Theorem If T is a tree, then $|E(T)| = |V(T)| - 1$.

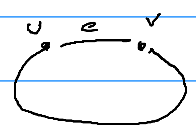
Proof. We prove this result by induction on $|V(T)|$. The claim holds if $|V(T)| = 1$, since the only tree with one vertex has no edges.

Now consider a tree T with $|V(T)| = n+1 \geq 2$. By the tree induction lemma T has a leaf v, and by removing v from T along with its single incident edge we get a tree T' with $|V(T')| = n$. By the induction hypothesis T' has $|V(T')| - 1 = n-1$ edges. Since T has exactly one more edge than T', namely the single edge incident to v, T must have $n-1+1 = n$ edges. This implies the claim. $\square$

2.3 Characterisations of Trees

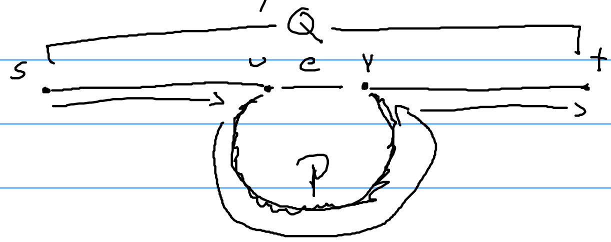
Theorem A graph G is a tree if and only if it is a minimal connected graph, i.e., if G is connected and removing any edge from $E(G)$ yields a graph that is not connected.

Proof. For the direction from left to right assume that G is a tree, i.e., connected and acyclic. Assume for contradiction that G is not minimal, so there exists an edge $e \in E(G)$ with endpoints $u, v \in V(G)$ such that the graph G' with $V(G') = V(G)$ and $E(G') = E(G) \setminus \{e\}$ is connected. G' then contains a u-v-path, which does not contain e. This path, together



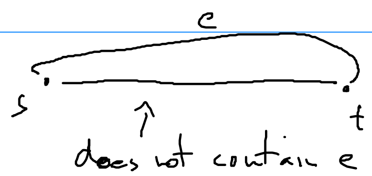
with the edge e, forms a cycle in G, which is a contradiction to acyclicity of G.

For the direction from right to left, assume that G is a minimal connected graph. Assume for contradiction that G is not a tree, i.e., that it contains a cycle. Let $e \in E(G)$ be an edge on that cycle, and $u, v \in V(G)$ its endpoints. Let P be the u-v-path obtained by removing e from the cycle.



We claim that the graph G' with $V(G') = V(G)$ and $E(G') = E(G) \setminus \{e\}$ is connected, which would contradict minimality of G.
To see that G' is connected, consider $s, t \in V(G')$. Since G is connected, there exists an s-t-path Q in G. If Q does not contain e, it is an s-t-path in G'. If Q does contain e, then there exists an s-t-walk in G' that first follows Q from s to the first endpoint of e, then follows P to the other endpoint of e, and finally continues along Q to t. It follows that G' is connected, which contradicts minimality of G. □

<u>Theorem</u> A graph G is a tree if and only if it is a maximal acyclic graph, i.e., if it is acyclic but adding any edge creates a cycle.
Proof. For the direction from left to right, assume that G is a tree, i.e., connected and acyclic. Let e be an edge <u>not</u> in $E(G)$, $s, t \in V(G)$ its endpoints. Let G' be the graph with $V(G') = V(G)$ and $E(G') = E(G) \cup \{e\}$. Since G is connected, there exists a path $v_0 e_1 v_1 \dots e_m v_m$ with $v_0 = s$ and $v_m = t$ in G. Then $v_0 e_1 v_1 \dots e_m v_m e v_0$ is a cycle in G', so G was maximal acyclic.

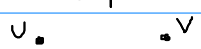


For the direction from right to left, assume that G is acyclic but adding any edge creates a cycle. Consider $u, v \in V(G)$. If G contains an edge with endpoints u and v, it contains a u-v-path. If G does not contain an edge with endpoints u and v, then by maximality of G, the graph G' with $V(G') = V(G)$ and $E(G') = E(G) \cup \{e\}$, where e is an edge with endpoints u and v, contains a cycle $ue_1v_1 \dots e_m v e u$. Then the path $ue_1v_1 \dots e_m v$ is a u-v-path in G. So G is connected. □

Theorem A graph G is a tree if and only if it does not contain any loops and contains a unique u-v-path for every $u, v \in V(G)$.

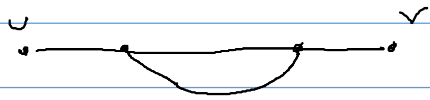
Proof idea:

"no path"



$\rightarrow$ not connected

"more than one path"



$\rightarrow$ cycle

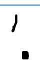
2.4 Counting Trees

recall: there are $2^{\binom{n}{2}}$ distinct simple graphs with n vertices

How many trees are there with n vertices?

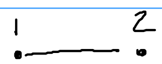
number of trees

$n=1$



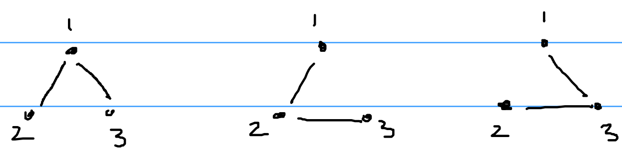
1

$n=2$

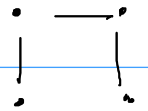
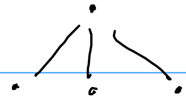


1

$n=3$



3

$n=4$   16

$\frac{4\cdot 3\cdot 2}{2} = 12$ of these   4 of these   oeis.org

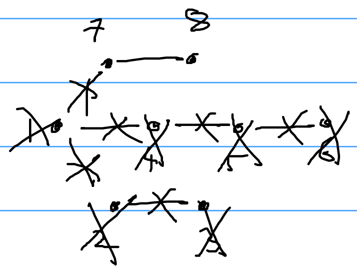
n^{n-2}

$[n]=$
$[4]=\{1,2,3,4\}$
$n-2=2$

<u>Algorithm</u> (Prüfer code) Let T be a tree with $V(T)=[n]$ where $n\geq 2$. Then the <u>Prüfer code</u> of T is a sequence of length $n-2$ over $[n]$ constructed in the following way. Start with the empty sequence, then repeat the following steps until two vertices remain in T:

1. Let v be the leaf of T with the smallest label.
2. Let u be the unique neighbour of v in T. Write down u as the next element in the sequence.
3. Remove v from T, along with its unique incident edge.

<u>Example</u>



T $|V(T)|=8$

Prüfer code of T: $2,1,5,4,1,7$

length of the code: is $|V(T)|-2=6$

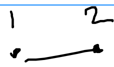
<u>Theorem</u> For any $n\geq 2$, there are n^{n-2} distinct trees with n vertices.

Proof. For any $n\geq 2$ the algorithm defines a function from the set of trees with vertex set $[n]$ to the set of sequences of length $n-2$ over $[n]$.

We prove the theorem by showing that the function is injective, i.e., the algorithm produces distinct sequences for distinct trees, and surjective, i.e., every sequence is the sequence of some tree.

We will use two simple observations about the algorithm. First, each vertex $v \in V(T)$ appears $d_T(v) - 1$ times in the Prüfer code of T. Second, once the first leaf v of T has been removed, the rest of the code will be equal to the code of the tree T' with $V(T') = V(T) \setminus \{v\}$ and $E(T') = E(T) \setminus \{uv\}$.

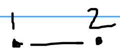
Let us first show that the algorithm yields different codes for different trees. We show this by induction on n. The claim is trivial for $n=2$: in that case there is a unique tree and a unique code, the empty sequence. Now assume $n \geq 3$ and consider two distinct trees T_1 and T_2 with vertex set $[n]$.



Let v_1 be the smallest leaf in T_1, u_1 its neighbour. Let v_2 be the smallest leaf in T_2, u_2 its neighbour.

If $v_1 \neq v_2$, and w.l.o.g., $v_1 < v_2$, then $1 = d_{T_1}(v_1) \neq d_{T_2}(v_1) > 1$. Since v_1 appears $d_{T_1}(v_1) - 1$ times in the code for T_1 and $d_{T_2}(v_1) - 1$ times in the code for T_2, the codes for T_1 and T_2 are distinct. If $v_1 = v_2$ but $u_1 \neq u_2$, the codes for T_1 and T_2 start with distinct labels and are therefore distinct. Finally, if $v_1 = v_2$ and $u_1 = u_2$, the codes for T_1 and T_2 start with the same label and continue with the respective codes for $T_1[[n] \setminus \{v_1\}]$ and $T_2[[n] \setminus \{v_2\}]$. These are distinct tree with $n-1$ vertices, so by the induction hypothesis they have distinct codes.

Let us now show that every sequence is the code of some tree. We again show this by induction. As before the claim is trivial for $n=2$. Let us now consider $n \geq 3$, and a



sequence S of length $n-2$ over $[n]$. Let v be the smallest element of $[n]$ <u>not</u> in S, u the first element of S, and S' the sequence obtained by removing the first element from S. S' is a sequence of length $n-3$ over $[n] \setminus \{v\}$, so by the induction hypothesis there exists a tree T' with $V(T') = [n] \setminus \{v\}$ whose code is S'. Moreover, none of the elements of S' is a leaf in T': when an element S' was written down, it was the neighbour of a leaf in a tree with at least 3 vertices and thus not itself a leaf. Let T be the tree obtained by adding vertex v and edge uv to T'. The code of T starts with u and continues with S', so it is equal to S. We have found a tree whose code is equal to S. $\square$

Because the algorithm defines a bijection, it must have an "inverse" that takes a code and outputs the tree with that code. Given a sequence of length $n-2$ over $[n]$ this "inverse" starts by adding "$*$" to the end of the sequence (to help us remember where the end of the sequence is), then repeats the following steps until the "$*$" is at the start of the sequence:

1. Let s be the first element in the sequence, and x the smallest element not currently in the sequence.
2. Add edge sx to the tree, along with any endpoint not already present.
3. Remove s from the beginning of the sequence, and add x to the end of the sequence

Finally, when $*$ is at the beginning of the sequence, two numbers $x, y \in [n]$ are missing from the sequence. Add the edge xy to the tree, along with any endpoints not already there.

$n-2=5$ $\qquad$ $[n]=\{1,2,3,\dots,7\}$

<u>Example</u> $\cancel{1}, \cancel{2}, \cancel{1}, \cancel{3}, \cancel{4}, *, 5, 6, 2, 1, 3$ tree with 7 vertices