Graph Theory

Prof. A.G. Thomason

Michaelmas 2005

LaTeXed by Sebastian Pancratz

These notes are based on a course of lectures given by Prof. A.G. Thomason in Part II of the Mathematical Tripos at the University of Cambridge in the academic year 2005–2006.

These notes have not been checked by Prof. A.G. Thomason and should not be regarded as official notes for the course. In particular, the responsibility for any errors is mine — please email Sebastian Pancratz (`sfp25`) with any comments or corrections.

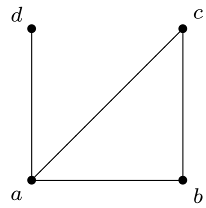
Contents

Chapter 1

Basic Definitions and Properties

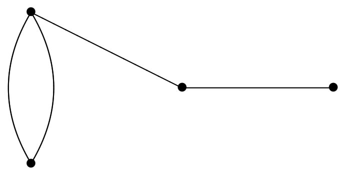
Definition. A *graph* is a pair $G = (V, E)$ where $E \subset V^{(2)} = \{\{x, y\} \subset V : x \neq y\}$.

Example. $V = \{a, b, c, d\}$, $E = \{\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}\}$. Informally, this graph can be represented as follows.



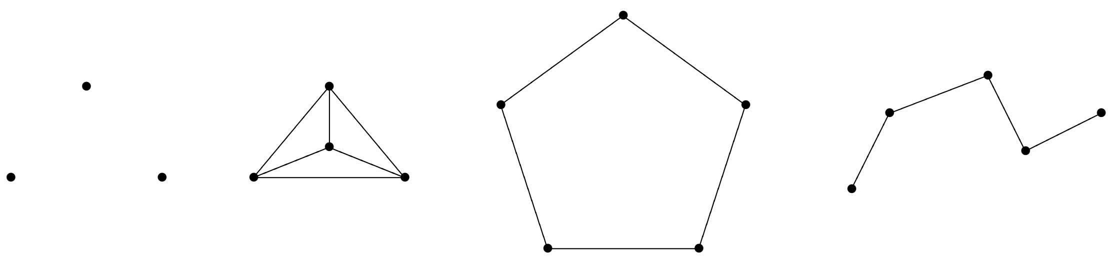
Definition. The set V is the set of *vertices* of G, also denoted $V(G)$. The set E is the set of *edges* of G, also denoted $E(G)$. All graphs considered here are finite. $|V|$ is called the *order* of G, often written $|G|$. $|E|$ is called the *size* of G, often written $e(G)$.

Our graphs have no multiple edges, but some authors allow these and call our graphs *simple*.



Definition. Two graphs G, H are *isomorphic* if there is a bijection $\psi : V(G) \to V(H)$ such that $ab \in E(G)$ if and only if $\psi(a)\psi(b) \in E(H)$.

Example.
(i) E_n, the graph with n vertices and no edges, called the empty graph of order n, e.g. E_3.
(ii) K_n, the complete graph of order n with $E(K_n) = V^{(2)}$, e.g. K_4.
(iii) P_n, the path of order n, e.g. P_5.
(iv) C_n, the cycle of order n, e.g. C_5.



Note that paths and cycles do not allow repetitions of vertices.

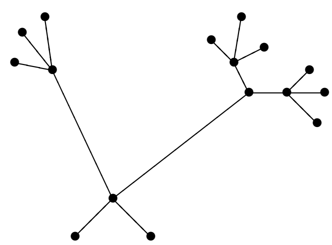
Definition. H is a *subgraph* of G if H is a graph with $V(H) \subset V(G)$ and $E(H) \subset E(G)$.

Every graph of order at most n is a subgraph of K_n. (Formally, every such graph is isomorphic to a subgraph of K_n, but we will not distinguish between distinct isomorphic graphs.)

Definition. H is an *induced subgraph* of G if $V(H) \subset V(G)$ and $E(H) = V^{(2)}(H) \cap E(G)$. If $W \subset V(G)$ we write $G[W]$ for the induced subgraph with vertex set W. Note that the only induced subgraphs of K_n are K_k for $k \leq n$.

Definition. A graph G is *connected* if for every pair of vertices $u, v \in V(G)$ there is a path in G from u to v (called a $u - v$ path). A *component* of G is a maximal connected subgraph, and it is necessarily induced. G is a disjoint union of its components.

Definition. A *forest* is an acyclic graph. A *tree* is a connected acyclic graph. The components of a forest are trees.



Theorem 1.1. The following are equivalent.

(i) G is a tree.
(ii) G is minimal connected, i.e. the removal of any edge destroys connectivity.
(iii) G is maximal acyclic, i.e. the addition of any new edge creates a cycle.

Proof. [$(i) \implies (ii)$] G is connected. Suppose $uv \in E(G)$ and $G - uv$ is still connected. Then there is a $u - v$ path P in $G - uv$ which with uv forms a cycle in G. So if G is acyclic and connected, it is minimal connected.

[$(ii) \implies (i)$] Suppose G is connected and has a cycle C, say. Let uv be an edge of the cycle. Let $x, y \in G$ be vertices such that some $x - y$ path P in G contains uv. Then $P - uv$ together with $C - uv$ still contains an $x - y$ path, so $G - uv$ is connected. Hence if G is minimal connected then G is acyclic (and connected).

[$(i) \implies (iii)$] G is a tree. If $uv \notin E(G)$ there is a $u - v$ path in G which with uv forms a cycle in $G + uv$.

[$(iii) \implies (i)$] G is maximal acyclic. If $uv \notin E(G)$ then $G + uv$ has a cycle containing uv, so G has a $u - v$ path. Hence G is connected, and acyclic by assumption. $\square$

Corollary 1.2. A graph G is connected if and only if it has a spanning tree, that is, a subgraph T such that $V(T) = V(G)$ and T is a tree.

Proof. Since T is connected and spanning, G is connected. Conversely, if G is connected, let T be a minimal connected spanning subgraph. By Theorem 1.1, T is a tree. $\square$

Definition. The set of *neighbours* of a vertex v is denoted $\Gamma(v) = \{w \in V(G) : vw \in E(G)\}$. A vertex $w \in \Gamma(v)$ is *adjacent* to v, an edge vw with $w \in \Gamma(v)$ is *incident* to v. The *degree* of v is $d(v) = |\Gamma(v)|$. The degrees of G written in some order form a *degree sequence*.

Lemma 1.3 (Handshaking Lemma)**.**

$$\sum_{v \in V} d(v) = 2e(G)$$

Definition. A *leaf* is a vertex of degree one.

Definition. The *minimum degree* of G is $\delta(G)$, the *maximum degree* is $\Delta(G)$.

Theorem 1.4. A tree of order at least 2 has at least 2 leaves. Note that this is the best possible bound by considering a path.

Proof. Let T be a tree and let x_1 be a vertex, which we choose to be a leaf if there is one. Let $x_1x_2\dots x_k$ be a path of maximal length. Since T is connected, we have $\delta(T) \geq 1$ (since $|G| \geq 2$), so $k \neq 1$. Then x_k is a leaf, else there exists $x_ku \in E(G)$ where $u \neq x_{k-1}$ and by maximality of the length of the path, $u = x_i$ for some $i \leq k-2$, contradicting that T is acyclic.

So we can assume x_1 is a leaf to obtain a second leaf x_k. ☐

Corollary 1.5. A tree of order n has size $n-1$.

Proof. We prove this by induction on n. (The cases $n \leq 2$ are trivial.) Let T be a tree of order $n \geq 3$. By Theorem 1.4, T has a leaf v. The graph $T - v$ has order $n-1$, is acyclic, and connected. (If v lies on an $x-y$ path in T, then $v = x$ or $v = y$.) So $T - v$ is a tree of order $n-1$, so has size $n-2$ by induction. ☐

Corollary 1.6. The following are equivalent.

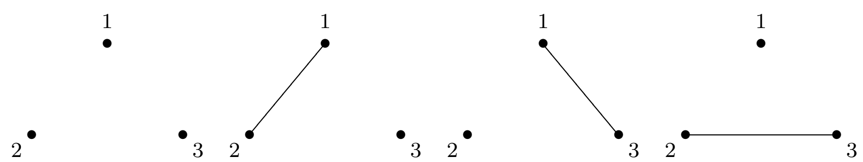
(i) G is a tree of order n.
(ii) G is connected, of order n and size $n-1$.
(iii) G is acyclic, of order n and size $n-1$.

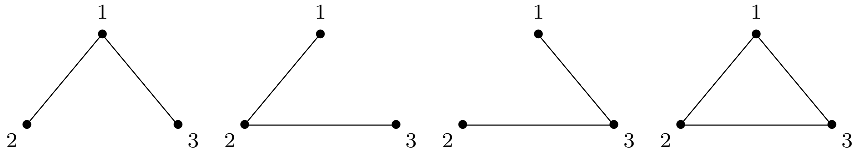
Proof. $[(i) \implies (ii), (i) \implies (iii)]$ This follows from the definition and Corollary 1.5.

$[(ii) \implies (i)]$ G contains a spanning tree T. By Corollary 1.5, $e(T) = n-1$, so $T = G$.

$[(iii) \implies (i)]$ Add edges to get a maximum acyclic graph G'. By Theorem 1.1, G' is a tree. Now $e(G') = n-1$, so $G' = G$. ☐

How many graphs of order n are there? Take the vertex set $[n] = \{1, \dots, n\}$. The next figure illustrates the case $n = 3$.





The number of labelled graphs, where we can distinguish vertices by name, is $2^{\binom{n}{2}}$. To count unlabelled graphs, we need to count the number of orbits in the set of labelled graphs under the action of some permutation group. We can use Burnside's lemma. This states that, for an action of a group G, (no. of G-orbits) $\times$ $|G|$ $=$ $\sum_{g\in G}$ (no. of fixed points of g).

order	3	4	5
no. of (unlabelled) graphs	4	11	34

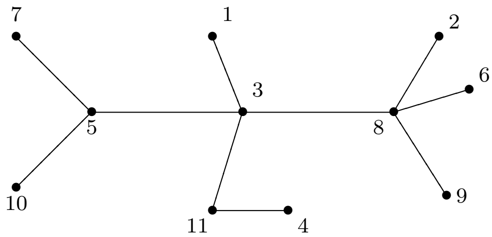
How many trees of order n are there?

order	2	3	4	5
no. of (labelled) trees	1	3	16	125

Theorem 1.7 (Caley)**.** There are n^{n-2} labelled trees of order n.

Proof (Prüfer). We construct a bijection between the set of labelled trees and the set of strings of length $n-2$ with alphabet $[n]$.

[Trees $\rightarrow$ Strings] Choose the smallest labelled leaf. Write down its neighbour. Remove the leaf. Repeat until one edge is left. Consider the following tree as an example.



Deleting 1,2,4,6,7,9,8,10,5, we obtain the string 3,8,11,8,5,8,3,5,3.

[Strings $\rightarrow$ Trees] Note that a vertex v appears $d(v)-1$ times. Amusingly,

$$\sum_v (d(v)-1) = \sum_v d(v) - n = 2(n-1) - n = n-2.$$

Take the least "available " vertex not in the sequence. Mark it "unavailable" and join it to the first number in the sequence. Delete the first number and repeat. When finished, join the two last available vertices. Continuiung the previous example, we start with the string 3,8,11,8,5,8,3,5,3 and mark the vertices 1,2,4,6,7,9,8,10,5 as "unavailable", and finally join 3 and 11. Observe that this process produces an acyclic graph of size $n-1$, so by Corollary 1.6, the result is a tree. $\square$

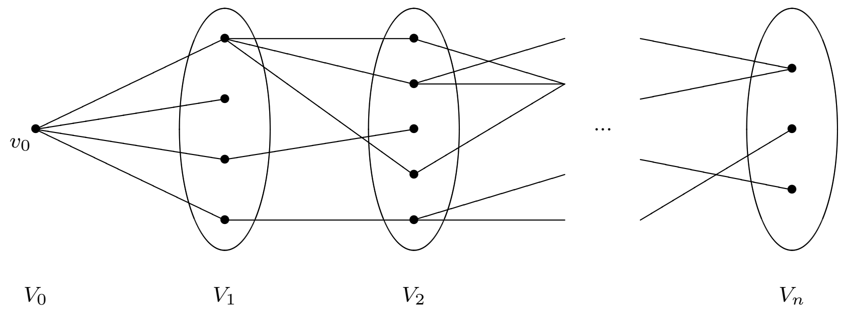
Remark. More formally, let f : Trees $\rightarrow$ Strings, g : Strings $\rightarrow$ Trees. Given T, let $f(T) = a = (a_1, \dots, a_{n-2})$; if $b \in \{1, \dots, n\}$ is minimal not in a, then $g(a)$ has an edge ba_1, and b is the smallest labelled leaf of $g(a)$. Since $f(T-b) = (a_2, \dots, a_n)$, by induction $g(a_2, \dots, a_n) = T - b$, so $g(a) = T$, $gf = \iota$, hence f is injective. By removing the smallest leaf, we similarly obtain $fg = \iota$ so f is surjective, hence f is bijective.

Definition. A graph is *r-partite* if its vertex set can be partitioned into r classes so no edge lies within a class. *Bipartitite* means 2-partite. Remarkably, we can characterise bipartite graphs.

Theorem 1.8. A graph is bipartite if and only if it has no odd cycles.

Proof. If a graph G is bipartite the result is clear since cycles alternate between the two classes.

Conversely, we may assume G is connected by considering components. The result is trivial for the empty graph, so suppose G is not the empty graph. Pick a vertex v_0, and let $V_i = \{w \in G : d(v_0, w) = i\}$ where $d(u, v)$ is the distance from u to v, i.e. the length of the shortest $u - v$ path.



In general, edges lie within classes V_i and between V_i and V_{i+1} but nowehere else, by definition of V_i.

Suppose there is an edge inside some V_i, uv say. Trace back paths of length i from each of u and v to v_0. The first time they meet together with the edge uv we obtain an odd cycle. Since G has no odd cycle, each V_i contains no edges. So $X = \bigcup_{i \text{ even}} V_i$, $Y = \bigcup_{i \text{ odd}} V_i$ gives a bipartition. $\square$

Remark. This gives an algorithm for 2-colouring graphs with no odd cycle.

Definition. An *eulerian tour* of a graph is a walk along edges covering each edge exactly once and returning to the start. A graph is *eulerian* if it has no isolated vertices (i.e. vertices with degree 0) and has an eulerian tour.

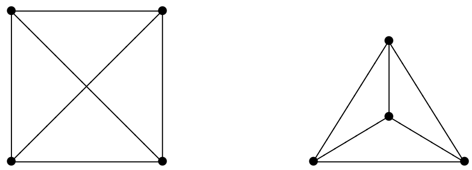
Theorem 1.9. G is eulerian if and only if $|G| > 1$, G is connected and all degrees are even.

Proof. If G is eulerian, the result is clear. We prove the converse by induction on the size of G. The conditions imply $\delta(G) \geq 2$, so Theorem 1.4 shows that there is a cycle C (since the order of G is at least 2 and G is connected, if G is acyclic there exists a leaf, contradicting $\delta(G) \geq 2$). $G - E(C)$ has all components isolated vertices or connected even degree components, each having an eulerian tour by induction. Go around C, taking in "sub-eulerian tours" when first encountered. $\square$

Remark. Analogeously, when starting and finishing at distinct vertices, require the existence of exactly two vertices of odd degree.

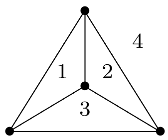
Definition. A graph is *planar* if it can be drawn in the plane without edges crossing.

There are no essential difficulties of analysis here. In particular, we take the Jordan Curve Theorem for granted in this course. Indeed, it can be shown that any such drawing can be done with straight line edges. (See Example Sheet 1.)



Definition. A *plane* graph is a planar graph drawn in the plane. A *face* is a simply-connected region, including the ∞-face.

Example. K_4 has four faces.



Lemma 1.10. Let G be a graph with $d(v)$ even for all $v \in V(G)$. Then $E(G)$ can be partioned into cycles.

Proof. We may ignore isolated vertices and assume $d(v) \geq 2$. If G contains a cycle we can remove it and note that the all degrees remain even. If there is no cycle then G contains a tree and hence a leaf, i.e. a vertex of degree 1, contradiction. ☐

Lemma 1.11. Let e be an edge of a plane graph G. Then e is in a boundary of two faces if and only if there exist a cycle C in G containing e.

Proof. The drawing of C is a simple polygonal closed curve, separating the plane.

Let F be one of the faces with e in its boundary. Let H be the spanning subgraph of G consisting of all edges h with F on one side, not F on the other. Going around a vertex v in a small loop, we observe that we enter and leave F the same number of times, so $d_H(v)$ is even. By Lemma 1.10, all edges of H are in cycles. ☐

Theorem 1.12 (Euler)**.** If G is a connected plane graph of order n and size m with f faces then $n - m + f = 2$.

Proof. Prove this by induction on m. Connectivity implies $m \geq n - 1$ as trees are minimal connected and have size $n - 1$. If $m = n - 1$ then G is a tree, so $f = 1$. If $m \geq n$, pick an edge e in some cycle. $G - e$ has n vertices, $m - 1$ edges and $f - 1$ faces since e lies between two faces. ☐

Definition. A *bridge* in a graph is an edge whose removal increases the number of components.

Equivalently, the edge does not lie in a cycle. Observe that if e is not a bridge and lies in a plane graph then it borders two distinct faces. Thus if G is a bridgeless plane graph with f_i faces of length i, then $\sum_i i f_i = 2m$. Note here that the ∞-face has a well-defined length for a bridgeless graph.

Definition. The *girth* $g(G)$ of a graph G is the length of the shortest cycle, or ∞ is G is acyclic.

Theorem 1.13. Let G be a connected bridgeless planar graph of order n and girth g. Then $e(G) \leq \frac{g}{g-2}(n-2)$. In particular, a planar graph of order n satisfies $e(G) \leq 3n-6$.

Proof. Draw G in the plane. Then $2m = \sum_i if_i \geq g\sum_i f_i = gf$. By Theorem 1.12, $n - m + \frac{2m}{g} \geq 2$, i.e. $m(g-2) \leq g(n-2)$. In particular, if G is a bridgeless connected planar graph then $e \leq 3n-6$, this holds for any planar graph G with $n \geq 3$, either by induction or by adding edges till G is bridgeless connected. $\square$

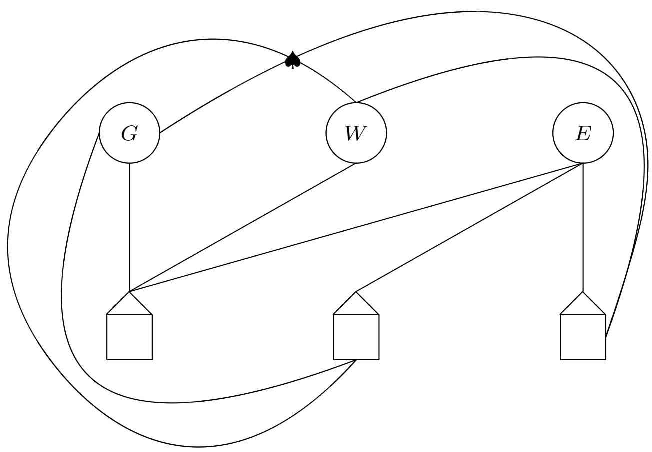
Remark. Adding edges leads to problems since we may decrease g. Instead, note that we have solved the bridgeless case. If G contains a bridge ab, $G - ab$ has two vertex-disjoint subgraphs satisfying the formula by induction. Then

$$\begin{aligned} e(G) &= e(G_1) + e(G_2) + 1 \\ &\leq \frac{g}{g-2}[(n_1-2)+(n_2-2)] + 1 \\ &= \frac{g}{g-2}(n-2) + 1 - 2\frac{g}{g-2} \\ &\leq \frac{g}{g-2}(n-2). \end{aligned}$$

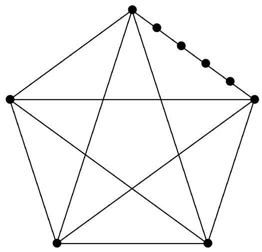
Example. Note $e(K_5) = 10 > 3(5-2)$, so K_5 is non-planar.

Definition. The *complete bipartite graph* $K_{p,q}$ is bipartite with p vertices in one class, q vertices in the other, and all pq possible edges between them.

Example. $e(K_{3,3}) = 9 > \frac{4}{2}(6-2)$ as $g(K_{3,3}) = 4$, so $K_{3,3}$ is non-planar.



Observe that having too many edges is not the only reason graphs fail to be planar. For example, no subdivision of K_5 is planar, as can be seen on replacing edges by disjoing paths, but these may satisfy the conclusion of Theorem 1.13 with sufficiently long paths.



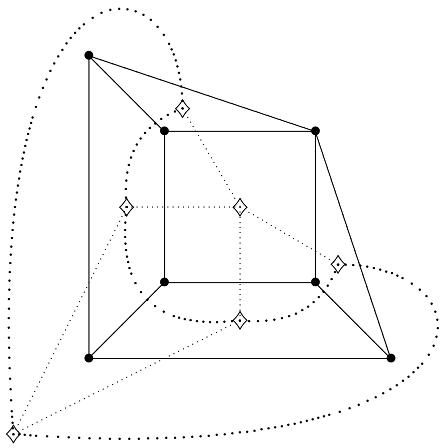
Remarkably, plane graphs can be characterised.

Theorem 1.14 (Kuratowski, 1930)**.** A graph is planar if and only if it contains no subdivision of K_5 or of $K_{3,3}$.

Proof. Proof omitted. □

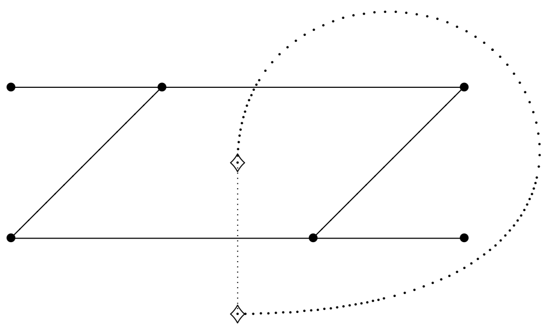
Definition. Given a plane graph G, we can construct the *dual* graph G^*. Place a dual vertex inside each face, and, for each edge, draw a dual edge joining the corresponding dual vertices.

Example. In this example, the original graph has 7 vertices, 11 edges, 6 faces and the dual has 6 vertices, 11 edges, 7 faces.



Note that usually the dual of the dual is the original graph.

Remark. If a graph is not 3-connected (see Chapter 2), the dual might not be simple.

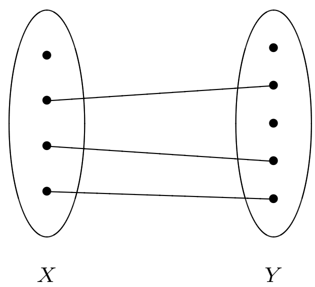


But if G is a 3-connected simple graph then so is G^*.

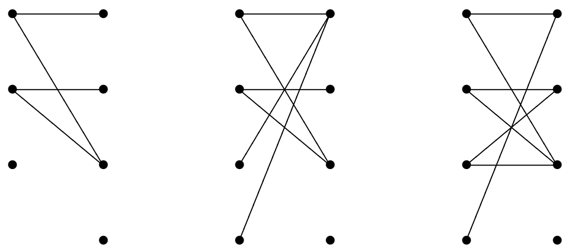
Chapter 2

Matchings and Connectivity

Definition. Let G be a bipartite graph with vertex classes X, Y. A *matching* from X to Y is a set of $|X|$ independent, i.e. pairwise non-incident, edges. If $|X| = |Y|$, this is also a matching from Y to X, also called a 1-factor, i.e. a 1-regular spanning subgraph, where *n-regular* means every vertex has degree n.



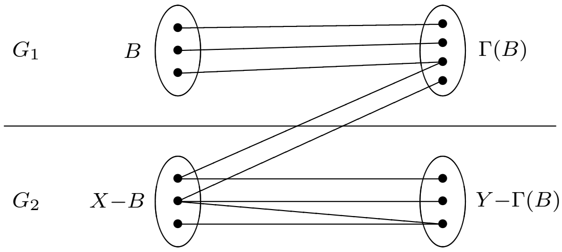
Consider X a set of men, Y a set of women, and $xy \in E(G)$ if x can marry y. Finding a matching from X to Y is to find wives for all men. We need $|Y| \geq |X|$ and that every man knows a woman. In general, every set of k men needs to know at least k women. For $A \subset X$ define $\Gamma(A) = \bigcup_{x \in A} \Gamma(x)$. Clearly we need $|\Gamma(A)| \geq |A|$ for all $A \subset X$. The following figure illustrates cases in which this fails.



Theorem 2.1 (Hall)**.** *Let G be a bipartite graph with vertex classes X, Y. Then G has a matching from X to Y if and only if*

$$\forall A \subset X \quad |\Gamma(A)| \geq |A| \tag{HC}$$

Proof 1. The necessity is clear. We prove sufficiency by induction on $|X|$. If for every $\emptyset \neq A \subsetneq X$ we have $|\Gamma(A)| > |A|$, then pick any edge $xy \in E(G)$ and in $G' = G - xy$ Hall's condition holds. The matching in G', by induction, together with xy gives a matching in G. Otherwise, there exists a *critical* set $\emptyset \neq B \subsetneq X$ such that $|\Gamma(B)| = |B|$. Let $G_1 = G[B \cup \Gamma(B)]$, $G_2 = G[(X - B) \cup (Y - \Gamma(B))]$.



For $A \subset B$, we have $\Gamma_1(A) = \Gamma(A)$, where $\Gamma_1(A)$ are the neighbours of A in G_1, hence by (HC) in G

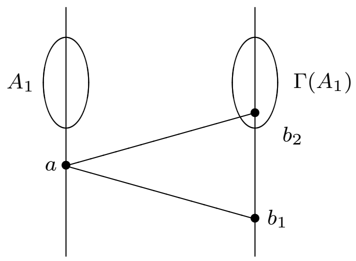
$$|\Gamma_1(A)| = |\Gamma(A)| \geq |A|,$$

so (HC) holds in G_1. For $A \subset X - B$, we have

$$\begin{aligned}|\Gamma_2(A)| &= |\Gamma(A \cup B)| - |\Gamma(B)| \\ &= |\Gamma(A \cup B)| - |B| \geq |A \cup B| - |B| = |A|\end{aligned}$$

using (HC) in G to establish the inequality. So (HC) holds in G_2. Thus G_1, G_2 both have matchings, hence so does G. $\square$

Proof 2. Consider a minimal (with respect to removing edges) subgraph in which (HC) still holds. If in the resulting graph $d(x) = 1$ for all $x \in X$ then what is left is a matching. If not, there exists $a \in X$ joined to $b_1, b_2 \in Y$ and sets $A_1, A_2 \subset X - \{a\}$ such that $|\Gamma(A_i)| = |A_i|$, $b_i \neq \Gamma(A_i)$ and $\Gamma(A_i \cup \{a\}) = \Gamma(A_i) \cup \{b_i\}$, for $i = 1, 2$.



Hence $\Gamma(A_1 \cup A_2 \cup \{a\}) = \Gamma(A_1 \cup A_2)$. Thus

$$\begin{aligned}|\Gamma(A_1 \cup A_2 \cup \{a\})| &= |\Gamma(A_1 \cup A_2)| \\ &= |\Gamma(A_1) \cup \Gamma(A_2)| \\ &= |\Gamma(A_1)| + |\Gamma(A_2)| - |\Gamma(A_1) \cap \Gamma(A_2)| \\ &\leq |\Gamma(A_1)| + |\Gamma(A_2)| - |\Gamma(A_1 \cap A_2)| \\ &= |A_1| + |A_2| - |\Gamma(A_1 \cap A_2)| \\ &\leq |A_1| + |A_2| - |A_1 \cap A_2| \\ &= |A_1 \cup A_2| \\ &= |A_1 \cup A_2 \cup \{a\}| - 1,\end{aligned}$$

violating (HC), a contradiction. Here we have used (HC) on $A_1 \cap A_2$ to establish the second inequality. $\square$

Corollary 2.2 (Defect form)**.** Let G be as above and $d \in \mathbb{N}$. There exists $|X| - d$ independent edges in G if and only if $|\Gamma(A)| \geq |A| - d$ for all $A \subset X$.

Proof. Introduce d new members of Y joined to all members of X. In the new graph (HC) holds, so there is a matching. Now remove the d new vertices. □

Corollary 2.3 (Polygamous version)**.** We can give every man 2 wives if and only if $|\Gamma(A)| \geq 2|A|$ for all $A \subset X$.

Proof. Replace each man by 2 clones of himself, i.e. with the same neighbours. In the new society (HC) holds, so marry off all men. Now remove the clones and give their wives to the original men. □

Tutte's theorem gives a necessary and sufficient condition for a 1-factor in a general, not necessarily bipartite, graph.

Definition. Given a collection $Y_1, \ldots, Y_n$ of subsets of a set Y, a set of *distinct representatives* is a set $\{y_1, \ldots, y_n\}$ with $y_i \in Y_i$ and $y_i \neq y_j$ for $i \neq j$.

Corollary 2.4. There is a set of distinct representatives if and only if

$$\forall S \subset [n] \quad \Big|\bigcup_{i \in S} Y_i\Big| \geq |S|.$$

Proof. Necessity is clear. To prove sufficiency, construct a bipartite graph with $X = \{Y_1, \ldots, Y_n\}$ and edges from $Y_i \in X$ to $y \in Y$ if $y \in Y_i$. Now apply Theorem 2.1. □

Definition. A graph G is *k-connected* if $|G| > k$ and $G - S$ is connected for every set $S \subset V(G)$ with $|S| < k$.

Definition. Define the *vertex connectivity* to be

$$\kappa(G) = \max\{k : G \text{ is } k\text{-connected}\}.$$

If G is not complete, $\kappa(G) = \min\{|S| : \exists S \subset V(G)\ G - S \text{ is disconnected}\}$.

Definition. Given $a, b \in V(G)$, $ab \notin E(G)$, the *local connectivity* is

$$\kappa(a, b; G) = \min\{|S| : \exists S \subset V(G) - \{a, b\} \text{ with no } a - b \text{ path in } G - S\}.$$

Clearly, $\kappa(G) = \min_{ab \notin E} \kappa(a, b; G)$ if G is not complete. There are edge connectivity analogues where a set $F \subset E(G)$ is removed.

$$\begin{aligned} \lambda(G) &= \min\{|F| : \exists F \subset E(G)\ G - F \text{ is disconnected}\} \\ \lambda(a, b; G) &= \min\{|F| : \exists F \subset E(G) \text{ with no } a - b \text{ path in } G - F\} \\ \lambda(G) &= \min_{a,b} \lambda(a, b; G) \end{aligned}$$

Note that in the case of edge connectivity, no special care is required for complete graphs. The following holds.

$$\kappa(G) \leq \lambda(G) \leq \delta(G)$$

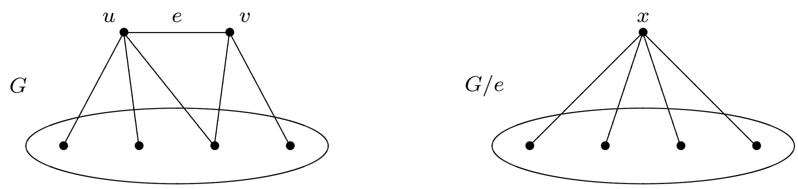
For the first inequality, remove one endvertex for each edge in our set of $\lambda(G)$ edges whose removal disconnects G. For the second inequality, remove $\delta(G)$ edges from a vertex of least degree. (See Example Sheet 1.)

Definition. A set of $a-b$ paths is *vertex disjoint* if the only vertices in more than one path are a, b.

Clearly, the size of any such set is at most $\kappa(a,b;G)$, since to separate a from b, we need to remove at least one vertex from each path. Remarkably, there is a set of $\kappa(a,b;G)$ paths.

Theorem 2.5 (Menger)**.** Let $a, b \in V(G)$ with $ab \notin E(G)$. Then there exists a set of $\kappa(a,b;G)$ vertex-disjoint $a-b$ paths.

Definition. We use the notion of *graph contraction.* If $e \in E(G)$, the graph G/e derived from G by contracting e is obtained by removing both endvertices u, v of e and introducting a new vertex x joined to $\Gamma(u) \cup \Gamma(v)$.

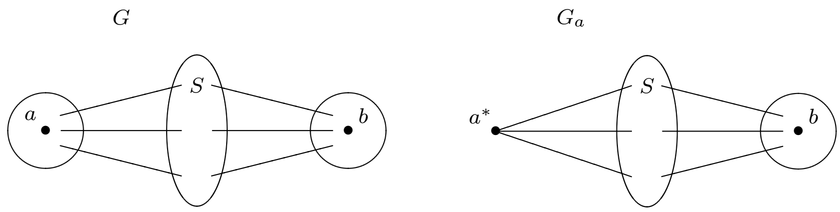


A *contraction of* G is obtained by a sequence of these operations.

Proof. Suppose not, and let G, a, b be a minimal counterexample, i.e. G has minimal order. Let $k = \kappa(a,b;G)$ and define a minimal cutset to be any set $S \subset V - \{a,b\}$ with $|S| = k$ such that $G - S$ has no $a-b$ path.

Claim (i): Every edge e not meeting a or b lies inside a minimal cutset. For otherwise, $\kappa(a,b;G/e) \geq k$ and a set of k vertex-disjoint $a-b$ paths in G/e would yield a set in G, contradiction.

Claim (ii): If S is a minimal cutset, then $S = \Gamma(a)$ or $S = \Gamma(b)$. For otherwise, let G_a be the graph obtained by contracting the component of $G - S$ containing a to a single vertex a^*.



Since $S \neq \Gamma(a)$, $|G_a| < |G|$. Clearly $\kappa(a^*, b; G_a) \geq k$. Thus there exists a set of k $a^* - b$ paths in G_a. Likewise define G_b and find a set of k $a - b^*$ paths in G_b. These paths would yield k vertex-disjoint $a-b$ paths in G.

If $\Gamma(a) \neq \Gamma(b)$ then $|\Gamma(a) \cap \Gamma(b)| < k$, then there exists an edge (in the cutset) lying neither inside $\Gamma(a)$ nor in $\Gamma(b)$, contradicting the claims. But if $\Gamma(a) = \Gamma(b)$ we immediately have k $a-b$ paths. □

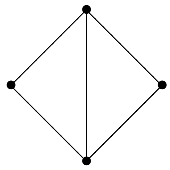
Corollary 2.6. Let $\kappa(G) \geq k$, let X, Y be disjoint subsets of $V(G)$ with $|X|, |Y| \geq k$. Then there exists a set of k completely vertex disjoint $X - Y$ paths.

Proof. Add new vertices x joined to all vertices in X, y joined to all vertices in Y, to form G^*. Then $\kappa(x, y; G^*) \geq k$, so the desired paths exist by Menger's theorem. $\square$

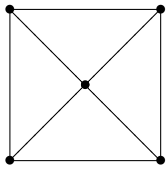
Theorem 2.7 (Edge form of Menger's theorem)**.** If $a, b \in V(G)$ then there exists a set of $\lambda(a, b; G)$ edge-disjoint $a - b$ paths. Note that, trivially, no larger such set exists.

Proof. Either we mimic the proof of the vertex form, or we construct the *line graph* $L(G)$ of G, whose vertex set is the set of edges of G, where $ef \in E(L(G))$ if e, f are incident edges of G.

G



$L(G)$



Now form H from $L(G)$ by joining a new vertex a^* to all vertices of $L(G)$ that correspond to edges of G meeting a, and introduce b^* likewise.

Note an $a - b$ path in G gives an $a^* - b^*$ path in H, and an $a^* - b^*$ path in H gives a set of edges containing an $a - b$ path in G. We have $\kappa(a^*, b^*; H) = \lambda(a, b; G)$, so there is a set of $\lambda(a, b; G)$ vertex-disjoint $a^* - b^*$ paths in H by Theorem 2.5, so there is a set of $\lambda(a, b; G)$ edge-disjoint $a - b$ paths in G. $\square$

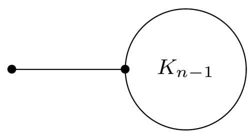
Remark (Menger's theorem implies Hall's theorem)**.** Introduce new vertices x joined to all vertices in X and y joined to all vertices in Y to obtain G' from G. Then $\kappa(x, y, G') = |X|$ if and only if (HC) holds in G. ((HC) fails if and only if there xists $A \subset X$ such that $|A| > |\Gamma(A)|$ if and only if we can remove $(X - A) \cup \Gamma(A)$ of size less than X.)

Chapter 3

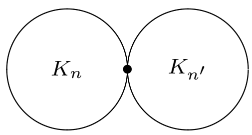
Extremal Graph Theory

Definition. A *Hamiltonian cycle* in a graph is a spanning cycle, i.e. it meets every vertex exactly once.

Q: How many edges are needed to guarantee a Hamiltonian cycle?
A: We need more than $\binom{n}{2} - (n-2)$ edges. (See Example Sheet 2.)

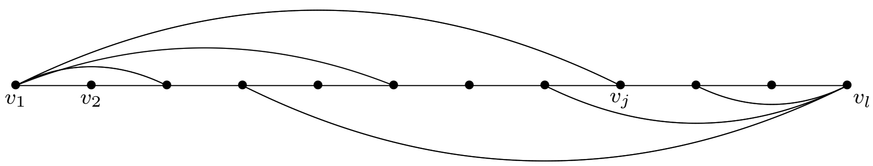


Q: How large must $\delta(G)$ be to guarantee a Hamiltonian cycle?
A: We need $\delta(G) \geq \frac{n}{2}$.



Theorem 3.1. Let G be a graph of order $n \geq 3$ such that every pair of non-adjacent vertices x, y satisfies $d(x) + d(y) \geq k$. If $k < n$ and G is connected then G has a path of length k. If $k = n$ then G has a Hamiltonian cycle.

Proof. Observe that if $k = n$ then G must be connected, since any two non-adjacent vertices have a common neighbour. Suppose that G has no Hamiltonian cycle, for otherwise we are done. Let $P = v_1 v_2 \dots v_l$ be a path of maximum length. Notice that G has no l-cycle, because if $l = n$ this would be a Hamiltonian cycle and if $l < n$ by connectivity there would be a path of length $l + 1$. In particular, $v_1 v_l \notin E(G)$, so $d(v_1) + d(v_l) \geq k$. Note that all neighbours of v_1, v_l lie in P, by maximality of P.



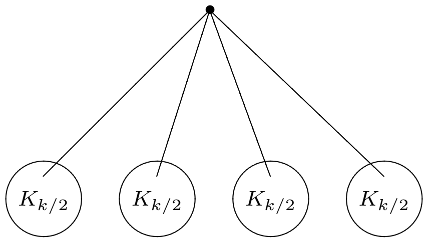
Let $S = \{i : v_1 v_i \in E(G)\}$, $|S| = d(v_1)$, $T = \{i : v_l v_{i-1} \in E(G)\}$, $|T| = d(v_l)$. Now $S \cup T \subset \{2, \dots, l\}$ and $S \cap T = \emptyset$, for if $j \in S \cap T$ then $v_1 v_2 \dots v_{j-1} v_l v_{l-1} \dots v_j$ would be an l-cycle. Hence

$$l - 1 \geq |S \cup T| = |S| + |T| = d(v_1) + d(v_l) \geq k.$$

If $k = n$, this is impossible, so we have a Hamiltonian cycle. If $k < n$, then P has length at least k. □

Corollary 3.2 (Dirac)**.** If G has order n and $\delta(G) \geq \frac{n}{2}$ then G has a Hamiltonian cycle.

Remark. Note if $2 \mid n$ and $\frac{k}{2} \mid n-1$ then Theorem 3.1 is best possible.



In this example, the longest path has length $l = k+1$.

Theorem 3.3. Let G be a graph of order n with no path of length k. Then $e(G) \leq \frac{k-1}{2}n$. Moreover, equality holds only if $k \mid n$ and G is a disjoint of copies of K_k.

Proof. By induction on n. The result is easily true for $n \leq k$. In general, if G is disconnected, we are home at once by the hypothesis applied to each component. If G is connected, $\delta(G) \leq \frac{k-1}{2}$ by Theorem 3.1. Let x have degree at most $\frac{k-1}{2}$. Since G is connected, K_k is not a subgraph of G, so

$$e(G-x) < \frac{k-1}{2}(n-1)$$

by the inductive hypothesis. Then

$$e(G) \leq e(G-x) + \frac{k-1}{2} < \frac{k-1}{2}n.$$ □

Definition. Given a fixed graph F, define

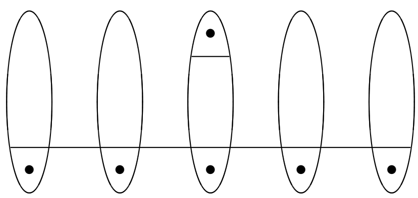
$$\mathrm{ex}(n, F) = \max\{e(G) : |G| = n, F \not\subset G\}.$$

Q: How many edges are there in a K_{r+1}-free graph?

Observe that r-partite graphs contain no K_{r+1}. To obtain an r-partitite graph of maximum size and of order n, it should be complete r-partite. Moreover, if we have two classes X, Y with $|X| \geq |Y| + 2$, changing to class sizes $|X| - 1$ and $|Y| + 1$ gains us $-|Y| + (|X| - 1) > 0$ edges. Hence, there is a unique r-partite graph of order n and maximum size. The classes have size $\lfloor \frac{n}{r} \rfloor$ or $\lceil \frac{n}{r} \rceil$ and it is complete r-partite. It is called the r-partite Turán graph of order n. It is denoted by $T_r(n)$ and its size is $t_r(n)$.

The value of $t_r(n)$ can be written explicity in terms of the remainder after n divided by r but this is awkward to work with. When working with $t_r(n)$, it is more convenient to use some observations about it derived from the structure of $T_r(n)$.

Consider the case $r = 5$, remainder 3. Remove the vertex at the top for ($*$), remove the vertices at the bottom for ($**$).



First, observe that vertices of least degree in $T_r(n)$ lie in the largest classes, and if we remove one such vertex, we get $T_r(n-1)$; hence

$$t_r(n) - \delta(T_r(n)) = t_r(n-1) \tag{$*$}$$

Moreover, removing one vertex from each class (i.e. a K_r) and noting we have now a $T_r(n-r)$ in which each vertex has $r-1$ neighbours in K_r,

$$t_r(n) = \binom{r}{2} + (n-r)(r-1) + t_r(n-r) \tag{$**$}$$

Finally, note $\Delta(T_r(n)) \leq \delta(T_r(n)) + 1$, so if we have a graph with $|G| = n$ and $\delta(G) > \delta(T_r(n))$ then $e(G) > t_r(n)$. More explicitly, by comparing the degree sequence of G with that of the Turán graph, we have the following.

If $|G| = n$ and $\delta(G) > \delta(T_r(n))$ then

$$e(G) \geq t_r(n) + \tfrac{1}{2}M \tag{$\dagger$}$$

where M is the number of vertices of minimum degree in $T_r(n)$.

Q: Can we do better than $T_r(n)$ and still be K_{r+1}-free?

Theorem 3.4 (Turán, 1941)**.** Let G be K_{r+1}-free of order n with $e(G) \geq t_r(n)$. Then $G = T_r(n)$.

Proof 1. By induction on n. The case $n \leq r$ is trivial as then $T_r(n) = K_n$. In general, given G, remove edges to obtain G' with $e(G') = t_r(n)$. By ($\dagger$), $\delta(G') \leq \delta(T_r(n))$. Let x be a vertex of minimum degree in G'. Then $G' - x$ has order $n-1$, is K_{r+1}-free and

$$\begin{aligned} e(G'-x) &= e(G') - \delta(G') \\ &= t_r(n) - \delta(G') \\ &\geq t_r(n) - \delta(T_r(n)) \\ &= t_r(n-1) \end{aligned}$$

using ($*$). By the induction hypothesis, $G' - x = T_r(n-1)$. Since there must be some class of G' in which x has no neighbour (otherwise $K_{r+1} \subset G'$), G' is r-partite. But $e(G') = t_r(n)$, so $G' = T_r(n)$. Since $T_r(n)$ is maximal K_{r+1}-free, $G = G' = T_r(n)$. □

Proof 2. By induction on n. Obtain G' from G by adding edges until the graph is maximal K_{r+1}-free. Certainly G' contains some K_r, K say. Each vertex of G' has at most $r-1$ neighbours in K_r, so

$$e(G') \leq \binom{r}{2} + (n-r)(r-1) + e(G'-K).$$

By ($**$), $e(G'-K) \leq t_r(n-r)$, so $G' - K = T_r(n-r)$ by induction and equality holds throughout, so every vertex of $G' - K$ is joined to all vertices but one of K. Since vertices in different classes of $G' - K$, i.e. $T_r(n-r)$, miss different vertices of K, we have $G' = T_r(n)$ and $G = G'$. □

It is natural to ask the bipartite analogue of the previous question. What is the maximum size of an $n \times n$ bipartite graph, i.e. with n vertices in each class, that contains no complete bipartite subgraph $K_{t,t}$? This is known as the problem of Zarankiewicz. Denote the maximum size by $z(n,t)$. The value of $z(n,t)$ is unkown, even approximately. The following simple idea is thought to be accurate.

Theorem 3.5.

$$\begin{aligned} z(n,t) &\leq (t-1)^{1/t}(n-t+1)n^{1-1/t} + (t-1)n \\ &= \mathcal{O}(n^{2-1/t}) \end{aligned}$$

if t is fixed and $n \to \infty$.

Remark. Even the proper rate of growth, as a power of n, is unknown.

Proof. Let G be a maximal $n \times n$ $K_{t,t}$-free graph of size $z(n,t)$ with bipartition X, Y, $|X| = |Y| = n$. Let the vertices in X have degrees $d_1, \ldots, d_n$; note by maximality of G that $d_i \geq t-1$ for all $i = 1, \ldots, n$. Let $nd = \sum_{i=1}^n d_i = z(n,t)$. Let

$$S = \{(x,T) : x \in X, T \subset Y \text{ such that } |T| = t \ \wedge \ T \subset \Gamma(x)\}.$$

If $x \in X$ has degree d_i, there are $\binom{d_i}{t}$ pairs (x,T) in S with this choice of x. If $T \subset Y$ with $|T| = t$ there are at most $t-1$ x's with (x,T) in S. Thus

$$\sum_{i=1}^n \binom{d_i}{t} = |S| \leq (t-1)\binom{n}{t} \tag{A}$$

Since the polynomial $\binom{w}{t}$ in w is convex if $w \geq t-1$,

$$n\binom{d}{t} \leq \sum_{i=1}^n \binom{d_i}{t} \leq (t-1)\binom{n}{t} \tag{B}$$

Therefore,

$$\left(\frac{d-t+1}{n-t+1}\right)^t \leq \frac{d(d-1)(d-2)\cdots(d-t+1)}{n(n-1)(n-2)\cdots(n-t+1)} \leq \frac{t-1}{n} \tag{$*$}$$

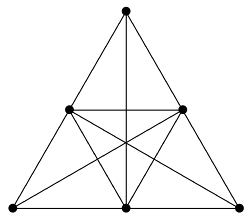
The result follows. □

Theorem 3.6. $z(n,2) \leq \frac{1}{2}n(1+\sqrt{4n-3})$ and equality holds for infinitely many n.

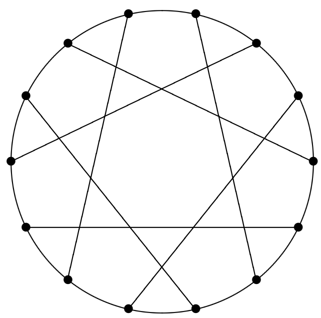
Proof. The above proof shows $d(d-1) \leq n-1$ by ($*$) for $t = 2$. Hence $d \leq \frac{1}{2}(1+\sqrt{4n-3})$. To obtain equality, we need all above inequalities to holds exactly. Thus, from (B), all degrees in X are equal to d, which hence must be an integer, and, from (A) with $t = 2$, every two vertices of Y have exactly one common neighbour in X, and vice versa by arguing with X, Y transposed.

The existence of this graph is equivalent to the existence of a projective plane of order p, where $n = p^2 + p + 1$. This is a set of n points, the vertices of Y, together with n subsets of points called lines, the sets $\Gamma(x)$. Each point is in the same number of lines $(p+1)$, each line has the same number of points $(p+1)$, each pair of points is in exactly one common line, and each two lines have exactly one common point. It is known that there exists a projective plane for every prime power order. □

Example. For example, the following image shows the Fano plane, i.e. the projective plane of order 2.



This gives rise to the Heawood graph.



It is known that there are no such planes of order 6 (easy) or 10 (hard).

Theorem 3.7. $\mathrm{ex}(n, K_{2,2}) \leq \frac{n}{4}(1 + \sqrt{4n-3})$.

Proof. Suppose $|G| = n$, $G \not\supset K_{2,2}$. As before we count vertices $x \in V(G)$ and covering sets $S \subset V(G)$ with $|S| = t$, $x \notin S$. We obtain

$$n\binom{d}{t} \leq \sum_{v \in V} \binom{d(v)}{t} \leq (t-1)\binom{n}{t}$$

where d is the average degree. The result follows by observing $e(G) = \frac{1}{2}dn$. □

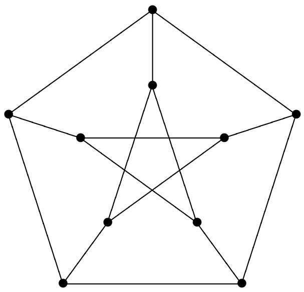
Remark. Not surprisingly, there is no nice exact description of $\mathrm{ex}(n, F)$ in general. Usually, the case of n small relative to $|F|$ is a mess, but sometimes things get nicer for larger n. We have the following examples.

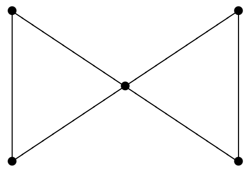
$$\mathrm{ex}(n, K_{r+1}) = t_r(n) \sim (1 - \tfrac{1}{r})\binom{n}{2}$$

This is clear by noting that any $x \in T_r(n)$ is joined to a share of approximately $\frac{r-1}{r}$ vertices. For the next examples, see Example Sheet 2.

$$\begin{aligned} \mathrm{ex}(n, K_3) &= \left\lfloor \frac{n^2}{4} \right\rfloor \\ \mathrm{ex}(n, C_5) &= t_2(n) \text{ for } n \geq 6 \\ \mathrm{ex}(n, F) &= \left\lfloor \frac{n^2}{4} \right\rfloor + 2 \text{ for } n \geq 5 \\ \mathrm{ex}(n, P) &= t_2(n) + n - 2 \text{ for large } n \end{aligned}$$

where P is the Petersen graph, and both P and F are shown below.





For general graphs F, can we find $\mathrm{ex}(n, F)$ approximately? Can we find $\lim_{n\to\infty} \mathrm{ex}(n, F)/\binom{n}{2}$?

Definition. Define $K_r(t)$ to be the complete r-partite graph with t vertices in each class, i.e. $K_r(t) = T_r(rt)$.

The following lemma is the heart of the matter.

Lemma 3.8. Let $r, t \geq 1$ be integers and let $\varepsilon > 0$ be real. Then, if n is sufficiently large (i.e. there exists $n_1(r, t, \varepsilon)$ such that if $n > n_1$), every graph G with $|G| = n$ and $\delta(G) \geq (1 - \frac{1}{r} + \varepsilon)n$ contains $K_{r+1}(t)$.

Note that $K_{r+1}(1) = K_{r+1}$ and $\delta(T_r(n)) \approx (1 - \frac{1}{r})n$.

Proof. We proceed by induction on r, proving the base case $r = 1$ and the general case $r > 1$ simultaneously. (Also note that the case $r = 1$ can be derived from Theorem 3.5.) Let $T = \lceil \frac{2t}{\varepsilon r} \rceil$. We proceed in three simple steps.

(i) G contains $K_r(T)$; call it K. (This part uses induction.)
(ii) $G - K$ contains a large set of vertices U, each joined to at least t in each class of K.
(iii) Many vertices in U (certainly at least t) are joined to the same t in each class of K. This gives $K_{r+1}(t)$.

It will be evident that each step holds if n is sufficiently large. For exactness, we shall use only (i) $n_1(1, t, \varepsilon) \geq T$ and $n_1(r, t, \varepsilon) \geq n_1(r-1, T, 1/r(r-1))$ for $r \geq 2$, (ii) $n_1(r, t, \varepsilon) \geq \frac{6rT}{\varepsilon}$, and (iii) $n_1(r, t, \varepsilon) \geq \frac{3t}{\varepsilon r}\binom{T}{t}^r$.

(i) This is trivial if $r = 1$, provided $n \geq T$. If $r > 1$ and since

$$\delta(G) \geq (1 - \tfrac{1}{r-1} + \tfrac{1}{r(r-1)})n$$

we have that $G \supset K_r(T)$ if n is large, by the induction hypothesis.

(ii) Let U be the vertices in $G - K$ having at least $(1 - \frac{1}{r} + \frac{\varepsilon}{2})|K|$ neighbours in K. Since each vertex of K has degree at least $(1 - \frac{1}{r} + \varepsilon)n$, writing $e(G - K, K)$ for the number of edges between $G - K$ and K, we have

$$\begin{aligned} |K|[(1 - \tfrac{1}{r} + \varepsilon)n - |K|] &\leq e(G - K, K) \\ &\leq |U||K| + (n - |U|)(1 - \tfrac{1}{r} + \tfrac{\varepsilon}{2})|K|, \end{aligned}$$

so

$$\tfrac{\varepsilon n}{2} - |K| \leq |U|(\tfrac{1}{r} - \tfrac{\varepsilon}{2}).$$

Since $|K| \leq \frac{\varepsilon n}{6}$ if n is large, this gives $\frac{|U|}{r} \geq \frac{\varepsilon n}{3}$, so $|U| \geq \frac{\varepsilon r n}{3}$.

(iii) Each vertex of U has at least

$$\begin{aligned}(1-\tfrac{1}{r}+\tfrac{\varepsilon}{2})|K| = (1-\tfrac{1}{r}+\tfrac{\varepsilon}{2})rT &= (r-1)T+\tfrac{\varepsilon rT}{2}\\ &\geq (r-1)T+t\end{aligned}$$

neighbours in K, and hence to at least t in each class of K. Thus, for each $u \in U$, we can pick a $K_r(t) \subset K$ that is in the neighbourhood of u. Since there are only $\binom{T}{t}^r$ possible choices for $K_r(t)$, and since $|U| \geq \frac{\varepsilon rn}{3} \geq t\binom{T}{T}^r$ if n is large, there exist t vertices in U for which the same choice was made, i.e. we have $K_{r+1}(t)$. □

Theorem 3.9 (Erdős–Stone, 1946)**.** Let $r, t \geq 1$ be integers and $\varepsilon > 0$ be real. If n is sufficiently large (i.e. there exists $n_0(r,t,\varepsilon)$ such that if $n > n_0$) then every graph G with $|G| = n$ and $e(G) \geq (1-\frac{1}{r}+\varepsilon)\binom{n}{2}$ contains $K_{r+1}(t)$.

Proof. It is enough to show that G contains a large subgraph H with $\delta(H) \geq (1-\frac{1}{r}+\frac{\varepsilon}{2})|H|$. To be precise, we find H with $|H| > S = \lfloor \varepsilon^{1/2} n \rfloor$. Then if $n_0 > 2\varepsilon^{-1/2} n_1(r,t,\frac{\varepsilon}{2})$, we have $|H| > n_1(r,t,\frac{\varepsilon}{2})$ and Lemma 3.8 shows $K_{r+1}(t) \subset H$. For technical reasons, we also require $\binom{s+1}{2} \geq 2n$ which is possible since the left-hand side is of order n^2.

Suppose no such H exists. Then we can construct a sequence of graphs $G = G_n \supset G_{n-1} \supset G_{n-2} \supset \cdots \supset G_s$ where $|G_j| = j$, and the vertex in G_j but not in G_{j-1} has degree at most $(1-\frac{1}{r}+\frac{\varepsilon}{2})j$ in G_j. Then

$$\begin{aligned}e(G_s) &\geq (1-\tfrac{1}{r}+\varepsilon)\binom{n}{2} - \sum_{j=s+1}^{n}(1-\tfrac{1}{r}+\tfrac{\varepsilon}{2})j\\ &= (1-\tfrac{1}{r}+\tfrac{\varepsilon}{2})\binom{s+1}{2} + \tfrac{\varepsilon}{2} - (1-\tfrac{1}{r}+\tfrac{\varepsilon}{2})n\\ &> \frac{\varepsilon n^2}{4} > \binom{s}{2}\end{aligned}$$

for sufficiently large n. This is a contradiction since $|G_s| = s$, so $e(G_s) \leq \binom{s}{2}$. □

Definition. The *chromatic number* $\chi(F)$ of a graph F is the smallest k such that F is k-partite.

Corollary 3.10.

$$\lim_{n\to\infty} \frac{\mathrm{ex}(n,F)}{\binom{n}{2}} = 1 - \frac{1}{\chi(F)-1}.$$

Proof. Let $r+1 = \chi(F)$. Then $F \not\subset T_r(n)$, so $\mathrm{ex}(n,F) \geq t_r(n)$, whence

$$\liminf_{n\to\infty} \frac{\mathrm{ex}(n,F)}{\binom{n}{2}} \geq \lim_{n\to\infty} \frac{t_r(n)}{\binom{n}{2}} = 1 - \frac{1}{r}.$$

Conversely, given $\varepsilon > 0$, if $|G| > n_0(r,|F|,\varepsilon)$ and $e(G) \geq (1-\frac{1}{r}+\varepsilon)\binom{|G|}{2}$ then, by Theorem 3.9, $G \supset K_{r+1}(|F|) \supset F$. Therefore,

$$\limsup_{n\to\infty} \frac{\mathrm{ex}(n,F)}{\binom{n}{2}} \leq (1-\tfrac{1}{r}+\varepsilon),$$

for all $\varepsilon > 0$. □

Chapter 4

Colouring

Definition. A *(vertex) colouring* of a graph G with k colours is a map $c : V(G) \to [k]$ such that $c(u) \neq c(v)$ if $uv \in E(G)$.

The chromatic number of G is the smallest k such that G can be coloured with k colours, denoted $\chi(G)$. Unlike the case $k = 2$, there is no nice characterisation of k colourable graphs for $k \geq 3$. Likewise, there is no known good way of finding $\chi(G)$.

The greedy algorithm runs through the vertices of a graph using some pre-arranged order. It assigns the least colour to a vertex that is not used on its already coloured neighbours.

$$c(v_j) = \min(\mathbb{N} - \{c(v_i) : i < j \ \wedge \ v_i v_j \in E(G)\}).$$

It is important to realise that the number of colours used depends on the ordering.

Theorem 4.1. $\chi(G) \leq 1 + \max_H \delta(H)$, the maximum taken over all subgraphs of G.

Proof. Let v_n be a vertex of minimum degree in G, let v_{n-1} be a vertex of minimum degree in $H_{n-1} := G[V(G) - \{v_n\}]$, let v_{n-2} be a vertex of minimum degree in $H_{n-2} := G[V(G) - \{v_n, v_{n-1}\}]$, etc.

Let $d = \max_j \delta(H_j)$, taking $H_n = G$. Then each v_j has at most d neighbours v_i with $i < j$. Then the greedy algorithm uses at most $1 + d$ colours when run on this ordering of the vertices.

It looks like $d \leq \max_H \delta(H)$, but in fact equality holds. For, given $H \subset G$, let $j = \max\{i : v_i \in H\}$. Then $H \subset H_j$, so $\delta(H) \leq \delta(H_j) \leq d$ since v_j is of minimum degree in H_j and $v_j \in H$. $\square$

Corollary 4.2. $\chi(G) \leq \Delta(G) + 1$.

Remark. Note that equality holds if G is complete. Also note that, in fact, the greedy algorithm never uses more than $1 + \Delta$ colours.

Clearly, $\chi(G)$ is equal to the maximum of the chromatic numbers of its components. Indeed, if $\kappa(G) = 1$ and x is a cutvertex, i.e. $\kappa(G - x) = 0$, then $\chi(G)$ is the maximum of the chromatic numbers of the pieces meeting at x.

We make the following observation. If G is connected and $v_n \in V(G)$, then we can order the remaining vertices s.t. each has at least one later neighbour, e.g. by decreasing distance from v_n. As a consequence, if G is connected and not regular, then $\chi(G) \leq \Delta(G)$. Taking any vertex v_n with $d(v_n) < \Delta$, we can apply the greedy algorithm with the above ordering.

Definition. A *block* of a graph G is a maximal 2-connected subgraph.

Recall that a bridge is an edge in no cycle, and note all other edges lie in blocks. G is a "tree" of blocks and bridges. In particular, two blocks pairwise intersect in at most one vertex and there exists at least two endblocks.

Theorem 4.3 (Brooks' theorem). Let G be a connected graph with $\chi(G) = \Delta(G) + 1$. Then G is complete or an odd cycle.

Proof. If $\Delta = 2$ then G is a path or a cycle, and the theorem is easily verified, so we may assume $\Delta \geq 3$. We assume G is connected, not complete, and we colour with at most Δ colours.

Also, if $\kappa(G) = 1$ then no block is $K_{\Delta+1}$ and since, by induction on the order of the graph, each block needs at most Δ colours, so does G. We shall find vertices v_1, v_2, v_n such that

(i) $v_1 v_2 \notin E(G)$;
(ii) $v_1, v_2 \in \Gamma(v_n)$;
(iii) $G - \{v_1, v_2\}$ is connected.

Given these vertices, pick v_{n-1} joined to v_n (there will be at least one v_j with $3 \leq j \leq n-1$ since $\Delta \geq 3$ implies $|G| \geq 4$), pick v_{n-2} joined to one of $\{v_{n-1}, v_n\}$, pick v_{n-3} joined to one of $\{v_{n-2}, v_{n-1}, v_n\}$ etc. We can do this because of (iii). We end up with $v_1, v_2, v_3, \ldots, v_n$, i.e. all vertices of G, where every vertex v_j for $3 \leq j \leq n-1$ has a neighbour v_i with $i > j$.

Let us use the greedy algorithm. Then $c(v_1) = c(v_2) = 1$ by (i). Also $c(v_j) \leq \Delta$ for $3 \leq j \leq n-1$ by the preceding observation. Also, since v_n has two neighbours of the same colour by (ii), we have $c(v_n) \leq \Delta$.

To find v_1, v_2, v_n if $\kappa(G) \geq 3$, take v_n of degree $\Delta(G)$ and since $G \neq K_{\Delta+1}$, we can take two non-adjacent neighbours v_1, v_2 of v_n. (Suppose all neighbours of v_n are adjacent, then $K_{\Delta+1} \subset G$, but G is connected and Δ is maximal, hence $G = K_{\Delta+1}$, contradiction.)

In the case $\kappa(G) = 2$, take v_n s.t. $\kappa(G - v_n) = 1$. Then every endblock of $G - v_n$ contains a non-cutvertex joined to v_n. Let v_1, v_2 be two such vertices in different endblocks. $\square$

Definition. The *clique number* of G is $\omega(G) = \max\{r : G \supset K_r\}$.

Definition. The *independence number* $\alpha(G)$ of G is the maximum size of any independent set, i.e. a set of vertices with no edges between them, so $\alpha(G) = \omega(\bar{G})$.

We have the following trivial lower bound for $\chi(G)$.

$$\max\left\{\omega(G), \frac{|G|}{\alpha(G)}\right\} \leq \chi(G).$$

Definition. Given a graph G, let $p_G(x)$ be defined as the number of ways to colour G with colours $1, 2, \ldots, x$.

Example. (i) If $\bar{K}_n$ is the complement of K_n then $p_{\bar{K}_n}(x) = x^n$, by choosing any of x colours for each of the n vertices.
(ii) For a tree T, $p_T(x) = x(x-1)^{n-1}$.

(iii) For a complete graph, we have $p_{K_n}(x) = x(x-1)(x-2)\cdots(x-n+1)$ for $x \geq n$ as we have x choices for the first vertex, $x-1$ choices for the second vertex, etc. If $1 \leq x \leq n-1$ then $p_{K_n}(x) = 0$.

Theorem 4.4. For all edges $e \in E(G)$, $p_G(x) = p_{G-e}(x) - p_{G/e}(x)$.

Proof. The colourings of G are these colourings of $G-e$ where the ends of e get different colours. But the colourings of $G-e$ where ends of e have the same colour are precisely colourings of G/e. □

Observe, for example, that $p_G(x) = \prod_C p_C(x)$ where C runs over the components of G. But nearly all other information about $p_G(x)$ is derived given Theorem 4.4 and by induction on $|E(G)|$. In particular, $p_G(x)$ is a polynomial called the chromatic polynomial. (Although this can be seen directly as well, see Example Sheet 3.)

Corollary 4.5. $p_G(x) = x^n - a_{n-1}x^{n-1} + \cdots + (-1)^n a_0$, where $n = |G|$, $a_{n-1} = e(G)$, $a_j \geq 0$ for all j, and $\min\{j : a_j \neq 0\} = k$ where k is the number of components of G.

Proof. This is left as an exercise via Theorem 4.4 and induction. □

Remark. Note that G is not specified by $p_G(x)$ up to isomorphism.

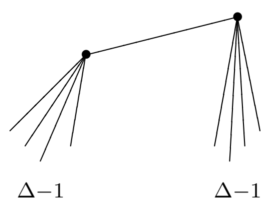
Definition. A *k-edge-colouring* of the graph G is a map $c : E(G) \to [k]$ where $c(e) \neq c(f)$ if the two edges e, f share an endvertex.

Remark. Note an edge-colouring of G is a vertex-colouring of the line graph $L(G)$. But edge-colourings enjoy special properties that merit attention.

Definition. The minimum number of colours needed to edge-colour G is the *chromatic index* denoted $\chi'(G)$.

Clearly $\chi'(G) \geq \Delta(G)$ and

$$\begin{aligned} \chi'(G) &\leq 1 + \Delta(L(G)) \\ &\leq 1 + (2\Delta(G) - 2) \\ &= 2\Delta(G) - 1. \end{aligned}$$

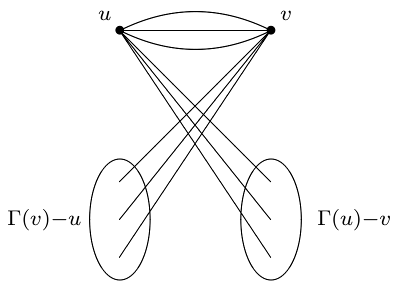


By Brooks' theorem 4.3, $\chi'(G) \leq 2\Delta(G) - 2$ if G is connected and not an odd cycle or an edge.

Theorem 4.6. Let G be a bipartite multigraph. Then $\chi'(G) = \Delta(G)$.

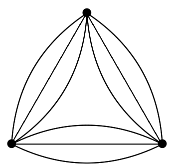
Proof. This can be proved by applying Hall's theorem (see Example Sheet 3). Here is a direct proof. First, observe we may assume G is Δ-regular. If not, replace G by the graph formed from G and a copy G' of G, joining $v \in G$ to $v' \in G'$ by $\Delta - d(v)$ edges.

We prove the theorem for Δ-regular multigraphs by induction on $|G| + \Delta(G)$. Note it is true if $|G| = 2$. It is also true if $\Delta = 0$, so we may assume $\Delta \geq 1$. Pick an edge uv of multiplicity $m \geq 1$. Make $G - \{u, v\}$ Δ-regular by adding $\Delta - m$ edges between $\Gamma(u)$ and $\Gamma(v)$, i.e. between $\Gamma(u) - v$ and $\Gamma(v) - u$.



Colour this multigraph with Δ colours. Some colour, red say, does not appear on a new edge. Thus the red edges together with uv (one copy) form a 1-factor in G. Colour one uv-edge red, then the red edges do not meet each other, but meet every vertex. Remove them from G and get a $(\Delta - 1)$-regular graph. Colour this by induction. □

Note Theorem 4.6 fails for non-bipartite graphs, e.g. K_3, and even more so for multigraphs. For example, the following graph has $\Delta = 6$ but $\chi' = 9$.

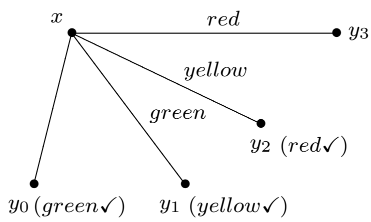


Theorem 4.7 (Vizing, 1965)**.** Let G be a graph. Then

$$\Delta(G) \leq \chi'(G) \leq \Delta(G) + 1.$$

This is clearly the best bound possible. For K_3, $\Delta = 2$, $\chi' = 3$. Note $\chi'(K_n) = n$ only for $n \geq 3$ odd. $\chi'(K_{2n}) = 2n - 1$ according to Bollobás.

Proof. It is enough to show that if G has a colouring with $\Delta + 1$ colours leaving one edge uncoloured, then it has a colouring with every edge coloured. Let xy_0 be the uncoloured edge. Note that at every vertex at least one colour is unused.



Construct a sequence of edges $xy_0, xy_1, \ldots, xy_k$ such that a colour c_i available at y_i is the colour of xy_{i+1}. Do this maximally with distinct c_i. We must stop either (i) because c_k does not appear at x or (ii) because $c_k = c_j$ for some $0 \leq j < k$.

In case (i), recolour xy_i with c_i, $0 \le i \le k$, and we are done.

In case (ii), recolour xy_i, $0 \le i < j$ and uncolour xy_j. Now let c be a colour not used at x and let H be the subgraph consisting of edges coloured c and c_k. We can swap c and c_k in any component of H and still have a proper colouring.

If x and y_j lie in different components of H, swap c and c_k in the component containing x. This frees c_k at x and leaves y_j unaffected, so colour xy_j with $c_k = c_j$, so we are done.

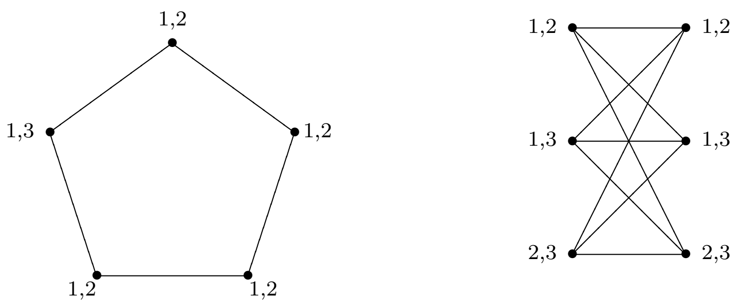
Thus we may assume x, y_j are in the same component of H. But $\Delta(H) \le 2$, so H consists of paths and cycles. But $d_H(x), d_H(y_j), d_H(y_k) \le 1$, so x, y_j, y_k lie at the ends of paths. Thus y_k is in a different component of H from x and y_j. Swap colours in the component of H at y_k. Recolour xy_i with c_i, $j \le i < k$, and colour xy_k with c. $\square$

Definition. A *list colouring* of a graph G is a colouring $c : V(G) \to \mathbb{N}$ (as usual $c(v) \neq c(u)$ if $uv \in E(G)$) with $c(v) \in L(v)$ where $L(v) \subset \mathbb{N}$ is a list of colours available at v.

For the usual colouring we have $L(v) = [k]$. Define

$$\chi_l(G) = \min\{k : \exists \text{ list colouring whenever } |L(v)| \ge k \ \forall v \in G\}.$$

Clearly $\chi_l(G) \ge \chi(G)$. In general, χ_l can be much bigger than χ, even for a bipartite graph G (e.g. $\chi_l(K_{3,3}) = 3$).



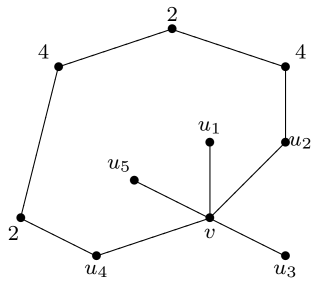
However, $\chi_l(G) \le 1 + \max_H \delta(H)$ holds by the previous proof.

We can define $\chi_l'(G)$ analogously for edge colourings. Amazingly, $\chi_l'(G) = \chi'(G) = \Delta(G)$ for bipartite graphs G (Galvin, 1994). It is conjectured $\chi_l'(G) = \chi'(G)$ for all graphs G.

Consider $\chi(G)$ for G planar. Since $e(G) \le 3|G| - 6$, it follows that $\delta(G) \le 5$. Also, if $H \subset G$ then H is planar, $\delta(H) \le 5$. Thus $\chi(G) \le 6$ by Theorem 4.1. We can improve this.

Theorem 4.8 (Heawood, 1890, "Five Colour Theorem")**.** Let G be a planar graph. Then $\chi(G) \le 5$.

Proof. Suppose otherwise and let G be a minimal counterexample, drawn in the plane. Let v be a vertex with $d(v) \le 5$. Colour $G - v$ with 5 colours. Then v must have a neighbour of each colour, or else we could colour v too; let u_i be the neighbour of v with colour i, $1 \le i \le 5$.



If u_2, u_4 are in different components of the 2/4 coloured subgraph, i.e. subgraph induced by vertices of colours 2/4, then swap 2 and 4 in the component at u_2. This makes u_2 colour 4 and leaves colours at u_i, $i \neq 2$, unchanged. Then colour v with 2, contradiction. So there exists a path coloured 2/4 from u_2 to u_4.

But likewise there is a 3/5 coloured path from u_3 to u_5, contradicting planarity. □

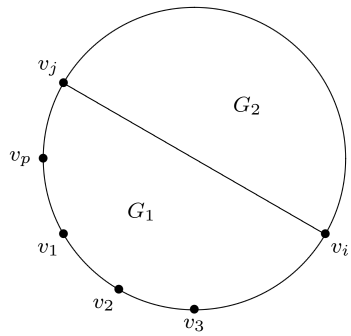
Remarkably, we can get more for less.

Theorem 4.9 (Thomasson, 1993)**.** $\chi_l(G) \le 5$ for planar G.

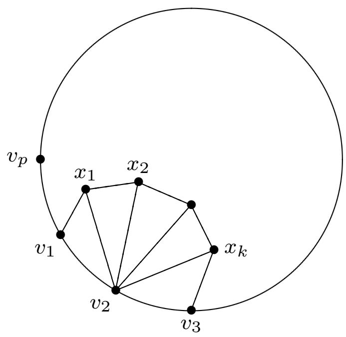
Proof. We prove the following by induction on $|G|$.

Let G have an outer cycle $v_1v_2 \dots v_p$ and have triangular faces inside. Let $|L(v_1)| = |L(v_p)| = 1$, $L(v_1) \neq L(v_p)$. Let $|L(v_i)| \ge 3$ for $2 \le i \le p-1$ and $|L(v)| \ge 5$ (†) elsewhere. Then G can be coloured.

If there is a chord v_iv_j, where v_i, v_j are not successive elements of the cycle, let G_1, G_2 be the subgraphs with boundaries $v_1v_2 \dots v_iv_j \dots v_p$ and $v_iv_{i+1} \dots v_j$, respectively. G_1 can be coloured by (†). Colours at v_i, v_j are now forced, so colour G_2 using (†).



If there is no chord, let $v_1x_1x_2 \dots x_kv_3$ be the neighbours of v_2.

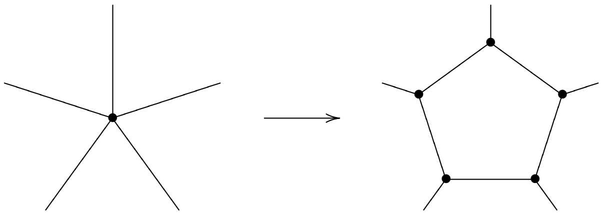


Let $G' = G - v_2$. Pick $i, j \in L(v_2) - L(v_1)$. Let $L'(x_m) = L(x_m) - \{i, j\}$. Colour G' using L' by (†). One of i or j is not used to colour v_3, so colour v_2 with it. $\square$

Remark. There exist (non-trivial) planar graphs G with $\chi_l(G) = 5$.

But the Four Colour Problem (Guthrie, 1852) is to show $\chi(G) \leq 4$ for planar graphs G. It was "proved" by Kempe in 1879 for which he was made FRS. In 1890 Heawood found a mistake. Kempe's proof was similar to the proof of Theorem 4.8, and the paths used are still known as "Kempe chains".

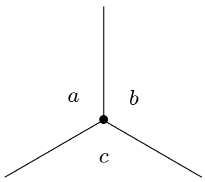
The Four Colour Problem is stated in dual form. The faces of any *map*, i.e. connected bridgeless plane multigraph, can be coloured with four colours so that no two contiguous faces have the same colour. Tait (1880) found a beautiful equivalent form of this. First observe that it is enough to colour cubic maps: see this either by triangulating the original graph or making the following replacement.



Theorem 4.10 (Tait, 1880)**.** The Four Colour Theorem holds if and only if $\chi'(G) = 3$ for every cubic bridgeless planar G.

Proof. We must show such a G is 4-face-colourable if and only if it is 3-edge-colourable. We use colours $\mathbb{Z}_2 \times \mathbb{Z}_2$ on faces, i.e. $\{00, 01, 10, 11\}$ with addition coordinatewise modulo 2, and the same without 00 on edges.

Suppose G is 4-face-coloured. Give an edge the colour which is the sum of adjacent face colours. Since G is bridgeless, 00 does not appear on an edge.



Note $a + c \neq b + c$ as $a \neq b$, so the edge colouring is proper, i.e. no two incident edges have the same colour.

Suppose conversely G is 3-edge-coloured. Pick a face F_0 and, for any other face F, walk from F_0 to F, adding up the colours of edges when crossed, and give the result to F. Note this gives adjacent faces different colours since 00 is on no edge. But we must check the colour of F is independent of the route chosen.

This is equivalent to checking that, if we go for a circular walk from F_0, returning to F_0, the sum of edges crossed is 00.

Consider the dual. It is a triangulated plane graph. Label each dual edge with the colour of the original edge that it crosses. We must show the labels around any cycle sum to 00. Since G was properly coloured, the edges on each triangular face are 01, 10 and 11. Hence, the sum around any face is 00. But if we have a cycle,

$$\sum_{\substack{\text{faces}\\\text{inside}\\\text{cycle}}} \sum \text{edges around face} \equiv \sum \text{edges around cycle} \pmod 2. \qquad \square$$

Remark. Tait further conjectured, and indeed believed he had proved, that every cubic plane bridgeless graph has a Hamiltonian cycle. Since a cubic graph has even order by the Handshaking Lemma, colour cycle edges red and blue, and colour the remaining edges green, i.e. $\chi'(G) = 3$.

But in 1946 Tutte found a counterexample of order 46 (see Example Sheet 3). The smallest counterexample has order 38.

Wagner (1935) proved that the Four Colour Theorem holds if and only if $\chi(G) \geq 5 \implies G \succ K_5$.

Haderinger (1943) conjectured $\chi(G) \geq k \implies G \succ K_k$. This is easy for $k = 4$. The only case known is $k = 6$ (equivalent to the Four Colour Theorem), $k \geq 7$ is unknown.

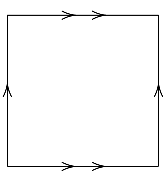
In 1976, Appel and Haken, using ideas of Heesch, announced a computer-based proof. Few people have read it. In 1997, Robertson, Sanders, Seymour, Thomas gave a new simpler proof based on the same ideas.

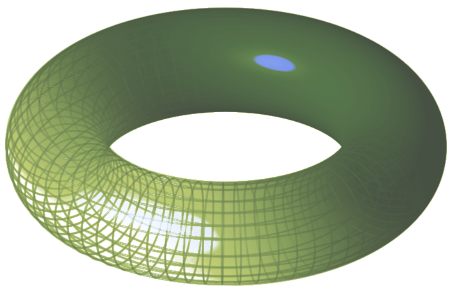
Graphs on other surfaces

Definition. A surface (2-dimensional, closed, compact) has an *Euler characteristic* $E \leq 2$ such that a graph drawn on this surface in such a way that each region, i.e. face, is simply connected satisfies $n - m + f = E$ where n is the order, m is the size and f is the number of faces.

An orientable surface, i.e. a surface with an inside and outside, is a sphere with some number $g \geq 0$ of handles. Note $E = 2 - 2g$.

(i) $g = 1$ yields the torus ($E = 0$).

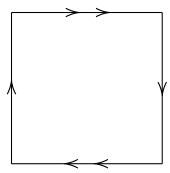




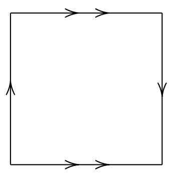
(ii) $g = 2$ yields the double torus ($E = -2$).

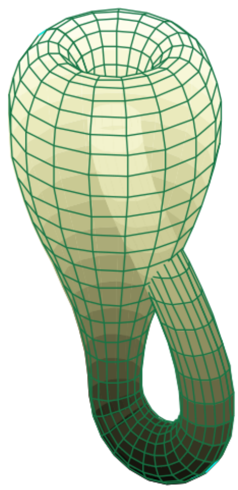
There are also non-orientable surfaces, one for each value of $E \leq 1$.

(i) The projective plane ($E = 1$).



(ii) The Klein bottle ($E = 0$).





If we have a maximal planar graph on a surface of characteristic E, then every face is a triangle, so $2m = 3f$ and $n - m + f = E$, so $m = 3(n - E)$. Hence every graph on the surface satisfies $m \leq 3(n - E)$.

Consider the projective plane. Then $m \leq 3n - 3$, so $\delta(G) \leq 5$ for every graph on this surface. By Theorem 4.1, $\chi(G) \leq 6$. We can draw K_6 on the projective plane (see Example Sheet 3), so in fact $\chi(G) = 6$ is attained.

Theorem 4.11 (Heawood, 1890)**.** If G is a graph drawn on a surface of characteristic $E \leq 1$, then

$$\chi(G) \leq H(E) = \left\lfloor \frac{7 + \sqrt{49 - 24E}}{2} \right\rfloor.$$

Proof. We already considered $E = 1$, so assume $E \leq 0$. Let G be a minimal graph on the surface having chromatic number k. Then $|G| \geq k$ and $\delta(G) \geq k - 1$, else we can remove a vertex and colour the graph with $k - 1$ colours. Hence

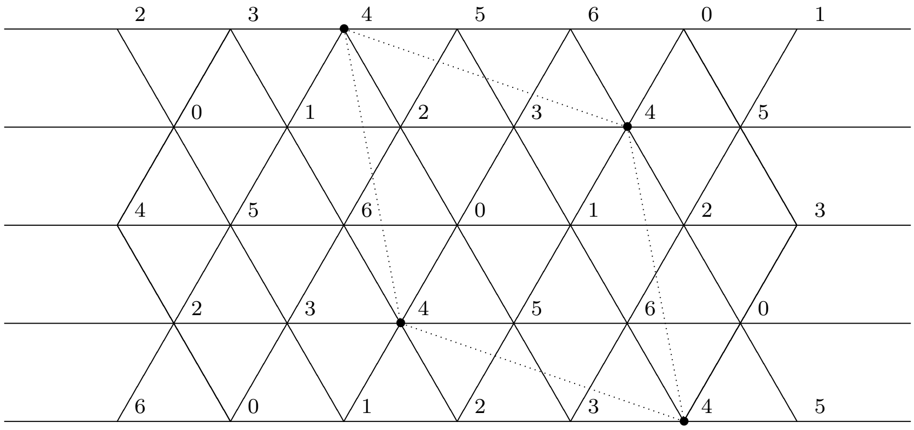
$$k - 1 \leq \delta(G) \leq \frac{2(3(|G| - E))}{|G|} = 6 - 6\frac{E}{k}$$

as $E \leq 0$, or $k^2 - 7k + 6E \leq 0$. □

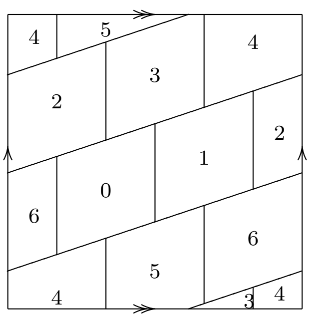
Remark. $H(2) = 4$.

Can equality hold? Consider $E = 0$, i.e. the torus or Klein bottle, when $H(0) = 7$. Let G be minimal of chromatic number 7, if equality is possible, then equality holds above. Note $\delta(G) \geq 6$ but $e(G) \leq 3|G|$ so G is 6-regular. By Brooks' theorem 4.3, $G = K_7$. Therefore, equality is attainable if and only if K_7 can be drawn on the surface.

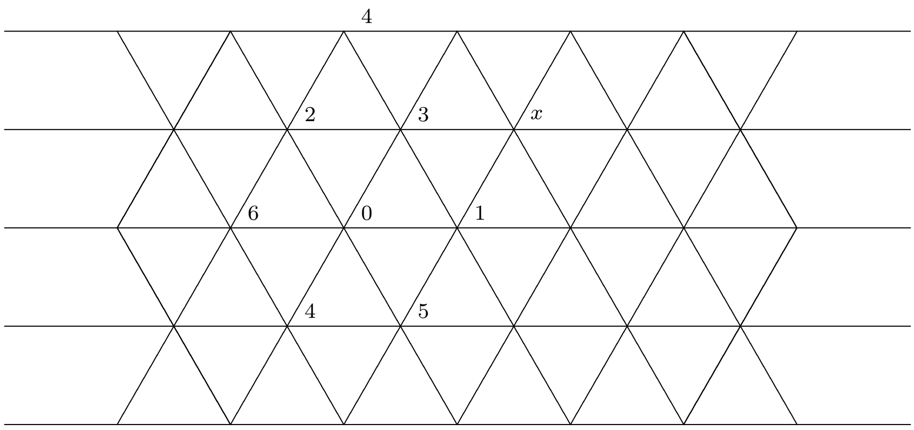
For the torus this is possible. Tile the plane using the shaded quadrangle forming a torus.



Dually, we have the following.



How about the Klein bottle? We must embed K_7 so that every face is a triangle, since $21 = m = 3(n - E)$. Locally, it looks like a planar triangulation.



Without loss of generality, fix 0 and numbers $1, \ldots, 6$ around it. x meets 1 and 3, so cannot be 2,0 or 5, so x is 4 or 6. By symmetry $x = 4$. By the same argument y is 4 or 5 but y is joined to 3 which joins $x = 4$, so $y = 5$. Hence, the pattern is now determined: it is the previous pattern. But when we identify all 1s, all 2s etc., we get a torus. Hence, K_7 can not be embedded on a Klein bottle.

Amazingly, the Klein bottle is the only exception to

$$\max\{\chi(G) : G \text{ embeds on surface } S \text{ of characteristic } E\} = H(E).$$

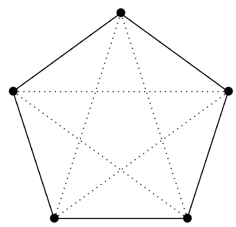
For $E \leq 1$, this is equivalent to embedding $K_{H(E)}$ on S. Heawood thought he had proved this; it was completed by Ringel and Youngs (1945).

Chapter 5

Ramsey Theory

The simplest example is the pigeon-hole principle.

Suppose the edges of K_6 are coloured red and blue. Then there is a monochromatic K_3. This fails for K_5.



Is it true if n is large enough and edges of K_n are coloured red and blue then there is a monochromatic K_4?

Definition. In general, if $s \in \mathbb{N}$ let $R(s)$ be the smallest n, if it exists, such that, if the edges of K_n are coloured red and blue, there is a monochromatic K_s.

For example, $R(3) = 6$ by the above. Also $R(2) = 2$, trivially.

Definition. It is convenient to define $R(s,t)$ to be the smallest n, if it exists, such that, if the edges of K_n are coloured red and blue, there must be a red K_s or a blue K_t.

Then $R(s) = R(s,s)$, $R(s,t) = R(t,s)$, $R(s,2) = s$, $R(\min\{s,t\}) \leq R(s,t) \leq R(\max\{s,t\})$.

Theorem 5.1 (Ramsey, 1930, Erdős–Szekares, 1935)**.** $R(s,t)$ exists. Moreover, $R(s,t) \leq R(s-1,t) + R(s,t-1)$.

Proof. It is enough to verify the inequality. Let $a = R(\text{s-1},t)$, $b = R(s,\text{t-1})$, $n = a+b$ and colour K_n red and blue. Pick a vertex v. Amongst the $a+b-1$ edges meeting v, there are either at least a red or b blue. We may assume the latter. So v is joined to a K_b by blue edges. Now $b = R(s,t-1)$, so this K_b either contains a red K_s, or it contains a blue K_{t-1} which, with v, forms a blue K_t. □

Remark. (i) The proof shows that $R(s,t) \leq \binom{s+t-2}{s-1}$.

(ii) This is not exact, e.g. equality can hold in the above proof only if there is an $(a-1)$-regular graph of order $a+b-1$ which is false if a,b both even.

(iii) We only know $R(2) = 2$, $R(3) = 6$, $R(4) = 18$, no other values are known.

Definition. Suppose we use $k \geq 2$ colours. Let $R_k(s_1, s_2, \ldots, s_k)$ be the smallest n, if it exists, such that, if edges of K_n are coloured $1, 2, \ldots, k$, there is a K_{s_i} coloured i, for some i.

Note $R(s,t) = R_2(s,t)$.

Theorem 5.2. Let $k \geq 2$, let $s_1, \ldots, s_k \in \mathbb{N}$. Then $R_k(s_1, \ldots, s_k)$ exists.

Proof. We can either generalise the proof for Theorem 5.1 or proceed as follows.

Use induction on k. The case $k = 2$ is Theorem 5.1. In general, let K_n be coloured with k colours $1, \ldots, k$. Go colourblind, so that colours 1 and 2 are indistinguishable. Suppose $n \geq R_{k-1}(R(s_1, s_2), s_3, \ldots, s_k)$. Either there is a K_{s_i} coloured i for some $i \geq 3$, whence we are done, or there is a $K_{R(s_1,s_2)}$ coloured $1/2$. In the latter case, be healed, so there is a K_{s_1} coloured 1 or a K_{s_2} coloured 2. □

The proof shows $R_k(s_1, \ldots, s_k) \leq R_{k-1}(R(s_1, s_2), s_3, \ldots, s_k)$. A modification of the proof of Theorem 5.1 gives

$$\begin{aligned} R_k(s_1, \ldots, s_k) \leq R_k(s_1 - 1, s_2, \ldots, s_k) + R_k(s_1, s_2 - 1, s_3, \ldots, s_k) + \cdots \\ + R_k(s_1, \ldots, s_{k-1}, s_k - 1) - k + 2. \end{aligned}$$

Note that the second bound is much better.

Definition. Given $r \in \mathbb{N}$, an *r-uniform hypergraph* is a pair (V, E) where $E \subset V^{(r)} = \{Y \subset V : |Y| = r\}$.

Example. (i) A graph is a 2-uniform hypergraph.

(ii) The complete r-uniform hypergraph of order n is $K_n^{(r)}$, has vertex set $[n]$ and $E = V^{(r)}$. So $|E| = \binom{n}{r}$.

Does there exist a smallest number $n = R^{(r)}(s,t)$ such that if we colour the edges of $K_n^{(r)}$ red and blue then there is a red $K_s^{(r)}$ or a blue $K_t^{(r)}$? For example, $R^{(2)}(s,t) = R(s,t)$ and $R^{(1)}(s,t) = s + t - 1$, which again is the pigeon-hole principle.

Theorem 5.3 (Ramsey for r-sets)**.** $R^{(r)}(s,t)$ exists for all $r \geq 1$, $s, t \geq r$.

Proof. We show that if $a = R^{(r)}(s-1, t)$, $b = R^{(r)}(s, t-1)$ and $n = 1 + R^{(r-1)}(a, b)$ exist then $R^{(r)}(s,t) \leq n$. Let the edges of $K_n^{(r)}$ be coloured red and blue and let $v \in V(K_n^r)$. Consider the $K_{n-1}^{(r-1)}$ with vertex set $V(K_n^{(r)}) - \{v\}$. Given an edge Z of $K_{n-1}^{(r-1)}$, colour it in the same colour as we gave to $Z \cup \{v\}$. Since $n - 1 = R^{(r-1)}(a, b)$ this means we get a red $K_a^{(r-1)}$ or a blue $K_b^{(r-1)}$, without loss of generality assume the latter. Hence in the original colouring there are b vertices such that every r-set formed from v and $r - 1$ of these b vertices is blue. Since $b = R^{(r)}(s, t-1)$, these vertices either contain a red $K_s^{(r)}$ or else a blue K_{t-1}^r which with v forms a blue $K_t^{(r)}$. □

Colour blindness shows $R_k^{(r)}(s_1, s_2, \ldots, s_k)$ exists with the natural definition. The bounds obtained on $R^{(r)}(s,t)$ are huge.

What if we colour infinite graphs? If we colour $\mathbb{N}^{(2)}$ red and blue, do we get an infinite monochromatic set M, i.e. an infinite $M \subset \mathbb{N}$ with $M^{(2)}$ same colour?

(i) Colour $\{i,j\}$ red if $i+j$ is odd. Then M the set of even numbers works.
(ii) Colour $\{i,j\}$ red if $\max\{k : 2^k \mid i+j\}$ is odd. Then $M = \{4^l : l \geq 1\}$ works.
(iii) Colour $\{i,j\}$ red if $i+j$ has an odd number of distinct prime factors. No explicit set M is known.

Theorem 5.4 (Infinite Ramsey)**.** Let $\mathbb{N}^{(r)}$ be coloured with k colours. Then there is an infinite $M \subset \mathbb{N}$ such that $M^{(r)}$ is monochromatic.

Proof. By induction on r. The case $r = 1$ is just the pigeon-hole principle. We shall find a sequence $v_1 < v_2 < v_3 < \cdots$ of elements of $\mathbb{N}$ and a sequence of colours $c_1, c_2, c_3, \ldots$ such that if $Y \subset \{v_1, v_2, v_3, \ldots\}^{(r)}$ and $\min Y = v_i$ then Y has colour c_i.

Given such a sequence we are home, because there is a subsequence $v_{i_1}, v_{i_2}, v_{i_3}, \ldots$ for which $c_{i_1}, c_{i_2}, c_{i_3}, \ldots$ is constant and then we take $M = \{v_{i_1}, v_{i_2}, v_{i_3}, \ldots\}$.

To construct these sequences, suppose we so far found $v_1 < v_2 < \cdots < v_j$ and colours $c_1, c_2, \ldots, c_j$ and an infinite set $N_j \subset \{n \in \mathbb{N} : n > v_j\}$ such that if Y is an r-subset of $\{v_1, \ldots, v_j\} \cup N_j$ and $\min Y = v_i$ then Y has colour c_i. Starting with $N_0 = \mathbb{N}$ it suffices to show we can get from j to $j+1$. Let $v_{j+1} = \min N_j$. For each subset $Z \subset (N_j - \{v_j\})^{(r-1)}$, colour it the same colour as $\{v_{j+1}\} \cup Z$. By Ramsey for $r-1$, there is an infinite subset $N_{j+1} \subset N_j - \{v_{j+1}\}$ where every Z has the same colour c_{j+1}. □

Definition. A set $Y \subset \mathbb{N}$ is big if $|Y| \geq \min Y$.

Example. So $\{3, 9, 24\}$ is big, $\{5, 37, 36, 209\}$ is not.

Theorem 5.5. Given $s, k, r \geq 1$ there exists $B(s,k,r)$ such that if $n \geq B(s,k,r)$ and $\{s, s+1, \ldots, n\}^{(r)}$ are coloured with k colours then there is a big subset $Y \subset \{s, s+1, \ldots, n\}$ with $Y^{(r)}$ monochromatic.

Proof. Trivially, if the assertion holds for some N, it will hold for all $n \geq N$. So if the theorem is false, the assertion fails for all n.

Suppose not. Then for each $n \geq s$ define a colouring

$$c_n : \{s, s+1, \ldots, n\}^{(r)} \to [k]$$

with no monochromatic big subset. Enumerate $\{s, s+1, \ldots, n\}^{(r)}$ as $Z_1, Z_2, \ldots$. We can pick an infinite subsequence $c_1^{(1)}, c_2^{(1)}, \ldots$ of $c_s, c_{s+1}, \ldots$ on which the colour of Z_1 is constant, call it $c(Z_1)$. Then pick an infinite subsequence $c_1^{(2)}, c_2^{(2)}, \ldots$ on which the colour of Z_2 is constant, $c(Z_2)$ say. Repeat, picking a subsequence $c_1^{(j)}, c_2^{(j)}, \ldots$ of $c_1^{(j-1)}, c_2^{(j-1)}, \ldots$ on which the colour of Z_j is $c(Z_j)$. Hence we get a colouring $c : \{s, s+1, \ldots, n\}^{(r)} \to [k]$ with no monochromatic big subset: for if c were constant on $Y^{(r)}$, let $l = \max\{j : Z_j \in Y^{(r)}\}$, then c agrees with $c_1^{(l)}$ on $Y^{(r)}$ and $c_1^{(l)} = c_n$, some n, which has no big monochromatic Y.

But by Ramsey's Infinite Theorem there is an infinite $M \subset \{s, s+1, \ldots\}$ such that c is monochromatic on $M^{(r)}$. Now let $m = \min M$ and let Y be the first m elements of M. Then Y is big and c is monochromatic on $Y^{(r)}$, contradiction. □

Remark. Theorem 5.5 is a purely finite statement, though we proved it via an infinite argument, actually using compactness. Amazingly, this theorem cannot be proved without recourse to such argument (Paris, Harrington 1977). This is the first natural example of Gödel's Incompleteness Theorem.

Other structures support Ramsey-like theorems. For example, van der Waerden's theorem states if we colour the integers $1, 2, \ldots, n$ with k colours there exists an l-term arithmetic progression in one colour, for sufficiently large n.

The following Ramsey numbers are known, with $R(3,9), R^{(3)}(4,4,4), R(4,5)$ having been found in 1982, 1991, 1992, respectively.

$$\begin{array}{cccc}
R(3,3) = 6 & R(3,4) = 9 & R(3,5) = 14 & R(3,6) = 18 \\
R(3,7) = 23 & R(3,8) = 28 & R(3,9) = 36 & R(4,4) = 18 \\
R(4,5) = 25 & R_3(3,3,3) = 17 & R^{(3)}(4,4,4) = 13 &
\end{array}$$

The following bounds have been established.

- (Upper bounds) We have

$$R(s) \leq \binom{2s-2}{s-1} \sim \frac{2^{2s-2}}{\sqrt{2\pi s}},$$

  and the only improvement made in the last 70 years is an extra factor of $\frac{1}{\sqrt{s}}$.
- (Lower bounds) Trivially,

$$R(s) > (s-1)^2.$$

  Consider the Turán graph $T_{s-1}((s-1)^2)$. There is no K_s, so in particular no red K_s, and no s vertices with no edges between them, so no blue K_s. The bound

$$R(s) > s^3$$

  is more tricky. The best known construction gives

$$R(s) > e^{\frac{\log^2 s}{4 \log \log s}}.$$

  However, this is still basically zero compared to the upper bound.

Chapter 6

Probabilistic Methods

Theorem 6.1 (Erdős, 1947)**.** Let $s \geq 3$. Then $R(s) \geq 2^{(s-1)/2}$.

Proof. Colour the edges of K_n independently red and blue with probability $\frac{1}{2}$. There are $\binom{n}{s}$ K_s's in K_n. Each is monochromatic with probability $2\left(\frac{1}{2}\right)^{\binom{s}{2}}$. So the expected number of monochromatic K_s is $\binom{n}{s}2^{1-\binom{s}{2}}$. But if $n = \lfloor 2^{(s-1)/2} \rfloor$, then

$$\begin{aligned}\binom{n}{s}2^{1-\binom{s}{2}} &\leq \frac{2}{s!}n^s 2^{-\binom{s}{2}} \\ &< (n2^{-(s-1)/2})^s \leq 1.\end{aligned}$$

But the expected number can be less than 1 only if there is a colouring with no monochromatic K_s. □

We could derandomise this argument by observing that all we did was to prove the average number of monochromatic K_s over all colourings of K_n is less than 1. But this would be to lose the potential of the pobabilistic approach.

Recall the following from *Probability.* Let Ω be a finite probability space, e.g. in the above argument Ω is the set of all $2^{\binom{n}{2}}$ red and blue colourings of K_n, with uniform distribution. An event is a subset $A \subset \Omega$. A random variable is a function $X : \Omega \to \mathbb{R}$. The expectation of X is $\mathbb{E}\,X = \sum_{\omega\in\Omega} X(\omega)\,\mathbb{P}(\omega)$. It is important to note that expectation is linear.

$$\begin{aligned}\mathbb{E}(X+Y) &= \sum_{\omega\in\Omega}(X(\omega)+Y(\omega))\,\mathbb{P}(\omega) \\ &= \sum_{\omega\in\Omega} X(\omega)\,\mathbb{P}(\omega) + \sum_{\omega\in\Omega} Y(\omega)\,\mathbb{P}(\omega) \\ &= \mathbb{E}\,X + \mathbb{E}\,Y.\end{aligned}$$

The indicator function of an event A is $I_A : \Omega \to \{0,1\}$,

$$I_A(\omega) = \begin{cases} 0 & \text{if } \omega \notin A \\ 1 & \text{if } \omega \in A \end{cases}.$$

Note that $\mathbb{E}\,I_A = \sum_{\omega\in\Omega} I_A(\omega)\,\mathbb{P}(\omega) = \sum_{\omega\in A} \mathbb{P}(\omega) = \mathbb{P}(A)$. Variables X that count can often be written as sums of indicators, e.g. if $X(\omega)$ is the number of monochromatic K_s

in the above proof, then for each $\alpha \in [n]^{(s)}$, let I_α indicate that K_s on the vertex set α is monochromatic. Then $X = \sum_\alpha I_\alpha$, so

$$\begin{aligned}\mathbb{E}\,X &= \sum_\alpha \mathbb{E}\,I_\alpha = \sum_\alpha \mathbb{P}(K_s \text{ on } \alpha \text{ is monochromatic})\\ &= \sum_\alpha 2^{1-\binom{s}{2}} = \binom{n}{s} 2^{1-\binom{s}{2}}.\end{aligned}$$

If X is a counting variable, $X = \sum_\beta I_\beta$ then $\mathbb{E}\,X = \sum_\beta \mathbb{E}\,I_\beta = \sum_\beta \mathbb{P}(I_\beta = 1)$.

Theorem 6.2. The graph G has an independent set, i.e. an edgeless set, of size at least

$$\sum_{v \in G} \tfrac{1}{d(v)+1} \geq \tfrac{|G|}{d+1},$$

where d is the average degree, i.e. $e(G) = |G|d/2$.

Proof. Take a random ordering of the vertices, use the greedy colouring algorithm, and let X be the number of vertices coloured 1. Then $X = \sum_{v \in G} I_v$ where I_v is the indicator that v gets colour 1. Then

$$\mathbb{P}(I_v = 1) \geq \mathbb{P}(v \text{ precedes all its neighbours}) = \frac{1}{d(v)+1}.$$

Thus

$$\mathbb{E}\,X = \sum_{v \in G} \mathbb{P}(I_v = 1) \geq \sum_{v \in G} \frac{1}{d(v)+1}.$$

So there is some ordering where

$$X \geq \sum_{v \in G} \frac{1}{d(v)+1}.$$

Since $\frac{1}{X-\varepsilon} + \frac{1}{X+\varepsilon} \geq \frac{2}{X}$, we get

$$\sum_{v \in G} \frac{1}{d(v)+1} \geq \frac{|G|}{d+1}. \qquad \square$$

Remark. This is equivalent to Turán's theorem when applied to the complement. The full theorem can be recovered by examining cases of equality.

We keep using $\mathbb{P}(X > \mathbb{E}\,X) < 1$, which is a special case of Markov's inequality. Since $I_{\{|X|>t\}} < \frac{|X|}{t}$, take expectations to obtain

$$\mathbb{P}(|X| > t) < \frac{\mathbb{E}|X|}{t}.$$

Definition. Let $\mathcal{G}(n,p)$ be the space of all $2^{\binom{n}{2}}$ labelled graphs with vertex set $[n]$, where a given graph with m edges has probability $p^m(1-p)^{\binom{n}{2}-m}$. This is equivalent to inserting edges independently with probability p.

In our earlier Ramsey example the red graph was an element of $\mathcal{G}(n, \frac{1}{2})$. We showed the clique size, i.e. the size of a maximum complete subgraph, is at most $2\log_2 n + 1$ and the independent set size is at most $2\log_2 n + 1$, in one graph at least. Hence for such a graph G,

$$\chi(G) \geq \frac{n}{2\log_2 n + 1}.$$

This shows that the chromatic number can be far larger than the clique number, a perhaps counterintuitive fact. In fact, there are constructions of triangle-free graphs with large chromatic number. (See Example Sheet 3.)

Theorem 6.3 (Erdős, 1959)**.** *Let $k, g \in \mathbb{N}$ with $g \geq 3, k \geq 2$. Then there exists a graph G with $\chi(G) \geq k$ and girth at least g.*

Proof. Consider a random graph $G \in \mathcal{G}(n,p)$ where $p = n^{-1+1/g}$. Let X_l be the number of cycles of length l and let $X = X_3 + X_4 + \cdots + X_{g-1}$. Then

$$\begin{aligned} \mathbb{E}\, X &= \sum_{l=3}^{g-1} \mathbb{E}\, X_l \leq \sum_{l=3}^{g-1} n^l p^l = \sum_{l=3}^{g-1} n^{\frac{l}{g}} \\ &\leq g n^{\frac{g-1}{g}} < \frac{n}{4} \end{aligned}$$

if n is sufficiently large. Let A be the event $\{X > \frac{n}{2}\}$, then by Markov $\mathbb{P}(A) < \frac{1}{2}$ if n is sufficiently large.

Let Y be the number of independent sets of size $t = \lceil \frac{n}{2k} \rceil$. Then

$$\begin{aligned} \mathbb{E}\, Y &= \binom{n}{t}(1-p)^{\binom{t}{2}} \\ &\leq n^t (e^{-p})^{\binom{t}{2}} \\ &\leq n^t e^{(-p\frac{n^2}{9k^2})} \\ &= \exp\left\{t \log n - p\frac{n^2}{9k^2}\right\} \\ &= \exp\left\{t \log n - \frac{n^{1+\frac{1}{g}}}{9k^2}\right\} \\ &< \frac{1}{2} \end{aligned}$$

if n is sufficiently large, using approximations which are explained separately later. Let B be the event $Y \geq 1$, then by Markov $\mathbb{P}(B) < \frac{1}{2}$ for sufficiently large n.

Since $\mathbb{P}(A \cup B) < 1$ for sufficiently large n, there exists a graph G where neither A nor B happens, so it has at most $\frac{n}{2}$ short cycles and no independent set of size $\lceil \frac{n}{2k} \rceil$. Form G' by removing a vertex from each short cycle. Then G' has girth at least g and $|G'| \geq \frac{n}{2}$. Since G' has no independent set of size at least $\frac{n}{2k}$, $\chi(G') \geq \frac{|G|}{n/2k} \geq k$. □

We have made the following approximations.

(i) $1 - p \leq e^{-p}$ true for all p.

(ii) Let $t = \left\lceil \frac{n}{2k} \right\rceil$, then $\binom{t}{2} = \frac{n^2}{8k^2} + \mathcal{O}(n) > \frac{n^2}{9k^2}$ for sufficiently large n.

(iii) For fixed L,

$$\frac{e^n}{n^L} > \frac{n}{(L+1)!} \to \infty \text{ as } n \to \infty.$$
$$n = \log m, \ \frac{m}{(\log m)^L} \to \infty \text{ as } m \to \infty.$$
$$\frac{m^\varepsilon}{(\log m)^L} = \left[\frac{m}{(\log m)^{\frac{L}{\varepsilon}}} \right]^\varepsilon \to \infty.$$

Recall that $z(n,t) = \mathcal{O}(n^{2-1/t})$. To get a lower bound we would consider random $n \times n$ bipartite graphs with p as large as possible, so the expected number of $K_{t,t}$ is less than 1. The expected number of $K_{t,t}$'s is $\binom{n}{t}^2 p^{t^2}$, so $p \approx n^{-2/t}$, giving an expected size of $pn^2 \approx n^{2-2/t}$. We can do better by making use of the linearity of expectation.

Theorem 6.4. Let $n \geq t \geq 2$. Then $z(n,t) > \frac{3}{4} n^{2-2/(t+1)}$.

Proof. Let X, Y be classes of n vertices and generate random $n \times n$ bipartite graphs by inserting edges independently with probability p. Let J be the size of this graph and K the number of $K_{t,t}$'s. By throwing out an edge from each $K_{t,t}$ we get a graph with no $K_{t,t}$, so $z(n,t) > J - K$. Now $\mathbb{E}\, J = pn^2$ and $\mathbb{E}\, K = \binom{n}{t}^2 p^{t^2}$, so

$$\begin{aligned} \mathbb{E}(J-K) &= pn^2 - \binom{n}{t}^2 p^{t^2} \\ &\geq pn^2 - \tfrac{1}{4} n^{2t} p^{t^2}. \end{aligned}$$

Taking $p = n^{-2/(t+1)}$, we obtain $\mathbb{E}(J-K) \geq \frac{3}{4} n^{2-2/(t+1)}$, so there is a graph with $J - K \geq \frac{3}{4} n^{2-2/(t+1)}$. □

So far we used Markov's inequality to show that some random variable is unlikely to be large. What if we want to say a variable is likely to be large? The variance of X is $\mathbb{E}(X - \mathbb{E}\, X)^2$. Then by Markov's inequality,

$$\begin{aligned} \mathbb{P}(|X - \mathbb{E}\, X| > t) &= \mathbb{P}((X - \mathbb{E}\, X)^2 > t^2) \\ &< \frac{\mathbb{E}(X - \mathbb{E}\, X)^2}{t^2} = \frac{\operatorname{Var}(X)}{t^2} \end{aligned}$$

which is Chebychev's inequality.

Lemma 6.5. Let $X_n \colon \Omega_n \to \mathbb{R}$ be a sequence of random variables. Suppose that $\mathbb{E}(X_n^2)/(\mathbb{E}\, X_n)^2 \to 1$ as $n \to \infty$, or equivalently $\operatorname{Var}(X_n)/(\mathbb{E}\, X_n)^2 \to 0$. Then for any constant $c > 0$ we have

$$\mathbb{P}(|X_n - \mathbb{E}\, X_n| \geq c\, \mathbb{E}\, X_n) \to 0$$

as $n \to \infty$. In particular, $\mathbb{P}(X_n = 0) \to 0$.

Proof. Apply Chebyshev's inequality with $t = c\, \mathbb{E}\, X_n$. For the second part, take $c = 1$. □

Note if X is counting,

$$\begin{aligned}\mathbb{P}(X=0) &\leq \mathbb{P}(|X-\mathbb{E}\,X| \geq \mathbb{E}\,X) \\ &\leq \mathbb{P}(|X-\mathbb{E}\,X| > t) \qquad \text{for all } t < \mathbb{E}\,X \\ &< \frac{\operatorname{Var}(X)}{t^2}.\end{aligned}$$

So $\mathbb{P}(X=0) \leq \operatorname{Var}(X)/(\mathbb{E}\,X)^2$. We compute $\operatorname{Var}(X)$ as follows.

$$\operatorname{Var}(X) = \mathbb{E}(X-\mathbb{E}\,X)^2 = \mathbb{E}(X^2 - 2X\,\mathbb{E}\,X + (\mathbb{E}\,X)^2) = \mathbb{E}\,X^2 - (\mathbb{E}\,X)^2.$$

Suppose X is a sum of indicators, $X = \sum_A I_A$. Then

$$\mathbb{E}\,X = \sum_A \mathbb{E}\,I_A = \sum_A \mathbb{P}(A),$$

so

$$(\mathbb{E}\,X)^2 = \Big(\sum_A \mathbb{P}(A)\Big)^2 = \sum_{A,B} \mathbb{P}(A)\,\mathbb{P}(B).$$

Now

$$X^2 = \Big(\sum_A I_A\Big)^2 = \sum_{A,B} I_A I_B = \sum_{A,B} I_{A\cap B}.$$

So

$$\mathbb{E}\,X^2 = \sum_{A,B} \mathbb{E}\,I_{A\cap B} = \sum_{A,B} \mathbb{P}(A\cap B) = \sum_{A,B} \mathbb{P}(A)\,\mathbb{P}(B|A).$$

So

$$\operatorname{Var}(X) = \sum_{A,B} \mathbb{P}(A)\left[\mathbb{P}(B|A) - \mathbb{P}(B)\right].$$

Note there is no contribution from a pair A, B that is independent.

Theorem 6.6. Let $\omega(n) \to \infty$. Let $G \in \mathcal{G}(n,p)$. If $p = \frac{\log n - \omega(n)}{n}$ then G has isolated vertices almost surely. If $p = \frac{\log n + \omega(n)}{n}$, then G has no isolated vertices almost surely.

Definition. Here an event A happens *almost surely* if $\mathbb{P}(A) \to 1$ as $n \to \infty$.

Proof. Let X be the number of isolated vertices. Then $X = \sum_v I_v$ where I_v indicates whether v is isolated. So $\mathbb{E}\,X = n(1-p)^{n-1}$. Also

$$\begin{aligned}\operatorname{Var}(X) &= \sum_{u,v} \mathbb{P}(u \text{ isolated})\left[\mathbb{P}(v \text{ isolated}|u \text{ isolated}) - \mathbb{P}(v \text{ isolated})\right] \\ &= n(1-p)^{n-1}\left[1-(1-p)^{n-1}\right] \\ &\quad + n(n-1)(1-p)^{n-1}\left[(1-p)^{n-2}-(1-p)^{n-1}\right] \\ &\leq \mathbb{E}\,X + n^2(1-p)^{n-1}p(1-p)^{n-2}\end{aligned}$$

$$= \mathbb{E}\,X + \tfrac{p}{1-p}(\mathbb{E}\,X)^2.$$

If $p = \frac{\log n + \omega(n)}{n}$ then

$$\begin{aligned}\mathbb{E}\,X &= \tfrac{1}{1-p} n(1-p)^n \\ &\le \tfrac{1}{1-p} n e^{-pn} \to 0,\end{aligned}$$

so $X = 0$ almost surely by Markov's inequality. If $p = \frac{\log n - \omega(n)}{n}$ then

$$\mathbb{E}\,X \approx \frac{1}{1-p} n e^{-pn} \to \infty.$$

(Here we could also use that $1 - p \ge e^{-p-p^2}$ if $p < \frac{1}{4}$). So

$$\frac{\operatorname{Var} X}{(\mathbb{E}\,X)^2} \le \frac{1}{\mathbb{E}\,X} + \frac{p}{1-p} \to 0,$$

so $X \ne 0$ almost surely by Chebyshev's inequality. □

Remark. This is a typical *threshold phenomenon*.

As a final example, consider $K_d \subset G \in \mathcal{G}(n,p)$. Let X_d be the number of K_d's in $G \in \mathcal{G}(n,p)$, then $\mathbb{E}\,X_d = \binom{n}{d} p^{\binom{d}{2}} = \mu(d)$, say. Now

$$\mu(d) \approx \frac{[np^{(d-1)/2}]^d}{d!}$$

so if $d > (2+\varepsilon)\log_{1/p} n$ then $\mu(d) \to 0$, whereas if $d < (2-\varepsilon)\log_{1/p} n$ then $\mu(d) \to \infty$.

Theorem 6.7. Let $0 < p < 1$ be fixed. Let $\mu(d) = \binom{n}{d} p^{\binom{d}{2}}$.

(i) If $\mu(d) \to 0$ then $G \in \mathcal{G}(n,p)$ almost certainly does not contain a K_d.

(ii) If $\mu(d) \to \infty$ then $G \in \mathcal{G}(n,p)$ almost certainly contains a K_d.

Proof. The first assertion follows from Markov's inequality and for the second we need only show $\operatorname{Var}(X)/(\mathbb{E}\,X)^2 \to 0$ by Chebychev's inequality.

Now let $X_d = \sum_\alpha I_\alpha$, where α runs over $[n]^{(d)}$ and I_α indicates $G[\alpha]$ is complete. We know

$$\operatorname{Var} X = \sum_{\alpha,\beta} \mathbb{P}(I_\alpha = 1)\left[\mathbb{P}(I_\beta = 1 | I_\alpha = 1) - \mathbb{P}(I_\beta = 1)\right].$$

Now I_α and I_β are independent if $|\alpha \cap \beta| \le 1$, because there is no common edge, and in general $\mathbb{P}(I_\beta = 1|I_\alpha = 1)$ depends only on the value of $l = |\alpha \cap \beta|$. Hence

$$\operatorname{Var} X = \binom{n}{d} p^{\binom{d}{2}} \sum_{l=2}^{d} \binom{d}{l}\binom{n-d}{d-l}\left[p^{\binom{d}{2}-\binom{l}{2}} - p^{\binom{d}{2}}\right]$$

where $\binom{d}{l}\binom{n-d}{d-l}$ is the number of β's meeting α in exactly l vertices. Ignoring the final subtracted term, we have

$$\frac{\operatorname{Var} X}{(\mathbb{E}\,X)^2} \le \frac{1}{\mu(d)} \sum_{l=2}^{d} \binom{d}{l}\binom{n-d}{d-l} p^{\binom{d}{2}-\binom{l}{2}}$$

$$
\begin{aligned}
&= \frac{1}{\binom{n}{d}} \sum_{l=2}^{d} \binom{d}{l} \binom{n-d}{d-l} p^{-\binom{l}{2}} \\
&\leq \sum_{l=2}^{d} \frac{\binom{d}{l}\binom{n-l}{d-l}}{\binom{n}{d}} p^{-\binom{l}{2}} \\
&= \sum_{l=2}^{d} \frac{\binom{d}{l}^2}{\binom{n}{l}} p^{-\binom{l}{2}} =: \sum_{l=2}^{d} a_l
\end{aligned}
$$

using $\binom{n}{d}\binom{d}{l} = \binom{n-l}{d-l}\binom{n}{l}$. By looking at the ratios $\frac{a_{l+1}}{a_l}$, we see a_l decreases then increases and $a_l \leq \max\{a_2, a_d\}$ and indeed $\sum a_l \leq da_2 + Ca_d$ where C is some constant. Since

$$da_2 \leq \tfrac{2d^5}{n^2}\tfrac{1}{p} \to 0$$

because $d \leq (2+\varepsilon)\log_{1/p} n$ when $\mu(d) \to \infty$, and $a_d = \frac{1}{\mu(d)} \to 0$, we are done. $\square$

Remark.

$$\frac{a_{l+1}}{a_l} = \frac{(d-l)^2}{(l+1)(n-l)} \left(\tfrac{1}{p}\right)^l.$$

We may assume $d \approx 2\log_{1/p} n$. If $l = o(\log n)$ then $\frac{a_{l+1}}{a_l} < 1$. For larger l, choose $1 < R < \frac{1}{p}$ and C' so if $d - l < C'$ then $\frac{a_{l+2}}{a_{l+1}} > R\frac{a_{l+1}}{a_l}$. Therefore

$$\sum_{l=2}^{d} a_l \leq da_2 + (\tfrac{1}{R} + C')a_d.$$

Corollary 6.8. Let $0 < p < 1$ be fixed. Then the clique number, i.e. the size of the largest complete subgraph, of $G \in \mathcal{G}(n,p)$ is $(2+o(1))\log_{1/p} n$ almost surely.

Remark. With more care, we can compute the clique number to within 1.

Chapter 7

Eigenvalue Methods

Definition. Let G be a graph with vertex set $[n]$. The *adjacency matrix* of G is the $n \times n$ matrix $A = (a_{ij})$,

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

Since A is real symmetric, it is diagonalisable and has n real eigenvalues, which are easy to compute and give useful information about G. For example, $(A^2)_{ij} = \sum_k a_{ik}a_{kj}$ is the number of walks of length 2 from i to j. In general, $(A^d)_{ij}$ is the number of walks of length d from i to j.

Definition. If G is a connected graph then the *diameter* $\operatorname{diam}(G)$ is the maximum distance between any two vertices.

The set $\{I, A, A^2, \dots, A^{\operatorname{diam}(G)}\}$ is linearly independent, so in particular G has at least $\operatorname{diam}(G) + 1$ distinct eigenvalues. (Note that changing the labelling of the vertices of G is effectively just a change of basis and does not change the eigenvalues.)

Since A is real symmetric, take an orthonormal basis $e_1, \dots, e_n$ of eigenvectors. Take a unit vector x, so $x = \xi_1 e_1 + \dots + \xi_n e_n$ and $|x|^2 = \xi_1^2 + \dots + \xi_n^2 = 1$. Now $x^T A x = \lambda_1 \xi_1^2 + \dots + \lambda_n \xi_n^2$ where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A. In particular, if $\lambda_{\min} := \min_i \lambda_i$, $\lambda_{\max} := \max_i \lambda_i$ then $\lambda_{\min} = \min_{|x|=1} x^T A x$, $\lambda_{\max} = \max_{|x|=1} x^T A x$.

In particular, if $E(G) \neq \emptyset$, say $\{1, 2\} \in E(G)$, taking $x = \frac{1}{\sqrt{2}}(1, 1, 0, \dots, 0)$ then $\lambda_{\max} \geq 1$; and taking $x = \frac{1}{\sqrt{2}}(1, -1, 0, \dots, 0)$ then $\lambda_{\min} \leq -1$.

Let $W \subset [n]$ and let H be the induced subgraph $G[W]$. Let y be a vector in $\mathbb{R}^{|W|}$ and let x be the vector corresponding to y with zeros in the coordinates $[n] - W$. Pick y' corresponding to $\lambda_{\min}(H)$, then $(x')^T A (x') \geq \lambda_{\min}(G)$, and likewise, choosing y'' corresponding to $\lambda_{\max}(H)$, then $(x'')^T A (x'') \leq \lambda_{\max}(G)$. Hence

$$\lambda_{\min}(G) \leq \lambda_{\min}(H) \leq \lambda_{\max}(H) \leq \lambda_{\max}(G).$$

If G is bipartite then $A = \left(\begin{smallmatrix} 0 & * \\ * & 0 \end{smallmatrix}\right)$, after an appropriate relabelling of vertices. Consider the characteristic polynomial $\det(A - tI)$. A term in the expansion is obtained by taking j t's from top left block, involving $k - j$ entries taken from bottom left, and hence involving $n - k - (k - j)$ entries from bottom right, so power of t is $n - 2k + 2j$. Thus if n is even, the characteristic polynomial is a polynomial in t^2; if n is odd, it is t times such a polynomial. In particular, if λ is a root, so is $-\lambda$ with the same multiplicity.

Theorem 7.1. Let G be a graph. Then

(i) $\delta(G) \leq \lambda_{\max} \leq \Delta(G)$;
(ii) $|\lambda| \leq \Delta(G)$ for all λ, i.e. $\lambda_{\min} \geq -\Delta(G)$.

If G is connected, then also

(iii) $\lambda_{\max} = \Delta(G)$ if and only if G is regular, in which case $\lambda_{\max}$ has multiplicity 1;
(iv) $\lambda_{\min} = -\Delta(G)$ if and only if G is regular and bipartite, in which case $\lambda_{\min}$ has multiplicity 1.

Proof. Let $y = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$, then $(Ay)_i = \sum_{j\in\Gamma(i)} a_j$ and $y^T Ay = 2\sum_{ij\in E(G)} a_i a_j$.

Take $y = \frac{1}{\sqrt{n}}(1, 1, \ldots, 1)$, then $y^T Ay = \frac{2e(G)}{n} = d \geq \delta(G)$ where d is the average degree, showing the first part of (i). Let $m \in [n]$ be such that $a_m \geq |a_i|$ for all i. We may assume $a_m = 1$. Then $\lambda = (Ay)_m = \sum_{j\in\Gamma(m)} a_j$. In particular,

$$|\lambda| \leq \sum_{j\in\Gamma(m)} |a_j| \leq d(m) \leq \Delta(G)$$

proving (ii) and the second half of (i).

For (iii) and (iv) note that we can recover any information about eigenvalues by considering components of G, e.g. if G has two disjoint components then the adjacency matrix is $\left(\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right)$.

Suppose now $\lambda = \Delta$. Then $\sum_{j\in\Gamma(m)} a_j = \Delta$ which means $d(m) = \Delta$ and $a_j = 1$ for all $j \in \Gamma(m)$. Applying this argument with one of these j's instead of m shows $d(j) = \Delta$ and $a_i = 1$ for all $i \in \Gamma(j)$. Since G is connected, we repeat to get that $y = (1, 1, \ldots, 1)$ and G is Δ-regular, proving (iii).

Finally, if $\lambda = -\Delta$ then $\sum_{j\in\Gamma(m)} a_j = -\Delta$, so $d(m) = \Delta$ and $a_j = -1$ for all $j \in \Gamma(m)$. Similarly, for all $j \in \Gamma(m)$, $d(j) = \Delta$ and $a_i = 1$ for all $i \in \Gamma(j)$. Thus G is Δ-regular, $a_i = \pm 1$ for all i and $a_i a_j = -1$ for all $ij \in E(G)$. So G is bipartite as we can partition G according to whether $a_i = \pm 1$. The converse is clear since $A(1, 1, -1, 1 \ldots, 1)^T = -\Delta(1, 1, -1, 1, \ldots)^T$. $\square$

Given a graph G, an orientation $\vec{G}$ of G is obtained by giving a direction to each edge. The incidence matrix B of $\vec{G}$ is the $n \times e(G)$ matrix $B = (b_{vl})$,

$$b_{vl} = \begin{cases} 1 & \text{if } l = uv \text{ for some } u \\ -1 & \text{if } l = vu \text{ for some } u \; . \\ 0 & \text{otherwise} \end{cases}$$

Definition. The *(combinatorial) Laplacian* on the graph G is $L = BB^T$. Note that $L = D - A$ where $D = \mathrm{diag}(d(1), \ldots, d(n))$, so L does not depend on the orientation.

Let the eigenvalues of L be $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_n$. Then since L is positive semi-definite, we have $\mu_1 \geq 0$. Indeed, $x^T LX = (B^T x)^T (B^T x) = \sum_{ij\in E(G)} (x_i - x_j)^2$, so $\mu_1 \geq 0$, and indeed taking $x = \frac{1}{\sqrt{n}}(1, 1, \ldots, 1)$ we see $\mu_1 = 0$.

Proposition 7.2. $\mu_2 = \min\{x^T Lx/\|x\|^2 : x \neq 0, \sum_{i=1}^n x_i = 0\}$.

Proof. We have an orthonormal basis $e_1, \ldots, e_n$ with $Le_i = \mu_i e_i$. We can take $e_1 = \frac{1}{\sqrt{n}}(1, \ldots, 1)$. Now suppose $\sum_{i=1}^n x_i = 0$. Writing $x = \sum_{i=1}^n \xi_i e_i$ we have $\xi_1 = x \cdot e_1 = 0$ and

$$x^T Lx = \sum_{i=1}^n \mu_i \xi_i^2 \geq \sum_{i=1}^n \mu_2 \xi_i^2 = \mu_2 \|x\|^2$$ □

Theorem 7.3. Let G be a graph and $U \subset V(G)$. Then there are at least $\mu_2 |U||V - U|/|G|$ edges between U and $V - U$.

Proof. We may assume $0 < k = |U| < n = |V(G)|$. Let $x = (x_1, \ldots, x_n)$ where $x_i = n-k$ if $i \in U$ and $x_i = -k$ if $i \in V - U$. Since L is symmetric, there is an orthonormal basis of eigenvectors $e_1, \ldots, e_n$ with $Le_i = \mu_i e_i$ where $e_1 = \frac{1}{\sqrt{n}}(1, \ldots, 1)$. Then $x = \sum_{i=2}^n \xi_i \mu_i e_i$ and

$$x^T Lx = \sum_{i=2}^n \xi_i^2 \mu_i \geq \mu_2 \sum_{i=2}^n \xi_i^2 = \mu_2 \|x\|^2 = \mu_2 kn(n-k).$$

But

$$x^T Lx = \sum_{ij \in E(G)} (x_i - x_j)^2 = |F| n^2$$

where F is the set of edges between U and $V - U$. □

Remark. If μ_2 is large, the graph "expands".

How big can $n = |G|$ be if $\Delta(G) \leq d$ and $\operatorname{diam}(G) \leq 2$? Pick any $v \in V(G)$. Then

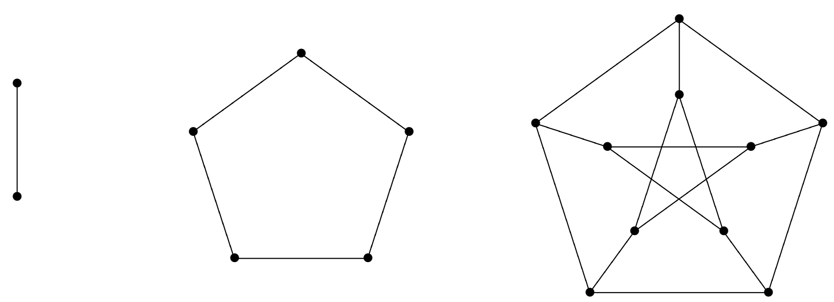
$$|\{w \in V : d(v, w) = 1\}| = |\Gamma(v)| \leq d$$
$$|\{w \in V : d(v, w) = 2\}| \leq d(d-1)$$

and hence $n \leq 1 + d + d(d-1) = d^2 + 1$.

If equality holds, G is d-regular and contains no C_3 or C_4, i.e. it has girth $g(G) = 5$.

Does there exist a d-regular graph of diameter 2 and order $d^2 + 1$? Note this is the maximum possible order and equivalent to asking for diameter 2, girth 5. Such a graph is called a *Moore graph*.

Example. Consider the cases $d = 1, d = 2, d = 3$. These lead to K_2, C_5, and the Petersen graph, respectively.



Note that all of these are unique solutions.

Definition. A *strongly regular* graph with parameters (d, a, b) is regular of degree d, every pair of adjacent vertices has exactly a common neighbours and every pair of non-adjacent vertices has exactly b common neighbours.

Remark. Therefore, a Moore graph is a strongly regular graph with parameters $(d, 0, 1)$.

Theorem 7.4. Let G be a strongly regular graph with parameters (d, a, b) and order n. Then

$$\tfrac{1}{2}\left\{n - 1 \pm \frac{(n-1)(b-a) - 2d}{\sqrt{(a-b)^2 + 4(d-b)}}\right\}$$

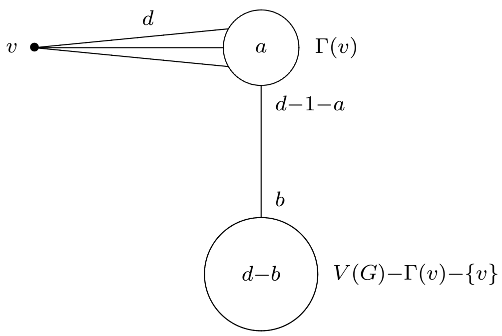
are natural numbers.

Proof. We may assume G is connected for otherwise $b = 0$ and each component is a K_{d+1}.

Let A be the adjacency matrix of G. Then

$$(A^2)_{ij} = \begin{cases} d & \text{if } i = j \\ a & \text{if } ij \in E(G) \\ b & \text{if } ij \in E(\bar{G}) \end{cases}$$

which is equivalent to G being strongly regular with parameters (d, a, b).



Let J be the matrix with all entries 1 and set $B = J - I - A$, so B is the adjacency matrix of $\bar{G}$. Equivalently to G being strongly regular with parameters (d, a, b), we have $A^2 = dI + aA + bB$. From the diagram, $BA = 0I + (d - 1 - a)A + (d - b)B$. So

$$\begin{aligned} A^3 &= dA + aA^2 + bBA \\ &= dA + aA^2 + b[(d-1-a)A] + (d-b)[A^2 - dI - aA], \end{aligned}$$

since $bB = A^2 - dI - aA$. Thus A satisfies a cubic,

$$A^3 - (d - b + a)A^2 - [d(b-a) + d - b]A + d(d-b)I = 0.$$

Now consider the eigenvalues of A. Since G is d-regular, d is an eigenvalue of multiplicity 1. But A satisfies a cubic, so has only two other eigenvalues λ_1, λ_2 with multiplicities r, s. Hence d, λ_1, λ_2 are roots of

$$t^3 - (d - b + a)t^2 - [d(b-a) + d - b]t + d(d-b) = 0,$$

so

$$\lambda_1, \lambda_2 = \tfrac{1}{2}\{a - b \pm \sqrt{(a-b)^2 + 4(d-b)}\}.$$

Now $r + s = n - 1$ and $d + r\lambda_1 + s\lambda_2 = \operatorname{tr} A = 0$. Solve for r, s to obtain the result. □

Theorem 7.5. There exists a Moore graph only if $d \in \{1, 2, 3, 7\}$ and possibly $d = 57$.

Proof. Recall that $n = d^2 + 1$, so

$$\tfrac{1}{2}\left\{d^2 + \frac{d^2 - 2d}{\sqrt{4d-3}}\right\} \in \mathbb{N}.$$

Hence either $d = 2$, in which case $d^2 - 2d = 0$, or $4d - 3 = l^2$ where $l \in \mathbb{Z}$. So $r = \frac{1}{2}\{d^2 + \frac{d^2-2d}{l}\}$, so $l^5 + l^4 + 6l^3 - 2l^2 + (9 - 32r)l - 15 = 0$. So $l \mid 15$. Thus $l \in \{1, 3, 5, 15\}$, so $d \in \{1, 3, 7, 57\}$, and the special case $d = 2$. □