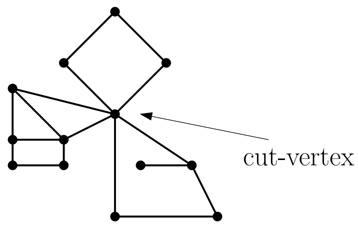
NDMI011: Combinatorics and Graph Theory 1

Lecture #9
2-connected graphs and the Ear lemma. Cayley's formula

Irena Penev

1 2-connected graphs and ear decomposition

A *cut-vertex* of a graph G is any vertex $v \in V(G)$ such that $G \setminus v$ has more components than G.

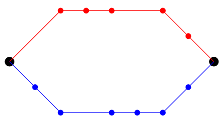


Recall that, for a non-negative integer k, a graph G is *k-connected* if $|V(G)| \geq k+1$ and for all $S \subseteq V(G)$ such that $|S| \leq k-1$, we have that $G \setminus S$ is connected. So, a graph is 2-connected if it has at least three vertices, is connected, and has no cut-vertices.

Lemma 1.1. *Let G be a graph on at least two vertices. Then G is 2-connected if and only if any two distinct vertices lie on a common cycle.*[1]

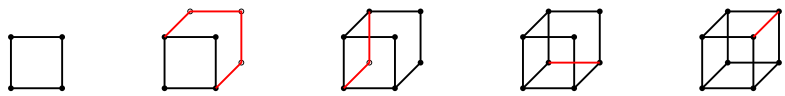
Proof. By Menger's theorem (global version), a graph on at least two vertices is 2-connected if and only if for any pair of distinct vertices, there are two internally disjoint paths between them. But obviously, two distinct vertices lie on a common cycle if and only if there are two internally-disjoint paths between them. The result now follows.

[1]Note that if G has at least two vertices, and any two distinct vertices lie on a common cycle, then in particular, G contains a cycle, and therefore, G has at least three vertices.



□

In this section, we give a full structural description of 2-connected graphs. A *path addition* (sometimes called *ear addition*) to a graph H is the addition to H of a path between two distinct vertices of H in such a way that no internal vertex and no edge of the path belongs to H. In the picture below, we show how the cube graph can be constructed by starting with a cycle of length four and then repeatedly adding ears (the path/ear added at each step is in red).



The Ear lemma. *A graph is 2-connected if and only if it is a cycle or can be obtained from a cycle by repeated ear addition.*

Proof. We first prove the "if" (i.e. "$\Longleftarrow$") part of the lemma. Clearly, cycles are 2-connected (indeed, every cycle has at least three vertices, is connected, and has no cut-vertices).[2] Further, if a graph G can be obtained from a 2-connected graph H by adding an ear, then G has at least three vertices (because H does), and it is easy to see that G is connected and has no cut-vertices;[3] so, G is 2-connected. It now follows by an easy induction (e.g. on the number of ears added) that any graph obtained from a cycle by repeated ear addition is 2-connected. This proves the "if" part of the lemma.

It remains to prove the "only if" (i.e. "$\Longrightarrow$") part of the lemma. Fix a 2-connected graph G. By Lemma 1.1, G contains a cycle.[4] Now, let H be a maximal subgraph of G that either is a cycle or can be obtained from a cycle by repeated ear addition.[5] We must show that $H = G$.

First, we claim that H is an induced subgraph of G.[6] If not, then there exist distinct vertices $u, v \in V(H)$ that are adjacent in G, but not in H; but then the graph obtained from H by adding the one-edge path u, v contradicts the maximality of H. So, H is indeed an induced subgraph of G.

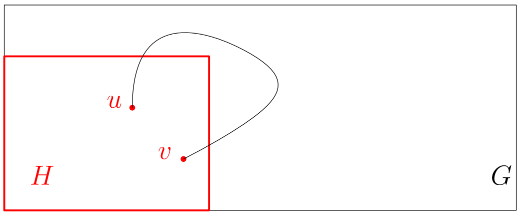
[2] Alternatively, this follows from Lemma 1.1.

[3] Check this!

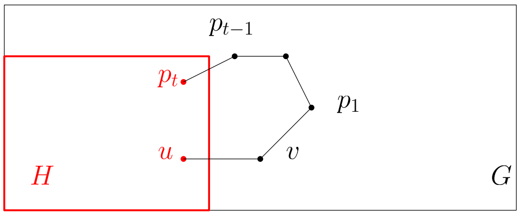
[4] Indeed, G has at least three vertices (because it is 2-connected), and by Lemma 1.1, any two of them lie on a common cycle. So, G contains a cycle.

[5] This means that no subgraph H^* of G that either is a cycle or can be obtained from a cycle by repeated ear addition contains H as a proper subgraph.

[6] A graph H is an *induced subgraph* of a graph G if $V(H) \subseteq V(G)$, and for all distinct $u, v \in V(H)$, we have that $uv \in E(H)$ if and only if $uv \in E(G)$.



It remains to show that $V(H) = V(G)$. Suppose otherwise. Then since G is connected, there is at least one edge between $V(H)$ and $V(G) \setminus V(H)$; fix adjacent vertices $u \in V(H)$ and $v \in V(G) \setminus V(H)$. Since both G is 2-connected, we know that $G \setminus u$ is connected; consequently, there is a path in $G \setminus u$ from v to some vertex in $V(H) \setminus \{u\}$; let $P = v, p_1, \dots, p_t$ $(t \geq 1)$ be a path in $G \setminus u$ with $p_t \in V(H) \setminus \{u\}$; we may assume that $p_1, \dots, p_{t-1} \in V(G) \setminus V(H)$.[7] But now the graph obtained from H by adding the path $u, v, p_1, \dots, p_t$ contradicts the maximality of H.



This proves that $V(H) = V(G)$. Since we already know that H is an induced subgraph of G, it follows that $H = G$. This proves the "only if" part of the lemma. □

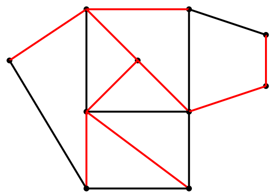
2 Cayley's formula for the number of spanning trees of a complete graph

Recall that a *forest* is an acyclic graph (i.e. a graph that has no cycles), and a *tree* is a connected forest. A *leaf* in a graph G is a vertex of degree one, i.e. a vertex that has exactly one neighbor. In what follows, we will use the well-known fact that every tree on at least two vertices has a leaf.[8] It is clear that if v is a leaf of a tree T, then $T \setminus v$ is still a tree.

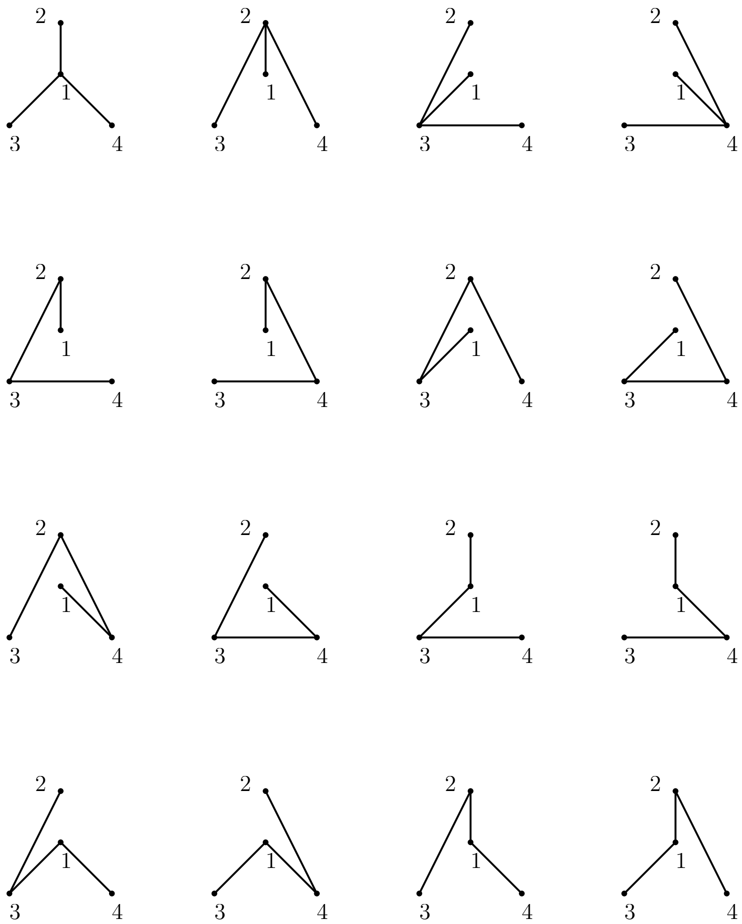
[7]Otherwise, we fix a minimal index $i \in \{1, \dots, t-1\}$ such that $p_i \in V(H)$, and we consider the path $v, p_1, \dots, p_i$ instead of $v, p_1, \dots, p_t$.

[8]In fact, every tree on at least two vertices has at least two leaves. Let us prove this. Suppose that T is a tree on at least two vertices, and let $P = p_1, \dots, p_t$ be a path of maximum length in T. Since T has at least one edge (because it is connected and has at least two vertices), we know that $t \geq 2$. We claim that p_1 and p_t are leaves of T; by symmetry, it suffices to show that p_1 is a leaf. Obviously, p_1 is adjacent to p_2 in T. Further, if p_1 were adjacent to some p_i with $i \in \{3, \dots, t\}$, then $p_1, p_2, \dots, p_i, p_1$ would be a cycle in T, contrary to the fact that T is a tree. Finally, if p_1 were adjacent to some vertex $v \in V(T) \setminus \{p_1, \dots, p_t\}$, then the path $v, p_1, \dots, p_t$ would contradict the maximality of P. So, p_2 is the only neighbor of p_1 in T, and it follows that p_1 is a leaf of T.

A *spanning tree* of a connected graph G is a tree T that is a subgraph of G, and satisfies $V(T) = V(G)$. An example is given below (the edges of the spanning tree are in red).



Now, suppose we are given a labeled complete graph on n ($n \geq 2$) vertices (say, with vertices labeled $1, \ldots, n$). We would like to count the number of spanning trees in this graph; equivalently, we would like to count the number of trees on the vertex set $\{1, \ldots, n\}$. There is only one spanning tree for K_2, and it is easy to see that there are three spanning trees for K_3. For K_4, there are 16 spanning trees, represented below (only the edges of the trees are represented; the remaining edges of the K_4 are not shown).



Our goal in this section is to prove the following.

Cayley's formula. *For all $n \geq 2$, the number of spanning trees of K_n is n^{n-2}.*

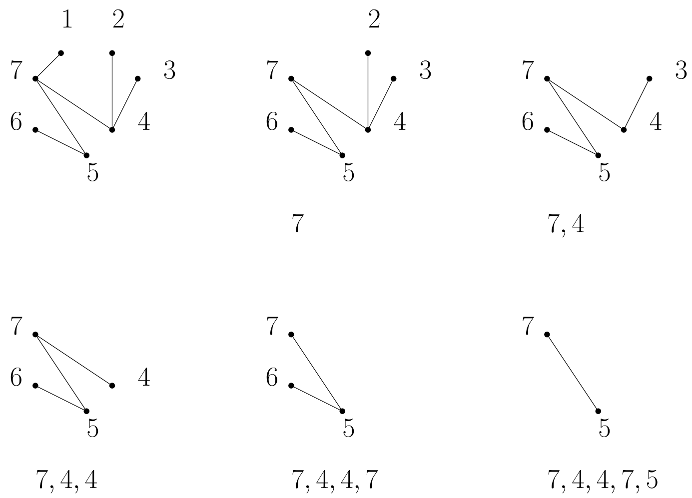
There are a number of known proofs of Cayley's formula; here, we give the one that uses the so called "Prüfer codes."

We will show that for all finite sets $S \subseteq \mathbb{N}$ with $|S| \geq 2$, the number of trees on the vertex set S is $|S|^{|S|-2}$ (see Lemma 2.4). Obviously, this will immediately imply Cayley's formula, since the number of spanning trees of K_n is equal to the number of vertices on the vertex set $\{1, \dots, n\}$.

To simplify terminology, we will say that a tree is an *integer tree* if all its vertices are positive integers.[9] We now define the *Prüfer code* of integer trees on at least two vertices recursively, as follows:

- for any integer tree T on exactly two vertices, the Prüfer code of T, denoted by $P(T)$, is the empty sequence;
- for any integer tree T on at least three vertices, we define the Prüfer code of T to be $P(T) := a_i, P(T \setminus i)$, where i is the smallest leaf of T, and a_i is the unique neighbor of i in T.[10]

An example is given below (the Prüfer code of the tree in the top left corner is $7, 4, 4, 7, 5$, and the procedure for finding it is shown below).



Now, our goal is to show that given a set $S \subseteq \mathbb{N}$ (with $|S| = n \geq 2$), the function $T \mapsto P(T)$ is a bijection between the set of all trees with vertex set S, and the set of all sequences of length $n-2$ with terms in S.[11]

[9] Note, however, that this is **not** standard terminology. (There is no standard terminology for such trees.) We simply use the term "integer tree" as convenient shorthand in this section.

[10] So, $P(T)$ is obtained by adding a_i to the front of $P(T \setminus i)$.

[11] Obviously, there are precisely n^{n-2} such sequences.

Lemma 2.1. *If T is an integer tree on at least two vertices, then every non-leaf of T appears in $P(T)$, and none of the leaves do.*

Proof. We prove the lemma by induction on the number of vertices of the integer tree T. If T is a 2-vertex integer tree, then both its vertices are leaves, and by definition, $P(T)$ is the empty sequence; so the lemma is true for 2-vertex integer trees. Now, fix an integer $n \geq 2$, and assume inductively that the lemma holds for integer trees on n vertices. Let T be an integer tree on $n+1$ vertices. Let i be the smallest leaf of T, and let a_i be the unique neighbor of i. Since T is connected and has at least three vertices, adjacent vertices cannot both be leaves of T, and so a_i is a non-leaf of T. By construction, $P(T) = a_i, P(T \setminus i)$, and so the non-leaf a_i of T appears in $P(T)$, whereas the leaf i of T does not. Note that for $v \in V(T) \setminus \{i, a_i\}$, we have that $d_T(v) = d_{T \setminus i}(v)$, and so each vertex of T other than i and a_i is a leaf in T if and only if it is a leaf in $P(T \setminus i)$. The result now follows from the induction hypothesis. □

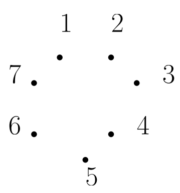
Lemma 2.2. *If two integer trees have the same vertex set and the same Prüfer code, then they are identical.*

Proof. We proceed by induction on the number of vertices. There is only one tree on a fixed two-element vertex set, and so the lemma clearly holds for 2-vertex integer trees. Now, fix an integer $n \geq 2$, and suppose inductively that the lemma is true for integer trees on n vertices. Let $S \subseteq \mathbb{N}$ with $|S| = n+1$, and let T_1 and T_2 be trees with vertex-set S and identical Prüfer code P. P is of length $n-1$, and so at least two members of S do not appear in P; let i be the smallest integer in S that does not appear in P. Let a_i be the first term of P, and let P_i be obtained from P by deleting its first term. By Lemma 2.1, i is the smallest leaf of both T_1 and T_2, and a_i is the unique neighbor of i in both T_1 and T_2. Further, we have that $P(T_1 \setminus i) = P(T_2 \setminus i) = P_i$, and so by the induction hypothesis, $T_1 \setminus i = T_2 \setminus i$. Since i has the same neighborhood in T_1 and in T_2, it follows that $T_1 = T_2$. □

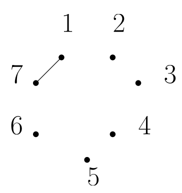
Lemma 2.3. *If $n \geq 2$ is an integer, and if $S \subseteq \mathbb{N}$ with $|S| = n$, then every sequence of length $n-2$, all of whose terms are in S, is the Prüfer code of some tree with vertex-set S.*

Proof. We proceed by induction on n. Suppose first that $S \subseteq \mathbb{N}$ satisfies $|S| = 2$, and let P be a sequence of length $2 - 2 = 0$, all of whose terms are in S. Then P is the empty sequence. Let T be the unique tree on the vertex-set S. Then $P(T) = P$. Now, fix an integer $n \geq 2$, and suppose inductively that the lemma is true for some $n \geq 2$. We need to show that it holds for $n + 1$. Let $S \subseteq \mathbb{N}$ be such that $|S| = n + 1$, and let P be a sequence of length $n - 1$, all of whose terms are in S. Let i be the smallest member of S that does not appear in P, and let a_i be the first term of P. Let P_i be the sequence obtained by deleting the first term from P. By the induction hypothesis, there is a tree T_i with vertex-set $S \setminus \{i\}$ and Prüfer code P_i. Let T be the tree obtained by adding the vertex i to T_i, and making i adjacent to a_i and to no other vertex of T_i. Now $P(T) = P$. This completes the induction. $\square$

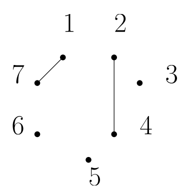
Note that the proof of Lemma 2.3 in fact gives us a recipe for "decoding" a given Prüfer code, i.e. for finding the tree to which the code corresponds. For an integer $n \geq 2$, an n-element set $S \subseteq \mathbb{N}$, and an $(n-2)$-term sequence P, with terms in S, we proceed as follows. If $n \geq 3$, then we let i be the smallest element of S that is not in P, and we let a_i be the first term of P. We make i and a_i adjacent, we delete i from S, and we delete the first term of P. We repeat the process until S only has two elements left, and P is the empty sequence. At this point, we make the last two remaining elements of S adjacent. An example is given below: the tree on the vertex set $S = \{1, 2, 3, 4, 5, 6, 7\}$ whose Prüfer code is $7, 4, 4, 7, 5$ is the tree on the bottom of the picture (e is the empty sequence).



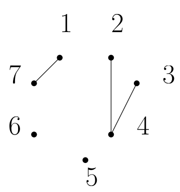
$P = 7, 4, 4, 7, 5$
$S = \{1, 2, 3, 4, 5, 6, 7\}$



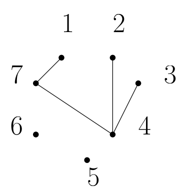
$P = 4, 4, 7, 5$
$S = \{2, 3, 4, 5, 6, 7\}$



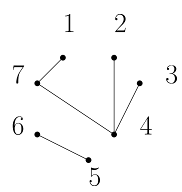
$P = 4, 7, 5$
$S = \{3, 4, 5, 6, 7\}$



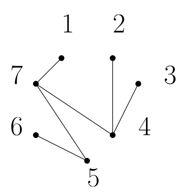
$P = 7, 5$
$S = \{4, 5, 6, 7\}$



$P = 5$
$S = \{5, 6, 7\}$



$P = e$
$S = \{5, 7\}$



Putting Lemmas 2.2 and 2.3 together, we obtain the following.

Lemma 2.4. *Let $n \geq 2$ be an integer, and let $S \subseteq \mathbb{N}$ be such that $|S| = n$. Then the number of trees on the vertex set S is n^{n-2}.*

Proof. By Lemmas 2.2 and 2.3, the mapping $T \mapsto P(T)$ is a bijection from the set of all integer trees on the vertex set S to the set of $(n-2)$-term sequences, all of whose terms are elements of S. There are precisely n^{n-2} sequences of length $n-2$, with terms in S, and it follows that there are precisely n^{n-2} trees on the vertex set S. □

Cayley's formula follows immediately from Lemma 2.4, since the number of spanning trees of K_n is precisely the number of trees on the vertex set $\{1, \ldots, n\}$.

2.1 Cayley's formula via determinants

In this subsection, we give (without proof) a formula for computing the number of spanning trees of **any** graph on the vertex set $\{1, \ldots, n\}$.

Suppose that $n \geq 2$ is an integer, and that G is a graph on the vertex set $\{1, \ldots, n\}$. Then the *Laplacian* of G is the matrix $Q = [q_{i,j}]_{n \times n}$ given by

$$q_{i,j} = \begin{cases} d_G(i) & \text{if } \; i = j \\ \\ -1 & \text{if } \; i \neq j \text{ and } ij \in E(G) \\ \\ 0 & \text{if } \; i \neq j \text{ and } ij \notin E(G) \end{cases}$$

We now need some notation. Suppose $A = [a_{i,j}]_{n \times m}$ is a matrix, and suppose $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$; then $A_{i,j}$ is the matrix obtained from A by deleting the i-th row and j-th column. In particular, $A_{1,1}$ is the matrix obtained from A by deleting the first row and first column.

Theorem 2.5. *Let $n \geq 2$ be an integer, let G be any graph on the vertex set $\{1, \ldots, n\}$, and let Q be the Laplacian of G. Then the number of spanning trees of G is precisely* $\det(Q_{1,1})$.

Proof. Omitted. □

Example 2.6. *Using Theorem 2.5, prove Cayley's formula.*

Solution. Fix an integer $n \geq 2$, and consider the complete graph on the vertex set $\{1, \ldots, n\}$. Then the Laplacian of this graph is the $n \times n$ matrix

$$Q = \begin{bmatrix} n-1 & -1 & -1 & \ldots & -1 \\ -1 & n-1 & -1 & \ldots & -1 \\ -1 & -1 & n-1 & \ldots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \ldots & n-1 \end{bmatrix}_{n \times n}.$$

The matrix $Q_{1,1}$ has exactly the same form, only it is of size $(n-1) \times (n-1)$:

$$Q_{1,1} = \begin{bmatrix} n-1 & -1 & -1 & \dots & -1 \\ -1 & n-1 & -1 & \dots & -1 \\ -1 & -1 & n-1 & \dots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \dots & n-1 \end{bmatrix}_{(n-1)\times(n-1)}.$$

We now compute the determinant of $Q_{1,1}$:

$$\det(Q_{1,1}) = \begin{vmatrix} n-1 & -1 & -1 & \dots & -1 \\ -1 & n-1 & -1 & \dots & -1 \\ -1 & -1 & n-1 & \dots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \dots & n-1 \end{vmatrix}_{(n-1)\times(n-1)}$$

$$\overset{(*)}{=} \begin{vmatrix} n-1 & -1 & -1 & \dots & -1 \\ -n & n & 0 & \dots & 0 \\ -n & 0 & n & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -n & 0 & 0 & \dots & n \end{vmatrix}_{(n-1)\times(n-1)}$$

$$\overset{(**)}{=} \begin{vmatrix} 1 & -1 & -1 & \dots & -1 \\ 0 & n & 0 & \dots & 0 \\ 0 & 0 & n & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & n \end{vmatrix}_{(n-1)\times(n-1)}$$

$$\overset{(***)}{=} n^{n-2},$$

where (*) is obtained by subtracting the first row from all the subsequent ones, (**) is obtained by adding to the first column the sum of all subsequent ones, and (***) is obtained by multiplying the diagonal entries of the upper triangular matrix that we obtained. By Theorem 2.5, we now have that the number of spanning trees of K_n is precisely n^{n-2}, which proves Cayley's formula. □