elements from S. In particular, if $S = \{1, 2, 3, \ldots, n\}$, then we write $\mathcal{T}_n$ instead of $\mathcal{T}_S$.

For example, the following tree is an element in $\mathcal{T}_9$.



Counting the number of labeled trees we get the following interesting pattern:

n	$\lvert\mathcal{T}_n\rvert$
1	$1 = 1^{-1}$
2	$1 = 2^0$
3	$3 = 3^1$
4	$16 = 4^2$
5	$125 = 5^3$

This strongly suggests our next theorem due to the British mathematician Arthur Cayley (1821 - 1895).

Theorem. *(Cayley's formula) Let $n \geq 1$ and let S be any n-element subset of $\mathbb{Z}$. Then $|\mathcal{T}_S| = n^{n-2}$.*

There are many known proofs of this famous theorem. The techniques used in these proofs range from clever bijections to a beautiful application of determinants from linear algebra. Here we will give a bijective proofs due to the German mathematician Heinz Prüfer (1896 - 1934).

To begin we define the following set of words.

Definition. With $n \geq 1$ and S any n-element subset of $\mathbb{Z}$, set W_S to be the set of all words with length $n - 2$ whose "letters" are elements of S. As before we write W_n when $S = \{1, 2, \ldots, n\}$.

As $|W_S| = n^{n-2}$, this suggests that we should try to establish a bijection between $\mathcal{T}_S$ and W_S as this implies

$$|\mathcal{T}_S| = |W_S| = n^{n-2}. \qquad (\star)$$

To do this we need a mapping between labeled trees and words. Prüfer found just a mapping!

Prüfer's map:

Input: $T \in \mathcal{T}_S$ for $|S| \geq 3$.
Set $T_1 = T$
for i from 1 to $n - 2$
   a) Let ℓ_i be the smallest leaf in T_i and let w_i be its (only) neighbor.
   b) Set $T_{i+1} = T_i - \ell_i$
Output: $w = w_1, w_2, \ldots, w_{n-2} \in W_S$

In this context we refer to w as the **Prüfer code** for T. To familiarize yourself with this algorithm, check that applying Prüfer's map to the tree above gives

$$1165556 \in W_9.$$

You should note, after tracing through this example, that the algorithm stops when we get down to only two remaining vertices.

It now only remains to prove that Prüfer's mapping is a bijection. To simplify this proof, we first a couple observations.

Observations. Let S be a subset of $\mathbb{Z}$ with $n \geq 3$ elements and fix $T \in \mathcal{T}_S$. Let $w = w_1, \ldots, w_{n-2}$ be the word obtained by applying Prüfer's map to T.

1. If T contains a vertex with label k and degree d, then w contains the letter k exactly $d-1$ times. Consequently, the only letters that do not appear in w are those that have degree 1 in T, i.e., the leaves of T.

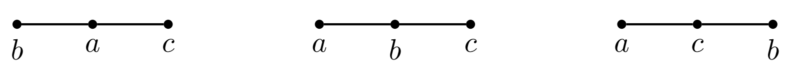
2. It follows from 1 that, the smallest letter ℓ not appearing in w is the smallest label among all the leaves in T. It follows from the definition of Prüfer's map that $\ell \sim_T w_1$ that the Prüfer code for $T - \ell$ is
$$w_2, w_3, \ldots, w_{n-2}.$$

With these two observations we are now ready to prove Cayley's formula.

Proof of Cayley's formula. As the table above shows, Cayley's formula holds for all sets with cardinality 1 or 2. As a result, we restrict our attention to sets whose cardinality is at least 3. (This is good since Prüfer's map is only defined for such sets!)

Instead of showing that Prüfer's map is separately 1-1 and onto it, we instead show that every $w \in W_S$ is the Prüfer code for a *unique* $T \in \mathcal{T}_S$. This will prove that Prüfer's map between $\mathcal{T}_S$ and W_S is a bijection. As mentioned in ($\star$), this is sufficient.

We proceed by induction on the cardinality of S. For our base case, let $S = \{a < b < c\}$ and observe that, in this case, $W_S = S$. Additionally, $\mathcal{T}_S$ consists of the trees



As the Prüfer codes for these trees are a, b, and c, respectively it follows that each word in W_S is the Prüfer code for a unique tree in $\mathcal{T}_S$.

Now for some $n \geq 3$, assume for induction that our statement holds for all n-element subsets of $\mathbb{Z}$. For the induction step, let S be an $(n+1)$-element subset of $\mathbb{Z}$ and consider an arbitrary word $w = w_1, w_2, \ldots, w_{n-1} \in W_S$. As w consists of $n-1$ elements drawn from the $(n+1)$-elements set S, there must be at least one letter in S that does not appear in w. Set ℓ to be the smallest such letter.

Now consider the n-element set $S' = S \setminus \{\ell\}$ and observe that $w' = w_2, \ldots, w_{n-1} \in S'$, since ℓ is not an letter in w. By induction we know there exists a unique tree $T' \in \mathcal{T}_{S'}$ whose Prüfer code is w'. Now define T to be the tree obtained from T' by inserting a leaf labeled ℓ so that its neighbor is the vertex labeled w_1. Certainly T has Prüfer code w.

It only remains to show that T is the unique such tree. To this end let Q be another labeled tree with Prüfer code w. By Observation 2, we know that the smallest label among all the leaves in Q is ℓ and that $Q - \ell$ has Prüfer code w'. As T' is the only tree with Prüfer code w' it follows that
$$Q - \ell = T' = T - \ell.$$
As the leaf labeled ℓ is adjacent to the vertex labeled w_1 in both Q and T we conclude that $T = Q$. This proves that T is the unique tree in $\mathcal{T}_S$ with Prüfer code w as needed. $\square$