Graph Theory

Oliver Janzer

December 30, 2025

Acknowledgement

Most of the material in these notes is based on the notes for Benny Sudakov's *Graph Theory* course at ETH Zürich, as well as on the books *Graph Theory* by Reinhard Diestel, *Introduction to Graph Theory* by Douglas West, and *Proofs from the book* by Martin Aigner and Günter Ziegler.

Contents

1 Basic notions

1.1 Graphs

Definition 1.1. A *graph* G is a pair $G = (V, E)$ where V is a set of *vertices* and E is a set of unordered pairs of vertices. The elements of E are called *edges*. We will often write (u, v) for the edge $\{u, v\}$. We write $V(G)$ for the set of vertices and $E(G)$ for the set of edges of a graph G. Also, $|G| = |V(G)|$ denotes the number of vertices and $e(G) = |E(G)|$ denotes the number of edges.

Remark 1.2. Very occasionally, we will consider graphs which may have loops or multiple edges. A *loop* is an edge (v, v) for some $v \in V$. An edge $e = (u, v)$ is a *multiple edge* if it appears multiple times in E. A graph is *simple* if it has no loops or multiple edges.

Unless explicitly stated otherwise, we will only consider simple graphs. General (potentially non-simple) graphs are also called *multigraphs*.

Definition 1.3.

- Vertices u, v are *adjacent* in G if $(u, v) \in E(G)$.
- An edge $e \in E(G)$ is *incident* to a vertex $v \in V(G)$ if $v \in e$.
- Edges e, e' are incident if $e \cap e' \neq \varnothing$.
- If $(u, v) \in E$ then v is a *neighbour* of u.

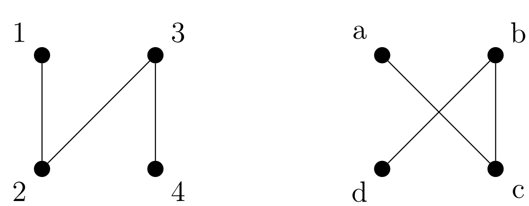
Example 1.4. Any symmetric relation between objects gives a graph. For example:

- let V be the set of people in a room, and let E be the set of pairs of people who met for the first time today;
- let V be the set of cities in a country, and let the edges in E correspond to roads connecting them;
- the internet: let V be the set of websites, and let the edges in E correspond to the links connecting them;
- social networks: let V be the users of Facebook, and let E be the pairs of users who are friends.

The usual way to picture a graph is to put a dot for each vertex and to join adjacent vertices with lines. The specific drawing is irrelevant, all that matters is which pairs are adjacent.

1.2 Graph isomorphism

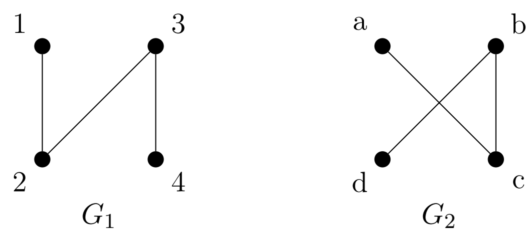
Question 1.5.



are these graphs in some sense the same?

Definition 1.6. Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be graphs. An *isomorphism* $\phi : G_1 \to G_2$ is a bijection (a one-to-one correspondence) from V_1 to V_2 such that $(u, v) \in E_1$ if and only if $(\phi(u), \phi(v)) \in E_2$. We say G_1 is *isomorphic* to G_2 if there is an isomorphism between them.

Example 1.7. Recall the graphs in Question 1.5:



The function $\phi : G_1 \to G_2$ given by $\phi(1) = a$, $\phi(2) = c$, $\phi(3) = b$, $\phi(4) = d$ is an isomorphism.

Remark 1.8. Isomorphism is an equivalence relation of graphs. This means that

- Any graph is isomorphic to itself
- if G_1 is isomorphic to G_2 then G_2 is isomorphic to G_1
- If G_1 is isomorphic to G_2 and G_2 is isomorphic to G_3, then G_1 is isomorphic to G_3.

Definition 1.9. An *unlabelled graph* is an isomorphism class of graphs. In the previous example G_1 and G_2 are different labelled graphs but since they are isomorphic they are the same unlabelled graph.

1.3 The adjacency and incidence matrices

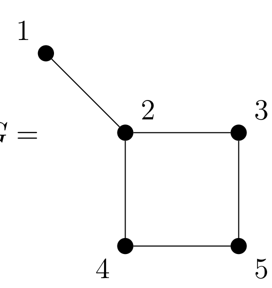
Let $[n] = \{1, \ldots, n\}$.

Definition 1.10. Let $G = (V, E)$ be a graph with $V = [n]$. The *adjacency matrix* $A = A(G)$ is the $n \times n$ symmetric matrix defined by

$$a_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Example 1.11.

$G =$



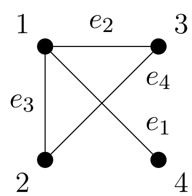
$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

Remark 1.12. Any adjacency matrix A is real and symmetric, hence the spectral theorem proves that A has an orthogonal basis of eigenvectors with real eigenvalues. This important fact allows us to use *spectral methods* in graph theory. Indeed, there is a large subfield of graph theory called *spectral graph theory*.

Definition 1.13. Let $G = (V, E)$ be a graph with $V = \{v_1, \dots, v_n\}$ and $E = \{e_1, \dots, e_m\}$. Then the *incidence matrix* $B = B(G)$ of G is the $n \times m$ matrix defined by

$$b_{ij} = \begin{cases} 1 & \text{if } v_i \in e_j, \\ 0 & \text{otherwise.} \end{cases}$$

Example 1.14.

$G =$ 

$$B = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

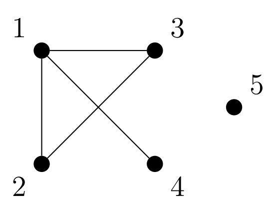
Remark 1.15. Every column of B has $|e| = 2$ entries equal to 1.

1.4 Degree

Definition 1.16. Given $G = (V, E)$ and a vertex $v \in V$, we define the *neighbourhood* $N(v)$ of v to be the set of neighbours of v. Let the *degree* $d(v)$ of v be $|N(v)|$, the number of neighbours of v. A vertex v is *isolated* if $d(v) = 0$.

Remark 1.17. $d(v)$ is the number of 1s in the row corresponding to v in the adjacency matrix $A(G)$ or the incidence matrix $B(G)$.

Example 1.18.



$d(1) = 3$, $d(2) = 2$, $d(3) = 2$, $d(4) = 1$, $d(5) = 0$;

5 is isolated.

Fact 1. *For any graph G on the vertex set $[n]$ with adjacency and incidence matrices A and B, we have $BB^T = D + A$, where*

$$D = \begin{pmatrix} d(1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d(n) \end{pmatrix}.$$

Notation 1.19. The *minimum degree* of a graph G is denoted $\delta(G)$; the *maximum degree* is denoted $\Delta(G)$. The *average degree* is

$$\bar{d}(G) = \frac{\sum_{v \in V(G)} d(v)}{|V(G)|}.$$

Note that $\delta \leq \bar{d} \leq \Delta$.

Definition 1.20. A graph G is *d-regular* if and only if all vertices have degree d.

Question 1.21. Is there a 3-regular graph on 9 vertices?

Proposition 1.22. *For every* $G = (V, E)$, $\sum_{v \in V(G)} d(v) = 2|E|$.

Proof. In the sum $\sum_{v \in V(G)} d(v)$ every edge $e = (u, v)$ is counted twice: once from u and once from v. $\square$

Corollary 1.23. *Every graph has an even number of vertices of odd degree.*

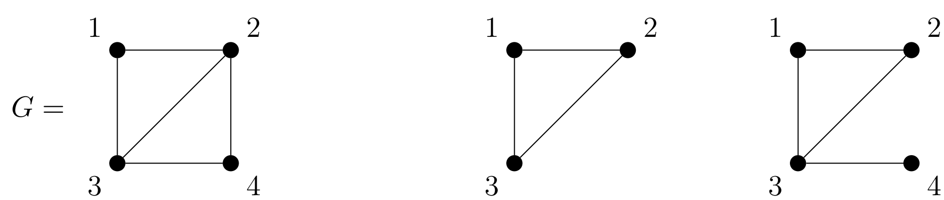
This shows that the answer to Question 1.21 is "no".

1.5 Subgraphs

Definition 1.24. A graph $H = (U, F)$ is a *subgraph* of a graph $G = (V, E)$ if $U \subseteq V$ and $F \subseteq E$. If $U = V$ then H is called *spanning*.

Definition 1.25. Given $G = (V, E)$ and $U \subseteq V$ $(U \neq \varnothing)$, let $G[U]$ denote the graph with vertex set U and edge set $E(G[U]) = \{e \in E(G) : e \subseteq U\}$. (We include all the edges of G which have both endpoints in U). Then $G[U]$ is called the subgraph of G *induced* by U.

Example 1.26.



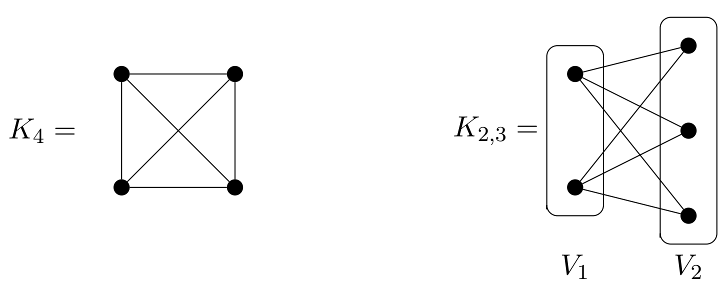
induced subgraph

not induced but spanning

1.6 Special graphs

- K_n is the *complete graph*, or a *clique*. Take n vertices and all possible edges connecting them.
- An *empty graph* has no edges.
- $G = (V, E)$ is *bipartite* if there is a partition $V = V_1 \cup V_2$ into two disjoint sets such that each $e \in E(G)$ intersects both V_1 and V_2.
- $K_{n,m}$ is the *complete bipartite graph*. Take $n + m$ vertices partitioned into a set A of size n and a set B of size m, and include every possible edge between A and B.

Example 1.27.



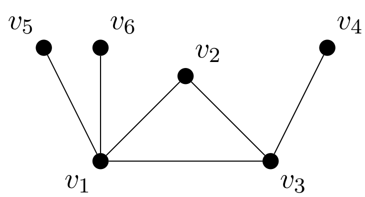
1.7 Walks, paths and cycles

Definition 1.28. A *walk* in G is a sequence of vertices $v_0, v_1, v_2, \ldots, v_k$, and a sequence of edges $(v_i, v_{i+1}) \in E(G)$. A walk is a *path* if all v_i are distinct. If for such a path with $k \geq 2$, (v_0, v_k) is also an edge in G, then $v_0, v_1, \ldots, v_k, v_0$ is a *cycle*. For multigraphs, we also consider loops and pairs of multiple edges to be cycles.

Remark 1.29. The above definition treats paths and cycles as having a start point and an endpoint (so reversing a path technically gives a different path). However, in practice, when we use language like "there is a unique path between u and v", we mean that the path or cycle is unique up to this choice.

Definition 1.30. The *length* of a path, cycle or walk is the number of edges in it.

Example 1.31.



$v_5v_1v_3v_4 \equiv$ path of length 3;

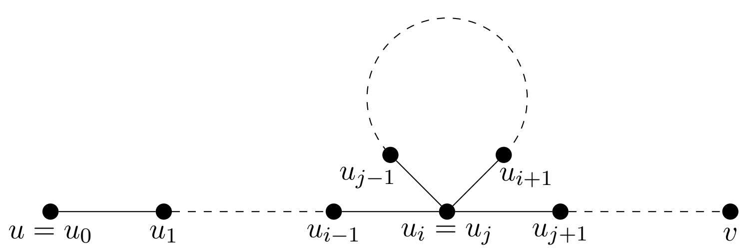
$v_1v_2v_3v_1 \equiv$ cycle of length 3;

$v_5v_1v_2v_3v_1v_6 \equiv$ walk of length 5.

Proposition 1.32. *Every walk from u to v in G contains a path between u and v.*

Proof. By induction on the length ℓ of the walk $u = u_0, u_1, \ldots, v_\ell = v$.

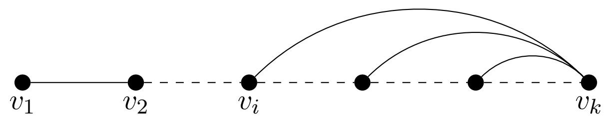
If $\ell = 1$ then our walk is also a path. Otherwise, if our walk is not a path there is $u_i = u_j$ with $i < j$, then $u = u_0, \ldots, u_i, u_{j+1}, \ldots, v$ is also a walk from u to v which is shorter. We can use induction to conclude the proof.



□

Proposition 1.33. *Every G with minimum degree $\delta \geq 2$ contains a path of length δ and a cycle of length at least $\delta + 1$.*

Proof. Let $v_1, \ldots, v_k$ be a longest path in G. Then all neighbours of v_k belong to $v_1, \ldots, v_{k-1}$ so $k - 1 \geq \delta$ and $k \geq \delta + 1$, and our path has at least δ edges. Let i $(1 \leq i \leq k-1)$ be the minimum index such that $(v_i, v_k) \in E(G)$. Then the neighbours of v_k are among $v_i, \ldots, v_{k-1}$, so $k - i \geq \delta$. Then $v_i, v_{i+1}, \ldots, v_k$ is a cycle of length at least $\delta + 1$.



□

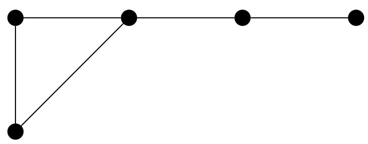
Remark 1.34. Note that we have also proved that a graph with minimum degree $\delta \geq 2$ contains cycles of at least $\delta - 1$ different lengths. This fact, and the statement of Proposition 1.33, are both tight; to see this, consider the complete graph $G = K_{\delta+1}$.

1.8 Connectivity

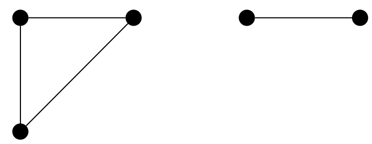
Definition 1.35. A graph G is *connected* if for all pairs $u, v \in G$, there is a path in G from u to v.

Note that it suffices for there to be a walk from u to v, by Proposition 1.32.

Example 1.36.



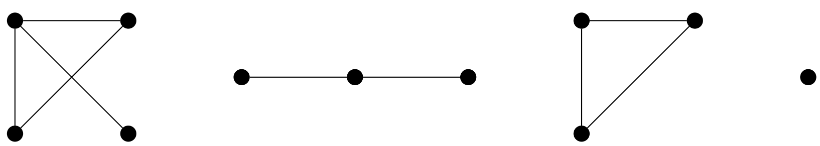
connected



not connected

Definition 1.37. A *(connected) component* of G is a connected subgraph that is maximal by inclusion. We say G is *connected* if and only if it has one connected component.

Example 1.38.

$G =$ 

has 4 connected components.

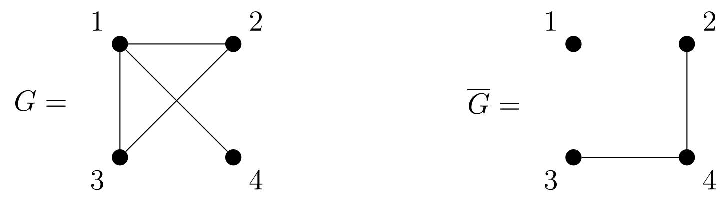
Proposition 1.39. *A graph with n vertices and m edges has at least $n - m$ connected components.*

Proof. Start with the empty graph (which has n components), and add edges one-by-one. Note that adding an edge can decrease the number of components by at most 1. □

1.9 Graph operations and parameters

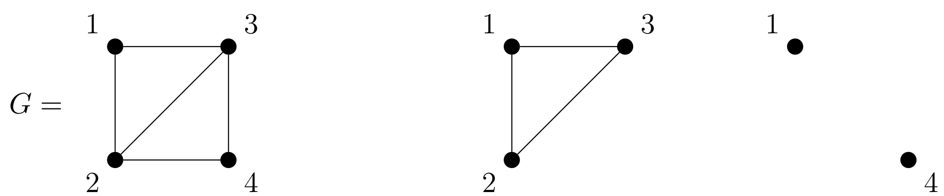
Definition 1.40. Given $G = (V, E)$, the *complement* $\overline{G}$ of G has the same vertex set V and $(u, v) \in E(\overline{G})$ if and only if $(u, v) \notin E(G)$.

Example 1.41.



Definition 1.42. A *clique* in G is a complete subgraph in G. An *independent set* is an empty induced subgraph in G.

Example 1.43.



clique in G independent set in G

Notation 1.44. Let $\omega(G)$ denote the number of vertices in a maximum-size clique in G; let $\alpha(G)$ denote the number of vertices in a maximum-size independent set in G.

Exercise 2. In Example 1.43, $\omega(G) = 3$ and $\alpha(G) = 2$.

Claim 1.45. A vertex set $U \subseteq V(G)$ is a clique if and only if $U \subseteq V(\overline{G})$ is an independent set.

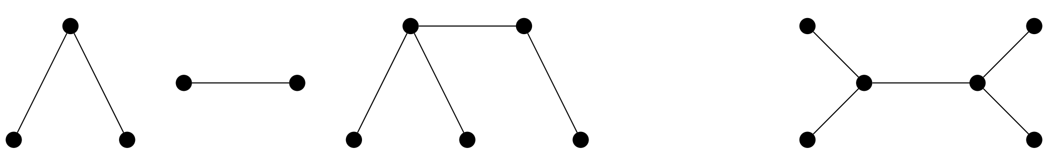
Corollary 1.46. *We have* $\omega(G) = \alpha(\overline{G})$ *and* $\alpha(G) = \omega(\overline{G})$.

2 Trees

2.1 Trees

Definition 2.1. A graph having no cycle is *acyclic*. A *forest* is an acyclic graph; a *tree* is a connected acyclic graph. A *leaf* (or *pendant vertex*) is a vertex of degree 1.

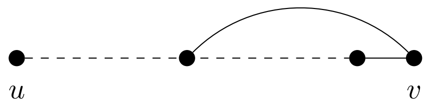
Example 2.2.



forest tree

Lemma 2.3. *Every finite tree with at least two vertices has at least two leaves. Deleting a leaf from an n-vertex tree produces a tree with $n - 1$ vertices.*

Proof. Every connected graph with at least two vertices has an edge. In an acyclic graph, the endpoints of a maximum path have only one neighbour on the path and therefore have degree 1. Hence the endpoints of a maximum path provide the two desired leaves.



(If v had multiple neighbours on the path we would get a cycle).

Suppose v is a leaf of a tree G, and let $G' = G \backslash v$. If $u, w \in V(G')$, then no u, w-path P in G can pass through the vertex v of degree 1, so P is also present in G'. Hence G' is connected. Since deleting a vertex cannot create a cycle, G' is also acyclic. We conclude that G' is a tree with $n-1$ vertices. □

2.2 Equivalent definitions of trees

Theorem 2.4. *For an n-vertex simple graph G (with $n \geq 1$), the following are equivalent (and characterize the trees with n vertices).*

(a) G is connected and has no cycles.

(b) G is connected and has $n-1$ edges.

(c) G has $n-1$ edges and no cycles.

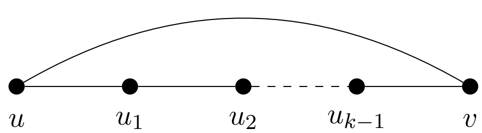
(d) For every pair $u, v \in V(G)$, there is exactly one u, v-path in G.

To prove this theorem we will need a small lemma.

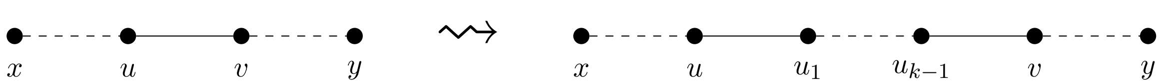
Definition 2.5. An edge of a graph is a *cut-edge* if its deletion disconnects the graph.

Lemma 2.6. *An edge contained in a cycle is not a cut-edge.*

Proof. Let (u, v) belong to a cycle.



Then any path $x \ldots y$ in G which uses the edge (u, v) can be extended to a walk in $G \backslash (u, v)$ as follows:



□

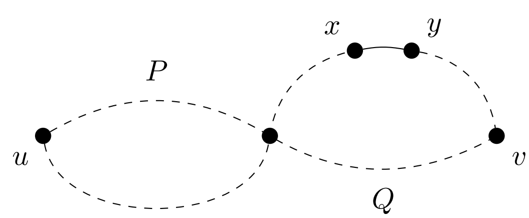
Proof of Theorem 2.4. We first demonstrate the equivalence of (a), (b), (c) by proving that any two of {connected, acyclic, $n-1$ edges} implies the third.

(a) $\implies$ (b), (c): We use induction on n. For $n = 1$, an acyclic 1-vertex graph has no edge. For the induction step, suppose $n > 1$, and suppose the implication holds for graphs with fewer than n vertices. Given G, Lemma 2.3 provides a leaf v and states that $G' = G \backslash v$ is acyclic and connected. Applying the induction hypothesis to G' yields $e(G') = n-2$, and hence $e(G) = n-1$.

(b) $\implies$ (a), (c): Delete edges from cycles of G one by one until the resulting graph G' is acyclic. By Lemma 2.6, G' is connected. By the paragraph above, G' has $n-1$ edges. Since this equals $|E(G)|$, no edges were deleted, and G itself is acyclic.

(c) $\implies$ (a), (b): Suppose G has k components with orders $n_1, \ldots, n_k$. Since G has no cycles, each component satisfies property (a), and by the first paragraph the ith component has $n_i - 1$ edges. Summing this over all components yields $e(G) = \sum(n_i - 1) = n - k$. We are given $e(G) = n - 1$, so $k = 1$, and G is connected.

(a) $\implies$ (d): Since G is connected, G has at least one u, v-path for each pair $u, v \in V(G)$. Suppose G has distinct u, v-paths P and Q. Let $e = (x, y)$ be an edge in P but not in Q. The concatenation of P with the reverse of Q is a closed walk in which e appears exactly once. Hence, $(P \cup Q) \backslash e$ is an x, y-walk not containing e. By Proposition 1.32, this contains an x, y-path, which completes a cycle with e and contradicts the hypothesis that G is acyclic. Hence G has exactly one u, v-path.



(d) $\implies$ (a): If there is a u, v-path for every $u, v \in V(G)$, then G is connected. If G has a cycle C, then G has two paths between any pair of vertices on C. $\square$

Definition 2.7. Given a connected graph G, a *spanning tree* T is a subgraph of G which is a tree and contains every vertex of G.

Corollary 2.8.

(a) Every connected graph on n vertices has at least $n-1$ edges and contains a spanning tree;

(b) Every edge of a tree is a cut-edge;

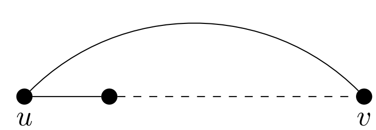
(c) Adding an edge to a tree creates exactly one cycle.

Proof.

(a) Delete edges from cycles of G one by one until the resulting graph G' is acyclic. By Lemma 2.6, G' is connected. The resulting graph is acylic so it is a tree. Therefore G had at least $n-1$ edges and contained a spanning tree.

(b) Note that deleting an edge from a tree T on n vertices leaves $n-2$ edges, so the graph is disconnected by (a).

(c) Let $u, v \in T$. There is a unique path in T between u and v, so adding an edge (u, v) closes this path to a unique cycle.



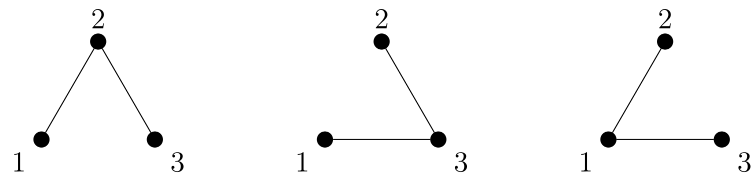
$\square$

2.3 Cayley's formula

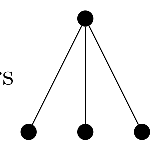
Question 2.9. What is the number of spanning trees in a labelled complete graph on n vertices?

Example 2.10.

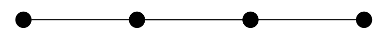
$n = 3$:



$n = 4$: 4 stars



$12 = 4!/2$ paths



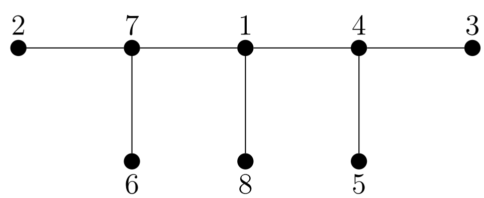
Theorem 2.11 (Cayley's Formula)**.** *There are n^{n-2} trees with vertex set $[n]$.*

We give two proofs of Cayley's formula. In our first proof, we establish a bijection between trees on $[n]$ and sequences in $[n]^{n-2}$.

Definition 2.12 (Prüfer code)**.** Let T be a tree on an ordered set S of n vertices. To compute the Prüfer sequence $f(T)$, iteratively delete the leaf with the smallest label and append the label of its *neighbour* to the sequence. After $n - 2$ iterations a single edge remains and we have produced a sequence $f(T)$ of length $n - 2$.

Example 2.13.

$T =$



We compute the Prüfer code for T as follows:

- 7 (delete 2)
- 4 (delete 3)
- 4 (delete 5)
- 1 (delete 4)
- 7 (delete 6)
- 1 (delete 7)

The edge remaining is $(1, 8)$. We then have $f(T) = (7, 4, 4, 1, 7, 1)$.

Proposition 2.14. *For an ordered n-element set S, the Prüfer code f is a bijection between the trees with vertex set S and the sequences in S^{n-2}.*

Proof. We need to show every sequence $(a_1, \ldots, a_{n-2}) \in S^{n-2}$ defines a unique tree T such that $f(T) = (a_1, \ldots, a_{n-2})$. We prove this by induction on n. If $n = 2$, then there is exactly one tree on 2 vertices and the algorithm defining f always outputs the empty sequence, the only sequence of length zero. So the claim clearly holds for $n = 2$.

Now, assume $n > 2$ and the claim holds for all ordered vertex sets S' of size less than n. Consider a sequence $(a_1, \ldots, a_{n-2}) \in S^{n-2}$. We need to show that $(a_1, \ldots, a_{n-2})$ can be uniquely produced by the algorithm.

Claim. Suppose that the algorithm produces $f(T) = (a_1, \ldots, a_{n-2})$ for some tree T. Then the vertices $\{a_1, \ldots, a_{n-2}\}$ are precisely those that are not leaves in T.

Proof of Claim. If a vertex v is a leaf in T then it can only appear in $f(T)$ if its neighbour gets deleted during the algorithm. But this would leave v as an isolated vertex, which is impossible. Conversely, if a vertex v is not a leaf then one of its neighbours must be deleted during the algorithm (it cannot be itself deleted before this happens). When this neighbour of v is deleted, v will be added to the Prüfer code for T, so is in $\{a_1, \ldots, a_{n-2}\}$.

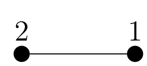
The claim implies that the label of the first leaf removed from T is the minimum element of the set $S \backslash \{a_1, \ldots, a_{n-2}\}$. Let v be this element. In other words, in every tree T such that $f(T) = (a_1, \ldots, a_{n-2})$ the vertex v is a leaf whose unique neighbour is a_1.

By the induction hypothesis, there is a unique tree T' with vertex set $S \backslash \{v\}$ such that $f(T') = (a_2, \ldots, a_{n-2})$. Adding the vertex v and the edge (a_1, v) to T' yields the only possible tree T with $f(T) = (a_1, \ldots, a_{n-2})$.

To see that we indeed have $f(T) = (a_1, \ldots, a_{n-2})$ for this choice, observe that (by the claim above) the set of leaves in T' is $(S \backslash \{v\}) \backslash \{a_2, \ldots, a_{n-2}\}$. After we attach the edge (a_1, v) to T', we get a new leaf v, and a_1 is no longer a leaf, even if it was in T'. Hence, the set of leaves in T is $S \backslash \{a_1, \ldots, a_{n-2}\}$. In this set, the minimal element is v by definition, so the first step of the Prüfer code algorithm on T removes v and adds a_1 to the sequence. This shows that $f(T) = (a_1, \ldots, a_{n-2})$. $\square$

Example 2.15. We use the idea of the above proof to compute the tree with Prüfer code 16631.

16631
2 is the smallest leaf



$\{1, 3, 4, 5, 6, 7\}$ left

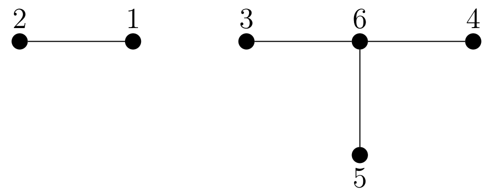
6631
4 is the smallest leaf

$\{1, 3, 5, 6, 7\}$ left

631
5 is the smallest leaf

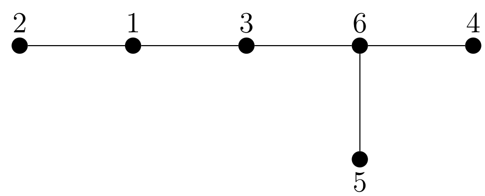
$\{1,3,6,7\}$ left

31
6 is the smallest leaf



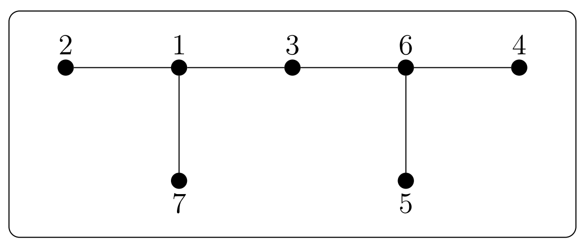
$\{1,3,7\}$ left

1
3 is the smallest leaf



$\{1,7\}$ left

Now add an edge between
the remaining vertices $\{1,7\}$.



To prove Cayley's formula, just apply Proposition 2.14 with the vertex set $[n]$ (note that there are n^{n-2} sequences in $[n]^{n-2}$).

For our second proof of Cayley's formula we need the following definition.

Definition 2.16. A *directed graph*, or *digraph* for short, is a vertex set and an edge (multi-)set of *ordered* pairs of vertices. Equivalently, a digraph is a (possibly not-simple) graph where each edge is assigned a direction. The *out-degree* (respectively *in-degree*) of a vertex is the number of edges incident to that vertex which point away from it (respectively, towards it).

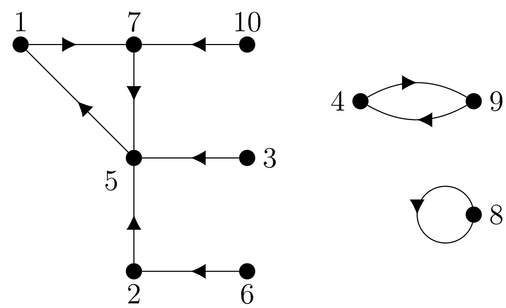
Proof of Cayley's formula (due to Joyal 1981). We count trees on n vertices which have two distinguished vertices called the "left end" L and the "right end" R, where L and R can coincide. Let t_n be the number of labelled trees on n vertices, and let T_n be the family of labelled trees with two distinguished vertices L, R. Clearly, $|T_n| = t_n n^2$, and it is thus enough to prove that $|T_n| = n^n$. We'll describe a bijection between the set of all mappings $f : [n] \to [n]$, and T_n. As the number of such mappings is clearly n^n, the result will follow.

So, let $f : [n] \to [n]$ be a mapping. We represent f as a directed graph G_f with vertex set $[n]$ and the set of directed edges $E(G_f) = \{(i, f(i)) : 1 \leq i \leq n\}$.

Example.

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 7 & 5 & 5 & 9 & 1 & 2 & 5 & 8 & 4 & 7 \end{pmatrix},$$

$G_f =$



Observe that G_f is a digraph in which the outdegree of every vertex is exactly one ($f(i)$ is the only out-neighbour of i).

Let us look at a component of G_f (ignoring edge directions for a moment). Since the out-degree of every vertex is exactly one, each such component contains as many edges as vertices and has therefore exactly one cycle (by Corollary 2.8). This is easily seen to be a directed cycle (just follow an edge leaving a current vertex until you hit a previously visited vertex).

Let M be the union of the vertex sets of these cycles. In order to create a tree, we will get rid of the union of these cycles and we will replace the union of the cycles by one path. It is easy to see that f restricted to M is a bijection; moreover, M is the unique maximal set on which f acts as a bijection.

Let us write

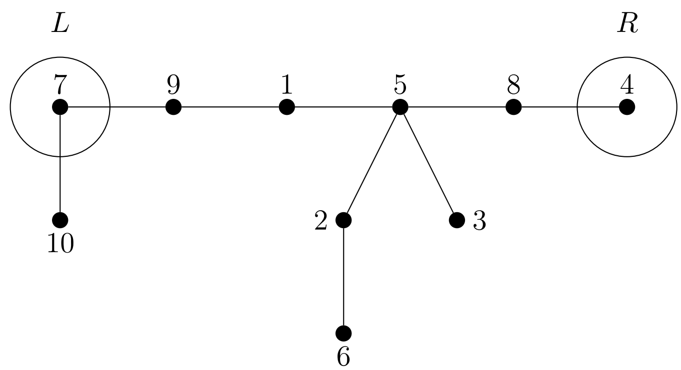
$$f|_M = \begin{pmatrix} v_1 & \dots & v_k \\ f(v_1) & \dots & f(v_k) \end{pmatrix},$$

where $v_1 < v_2 < \cdots < v_k$ (and $M = \{v_1, v_2, \dots, v_k\}$). This gives us the ordering $(f(v_1), \dots, f(v_k))$. Now we can choose $L = f(v_1)$, $R = f(v_k)$. The tree T corresponding to f is constructed as follows: Draw a (directed) path $f(v_1), f(v_2), \dots, f(v_k)$, and fill in the remaining vertices as in G_f (removing edge directions).

Example (continued)**.**

$$M = \{1, 4, 5, 7, 8, 9\},$$
$$f|_M = \begin{pmatrix} 1 & 4 & 5 & 7 & 8 & 9 \\ 7 & 9 & 1 & 5 & 8 & 4 \end{pmatrix},$$

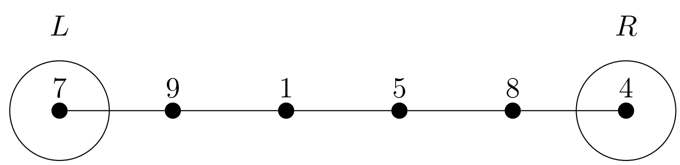
$T = T(f) =$



Reversing the correspondence is easy: given a tree T with two special vertices L and R, look at the unique path P of T connecting L and R. The vertices of P form the set M. Ordering the vertices of M gives us the first line of $f|_M$, the second line is given by the order of the vertices in P, from L to R.

Example (continued)**.**

$P =$



$$M = \{1, 4, 5, 7, 8, 9\},$$
$$f|_M = \begin{pmatrix} 1 & 4 & 5 & 7 & 8 & 9 \\ 7 & 9 & 1 & 5 & 8 & 4 \end{pmatrix}.$$

The remaining values of f are then filled in accordance with the unique paths from the remaining vertices to P (directing these paths towards P). It is not too hard to check, but we will not prove it rigorously, that this is the inverse of the previous procedure. $\square$

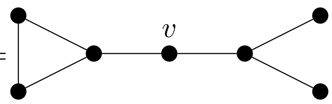
3 Connectivity

3.1 Vertex connectivity

Definition 3.1. A *vertex cut* in a connected graph $G = (V, E)$ is a set $S \subseteq V$ such that $G \backslash S := G[V \backslash S]$ has more than one connected component. A *cut vertex* is a vertex v such that $\{v\}$ is a cut.

Definition 3.2. G is called *k-connected* if $|V(G)| > k$ and $G \backslash X$ is connected for every set $X \subseteq V(G)$ with $|X| < k$. In other words, no two vertices of G are separated by fewer than k other vertices. Every (non-empty) graph is 0-connected and the 1-connected graphs are precisely the non-trivial connected graphs. The greatest integer k such that G is k-connected is the *connectivity* $\kappa(G)$ of G.

- $G = K_n$: $\kappa(G) = n - 1$

- $G = K_{m,n}$, $m \leq n$: $\kappa(G) = m$. Indeed, let G have bipartition $A \cup B$, with $|A| = m$ and $|B| = n$. Deleting A disconnects the graph. On the other hand, deleting $S \subset V$ with $|S| < m$ leaves both $A \backslash S$ and $B \backslash S$ non-empty and any $a \in A \setminus S$ is connected to any $b \in B \setminus S$. Hence $G \backslash S$ is connected.
- $G =$  : $\kappa(G) = 1$. Deleting v disconnects G, so v is a cut vertex.

Proposition 3.3. *For every graph G, $\kappa(G) \leq \delta(G)$.*

Proof. If G is a complete graph then trivially $\kappa(G) = \delta(G) = |G| - 1$. Otherwise let $v \in G$ be a vertex of minimum degree $d(v) = \delta(G)$. Deleting $N(v)$ disconnects v from the rest of G. □

Remark 3.4. High minimum degree does not imply connectivity. Consider two disjoint copies of K_n.

Theorem 3.5 (Mader 1972)**.** *Every graph of average degree at least $4k$ has a k-connected subgraph.*

Proof. For $k \in \{0, 1\}$ the assertion is trivial; we consider $k \geq 2$ and a graph $G = (V, E)$ with $|V| = n$ and $|E| = m$. For inductive reasons it will be easier to prove the stronger assertion that G has a k-connected subgraph whenever

(i) $n \geq 2k - 1$ and

(ii) $m \geq (2k - 3)(n - k + 1) + 1$.

(This assertion is indeed stronger, i.e. (i) and (ii) follow from our assumption of $\bar{d}(G) \geq 4k$: (i) holds since $n > \Delta(G) \geq 4k$, while (ii) follows from $m = \frac{1}{2}\bar{d}(G)n \geq 2kn$.)

We apply induction on n. If $n = 2k - 1$, then $k = \frac{1}{2}(n + 1)$, and hence

$$m \geq (n - 2)\frac{n + 1}{2} + 1 = \frac{1}{2}n(n - 1)$$

by (ii). Thus $G = K_n \supseteq K_{k+1}$, proving our claim. We therefore assume that $n \geq 2k$. If v is a vertex with $d(v) \leq 2k - 3$, we can apply the induction hypothesis to $G \backslash v$ and are done. So we assume that $\delta(G) \geq 2k - 2$. If G is itself not k-connected, then there is a separating set $X \subseteq V$ with less than k vertices, such that $G \backslash X$ is disconnected. Let V_1 be one component of $G \backslash X$ and let V_2 be the union of the other components. Note that there are no edges in G between V_1 and V_2. Let $G_i = G[V_i \cup X]$, so that $G = G_1 \cup G_2$, and every edge of G is either in G_1 or G_2 (or both). Each vertex in each V_i has at least $\delta(G) \geq 2k - 2$ neighbours in G and thus also in G_i, so $|G_1|, |G_2| \geq 2k - 1$. Note that each $|G_i| < n$, so by the induction hypothesis, if no G_i has a k-connected subgraph then each

$$e(G_i) \leq (2k - 3)(|G_i| - k + 1).$$

Hence,

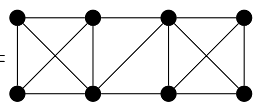
$$\begin{aligned} m &\leq e(G_1) + e(G_2) \\ &\leq (2k - 3)(|G_1| + |G_2| - 2k + 2) \\ &\leq (2k - 3)(n - k + 1) \quad \text{(since } |V(G_1) \cap V(G_2)| = |X| \leq k - 1\text{)}, \end{aligned}$$

contradicting (ii). □

3.2 Edge connectivity

Definition 3.6. A *disconnecting set* of edges is a set $F \subseteq E(G)$ such that $G \setminus F$ has more than one component. Given $S, T \subset V(G)$, the notation $[S, T]$ specifies the set of edges having one endpoint in S and the other in T. An *edge cut* is an edge set of the form $[S, \overline{S}]$, where S is a non-empty proper subset of $V(G)$. A graph is *k-edge-connected* if every disconnecting set has at least k edges. The *edge-connectivity* of G, written $\kappa'(G)$, is the minimum size of a disconnecting set. One edge disconnecting G is called a *bridge*.

Example 3.7.

- $G = K_n$: $\kappa'(G) = n - 1$.
- $G =$  : $\kappa'(G) = 3$, whereas $\kappa(G) = 2$.

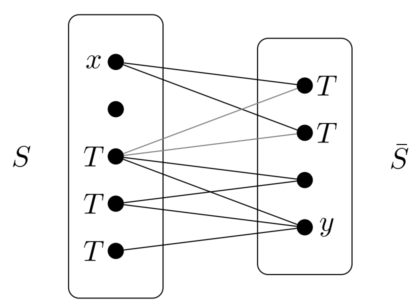
Remark 3.8. An edge cut is a disconnecting set but not the other way around. However, every minimal disconnecting set is a cut.

Theorem 3.9. $\kappa(G) \le \kappa'(G) \le \delta(G)$.

Proof. The edges incident to a vertex v of minimum degree form a disconnecting set; hence $\kappa'(G) \le \delta(G)$. It remains to show $\kappa(G) \le \kappa'(G)$. Suppose $|G| > 1$ and $[S, \overline{S}]$ is a minimum edge cut, having size $\kappa'(G)$.

If every vertex of S is adjacent to every vertex of $\overline{S}$, then $\kappa'(G) = |S||\overline{S}| = |S|(|G| - |S|)$. This expression is minimized at $|S| = 1$. By definition, $\kappa(G) \le |G| - 1$, so the inequality holds.

Hence we may assume there exists $x \in S$, $y \in \overline{S}$ with x not adjacent to y. Let T be the vertex set consisting of all neighbours of x in $\overline{S}$ and all vertices of $S \backslash x$ that have neighbours in $\overline{S}$ (illustrated below). Deleting T destroys all the edges in the cut $[S, \overline{S}]$ (but does not delete x or y), so T is a separating set. Now, by the definition of T we can injectively associate at least one edge of $\left[S, \overline{S}\right]$ to each vertex in T, so $\kappa(G) \le |T| \le \left|[S, \overline{S}]\right| = \kappa'(G)$.



∎

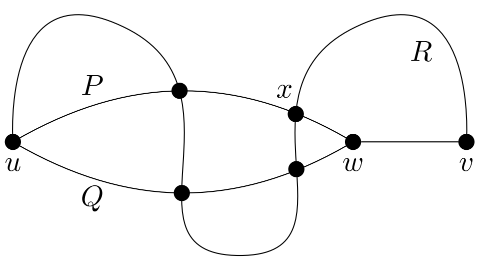
3.3 2-connected graphs

Definition 3.10. Two paths are *internally disjoint* if neither contains a non-endpoint vertex of the other. We denote the length of the shortest path from u to v (the *distance* from u to v) by $d(u, v)$.

Theorem 3.11 (Whitney 1932)**.** *A graph G having at least three vertices is 2-connected if and only if each pair $u, v \in V(G)$ is connected by a pair of internally disjoint u, v-paths in G.*

Proof. When G has internally disjoint u, v-paths, deletion of one vertex cannot separate u from v. Since this is given for every u, v, the condition is sufficient. For the converse, suppose that G is 2-connected. We prove by induction on $d(u, v)$ that G has two internally disjoint u, v paths. When $d(u, v) = 1$, the graph $G \backslash (u, v)$ is connected, since $\kappa'(G) \geq \kappa(G) = 2$. A u, v-path in $G \backslash (u, v)$ is internally disjoint in G from the u, v-path consisting of the edge (u, v) itself.

For the induction step, we consider $d(u, v) = k > 1$ and assume that G has internally disjoint x, y-paths whenever $1 \leq d(x, y) < k$. Let w be the vertex before v on a shortest u, v-path. We have $d(u, w) = k - 1$, and hence by the induction hypothesis G has internally disjoint u, w-paths P and Q. We may assume that neither P nor Q contains v, else we have two internally disjoint paths between u and v, using the two arcs of the cycle formed by P and Q. Since $G \backslash w$ is connected, $G \backslash w$ contains a u, v-path R. If this path avoids P or Q, we are finished, but R may share internal vertices with both P and Q. Let x be the last vertex of R belonging to $P \cup Q$. Without loss of generality, we may assume $x \in P$. We combine the u, x-subpath of P with the x, v-subpath of R to obtain a u, v-path internally disjoint from $Q \cup \{(w, v)\}$.



□

Corollary 3.12. *G is 2-connected if and only if every two vertices in G lie on a common cycle and $|G| \geq 3$.*

3.4 Menger's Theorem

Definition 3.13. Let $A, B \subseteq V(G)$. An A-B path is a path in which the first vertex is in A but the other vertices are not in A, and in which the last vertex is in B but the other vertices are not in B. Any vertex in $A \cap B$ is a trivial A-B path.

If $X \subseteq V(G)$ (or $X \subseteq E(G)$) is such that every A-B path in G contains a vertex (or an edge) from X, we say that X *separates* the sets A and B in G. This implies in particular that $A \cap B \subseteq X$.

Theorem 3.14 (Menger 1927)**.** *Let $G = (V, E)$ be a graph and let $S, T \subseteq V$. Then the maximum number of vertex-disjoint S-T paths is equal to the minimum size of an S-T separating vertex set.*

Proof. Obviously, the maximum number of disjoint paths does not exceed the minimum size of a separating set, because for any collection of disjoint paths, any separating set must contain a vertex from each path. So we just need to prove there is an S-T separating set and a collection of disjoint S-T paths with the same size.

We use induction on $|V| + |E|$. If $E = \varnothing$ then the only minimal S-T separating vertex set is $S \cap T$. Also, trivially, there are $|S \cap T|$ vertex-disjoint S-T paths, each being a single vertex from $S \cap T$. So the claim holds.

We now do the induction step, assuming $E \neq \varnothing$. We first consider the case where S and T are disjoint. Let k be the minimum size of an S-T separating vertex set. Choose $e = (u, v) \in E$. Let $G' = (V, E \setminus e)$. If each S-T separating vertex set in G' has size at least k, then inductively there exist k vertex-disjoint S-T paths in G', hence in G.

So we can assume that G' has an S-T separating vertex set C of size at most $k-1$. Then $C \cup \{u\}$ and $C \cup \{v\}$ are S-T separating vertex sets of G of size k.

Since C is a separating set for G', no component of $G' \backslash C$ has elements from both S and T. Let V_S be the union of components with elements from S, and let V_T be the union of components with elements in T. If we were to add the edge (u, v) to $G' \backslash C$ then there would be a path from S to T (because C does not separate S and T in G). So, without loss of generality $u \in V_S$ and $v \in V_T$.

Now we claim that each S-$(C \cup \{u\})$ separating vertex set of G' is S-T separating in G. Indeed, each S-T path P in G intersects $C \cup \{u\}$. Let P' be the subpath of P that goes from S to the first time it touches $C \cup \{u\}$. If P' ends with a vertex in C, then $u \notin P'$ so P' is an S-$(C \cup \{u\})$ path in G'. If P' ends in u, then it is disjoint from C and so by the above it contains only vertices in V_S. So $v \notin P'$ and again P' is an S-$(C \cup \{u\})$ path in G'. In both cases we showed that P' is an S-$(C \cup \{u\})$ path in G'. This implies that every S-T path in G contains an S-$(C \cup \{u\})$-path in G'. Hence, indeed each S-$(C \cup \{u\})$ separating vertex set of G' is S-T separating in G, so every such set has size at least k.

Therefore, by induction, G' contains k disjoint S-$(C \cup \{u\})$ paths. Similarly, G' contains k disjoint $(C \cup \{v\})$-T paths. Any path in the first collection intersects any path in the second collection only in C, since otherwise G' contains an S-T path avoiding C.

Hence, as $|C| = k - 1$, we can pairwise concatenate these paths to obtain $k - 1$ disjoint S-T paths. We can finally obtain a kth path by inserting the edge e between the path ending at u and the path starting at v.

It remains to consider the situation where S and T are not disjoint. Let $X = S \cap T$ and apply the induction hypothesis with the sets $S' = S \backslash X$ and $T' = T \backslash X$, in the graph $G' = G \backslash X$. Let k' be the size of a minimum separating set in G'. We can obtain a $(k' + |X|)$-vertex S-T separating set in G by adding every vertex in X to an S'-T' separating set in G'. Similarly we can obtain a collection of $k' + |X|$ vertex-disjoint S-T paths by adding each vertex in X as a trivial path to a collection of k' vertex-disjoint S'-T' paths in G'. □

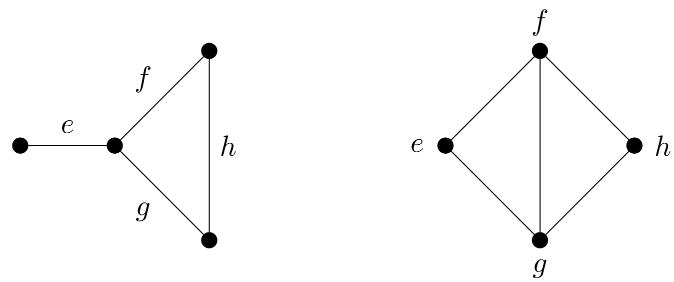
Corollary 3.15 (Fan lemma)**.** *For $S \subseteq V(G)$ and $v \in V(G) \setminus S$, the minimum number of vertices distinct from v separating v from S in G is equal to the maximum number of paths forming an v-S fan in G (that is, the maximum number of $\{v\}$-S paths which are disjoint except at v).*

Proof. Let $T = N(v)$. Note that the minimum number of vertices distinct from v separating v from S is equal to the minimum size of an S-T separating set. Now apply Menger's Theorem with S and T. Note that none of the resulting paths go through v; if one did, then it would contain two vertices of T, violating the definition of an S-T path. So we have a suitable number of vertex-disjoint S-T paths not including v, and we can append v to each path to give a v-S fan.

The reverse inequality is clear, because to separate v from S one needs to delete at least one vertex from each $\{v\}$-S path in a v-S fan. $\square$

Definition 3.16. The *line graph* of G, written $L(G)$, is the graph whose vertices are the edges of G, with $(e, f) \in E(L(G))$ when $e = (u, v)$ and $f = (v, w)$ in G (i.e. when e and f share a vertex).

Example 3.17.



G $\qquad$ $L(G)$

Note that a path in $L(G)$ corresponds to a sequence of distinct edges $e_0, \ldots, e_\ell$ in G such that every pair of consecutive edges is incident. This may not be a path in G, because the endpoints might not "line up": it might be that for some i the common endpoint of e_{i-1} and e_i is the same as the common endpoint of e_i and e_{i+1}. But we can always delete a few edges to obtain a path in G: if there exists i as described, we can simply delete the edge e_i, and repeatedly make such deletions until no conflicts remain.

Corollary 3.18. *Let u and v be two distinct vertices of G.*

(i) If $(u, v) \notin E$, then the minimum number of vertices different from u, v separating u from v in G is equal to the maximum number of internally vertex-disjoint u-v paths in G.

(ii) The minimum number of edges separating u from v in G is equal to the maximum number of edge-disjoint u-v paths in G.

Proof. For (i), apply Menger's Theorem with $S = N(u)$ and $T = N(v)$.

For (ii), apply Menger's Theorem to the line graph of G, with S as the set of edges adjacent to u and T as the set of edges adjacent to v. Filling the details is left as an exercise. $\square$

Theorem 3.19 (Global Version of Menger's Theorem)**.**

(a) A graph is k-connected if and only if it contains k internally vertex-disjoint paths between any two vertices (and has at least two vertices).

(b) A graph is k-edge-connected if and only if it contains k edge-disjoint paths between any two vertices.

Proof. First we prove (a). If a graph G contains k internally disjoint paths between any two vertices, then $|G| > k$ and G cannot be separated by fewer than k vertices; thus, G is k-connected.

Conversely, suppose that G is k-connected (and, in particular, has more than k vertices) but contains vertices u, v not linked by k internally disjoint paths. By Corollary 3.18, u and v are adjacent; let

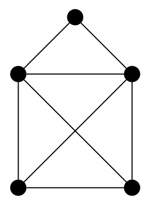
$G' = G\backslash(u,v)$. Then G' contains at most $k-2$ internally disjoint u,v-paths. By Corollary 3.18, we can separate u and v in G' by a set X of at most $k-2$ vertices. As $|G| > k$, there is at least one further vertex $w \notin X \cup \{u,v\}$ in G. Now X separates w in G' from either u or v (say, from u). But then $X \cup \{v\}$ is a set of at most $k-1$ vertices separating w from u in G, contradicting the k-connectedness of G.

(b) follows straight from Corollary 3.18. □

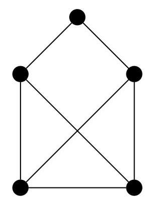
4 Eulerian and Hamiltonian cycles

4.1 Eulerian trails and tours

Question 4.1. Which of the two pictures below can be drawn in one go without lifting your pen from the paper?



or

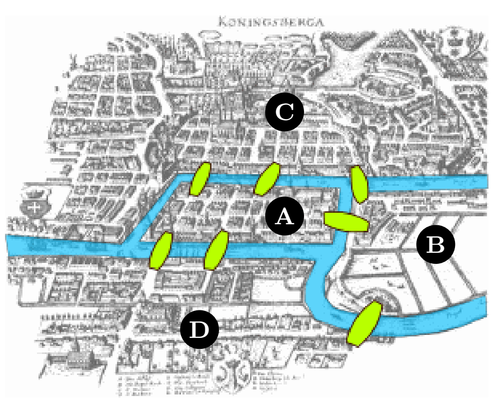


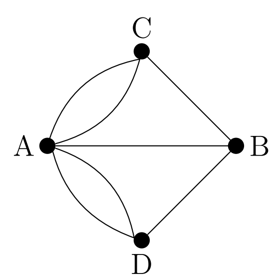
Definition 4.2. A *trail* is a walk with no repeated edges.

Definition 4.3. An *Eulerian trail* in a (multi)graph G is a walk in G passing through every edge exactly once. If this walk is closed (starts and ends at the same vertex) it is called an *Eulerian tour*.

One motivation for this concept is the "7 bridges of Königsberg" problem:

Question 4.4. Is it possible to design a closed walk passing through all the 7 bridges exactly once? Equivalently, does the graph on the right have an Eulerian tour?





Theorem 4.5. *A connected (multi)graph has an Eulerian tour if and only if each vertex has even degree.*

In order to prove this theorem we use the following lemma.

Lemma 4.6. *Every maximal trail in an even multigraph (i.e., a multigraph where all the vertices have even degree) is a closed trail.*

Proof. Let T be a maximal trail. If T is not closed, then T has an odd number of edges incident to the final vertex v. However, as v has even degree, there is an edge incident to v that is not in T. This edge can be used to extend T to a longer trail, contradicting the maximality of T. $\square$

Proof of Theorem 4.5. To see that the condition is necessary, suppose G has an Eulerian tour C. If a vertex v was visited k times in the tour C, then each visit used 2 edges incident to v (one incoming edge and one outgoing edge). Thus, $d(v) = 2k$, which is even.

To see that the condition is sufficient, let G be a connected graph with even degrees. Let $T = e_1e_2\ldots e_\ell$ (where $e_i = (v_{i-1}, v_i)$) be a longest trail in G. Then, by Lemma 4.6, T is closed, i.e., $v_0 = v_\ell$. If T does not include all the edges of G then, since G is connected, there is an edge e outside of T such that $e = (u, v_i)$ for some vertex v_i in T. But then $T' = ee_{i+1}\ldots e_\ell e_1 e_2 \ldots e_i$ is a trail in G which is longer than T, contradicting the fact that T is a longest trail in G. Thus, we conclude that T includes all the edges of G and so it is an Eulerian tour. $\square$

Corollary 4.7. *A connected multigraph G has an Eulerian trail if and only if it has either* 0 *or* 2 *vertices of odd degree.*

Proof. Suppose T is an Eulerian trail from vertex u to vertex v. If $u = v$ then T is an Eulerian tour and so by Theorem 4.5 it follows that all the vertices in G have even degree. If $u \neq v$, note that the multigraph $G \cup \{e\}$, where $e = (u, v)$ is a new edge, has an Eulerian tour, namely $T \cup \{e\}$. By Theorem 4.5 it follows that all the degrees in $G \cup \{e\}$ are even. Thus, we conclude that, in the original multigraph G, the vertices u, v are the only ones which have odd degree.

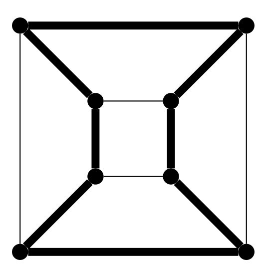
Now we prove the other direction of the corollary. If G has no vertices of odd degree then by Theorem 4.5 it contains an Eulerian tour which is also an Eulerian trail. Suppose now that G has 2 vertices u, v of odd degree. Then $G \cup \{e\}$, where $e = (u, v)$ is a new edge, only has vertices of even degree and so, by Theorem 4.5, it has an Eulerian tour C. Removing the edge e from C gives an Eulerian trail of G from u to v. $\square$

4.2 Hamilton paths and cycles

Definition 4.8. A Hamilton path/cycle in a graph G is a path/cycle visiting every vertex of G exactly once. A graph G is called Hamiltonian if it contains a Hamilton cycle.

Hamilton cycles were introduced by Kirkman in 1856, and were named after Sir William Hamilton, who produced a puzzle whose goal was to find a Hamilton cycle in a specific graph.

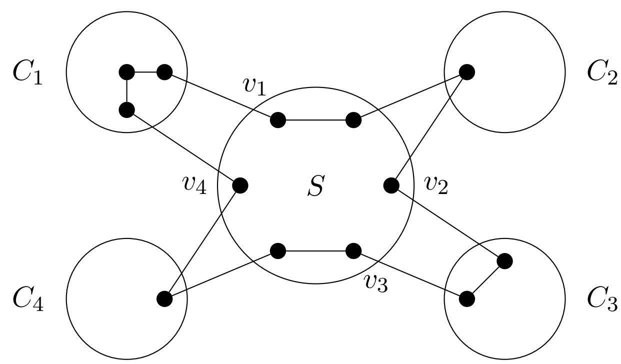
Example 4.9. Hamilton cycle in the skeleton of the 3-dimensional cube.



We give some necessary conditions for Hamiltonicity.

Proposition 4.10. *If G is Hamiltonian then for any non-empty set $S \subseteq V$ the graph $G \backslash S$ has at most $|S|$ connected components.*

Proof. Let $C_1, \ldots, C_k$ be the components of $G \backslash S$. Imagine that we are moving along a Hamilton cycle in some order, vertex-by-vertex (in the picture below, we are moving clockwise, starting from some vertex in C_1, say). We must visit each component of $G \backslash S$ at least once; when we leave C_i for the first time, let v_i be the subsequent vertex visited (which must be in S). Each v_i must be distinct because a cycle cannot intersect itself. Hence, S must have at least as many vertices as the number of connected components of $G \backslash S$.



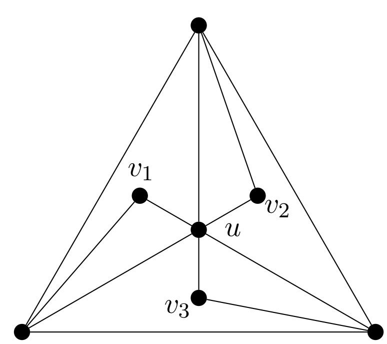
□

Corollary 4.11. *If a connected bipartite graph $G = (V, E)$ with bipartition $V = A \cup B$ is Hamiltonian then $|A| = |B|$.*

Proof. By deleting the vertices in A from G we get $|B|$ isolated vertices and so $G \backslash A$ has $|B|$ connected components. Thus, by Proposition 4.10 we conclude that $|A| \geq |B|$. By symmetry we can also show that $|B| \geq |A|$. Thus, we conclude that $|A| = |B|$. □

Example 4.12. The condition in Proposition 4.10 is not sufficient to ensure that a graph is Hamiltonian. The graph G below satisfies the condition of Proposition 4.10 but is not Hamiltonian. Indeed, one would need to include all the edges incident to the vertices v_1, v_2 and v_3 in a Hamilton cycle of G; however, in that case the vertex u would have degree at least 3 in that Hamilton cycle, which is impossible.

$G =$

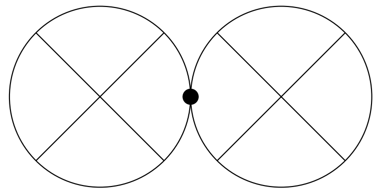


We also give some sufficient conditions for Hamiltonicity.

Theorem 4.13 (Dirac 1952)**.** *If G is a simple graph with $n \geq 3$ vertices and if $\delta(G) \geq n/2$, then G is Hamiltonian.*

Example 4.14. The minimum degree condition in Theorem 4.13 is best possible. Indeed:

- The graph consisting of two cliques of orders $\lfloor (n+1)/2 \rfloor$ and $\lceil (n+1)/2 \rceil$ sharing a vertex has minimum degree $\lfloor (n-1)/2 \rfloor$ but is not Hamiltonian (it is not even 2-connected).

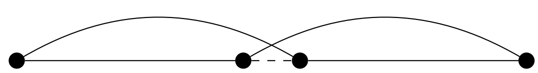


- If n is odd, then the complete bipartite graph $K_{(n-1)/2,(n+1)/2}$ has minimum degree $\frac{n-1}{2}$ but is not Hamiltonian.

Also, the condition that $n \geq 3$ must be included since K_2 is not Hamiltonian but satisfies $\delta(K_2) = |K_2|/2$.

Proof of Theorem 4.13. If there is a non-Hamiltonian graph satisfying the hypotheses, then adding edges cannot reduce the minimum degree, so we may restrict our attention to *maximal* non-Hamiltonian graphs G with minimum degree at least $n/2$. By "maximal" we mean that for every pair (u,v) of non-adjacent vertices of G, the graph obtained from G by adding the edge $e = (u,v)$ is Hamiltonian.

The maximality of G implies that G has a Hamilton path (since every Hamilton cycle in $G \cup \{e\}$ must contain a Hamilton path in G). Take a Hamilton path $v_1, v_2, \ldots, v_n$ in G. Let $u = v_1$ and $v = v_n$. We use most of this path with a small switch, to obtain a Hamilton cycle in G. If some neighbour of u immediately follows a neighbour of v on the path, say $(u, v_{i+1}) \in E(G)$ and $(v, v_i) \in E(G)$, then G has the Hamilton cycle $(u, v_{i+1}, v_{i+2}, \ldots, v_{n-1}, v, v_i, v_{i-1}, \ldots, v_2)$ shown below.



To prove that such a cycle exists, we show that there is a common index in the sets S and T defined by $S = \{1 \leq i \leq n-1 : (u, v_{i+1}) \in E(G)\}$ and $T = \{1 \leq i \leq n-1 : (v, v_i) \in E(G)\}$. Summing the sizes of these sets yields

$$|S \cup T| + |S \cap T| = |S| + |T| = d(u) + d(v) \geq n.$$

Neither S nor T contains the index n. This implies that $|S \cup T| < n$, and hence $|S \cap T| \geq 1$, as required. This is a contradiction. $\square$

Ore observed that this argument uses only that $d(u) + d(v) \geq n$. Therefore, we can weaken the requirement of minimum degree $n/2$ to require only that $d(u) + d(v) \geq n$ whenever u is not adjacent to v.

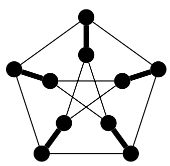
Theorem 4.15 (Ore 1960)**.** *If G is a simple graph with $n \geq 3$ vertices such that for every pair of non-adjacent vertices u, v of G we have $d(u) + d(v) \geq n$, then G is Hamiltonian.*

5 Matchings

Definition 5.1. A set of edges $M \subseteq E(G)$ in a graph G is called a *matching* if $e \cap e' = \varnothing$ for any pair of (distinct) edges $e, e' \in M$.

A matching is *perfect* if $|M| = \frac{|V(G)|}{2}$, i.e. it covers all vertices of G. We denote the size of the maximum matching in G, by $\nu(G)$.

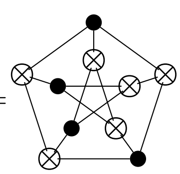
Example 5.2.

- $G = K_n$: $\nu(G) = \lfloor \frac{n}{2} \rfloor$
- $G = K_{s,t}$, $s \leq t$: $\nu(G) = s$
- $G =$  : $\nu(G) = 5$

Remark 5.3. A matching in a graph G corresponds to an independent set in the line graph $L(G)$

Definition 5.4. A set of vertices $T \subseteq V(G)$ of a graph G is called a *cover* of G if every edge $e \in E(G)$ intersects T ($e \cap T \neq \varnothing$), i.e., $G \setminus T$ is an empty graph. Then, $\tau(G)$ denotes the size of the minimum cover.

Example 5.5.

- $G = K_n$: $\tau(G) = n - 1$
- $G = K_{s,t}$, $s \leq t$: $\tau(G) = s$
- $G =$  : $\tau(G) = 6$.

  To see this, note that the graphs induced by the outer 5 vertices and inner 5 vertices are both 5-cycles C_5. Since $\tau(C_5) = 3$, at least 3 of the outer vertices and 3 of the inner vertices must be included in a vertex cover.

Proposition 5.6. $\nu(G) \leq \tau(G) \leq 2\nu(G)$.

Proof. Let M be a maximum matching in G. Since every cover has at least one vertex in each edge of M and edges are disjoint, we have $\nu(G) \leq \tau(G)$. Note also that since M is maximal, every edge $e \in E(G)$ intersects some edge $e' \in M$, otherwise we could get a larger matching. So the vertices covered by M form a cover for G, hence $\tau(G) \leq 2|M| = 2\nu(G)$. □

5.1 Real-world applications of matchings

Here are two situations where it is useful to think about and look for matchings.

- Suppose certain workers can operate certain machines, but only one at a time; this gives a bipartite graph between workers and machines. If we want to have many machines operating at the same time, we need a large matching in our bipartite graph.
- When students apply to semester projects, each student has a list of project preferences, and each professor also has a list of preferences for students. In order to decide which students should be assigned to which project, we need to find a bipartite matching that is somehow compatible with these preferences. This kind of situation is called the *stable matching* problem, and is extremely important in economics and operations research. The Gale-Shapley algorithm for efficiently computing a stable matching was worth the Nobel prize in economics in 2012.

5.2 Hall's Theorem

Theorem 5.7 (Hall 1935)**.** *A bipartite graph $G = (V, E)$ with bipartition $V = A \cup B$ has a matching covering A if and only if*

$$|N(S)| \geq |S| \quad \forall S \subseteq A. \tag{1}$$

Proof. It is easy to see that if G has such a matching then (1) holds.

To show the other direction, we apply induction on $|A|$. For $|A| = 1$ the assertion is true. Now let $|A| \geq 2$, and assume that (1) is sufficient for the existence of a matching covering A when $|A|$ is smaller.

If $|N(S)| \geq |S| + 1$ for every non-empty set $S \subsetneq A$, then we pick an edge $(a, b) \in G$ and consider the graph $G' = G \backslash \{a, b\}$ obtained by deleting the vertices a and b. Then every non-empty set $S \subseteq A \setminus \{a\}$ satisfies

$$|N_{G'}(S)| \geq |N_G(S)| - 1 \geq |S|,$$

so by the induction hypothesis G' contains a matching covering $A \setminus \{a\}$. Together with the edge ab, this yields a matching covering A in G.

Suppose now that A has a non-empty proper subset A' with neighbourhood $B' = N(A')$ such that $|A'| = |B'|$. It is easy to verify that $G' = G[A' \cup B']$ satisfies condition (1) and therefore, by the induction hypothesis, G' contains a matching covering A'. But $G \backslash (A' \cup B')$ satisfies (1) as well: for any set $S \subseteq A \setminus A'$ with $|N_{G \backslash (A' \cup B')}(S)| < |S|$ we would have $|N_G(S \cup A')| = |N_{G \backslash (A' \cup B')}(S)| + |B'| < |S \cup A'|$, contrary to our assumption. Again, by induction, $G \backslash (A' \cup B')$ contains a matching covering $A \setminus A'$. Putting the two matchings together, we obtain a matching in G covering A. $\square$

The next corollary gives a so-called *defect version* of Hall's theorem.

Corollary 5.8. *If in a bipartite graph $G = (A \cup B, E)$ we have $|N(S)| \geq |S| - d$ for every set $S \subseteq A$ and some fixed $d \in \mathbb{N}$, then G contains a matching of cardinality $|A| - d$.*

Proof. We add d new vertices to B, joining each of them to all the vertices in A. Call the resulting graph G'. Note that the new graph has

$$N_{G'}(S) \geq |N_G(S)| + d \geq |S| - d + d = |S|$$

for any $S \subseteq A$, so by Hall's theorem, G' contains a matching of A. At least $|A| - d$ edges in this matching must be edges of G. $\square$

Corollary 5.9. *If a bipartite graph $G = (A \cup B, E)$ is k-regular with $k \geq 1$, then G has a perfect matching.*

Proof. If G is k-regular, then clearly $|A| = |B|$, since the total number of edges is $k|A| = \sum_{x \in A} d(x) = \sum_{y \in B} d(y) = k|B|$. It thus suffices to show that G contains a matching covering A. Now every set $S \subseteq A$ is joined to $N(S)$ by a total of $k|S|$ edges, and these are among the $k|N(S)|$ edges of G incident with $N(S)$. Therefore $k|S| \leq k|N(S)|$, so G satisfies (1) and we can apply by Theorem 5.7. $\square$

Corollary 5.10. *Every regular graph of positive even degree has a 2-factor (a spanning 2-regular subgraph).*

Proof. We may assume that the graph is connected because it is enough to prove the statement for each connected component. So let G be any connected $2k$-regular graph. By Theorem 4.5, G contains an Euler tour. Define a new graph G' by splitting every vertex v into two vertices v^- and v^+. If an edge of the Euler tour goes from v to w, put an edge in G' from v^+ to w^-. So, the edges in G and in G' naturally correspond to each other. It is easy to see that G' is bipartite and k-regular so contains a perfect matching. Collapsing each pair of vertices v^-, v^+ back into a single vertex v, a perfect matching of G' corresponds to a 2-factor of G. (Each vertex v is incident to one edge which was incident to v^+ in G', and one edge incident to v^- in G'). $\square$

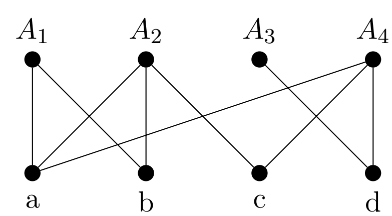
Remark 5.11. A 2-factor is a disjoint union of cycles covering all the vertices of the graph.

Definition 5.12. Let $A_1, \ldots, A_n$ be a collection of sets. A family $\{a_1, \ldots, a_n\}$ is called a *system of distinct representatives* (SDR) if all the a_i are distinct, and $a_i \in A_i$ for all i.

Corollary 5.13. *A collection $A_1, \ldots, A_n$ of finite sets has an SDR if and only if for all $I \subseteq [n]$ we have $|\bigcup_{i \in I} A_i| \geq |I|$.*

Proof. Define a bipartite graph with parts $A = [n]$ and $B = \bigcup_i A_i$ such that (i, a) is an edge if and only if $a \in A_i$. A matching covering $[n]$ in this graph corresponds exactly to an SDR, where an edge (i, a) in the matching means that $a_i = a$. But the condition $|\bigcup_{i \in I} A_i| \geq |I|$ for all $I \subseteq [n]$ is precisely Hall's condition for the existence of a matching covering A, so Hall's theorem provides the desired equivalence. $\square$

Example 5.14. $A_1 = \{a, b\}$, $A_2 = \{a, b, c\}$, $A_3 = \{d\}$, $A_4 = \{a, c, d\}$.

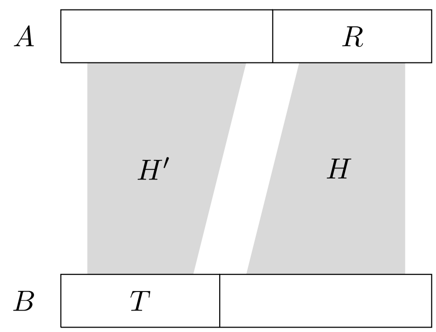


Theorem 5.15 (Kőnig 1931)**.** *If $G = (A \cup B, E)$ is a bipartite graph, then the maximum size of a matching in G equals the minimum size of a vertex cover of G.*

Proof. We have already seen that a minimum cover has at least the size of a maximum matching. Now take a minimum vertex cover U of G. We construct a matching of size $|U|$ to prove that equality can always be achieved.

Let $R = U \cap A$ and $T = U \cap B$. Let H, H' be the subgraphs of G induced by $R \cup (B \backslash T)$ and $T \cup (A \backslash R)$. We use Hall's theorem to show that H has a complete matching of R into $B \backslash T$ and H' has a complete matching of T into $A \backslash R$. Since these subgraphs are disjoint, the two matchings together form a matching of size $|U|$ in G.

Suppose $S \subseteq R$, and consider $N_H(S) \subseteq B \backslash T$. If $|N_H(S)| < |S|$, then we can substitute $N_H(S)$ for S in U and obtain a smaller vertex cover, since $N_H(S)$ covers all edges incident to S that are not covered by T. The minimality of U thus implies that Hall's condition holds in H, and hence H has a complete matching of R into $B \backslash T$. Applying the same argument to H' yields the rest of the matching.



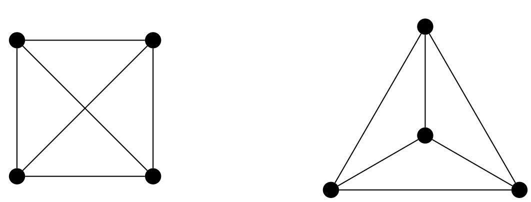
□

6 Planar Graphs

Definition 6.1. A *polygonal path* or *polygonal curve* in the plane is the union of finitely many line segments such that each segment starts at the end of the previous one and the segments are pairwise disjoint except that consecutive segments intersect in exactly one point. In a *polygonal* u, v*-path*, the beginning of the first segment is u and the end of the last segment is v.

A *drawing* of a graph G is a function that maps each vertex $v \in V(G)$ to a point $f(v)$ in the plane and each edge uv to a polygonal $f(u), f(v)$-path in the plane. The images of vertices are distinct. A point in $f(e) \cap f(e')$ other than a common end is a *crossing*. A graph is *planar* if it has a drawing without crossings. Such a drawing is a *planar embedding* of G. A *plane graph* is a particular drawing of a planar graph in the plane with no crossings.

Example 6.2.



K_4 planar drawing

Remark 6.3. We get the same class of graphs if we allow the images of edges to be continuous curves. This is because any continuous curve can be arbitrarily accurately approximated by a polygonal curve. Furthermore, by a theorem of Fáry, we get the same class of graphs if we only allow the images of edges to be straight line segments (rather than polygonal curves).

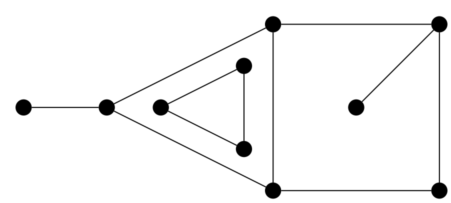
Definition 6.4. An *open set* in the plane is a set $U \subset \mathbb{R}^2$ such that for every $p \in U$, all points within some small distance from p belong to U. A *region* is an open set U that contains a polygonal u, v-path for every pair $u, v \in U$ (that is, it is "path-connected"). The *faces* of a plane graph are the maximal regions of the plane that are disjoint from the drawing.

Remark 6.5. The faces of G are pairwise disjoint (they are separated by the edges of G). Two points are in the same face if and only if there is a polygonal path between them which does not cross an edge of G. Also, note that a finite graph has a single unbounded face (the area "outside" of the graph).

Proposition 6.6. *A plane forest has exactly one face.*

Definition 6.7. The *length* of the face f in a planar embedding of G is the number of edges in the boundary of f, where an edge is counted twice if it is not contained in the boundary of any other face.

Example 6.8. The following graph has 4 faces of lengths 6, 6, 3 and 7.



For example, the outer face has length 7; the leftmost horizontal edge is counted twice as it only touches the outer face and no other.

Proposition 6.9. *If $l(f_i)$ denotes the length of a face f_i in a plane graph G, then $2e(G) = \sum_i l(f_i)$.*

Proof. In the sum $\sum_i l(f_i)$, every edge was counted twice. □

Theorem 6.10 (Euler's formula 1758)**.** *If a connected plane graph G has exactly n vertices, e edges and f faces, then $n - e + f = 2$.*

Proof. We use induction on the number of edges in G. If $e(G) = n - 1$ and G is connected, then G is a tree. We have $f = 1$, $e = n - 1$. Thus $n - e + f = 2$ holds.

If $e(G) \geq n$ and G is connected, G contains a cycle C. Choose any edge g on C. Let $G' = G \backslash g$. Then G' is connected so by the inductive hypothesis, for G', we have

$$n' - e' + f' = 2.$$

Here $n' = n$ and $e' = e - 1$. Also, deleting g unites two faces, so $f' = f - 1$. Thus,

$$n - e + f = 2.$$

□

Theorem 6.11. *If G is a planar graph with at least three vertices, then $e(G) \leq 3|G| - 6$. If G is also triangle-free, then $e(G) \leq 2|G| - 4$.*

Proof. It suffices to consider connected graphs; otherwise we could add edges to connect the graph. Also, we will assume that there are no degree 1 vertices. Indeed, the unique connected graph on 3 vertices which has a degree 1 vertex satisfies the theorem, and each time we delete a degree 1 vertex we only decrease $3|G| - 6 - e(G)$ (and the graph remains planar and connected).

Hence, we may assume that every vertex has degree at least 2. Now every face boundary contains at least three edges. Let $l(f_i)$ be the length of the ith face. Then $2e = \sum_i l(f_i) \geq 3f$. Hence $f \leq \frac{2}{3}e$. Substituting this into Euler's formula, we get

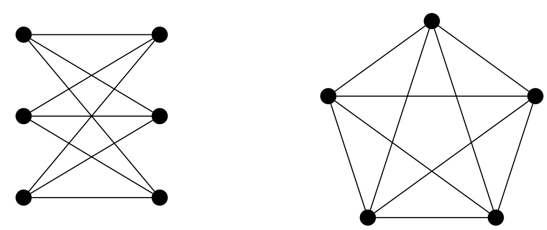
$$n - e + \frac{2}{3}e \geq 2,$$

and therefore $e \leq 3n - 6$.

When G is triangle-free, has at least 3 vertices and is connected, the faces have length at least 4. In this case $2e = \sum_i l(f_i) \geq 4f$, and we obtain $e \leq 2n - 4$ using Euler's formula. □

Corollary 6.12. *If G is a planar bipartite graph with $n \geq 3$ vertices then G has at most $2n - 4$ edges.*

Corollary 6.13. *K_5 and $K_{3,3}$ are not planar.*



$K_{3,3}$ K_5

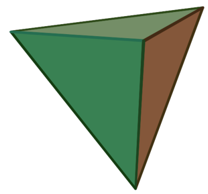
Proof. K_5 is a non-planar graph since $e = 10 > 9 = 3n - 6$. $K_{3,3}$ is a non-planar graph since $e = 9 > 8 = 2n - 4$. □

Remark 6.14 (Maximal planar graphs / triangulations)**.** The proof of Theorem 6.11 shows that having $3n - 6$ edges in a simple n-vertex planar graph requires $2e = 3f$, meaning that every face is a triangle. If G has some face that is not a triangle, then we can add an edge between non-adjacent vertices on the boundary of this face to obtain a larger plane graph. Hence the simple plane graphs with $3n - 6$ edges, the triangulations, and the *maximal* plane graphs are all the same family.

6.1 Platonic Solids

Definition 6.15. A *polytope* is a solid in 3 dimensions with flat faces, straight edges and sharp corners. Faces of a polytope are joined at the edges. A polytope is *convex* if the line segment connecting any two points of the polytope lies inside the polytope.

Example 6.16. The tetrahedron:

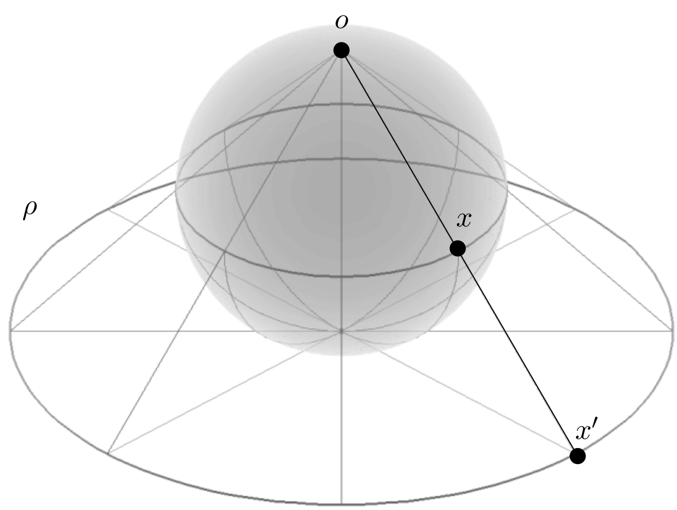


Definition 6.17. A *regular* or *Platonic* solid is a convex polytope which satisfies the following:

1. all of its faces are congruent regular polygons,
2. all vertices have the same number of faces adjacent to them.

We will now characterise all Platonic solids. The first step is to convert a convex polytope into a planar graph. To do this, we place the considered polytope inside a sphere. Then we project the polytope onto the sphere (imagine that the edges of the polytope are made from wire and we place a tiny lamp in the center). This yields a graph drawn on the sphere without edge crossings.

Now let us show that planar graphs are exactly graphs that can be drawn on the sphere. This becomes quite obvious if we use the *stereographic projection*. We place the sphere in the 3-dimensional space in such a way that it touches the considered plane ρ. Let o denote the point of the sphere lying farthest from ρ, the 'north pole'.



Then the stereographic projection maps each point $x \neq o$ of the sphere to a point x', where x' is the intersection of the line ox with the plane ρ. (For the point o, the projection is undefined.) This defines a bijection between the plane and the sphere without the point o. Given a drawing of a graph G on the sphere without edge crossings, where the point o lies on no arc of the drawing (which we may assume by a suitable choice of o), the stereographic projection yields a planar drawing of G. Conversely, from a planar drawing we get a drawing on the sphere by the inverse projection.

Corollary 6.18. *If K is a convex polytope with v vertices, e edges and f faces then $v - e + f = 2$.*

Suppose K is a Platonic solid. All its faces are congruent; assume that they have n vertices (and, thus, n edges). Let us assume moreover that each vertex is adjacent to m faces (and, thus, it has m

edges adjacent to it). Since each edge is adjacent to exactly two faces,

$$2e = nf. \tag{2}$$

Moreover, each edge is adjacent to two vertices, and one vertex belongs to m edges, thus

$$mv = 2e. \tag{3}$$

Expressing v and f in terms of e, and substituting to Euler's formula, we obtain that $\frac{2e}{m} - e + \frac{2e}{n} = 2$. Rearranging, we arrive at

$$\frac{1}{m} + \frac{1}{n} = \frac{1}{2} + \frac{1}{e}.$$

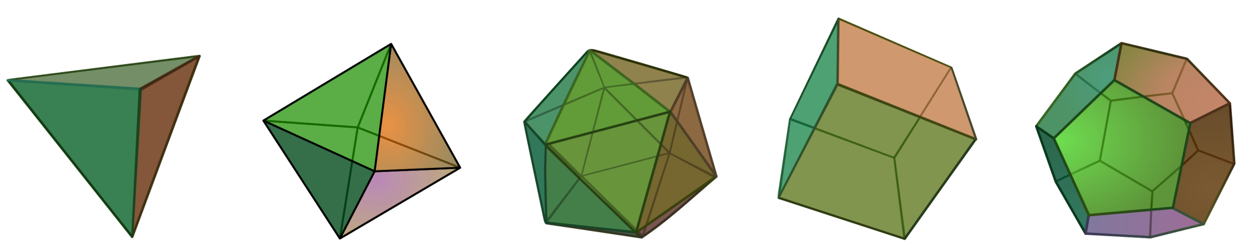
Note that since K is a 3-dimensional polytope, each of its faces is a polygon and thus has at least 3 vertices; that is, $n \geq 3$. Moreover, at each vertex, there are at least three faces meeting; $m \geq 3$. On the other hand, since $e \geq 1$, we must have

$$\frac{1}{m} + \frac{1}{n} > \frac{1}{2}. \tag{4}$$

These conditions do not leave too much leeway; there are only five possible (n, m) pairs for which the above inequality holds. These are $(3, 3), (3, 4), (3, 5), (4, 3), (5, 3)$.

A Platonic solid corresponds to each of these pairs. We list them below.

- **Tetrahedron**. Here $n = 3$ and $m = 3$. Thus, (4) yields that $e = 6$. By (3), $v = 4$, and by (2), $f = 4$. There are 4 vertices and 4 faces of the tetrahedron; the faces are regular triangles, and the vertices are adjacent to 3 edges.
- **Octahedron.** Here $n = 3$ and $m = 4$. Thus, (4) yields that $e = 12$. By (3), $v = 6$, and by (2), $f = 8$. There are 8 vertices and 8 faces of the octahedron; the faces are regular triangles, and the vertices are adjacent to 4 edges.
- **Icosahedron.** Here $n = 3$ and $m = 5$. Thus, (4) yields that $e = 30$. By (3), $v = 12$, and by (2), $f = 20$. There are 12 vertices and 20 faces of the icosahedron; the faces are regular triangles, and the vertices are adjacent to 5 edges.
- **Cube.** Here $n = 4$ and $m = 3$. Thus, (4) yields that $e = 12$. By (3), $v = 8$, and by (2), $f = 6$. There are 8 vertices and 6 faces of the tetrahedron; the faces are squares, and the vertices are adjacent to 3 edges.
- **Dodecahedron.** Here $n = 5$ and $m = 3$. Thus, (4) yields that $e = 30$. By (3), $v = 20$, and by (2), $f = 12$. There are 20 vertices and 12 faces of the tetrahedron; the faces are regular pentagons, and the vertices are adjacent to 3 edges.

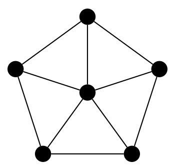


7 Graph colouring

7.1 Vertex colouring

Definition 7.1. A *k-colouring* of G is a labeling $f : V(G) \to \{1, \ldots, k\}$. It is a *proper k-colouring* if $(x, y) \in E(G)$ implies $f(x) \neq f(y)$. A graph G is *k-colourable* if it has a proper k-colouring. The *chromatic number* $\chi(G)$ is the minimum k such that G is k-colourable. If $\chi(G) = k$, then G is *k-chromatic.*

Example 7.2.

- $\chi(K_n) = n$
- $G =$  : $\chi(G) = 4$
- The chromatic number of an odd cycle is 3

Remark 7.3. The vertices having a given colour in a proper colouring must form an independent set, so $\chi(G)$ is the minimum number of independent sets needed to cover $V(G)$. Hence G is k-colourable if and only if G is k-partite. Multiple edges do not affect chromatic number. Although we define k-colouring using numbers from $\{1, \ldots, k\}$ as labels, the numerical values are usually unimportant, and we may use any set of size k as labels.

7.2 Some motivation

Example 7.4 (examination scheduling)**.** The students at a certain university have annual examinations in all the courses they take. Naturally, examinations in different courses cannot be held concurrently if the courses have students in common. How can all the examinations be organized in as few parallel sessions as possible? To find a schedule, consider the graph G whose vertex set is the set of all courses, two courses being joined by an edge if they give rise to a conflict. Clearly, independent sets of G correspond to conflict-free groups of courses. Thus, the required minimum number of parallel sessions is the chromatic number of G.

Example 7.5 (chemical storage)**.** A company manufactures n chemicals $C_1, C_2, \ldots, C_n$. Certain pairs of these chemicals are incompatible and would cause explosions if brought into contact with each other. As a precautionary measure, the company wishes to divide its warehouse into compartments, and store incompatible chemicals in different compartments. What is the least number of compartments into which the warehouse should be partitioned? We obtain a graph G on the vertex set $\{v_1, v_2, \ldots, v_n\}$ by joining two vertices v_i and v_j if and only if the chemicals C_i and C_j are incompatible. It is easy to see that the least number of compartments into which the warehouse should be partitioned is equal to the chromatic number of G.

7.3 Simple bounds on the chromatic number

Claim 7.6. If H is a subgraph of G then $\chi(H) \leq \chi(G)$.

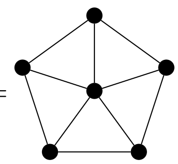
Proof. Note that a proper colouring of G is also a proper colouring of H. $\square$

Recall that $\omega(G)$ is the clique number of G, i.e., the order of the largest complete subgraph in G.

Corollary 7.7. $\chi(G) \geq \omega(G)$

Proof. Let $\omega(G) = t$. Then G contains a subgraph H which is isomorphic to K_t. Thus, by the claim above it follows that $\chi(G) \geq \chi(H) = t$. □

Example 7.8. Consider the following graph.

$G =$ 

In this case we have $\chi(G) = 4$ and $\omega(G) = 3$. Thus, the chromatic number can be bigger than the clique number.

Recall that $\alpha(G)$ is the independence number of G, i.e., the size of the largest independent set in G.

Proposition 7.9. $\chi(G) \geq \frac{|V(G)|}{\alpha(G)}$

Proof. Let $\chi(G) = k$. A proper k-colouring of $V(G)$ gives a partition $V(G) = V_1 \cup \ldots \cup V_k$ such that every V_i is an independent set. Hence, $|V_i| \leq \alpha(G)$. Therefore, $|V(G)| = \sum_{i=1}^{k} |V_i| \leq k\alpha(G)$. Thus, $k = \chi(G) \geq \frac{|V(G)|}{\alpha(G)}$ as claimed. □

Lemma 7.10. *For any graph $G = (V, E)$ and any $U \subseteq V$ we have $\chi(G) \leq \chi(G[U]) + \chi(G[V \setminus U])$.*

Proof. Properly colour U using $\chi(G[U])$ colours and properly colour $V \setminus U$ using $\chi(G[V \setminus U])$ other colours. This gives a proper colouring of G with $\chi(G[U]) + \chi(G[V \setminus U])$ colours. □

Lemma 7.11. *For any graphs G_1 and G_2 on the same vertex set, $\chi(G_1 \cup G_2) \leq \chi(G_1)\chi(G_2)$.*

Proof. Let c_1 and c_2 be proper colourings of G_1 and G_2 with the integers in $[\chi(G_1)]$ and $[\chi(G_2)]$ respectively. We colour the vertices of $G_1 \cup G_2$ with elements of the set $[\chi(G_1)] \times [\chi(G_2)]$, with the colouring c defined by $c(v) = (c_1(v), c_2(v))$. If v is adjacent to w in $G_1 \cup G_2$ then (v, w) is an edge in one of G_1 or G_2, so $c_1(v) \neq c_1(w)$ or $c_2(v) \neq c_2(w)$. This proves that $c(v) \neq c(w)$, so c is proper. □

Proposition 7.12.

(i) $\chi(G)\chi(\overline{G}) \geq |G|$

(ii) $\chi(G) + \chi(\overline{G}) \leq |G| + 1$

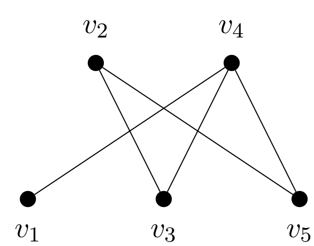
Proof. (i) follows from Lemma 7.11: we have $\chi(G)\chi(\overline{G}) \geq \chi(G \cup \overline{G}) = \chi(K_{|G|}) = |G|$.

(ii) can be proved by induction on $|G|$ (the case $|G| = 1$ is obvious). So, let $|G| = n+1$. Let $G_0 = G \backslash v$ for some vertex v. By induction we have $\chi(G_0) + \chi(\overline{G_0}) \leq n+1$. Let $c : V \to [k]$ be a proper colouring of G_0 and $f : V \to [\ell]$ be a proper colouring of $\overline{G_0}$, with $k + \ell = n+1$ (we might be using more colours than are necessary). If $d_G(v) < k$ then there is a colour $c_v \in [k]$ such that v has no neighbours coloured c_v. We can then colour v with c_v to extend c to a proper colouring of G with k colours. This would prove $\chi(G) \leq k$, and since $\chi(\overline{G}) = \chi(\overline{G_0} \cup \{v\}) \leq \ell + 1$ we have $\chi(G) + \chi(\overline{G}) \leq k + \ell + 1 \leq n + 2$. Otherwise $d_G(v) \geq k$ so $d_{\overline{G}}(v) \leq n - k = \ell - 1$. We can then use exactly the same reasoning as before to extend f to a proper colouring of $\overline{G}$ with ℓ colours, and since $\chi(G) \leq k+1$ we are done again. $\square$

7.4 Greedy colouring

Definition 7.13. The *greedy colouring* with respect to a vertex ordering $v_1, \ldots, v_n$ of $V(G)$ is obtained by colouring vertices in the order $v_1, \ldots, v_n$, assigning to v_i the smallest colour not already used on its lower-indexed neighbours.

Example 7.14.



This graph has chromatic number 2 but the greedy colouring needs 3 colours.

Definition 7.15. Let $G = (V, E)$ be a graph. We say that G is k-degenerate if every subgraph of G has a vertex of degree less than or equal to k.

Proposition 7.16. *G is k-degenerate if and only if there is an ordering $v_1, \ldots, v_n$ of the vertices of G such that each v_i has at most k neighbours among the vertices $v_1, \ldots, v_{i-1}$.*

Proof. If there is such an ordering, then for any subgraph H, consider the maximum vertex of H with respect to the ordering. This vertex has at most k neighbours in H, thus proving that G is k-degenerate.

Conversely, suppose G is k-degenerate. We prove the existence of a suitable ordering by induction on the number of vertices. Since G is k-degenerate, it has a vertex of degree at most k. Call this vertex v_n. Let $G' = G \backslash v_n$ and note that G' is still k-degenerate. Thus, there exists an ordering $v_1, \ldots, v_{n-1}$ of the vertices of G' satisfying the assertion of the proposition for G'. Then the ordering $v_1, \ldots, v_n$ satisfies the required conditions for G. $\square$

Definition 7.17. Define $\mathrm{dg}(G)$ to be the minimum k such that G is k-degenerate.

Remark 7.18. $\delta(G) \leq \mathrm{dg}(G) \leq \Delta(G)$.

Theorem 7.19. $\chi(G) \leq 1 + \mathrm{dg}(G)$.

Proof. Let $k = \mathrm{dg}(G)$. Fix an ordering $v_1, \ldots, v_n$ of $V(G)$ such that each v_i has at most k neighbours among $v_1, \ldots, v_{i-1}$. Use the greedy colouring on G with respect to this vertex ordering. This colouring uses at most $k+1$ colours, because when one colours v_i there are at most k colours which cannot be used. □

Corollary 7.20. $\chi(G) \leq \Delta(G) + 1$.

Note that $\mathrm{dg}(G)$ can be much smaller than $\Delta(G)$, for example if $G = K_{3,n-3}$ we have $\mathrm{dg}(G) = 3$ but $\Delta(G) = n - 3$.

Remark 7.21. This bound is tight if $G = K_n$ or if G is an odd cycle.

Theorem 7.22 (Brooks 1941)**.** *If G is a connected graph other than a clique or an odd cycle, then $\chi(G) \leq \Delta(G)$.*

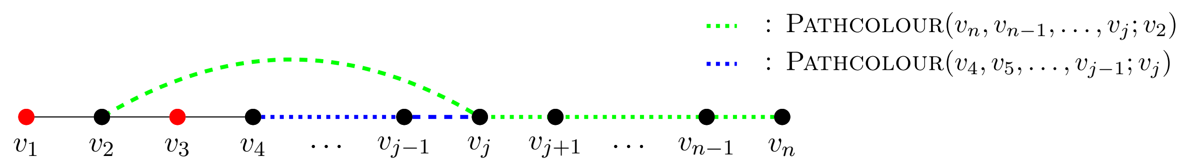
We will present a recent new proof due to Mariusz Zając which apart from being self contained and simpler than previous proofs has the advantage of being easily converted into an algorithm. The idea of the proof is to use induction on the number of vertices, however, the property of being connected is not very amenable to such arguments. The key idea of this proof is to show a slightly stronger result, which replaces the connectedness condition with that of not having a clique of certain size as a subgraph.

Theorem 7.23. *Let $k \geq 3$ be a natural number. Let G be a graph with $\Delta(G) \leq k$. If G does not contain a clique on $k+1$ vertices, then G is k-colorable.*

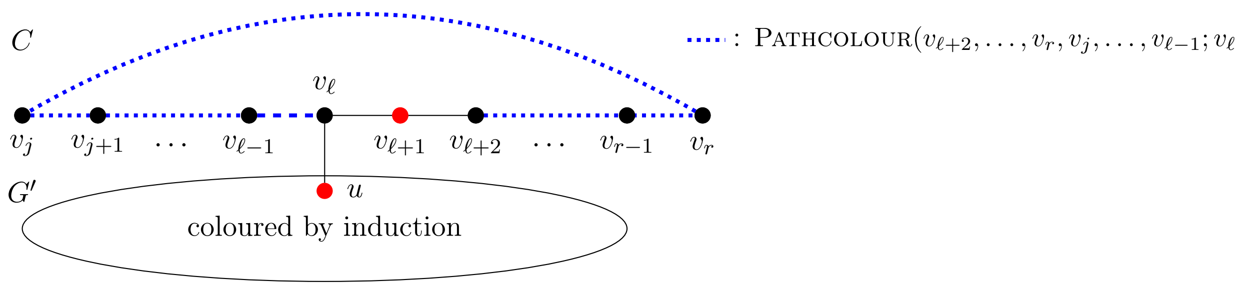
Proof. We will make use of the following easy observation. Suppose that G is partially coloured using at most k colours. Let $P = v_1 v_2 \ldots v_j$ be a path in G, and assume that the vertices of P are uncoloured. Then we may colour all vertices from v_1 up to v_{j-1} consecutively along P, since at the moment of colouring the vertex v_i its neighbour v_{i+1} is yet uncoloured, so v_i has at most $k-1$ coloured neighbours. We denote this sequential colouring procedure by $\textsc{PathColour}(v_1, v_2, \ldots, v_{j-1}; v_j)$. Note that after its execution the last vertex v_j of the path P remains uncoloured, in particular $\textsc{PathColour}$ does nothing if $j = 1$.

The proof now proceeds by induction on the number n of vertices of G. For $n \leq k$ the assertion holds trivially. We may further assume that G is k-regular, since otherwise we would delete a vertex of degree strictly less than k and apply induction. Let v be any vertex of G. Since G does not contain a clique on $k+1$ vertices, there exist two neighbours x, y of v that are not adjacent in G. Denote $v_1 = x$, $v_2 = v$, and $v_3 = y$. Let $P = v_1 v_2 v_3 \ldots v_r$ be a path starting with these three vertices and extending itself maximally, i.e. until some vertex v_r whose all neighbours are already on P.

Case 1. Suppose first that $r = n$, which means that P contains all vertices of G, and let v_j be any neighbor of v_2 other than v_1 and v_3 (it exists since $k \geq 3$). We start by giving the vertices v_1 and v_3 the same colour. Then apply procedures $\textsc{Pathcolour}(v_4, v_5, \ldots, v_{j-1}; v_j)$ and $\textsc{Pathcolour}(v_n, v_{n-1}, \ldots, v_j; v_2)$. Finally, colour the vertex v_2, which is possible because it has two neighbours in the same colour. The entire graph G is now k-coloured.



Case 2. Assume now that $r < n$. Recall that all neighbours of v_r are on the path P. Let v_j be the neighbor of v_r with the smallest index. So, $C = v_j v_{j+1} \dots v_r$ is a cycle in G. Consider the subgraph $G' = G - C$ obtained by deleting all vertices of C. We first colour G' using k colours by the induction hypothesis. If there is no edge between G' and C, then we are done by applying induction also to the subgraph induced by C. If, on the contrary, there is a vertex on C with a neighbor in G', then let v_ℓ be such vertex with the largest index, and let u be any of its neighbours in G'. Notice that $\ell < r$ because v_r has all of its neighbours on C. Since the vertex $v_{\ell+1}$ does not have neighbours in G', we may assign it the same colour as u. Now apply procedure PATHCOLOUR($v_{\ell+2}, \dots, v_r, v_j, \dots, v_{\ell-1}; v_\ell$) and finally colour v_ℓ, which is possible as it has two neighbours in the same colour. As previously, the entire graph G is coloured and the proof is complete. □



Proof of Theorem 7.22. If $\Delta(G) = 1$ the graph can only be a single edge, which is a clique. If $\Delta(G) = 2$, the graph is a union of disjoint paths and cycles and since it is connected it is a path or a cycle. For paths and even cycles, by colouring vertices alternately it is easy to see that they are 2-colourable. Let now $k = \Delta(G) \geq 3$. Theorem 7.23 implies that G can be k-coloured unless it contains K_{k+1} as a subgraph. If G contains K_{k+1}, then since $\Delta(G) = k$, this clique can send no edges to the rest of the graph, so since G is connected, we have $G = K_{k+1}$. □

7.5 Colouring planar graphs

Lemma 7.24. *A (simple) planar graph G contains a vertex of degree at most* 5.

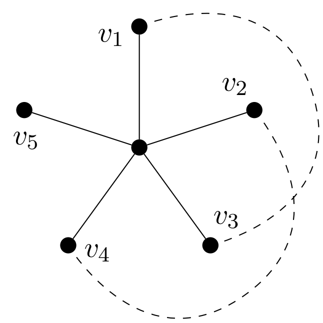
Proof. Recall that in a planar graph $|E(G)| \leq 3|V| - 6$. Thus, we have that $\sum_{v \in V(G)} d(v) = 2|E(G)| \leq 6|V| - 12 < 6|V|$ and so the lemma follows. □

Corollary 7.25. *A planar graph is* 5*-degenerate and thus* 6*-colourable.*

Proof. Every subgraph of a planar graph is planar, so by the above lemma, every subgraph has a vertex of degree at most 5, which means that the graph is 5-degenerate. It then follows from Theorem 7.19 that the graph is 6-colourable. □

Theorem 7.26 (5-colour theorem; Heawood 1890)**.** *Every planar graph G is* 5*-colourable.*

Proof. By induction of $|V(G)|$. For $|V(G)| \leq 5$ the statement is obvious. Assume $|V(G)| > 5$. Let v be a vertex of degree at most 5 in G. By induction, $G\backslash v$ is 5-colourable. Take a proper colouring $f : V(G) \setminus \{v\} \to \{1, \dots, 5\}$ of $G\backslash v$. If $d(v) < 5$, then f can be extended to a proper colouring of $V(G)$ by assigning $f(v) \in \{1, \dots, 5\} \setminus \{f(u) : uv \in E(G)\}$. Hence, we may assume that $d(v) = 5$. Fix a planar embedding of G in which the neighbours of v are coloured by f with the colours $1, \dots, 5$ in clockwise order (if f uses less than 5 colours on $N(v)$ then it can be extended to $V(G)$ as before). Let the corresponding vertices be $v_1, \dots, v_5$, i.e., $f(v_i) = i$ for $i = 1, \dots, 5$.



For $1 \leq i \neq j \leq 5$ let G_{ij} be the subgraph of $G\backslash v$ induced by the vertices of colours i and j. Switching the two colours in any connected component of G_{ij} again gives a proper 5-colouring of $G\backslash v$. If the component of G_{ij} containing v_i does not contain v_j, then we switch the colours in the component of G_{ij} which contains v_i, in order to remove the colour i from $N(v)$; we can then colour v with the colour i. We can therefore assume that for every pair $1 \leq i \neq j \leq 5$ the component of G_{ij} containing v_i also contains v_j. Let P_{ij} be a path in G_{ij} from v_i to v_j. Obviously, the vertices of P_{ij} are coloured alternatively by colours i and j. Note that the drawings of the paths P_{13} and P_{24} must intersect. Since the drawing is planar, they intersect in a vertex. But all the vertices of P_{13} are coloured 1 and 3 and all the vertices of P_{24} are coloured 2 and 4, which is a contradiction. $\square$

Theorem 7.27 (4-colour theorem, Appel-Haken 1977)**.** *Every planar graph is 4-colourable. (The countries of every plane map can be 4-coloured so that neighbouring countries get distinct colours).*

Remark 7.28. The only known proofs heavily use computers.

7.6 Large girth and large chromatic number

The bound $\chi(G) \geq \omega(G)$ can be tight, but (surprisingly) it can also be arbitrarily bad. There are graphs having arbitrarily large chromatic number, even though they do not contain K_3. Many constructions of such graphs are known. We give here one due to Alon, Krivelevich and Sudakov.

Example 7.29. Let G_k be a triangle-free graph with chromatic number k. Let us construct a new graph G_{k+1} as follows. We take initially G_{k+1} to consist of k vertex-disjoint copies of G_k denoted $H_1, H_2, \dots, H_k$. For each k-tuple of vertices $(v_1, v_2, \dots, v_k)$ with $v_i \in V(H_i)$, we add a new vertex u to G_{k+1} and add an edge between v_i and u for all $i \in [k]$.

Proposition 7.30. G_{k+1} *is triangle-free and has chromatic number* $k+1$.

Proof. Each H_i is triangle free so $H_1 \cup \cdots \cup H_k$ induces a triangle-free graph. In addition, since there are no edges between distinct H_i's, adding any vertex u joined to at most one vertex of each H_i will not create a triangle, so G_{k+1} is triangle-free.

Each H_i can be coloured using k colours (say $[k]$ is the set of colours). Then we can use colour $k+1$ to colour the vertices u added in the second step to get $\chi(G_{k+1}) \leq k+1$. On the other hand, assume there is a proper colouring of G_{k+1} using at most k colours. Since $\chi(H_i) = k$, each H_i must use all k different colours. In particular, for each $i \in [k]$, there is a vertex $v_i \in V(H_i)$ which got colour i. Then the vertex u adjacent to the k-tuple $(v_1, v_2, \ldots, v_k)$ must use a new colour, a contradiction showing that $\chi(G_{k+1}) \geq k+1$. $\square$

Definition 7.31. The *girth* of a graph is the length of its shortest cycle.

Theorem 7.32 (Erdős, 1959)**.** *Given $k \geq 3$ and $g \geq 3$, there exists a graph with girth at least g and chromatic number at least k.*

This result is especially surprising because if a graph has no short cycles then it "locally" looks like a tree, and all trees have chromatic number at most 2. This shows that the chromatic number really depends on the global structure of a graph, and cannot be estimated from local considerations.

The proof of Theorem 7.32 uses the so-called *probabilistic method*.

7.6.1 Digression: the probabilistic method

The general idea of the probabilistic method is to prove the existence of certain structures by showing that they exist with positive probability in some probability space. Let us introduce some probabilistic notation and results.

Definition 7.33. The *expectation* $\mathbb{E}[X]$ of a random variable X is the "average" value it takes: if X takes only countably many possible values, then

$$\mathbb{E}[X] = \sum_x x \Pr[X = x].$$

Remark 7.34. Expectation has the following properties:

- Expectation is linear: $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$, regardless of whether X and Y are independent.
- If X is a nonnegative integer valued random variable then

  $$\mathbb{E}[X] = \sum_{x=0}^{\infty} x \Pr[X = x] \geq \sum_{x=1}^{\infty} \Pr[X = x] = \Pr[X \geq 1].$$

  In particular, if $\mathbb{E}[X] < 1$ then $\Pr[X = 0] = 1 - \Pr[X \geq 1] \geq 1 - \mathbb{E}[X] > 0$.
- More generally, for any positive a,

  $$\mathbb{E}[X] = \sum_{x=0}^{\infty} x \Pr[X = x] \geq \sum_{x \geq a} a \Pr[X = x] = a \Pr[X \geq a].$$

  Dividing by a, we get *Markov's inequality*: $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

Before we proceed to the proof of Theorem 7.32, let us illustrate the method on a slightly easier problem.

A *tournament* is an orientation of the complete graph K_n. The name comes from observing that such a digraph corresponds to an n-player round-robin tournament, where every player plays against everyone else exactly once, and there are no draws. An edge from a to b means that player a beats player b.

Theorem 7.35. *For all sufficiently large n, there is a tournament on n vertices where any $\lfloor \frac{\log_2 n}{2} \rfloor$ vertices are beaten by some other vertex.*

Proof. Set $k = \lfloor \frac{\log_2 n}{2} \rfloor$ and consider the random tournament where each edge of K_n is given a random orientation (each direction with probability $1/2$) independently from the others. For a set K of k vertices, let X_K be the indicator random variable of the event that K is not beaten by any other vertex. Then $\Pr[X_K = 1] = (1 - \frac{1}{2^k})^{n-k}$. The variable $X = \sum_{|K|=k} X_K$ counts the number of such "bad" k-sets. Next we bound the expectation of X using the easy estimates $\binom{n}{k} \leq n^k$ and $1 - x \leq e^{-x}$:

$$\mathbb{E}[X] = \sum_{|K|=k} \mathbb{E}[X_K] = \binom{n}{k}(1 - \frac{1}{2^k})^{n-k} \leq n^k \cdot e^{-\frac{n-k}{2^k}} \leq \exp\Big(k \log n - \frac{n-k}{\sqrt{n}}\Big).$$

Here the exponent tends to $-\infty$ and hence $\mathbb{E}[X] \to 0$ as $n \to \infty$. So for n large enough, $\mathbb{E}[X] < 1$. Applying the observation above, we get $\Pr[X = 0] > 0$, i.e., the random tournament satisfies the desired property with positive probability. This proves the existence of such a tournament. □

7.6.2 Proof of Theorem 7.32

Proof. Given a large value of n (large enough to satisfy inequalities we will use later), we randomly generate a graph with vertex set $[n]$. Simply let each pair $\{x, y\}$ with $1 \leq x < y \leq n$ be an edge with probability p, independently. A graph with no large independent set has large chromatic number since $\chi(G) \geq n(G)/\alpha(G)$. We therefore choose p large enough to make the existence of large independent sets unlikely. We also want to choose p small enough that the expected number of short cycles (length less than g) is small. When we have such a graph satisfying both conditions, we can delete a vertex from each short cycle to obtain the desired graph.

To make it unlikely that the graph generated will have more than $n/2$ short cycles, we let $p = n^{\varepsilon-1}$, where $\varepsilon = 1/(2g)$. Let X denote the number of cycles of length less than g in the resulting random graph. For each possible cycle C, define a random variable X_C that takes the value 1 if C is present in our random graph, and 0 otherwise. We then have $X = \sum_C X_C$, where the sum is over all possible cycles C of length less than g. For any possible cycle C of length i, we have $\mathbb{E}[X_C] = p^i$, since we require i edges to be present, each of which is present, independently, with probability p.

Since there are at most n^i potential cycles of length i, linearity of expectation yields $\mathbb{E}[X] \leq \sum_{i=3}^{g-1} n^i p^i = \sum_{i=3}^{g-1} n^{\varepsilon i} < 2n^{\varepsilon g} = 2\sqrt{n}$. This implies that $\mathbb{E}[X]/n \to 0$ as $n \to \infty$. Hence, we have $\Pr[X \geq n/2] \leq \frac{\mathbb{E}[X]}{n/2} \to 0$ as $n \to \infty$ by Markov's inequality. In particular, we can take n large enough so that $\Pr[X \geq n/2] < 1/2$.

Since we will retain at least $n/2$ vertices, it suffices to show that there will be such graphs with $\alpha(G) \leq n/(2k)$; $\alpha(G)$ cannot grow when we delete vertices, and hence at least k independent sets will be needed to cover the remaining vertices. With $r = \lfloor n/(2k) \rfloor$, we have

$$\Pr[\alpha(G) > r] \leq \binom{n}{r+1}(1-p)^{\binom{r+1}{2}} < \left(ne^{-pr/2}\right)^{r+1}.$$

As $n \to \infty$ we have $ne^{-pr/2} \leq ne^{-n^{\varepsilon}/(6k)} \to 0$ (and $r \to \infty$), so $\Pr[\alpha(G) > r] \to 0$. If we also make n large enough that $\Pr[X \geq n/2] < 1/2$ and $\Pr[\alpha(G) > r] < 1/2$, there will exist an n-vertex graph G such that $\alpha(G) \leq n/(2k)$ and G has fewer than $n/2$ cycles of length less than g. We delete a vertex from each short cycle and retain a graph with girth at least g and chromatic number at least k. $\square$

8 Ramsey Theory

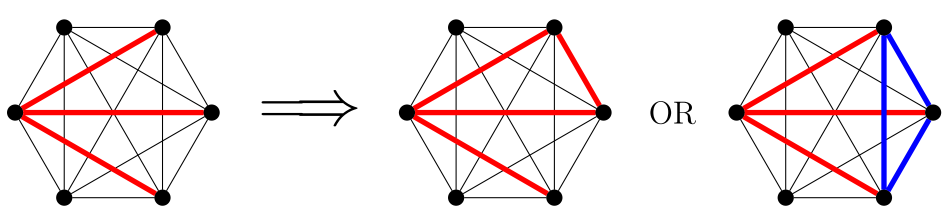
Ramsey theory refers to a large body of deep results in mathematics concerning partitions of large collections. An informal way of describing this is Motzkin's statement that "Complete disorder is impossible".

Proposition 8.1 (informal)**.** *Among six people it is possible to find three mutual acquaintances or three mutual non-acquaintances.*

To turn this into a mathematical statement, consider the complete graph whose vertices are the six people. Colour an edge between two people red if those people know each other, and blue otherwise. Proposition 8.1 then becomes the following statement.

Proposition 8.2. *In every red/blue edge-colouring of the complete graph on* 6 *vertices, there is a red triangle or a blue triangle (or both).*

Proof. Single out some vertex u. Out of the five edges incident to u, at least three are blue or at least three are red. Without loss of generality say there are three red edges (u, v_1), (u, v_2), (u, v_3). If there is a red edge (v_i, v_j), then u, v_i and v_j form a red triangle. Otherwise v_1, v_2 and v_3 form a blue triangle.



$\square$

As we shall see, given a natural number s, there is an integer R such that if $n \geq R$, then every colouring of the edges of K_n with red and blue contains either a red K_s or a blue K_s. Proposition 8.2 claims that for $s = 3$, $R = 6$ will do. More generally, we define the *Ramsey number* $R(s, t)$ as the smallest value of N for which every red-blue colouring of K_N yields a red K_s or a blue K_t. In particular, $R(s, t) = \infty$ if there is no such N such that in every red-blue colouring of K_N there is a red K_s or a blue K_t.

Example 8.3. We have $R(3,3) = 6$. Indeed, Proposition 8.2 shows that $R(3,3) \leq 6$. To see that $R(3,3) > 5$, note that if we colour a 5-cycle red and the remaining edges blue in a complete graph on 5 vertices, then both colours are triangle-free.

It is obvious that
$$R(s,t) = R(t,s)$$
for every $s, t \geq 2$ and $R(s,2) = R(2,s) = s$, since in a red-blue colouring of K_s either there is a blue edge or else every edge is red. The following result, due to Erdős and Szekeres, states that $R(s,t)$ is finite for every s and t, and at the same time it gives a bound on $R(s,t)$.

Theorem 8.4. *The function $R(s,t)$ is finite for all $s, t \geq 2$. Quantitatively, if $s > 2$ and $t > 2$ then*
$$R(s,t) \leq R(s-1,t) + R(s,t-1) \tag{5}$$
and
$$R(s,t) \leq \binom{s+t-2}{s-1}. \tag{6}$$

Proof. When proving (5) we may assume that $R(s-1,t)$ and $R(s,t-1)$ are finite. Let $N = R(s-1,t) + R(s,t-1)$ and consider a colouring of the edges of K_N with red and blue. We have to show that in this colouring there is either a red K_s or a blue K_t. To this end, let x be a vertex of K_N. Since $d(x) = N-1$, either there are at least $N_1 = R(s-1,t)$ red edges incident with x or there are at least $N_2 = R(s,t-1)$ blue edges incident with x. Without loss of generality assume that the first case holds. Consider a subgraph K_{N_1} of K_N spanned by N_1 vertices joined to x by red edges. If K_{N_1} has a blue K_t, we are done. Otherwise, by the definition of $R(s-1,t)$, the graph K_{N_1} contains a red K_{s-1} which forms a red K_s with x.

Inequality (6) holds if $s = 2$ or $t = 2$ (in fact, we have equality). Assume that $s > 2$, $t > 2$, and that (6) holds for every pair (s', t') with $2 \leq s' + t' < s + t$. Then by (5) we have
$$\begin{aligned} R(s,t) &\leq R(s-1,t) + R(s,t-1) \\ &\leq \binom{s+t-3}{s-2} + \binom{s+t-3}{s-1} = \binom{s+t-2}{s-1}. \end{aligned}$$
□

Remark. Note that $R(s,t) \leq N$ means that every graph on N vertices has either $\omega(G) \geq s$ or $\alpha(G) \geq t$.

It is customary to distinguish *diagonal Ramsey numbers* $R(s) = R(s,s)$ and *off-diagonal Ramsey numbers* $R(s,t)$ for $s \neq t$. It is not surprising that the diagonal Ramsey numbers are of greatest interest, and they are also the hardest to estimate.

We see from Theorem 8.4 that
$$R(s) \leq \binom{2s-2}{s-1} \leq \frac{2^{2s-2}}{\sqrt{s}}. \tag{7}$$

In an edge-colouring of a graph, we call a subgraph *monochromatic* if all of its edges have the same colour.

Theorem 8.5 (Erdős 1947)**.** *For s sufficiently large, we have $R(s, s) > 2^{s/2}$.*

Proof. Let $n = \lfloor 2^{s/2} \rfloor$. Take a random colouring of K_n in which each edge is red with probability 1/2 and blue with probability 1/2, independently over all edges. Let X be the random variable counting the number of monochromatic cliques of size s. For a set S of s vertices, let X_S be the indicator random variable for the event that the clique on vertex set S is monochromatic. Then $X = \sum_{|S|=s} X_S$. Note that $\mathbb{E}[X_S] = 2^{1-\binom{s}{2}}$ since this is the probability that all edges in S get the same colour. Since there are $\binom{n}{s}$ ways to choose S, by linearity of expectation, we have $\mathbb{E}[X] = \binom{n}{s}2^{1-\binom{s}{2}}$. Using that s is sufficiently large, we get $\binom{n}{s}2^{1-\binom{s}{2}} < (\frac{n}{2})^s 2^{-\binom{s}{2}} = (\frac{n}{2} \cdot 2^{-\frac{s-1}{2}})^s < 1$, so $\mathbb{P}(X = 0) \geq 1 - \mathbb{E}[X] > 0$. This means that there exists a red-blue colouring of the edges of K_n such that there is no monochromatic clique of size s. Thus, $R(s, s) > n$. □

Remark 8.6. We saw that $R(s, s) \leq 4^s$ (by (7)) and $R(s, s) \geq (\sqrt{2})^s$ (by Theorem 8.5). For the last 70 years, it was a major open problem to improve the base of the exponent (which is between $\sqrt{2}$ and 4) in either of these bounds. Very recently (in 2023), a breakthrough result of Campos–Griffiths–Morris–Sahasrabudhe improved the upper bound to $R(s, s) \leq (4 - c)^s$ for a small $c > 0$ (e.g., $c = 1/128$ works). Determining the correct base of the exponent, i.e. finding $\lim_{s\to\infty} R(s, s)^{1/s}$ (if it exists), remains one of the major open problems in Ramsey theory.

8.1 Multicolour Ramsey numbers

We now consider colourings with more than two colours.

Theorem 8.7. *Given $k \geq 3$ and $s_1, s_2, \ldots, s_k$, if N is sufficiently large, then every colouring of K_N with k colours is such that for some i, $1 \leq i \leq k$, there is a K_{s_i} coloured with the i-th colour. The minimal value of N for which this holds is usually denoted by $R_k(s_1, \ldots, s_k)$, and it satisfies*

$$R_k(s_1, \ldots, s_k) \leq R_{k-1}(R(s_1, s_2), s_3, \ldots, s_k).$$

Proof. In a k-colouring of K_N we replace the first two colours by a single new colour (say, purple). If $N = R_{k-1}(R(s_1, s_2), s_3, \ldots, s_k)$, then either there is a K_{s_i} coloured with the i-th colour for some i with $3 \leq i \leq k$, or else for $m = R(s_1, s_2)$ there is a purple K_m. In the first case, we are done. In the second case, in the original colouring this K_m is coloured with the first two original colours. Hence, as $m = R(s_1, s_2)$, for $i = 1$ or 2 we can find a K_{s_i} in K_m coloured with the i-th colour. This shows that $R_k(s_1, \ldots, s_k) \leq N = R_{k-1}(R(s_1, s_2), s_3, \ldots, s_k)$. □

Theorem 8.8. *We have*

$$R_k(3) \stackrel{def}{=} R_k(3, 3, \ldots, 3) \leq 3 \cdot k!.$$

Proof. We use induction on k. For $k = 2$, we have already shown $R_2(3, 3) \leq 6$ (here equality holds).

Suppose $k \geq 3$ and let x be any vertex of a k-coloured complete graph on $3 \cdot k!$ vertices. There are $3 \cdot k! - 1$ edges adjacent to this vertex, which are split into k colour classes. Note that $k \cdot (3(k-1)! - 1) < 3k! - 1$, so one of the k colours, say red, contains at least $3(k - 1)!$ edges adjacent to x. Let S be the set of those vertices joined to x by a red edge.

If S spans a red edge, this edge forms a red triangle together with x, and we are finished. If S spans no red edge, then it spans a complete graph with at least $3(k - 1)!$ vertices, whose edges are $(k - 1)$-coloured; thus by the induction hypothesis, one of these $k - 1$ colours contains a triangle and we are done again. □

Proposition 8.9. $R_k(3) > 2^k$.

Proof. We prove the stronger statement that the edges of K_{2^k} can be k-coloured such that each colour is a bipartite graph. To see this, split the vertices of K_{2^k} into two sets A, B of size 2^{k-1}, use colour k for all edges between A and B, and apply induction to colour the edges inside A and B with the colours $1, \ldots, k-1$ such that each colour is bipartite.

Since every bipartite graph is triangle-free, we get $R_k(3) > 2^k$. $\square$

Remark 8.10. Theorem 8.8 and Proposition 8.9 show that $R_k(3)$ grows at least exponentially in k and at most like $k! = \Theta(k)^k$. It is a major open problem to determine if $R_k(3) \leq C^k$ for a constant C independent of k.

8.2 Ramsey theory for integers

The following theorem, proved by Schur in 1916, became the starting point of an area that is still very active today.

Theorem 8.11. *For every $k \geq 1$ there is an integer m such that every k-colouring of $[m]$ contains integers x, y, z of the same colour such that $x + y = z$.*

Proof. We claim that we can choose $m = R_k(3) - 1$. Let $n = R_k(3) = m + 1$ and let $c : [m] \to [k]$ be a k-colouring. We define a k-edge-colouring c' of the complete graph with vertex set $[n]$ as follows. For $(i, j) \in E(K_n)$ set $c'(i, j) = c(|i - j|)$. By the definition of $n = R_k(3)$, there is a monochromatic triangle, say with vertex set $\{h, i, j\}$, so that $1 \leq h < i < j \leq n$ and $c'(h, i) = c'(i, j) = c'(j, h)$ ($= \ell$, say). But then $x = i - h, y = j - i$ and $z = j - h$ are such that $c(x) = c(y) = c(z) = \ell$ and $x + y = z$. $\square$

Remark 8.12. A similar result is known for many other linear equations, and (more generally) systems of linear equations. For example, the fact that $x_1, \ldots, x_r$ make an arithmetic progression can be expressed by the linear system $x_{i+2} - x_{i+1} = x_{i+1} - x_i$, $i = 1, \ldots, r - 2$. A famous theorem of Van der Waerden says that "arithmetic progressions are Ramsey", i.e., that for all r, k, there is an integer m such that every k-colouring of $[m]$ has a monochromatic r-term arithmetic progression. A more general theorem of Rado gives a characterization of all linear systems which are Ramsey.

8.3 Graph Ramsey numbers

Definition 8.13. Let G_1, G_2 be graphs. $R(G_1, G_2)$ is the minimal N such that any red/blue colouring of K_N contains either a red copy of G_1, or a blue copy of G_2.

Observe that $R(K_s, K_t) = R(s, t)$ by definition.

Remark 8.14. Note that $R(G_1, G_2) \leq R(|G_1|, |G_2|)$.

Theorem 8.15 (Chvátal 1977)**.** *If T is any m-vertex tree, then $R(T, K_n) = (m - 1)(n - 1) + 1$.*

Proof. For the lower bound, consider the colouring of $K_{(m-1)(n-1)}$ that has $n-1$ disjoint red cliques of order $m - 1$ and in which every other edge is blue. With red components of order $m - 1$, there is no red m-vertex tree. The blue edges form an $(n - 1)$-partite graph and hence cannot contain K_n.

For the upper bound, we proceed by induction on n (the case $n = 1$ is trivial). Given a two-colouring of $K_{(m-1)(n-1)+1}$, consider a vertex x. If x has more than $(m-1)(n-2)$ neighbours along blue edges, then among them is a red T or a blue K_{n-1}, which yields a blue K_n with x. Otherwise, every vertex has at most $(m-1)(n-2)$ incident blue edges and hence at least $m-1$ incident red edges. That is, the red subgraph R has minimum degree at least $m-1$. But this means we can embed a copy of T. Indeed, we claim that if a graph G has minimum degree at least $m-1$, then it contains every m-vertex tree T.

We prove this claim by induction on m. If $m = 1$ the claim is obvious, so suppose $m \geq 2$ and that the claim holds for all smaller m. Let ℓ be a leaf of T; by the induction hypothesis G contains $T \backslash \ell$ (which has $m-1$ vertices). Since every vertex in G has degree at least $m-1$, every vertex of the $T \backslash \ell$ we have found in G has a neighbour that is not in $T \backslash \ell$. Hence we can "reattach" ℓ to T using such a neighbour. $\square$

9 Extremal problems

Question 9.1. Let H be a fixed graph, and G be a graph on n vertices that contains no copy of H. How many edges can G have?

Definition 9.2. $\mathrm{ex}(n, H)$ is the maximal value of $e(G)$ among graphs G with n vertices containing no H as a subgraph.

Example 9.3. Consider the case where H is a triangle. Recall that bipartite graphs contain no triangles. So $K_{\frac{n}{2},\frac{n}{2}}$ gives a triangle-free graph with $\frac{n^2}{4}$ edges.

9.1 Turán's theorem

As a generalization of Example 9.3, notice that dense graphs not having K_{r+1} as a subgraph can be obtained by dividing the vertex set V into r pairwise disjoint subsets $V = V_1 \cup \ldots \cup V_r$ with $|V_i| = n_i$ and $n = n_1 + \ldots + n_r$, and joining two vertices if and only if they lie in distinct sets V_i, V_j. We denote the resulting graph by $K_{n_1,\ldots,n_r}$. It has $\sum_{i<j} n_i n_j$ edges. Assuming n is fixed, we get the maximal number of edges among such graphs when we divide the numbers n_i as evenly as possible, that is $|n_i - n_j| \leq 1$ for all i, j. Indeed, suppose $n_1 \geq n_2 + 2$. By shifting one vertex from V_1 to V_2, we obtain $K_{n_1-1,n_2+1,\ldots,n_r}$, and that contains $(n_1-1)(n_2+1) - n_1 n_2 = n_1 - n_2 - 1 \geq 1$ more edges than $K_{n_1,\ldots,n_r}$. In particular, if r divides n, then we may choose $n_i = \frac{n}{r}$ for all i, obtaining

$$\binom{r}{2}\left(\frac{n}{r}\right)^2 = \left(1 - \frac{1}{r}\right)\frac{n^2}{2}$$

edges. Turán's theorem states that this number is an upper bound for the edge-number of *any* graph on n vertices without an $(r+1)$-clique.

Definition 9.4. Let $n_1, \ldots, n_r$ be nonnegative integers such that $|n_i - n_j| \leq 1$ for all i, j and $n = n_1 + \cdots + n_r$. We call the graph $K_{n_1,\ldots,n_r}$ the n-vertex, r-partite *Turán graph*, and we denote it by $T_{n,r}$.

Theorem 9.5 (Turán 1941)**.** *Among all the n-vertex graphs with no $(r+1)$-clique, $T_{n,r}$ is the unique graph having the maximum number of edges.*

Proof. We use induction on r. The base case $r = 1$ is trivial. Let $r > 1$ and let $G = (V, E)$ have the maximum number of edges among n-vertex graphs containing no $(r + 1)$-cliques. Let $v_m \in V$ be a vertex of maximal degree $\Delta(G)$. Denote the set of neighbours of v_m by S, and let $T = V \setminus S$. As G contains no $(r + 1)$-clique and v_m is adjacent to all vertices of S, we note that S contains no r-clique.

We now construct the following graph H on V. H induces the same graph as G on S and contains all edges between S and T, but no edges within T. Note that T is an independent set in H, so any clique in H can contain at most one vertex from T. Hence, as $G[S]$ is K_r-free and $H[S] = G[S]$, it follows that H has no $(r + 1)$-cliques.

If $v \in S$, then we have $d_H(v) \geq d_G(v)$ by the construction of H, and for $v \in T$ we see $d_H(v) = |S| = \Delta(G) \geq d_G(v)$ by the choice of v_m. We infer $|E(H)| \geq |E(G)|$ and find that among all graphs with a maximal number of edges, there must be one of the form of H. By the induction hypothesis, the graph induced by S has at most as many edges as the $(r - 1)$-partite Turán graph $K_{n_1,\dots,n_{r-1}}$ on S. So $|E(G)| \leq |E(H)| \leq |E(K_{n_1,\dots,n_r})|$ with $n_r = |T|$. We have established that $|E(K_{n_1,\dots,n_r})|$ is maximized when $|n_i - n_j| \leq 1$ holds for all i, j, which implies that $T_{n,r}$ has the maximum possible number of edges among n-vertex K_{r+1}-free graphs.

To prove uniqueness, note that if G has a maximum number of edges, then $|E(G)| = |E(H)|$ and $|E(H)| = |E(K_{n_1,\dots,n_r})|$. The latter equation implies that $H[S]$ has the same number of edges as the Turán graph $K_{n_1,\dots,n_{r-1}}$ on S, and hence, by the induction hypothesis, $H[S] = G[S]$ is equal to this $(r - 1)$-partite Turán graph. The former equation implies that T touches exactly $\Delta(G)|T|$ edges in G. But this can only happen if T is an independent set in G. Indeed, the sum of the degrees of the vertices in T counts each edge spanned by T twice, and each edge connecting T and S once. As $\Delta(G)$ is the maximum degree in G, the sum of degrees of the vertices in T is at most $\Delta(G)|T|$, so T can only touch this many edges if it spans none of them. But then G is r-partite and since it has the maximum number of edges, $G = T_{n,r}$. □

We state the following theorem without proof.

Theorem 9.6 (Erdős-Stone). *Let H be a graph of chromatic number $\chi(H) = r + 1$. Then for every $\varepsilon > 0$ and large enough n,*

$$\left(1 - \frac{1}{r}\right)\binom{n}{2} \leq \mathrm{ex}(n, H) \leq \left(1 - \frac{1}{r}\right)\binom{n}{2} + \varepsilon n^2.$$

Note that if $\chi(H) = 2$ (that is, if H is bipartite) then the Erdős-Stone theorem simply says that $\mathrm{ex}(n, H) \leq \varepsilon n^2$ for any $\varepsilon > 0$ and large enough n, and does not give precise asymptotics. It is interesting to ask what more we can say in this case.

9.2 Bipartite Turán theorems

Theorem 9.7 (Kővári-Sós-Turán). *For any integers $r \leq s$, there is a constant c such that every $K_{r,s}$-free graph on n vertices contains at most $cn^{2-\frac{1}{r}}$ edges. In other words,*

$$\mathrm{ex}(n, K_{r,s}) \leq cn^{2-\frac{1}{r}}.$$

We will prove this theorem using (a special case of) Jensen's inequality, which says that if f is a convex function then for any $x_1, \dots, x_n$ we have

$$f(x_1) + \cdots + f(x_n) \geq nf\left(\frac{x_1 + \cdots + x_n}{n}\right).$$

Proof of Theorem 9.7. Consider a $K_{r,s}$-free graph $G = (V, E)$. Let $V = \{v_1, \dots, v_n\}$ and $d(v_i) = d_i$. Without loss of generality, we may assume that $d_i \geq r - 1$ for each i (else, we can add edges to the graph without creating a $K_{r,s}$). Let $\bar{d} = 2e(G)/n$ be the average degree of G. Since G contains no $K_{r,s}$, for any given r-tuple in V there are at most $s - 1$ vertices whose neighbourhood contains that r-tuple. The neighbourhood of v_i contains $\binom{d_i}{r}$ r-tuples, so

$$\sum_{i=1}^{n} \binom{d_i}{r} \leq (s-1)\binom{n}{r}.$$

We are going to estimate the left-hand side using Jensen's inequality. Note that the function $f : x \mapsto \binom{x}{r}$ is convex for $x \geq r - 1$ (one can verify this by computing the second derivative of f). Since $d_i \geq r - 1$ by assumption, we have, by Jensen's inequality,

$$\sum_{i=1}^{n} \binom{d_i}{r} = \sum_{i=1}^{n} f(d_i) \geq n \cdot f\left(\frac{d_1 + \ldots + d_n}{n}\right) = nf(\bar{d}).$$

Hence, we have

$$(s-1)\binom{n}{r} \geq \sum_{i=1}^{n} \binom{d_i}{r} \geq nf(\bar{d}) = n\binom{\bar{d}}{r}.$$

Note that

$$(s-1)\frac{n^r}{r!} > (s-1)\binom{n}{r} \qquad \text{and} \qquad \binom{\bar{d}}{r} > \frac{\left(\bar{d} - r + 1\right)^r}{r!},$$

so we obtain $(\bar{d} - r + 1)^r < (s-1)n^{r-1}$. But then

$$\bar{d} < (s-1)^{1/r} n^{1-1/r} + r - 1 < 2c \cdot n^{1-1/r}$$

for some c (that only depends on r and s), and hence $e(G) < cn^{2-1/r}$. □

Remark 9.8. The bound $\mathrm{ex}(n, K_{r,s}) = O(n^{2-1/r})$ in Theorem 9.7 is known to be tight for $r = 2, 3$ and when s is much larger than r, say $s \geq 1000^r$. It is an open problem to determine if the bound is tight for $(r, s) = (4, 4)$, namely, if there exist $K_{4,4}$-free n-vertex graphs with at least $cn^{7/4}$ edges, for some constant $c > 0$.