Graph Theory

Cambridge University Mathematical Tripos: Part II

12th June 2024

Contents

1 Introduction

1.1 Definitions

We use the notation $[n]$ for $\{1, \dots, n\}$. For a set X and $k \in \mathbb{N}$, we define $X^{(k)} = \{Y \subseteq X \mid |Y| = k\}$.

Definition. A *graph* is a pair (V, E), where V is a set of *vertices* and E is a set of *edges* where $E \subseteq V^{(2)}$. We use the notation $V(G)$ to denote the set of vertices and $E(G)$ to denote the set of edges, where $G = (V, E)$ is a graph. We define $|G| = |V(G)|$, and $e(G) = |E(G)|$.

Example. The complete graph on n vertices, denoted K_n, is the graph with $V = [n]$ and $E = V^{(2)}$.

Note that we sometimes use juxtaposition of names of vertices to denote an edge between them, so 13 represents the edge $\{1, 3\}$.

Remark. Edges are undirected. There are no edges from a vertex to itself. Edges between vertices are unique if they exist. Most of the graphs covered in this course are finite.

Example. The empty graph on n vertices, denoted $\overline{K}_n$, is the graph with vertex set $V = [n]$ and $E = \varnothing$.

Example. The path of length n, denoted P_n, is the graph with vertex set $V = [n+1]$ and edge set $E = \{\{1, 2\}, \dots, \{n, n+1\}\}$.

Example. The cycle of length n, denoted C_n, is the graph with vertex set $V = [n]$ and edge set $E = \{\{1, 2\}, \dots, \{n-1, n\}, \{n, 1\}\}$.

Definition. Let G be a graph, $x \in V(G)$. The *neighbourhood* of x in G is

$$N_G(x) = \{y \in V(G) \mid \{x, y\} \in E(G)\}$$

If y is a neighbour of x, we write $x \sim y$.

Note that $\sim$ is irreflexive and not transitive in general.

Definition. The *degree* of a vertex $x \in V(G)$ is defined as $\deg x = |N(x)|$.

Definition. Let G, H be graphs. A *graph isomorphism* is a bijection $\varphi \colon V(G) \to V(H)$ such that $\{u, v\} \in E(G) \iff \{\varphi(u), \varphi(v)\} \in E(H)$.

Definition. We say H is a *subgraph* of G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$.

If G is a graph, and $xy \in E(G)$, we define $G - xy$ to be the graph $(V(G), E(G) \setminus \{xy\})$. Similarly, for $x, y \in V(G)$, we define $G + xy$ to be the graph $(V(G), E(G) \cup \{xy\})$.

Definition. Let $x, y \in V(G)$. A *walk* from x to y in G is a sequence of vertices $(x, \dots, y)$ such that each consecutive pair of elements of the sequence is connected by an edge in G. A *path* from x to y in G is a walk where all the vertices are disjoint.

Definition. A graph is *connected* if every pair of vertices is connected with a path.

The concatenation of two paths or walks P and P' is written PP'.

Remark. The concatenation of two walks is a walk. The concatenation of two paths is not necessarily a path, if the two paths share a vertex.

Proposition. If W is a x–y walk for $x \neq y$, W contains a x–y path, where 'contains' denotes a subsequence.

Proof. Let W' be the minimal x–y walk in W. This is a path, because if there were a repeated vertex, we could find a shorter path by eliminating the detour. □

Definition. We define the *distance* between two vertices, denoted $d(x, y)$, to be the shortest length of a path between x and y. If G is connected, this turns G into a metric space on its vertices.

1.2 Trees

Definition. A graph G is a *acyclic* if it does not contain a cycle C_k as a subgraph. A graph G is a *tree* if it is acyclic and connected.

Proposition. The following are equivalent.
(i) G is a tree (acyclic and connected).
(ii) G is *minimally connected*: G is connected and for all $xy \in E(G)$, $G - xy$ is not connected.
(iii) G is *maximally acyclic*: G is acyclic and for all $xy \notin E(G)$, $G + xy$ contains a cycle.

Proof. *(i) implies (ii).* Let $xy \in E(G)$. Suppose $G - xy$ were connected. Then there exists an x–y path P in $G - xy$. We can then close up the path P into a cycle in G by adding the edge xy. This contradicts the fact that G is acyclic.

(ii) implies (i). Suppose G has a cycle C. Let $xy \in E(C)$ be an edge in the cycle. We claim that $G - xy$ is connected. Let P be a u–v path in G. If P contains the edge xy, replace the use of this edge with the remainder of the cycle, traversed in the opposite direction. This yields a u–v walk in $G - xy$ which contains a u–v path.

(i) implies (iii). Let $xy \notin E(G)$. By connectedness, there exists an x–y path P in G. Hence, adding xy to $E(G)$, we obtain a cycle by concatenating P with xy.

(iii) implies (i). Suppose G is not connected. Then there exist $x \neq y$ such that there is no x–y path in G. Hence, adding xy to $E(G)$ cannot yield a cycle. □

Definition. Let T be a tree. A *leaf* of T is a vertex $v \in V(T)$ where $\deg(v) = 1$.

Definition. Let G be a graph, and $X \subseteq V(G)$. Then the *graph induced on* X, denoted $G[X]$ is the graph $(X, \{xy \in E(G) \mid x \in X, y \in X\})$. If $x \in G$, we define $G - x$ to be the graph $G[V(G) \setminus \{x\}]$.

Proposition. Let T be a tree where $|T| \geq 2$. Then T has a leaf.

Proof. Let $P = x_1, \dots, x_k$ be a longest possible path in T. $N(x_k) \subseteq \{x_1, \dots, x_{k-1}\}$ by maximality of P. If $x_i \sim x_k$ for any $1 \leq i \leq k-2$, we have a cycle, which is a contradiction. Hence $N(x_k) = \{x_{k-1}\}$, so x_k is a leaf. □

Remark. This proof actually demonstrates that any tree has at least two leaves, by considering x_1. We could alternatively have proven the lemma by taking a non-backtracking walk in G, which exists assuming no leaf exists; then, since $V(G)$ is finite, we must return to a point somewhere on the graph.

Proposition. Let T be a tree with $n \geq 1$ vertices. Then $|E(T)| = e(t) = n - 1$.

Proof. We prove this by induction on n. The $n = 1$ case is trivial. Now, assume that all trees with n vertices have $n-1$ edges, and suppose T has $n+1$ vertices. T has a leaf x. Then $T - x$ is a tree with n vertices since it is still connected, and hence has $n-1$ edges. Since T has one more edge than $T - x$, namely the edge connecting the leaf x to $T - x$, T has n edges as required. □

Definition. Let G be a connected graph. Then a subgraph T of G is a *spanning tree* if $V(T) = V(G)$ and T is a tree.

Proposition. Every connected graph has a spanning tree.

Proof. Begin with G and remove edges of $E(G)$ such that the graph stays connected. When we can no longer remove edges, we must have a minimally connected subgraph of G, and hence a tree. □

1.3 Bipartite graphs

Definition. Let $G = (V, E)$ be a graph. G is *bipartite* if $V = A \cup B$ where $A \cap B = \varnothing$, such that all edges $(x, y) \in E$ satisfy $x \in A, y \in B$ or $x \in B, y \in A$.
The *complete bipartite graph* on n and m vertices, denoted $K_{n,m}$, is the bipartite graph with $|A| = n$, $|B| = m$ and with all possible edges.

Remark. Even cycles C_{2n} are bipartite, and odd cycles C_{2n+1} are not bipartite.

Definition. A *circuit* is a sequence $x_1, x_2, \dots, x_\ell, x_{\ell+1}$ where $x_i x_{i+1} \in E$ and $x_{\ell+1} = x_1$. In other words, a circuit is a closed walk. The *length* of this circuit is ℓ. A circuit is odd if its length is odd; a circuit is even if its length is even.

Proposition. Let C be an odd circuit in a graph G. Then C contains an odd cycle.

Proof. Let $x_1, \dots, x_\ell, x_1$ be an odd circuit. Either this is an odd cycle, or $x_i = x_j$ for $i < j$. Then $x_i, \dots, x_j$ is a circuit and $x_j, \dots, x_\ell, x_1, \dots, x_i$ is a circuit. Their lengths sum to ℓ, so one of them is odd. By induction, we can assume the odd circuit contains an odd cycle as required. □

Theorem. Let G be a graph. Then G is bipartite if and only if G does not contain an odd cycle.

Proof. If G contains an odd cycle, G is not bipartite because there exists a subgraph that is not bipartite. Suppose now that G contains no odd cycles. We may assume G is connected, because unions of disconnected bipartite graphs are bipartite. Let $x_0 \in V(G)$. Let $V_0 = \{x \in V(G) \mid d(x, x_0) \equiv 0 \bmod 2\}$ and $V_1 = \{x \in V(G) \mid d(x, x_0) \equiv 1 \bmod 2\}$. We show that this is a bipartition as required. Suppose $u, v \in V_i$ are connected. Then, u and v admit even (resp. odd) paths to x_0, so the circuit defined by the concatenation of these paths with the edge uv is an odd circuit, and hence contains an odd cycle. This contradicts our assumption. □

1.4 Planar graphs

Definition. A *plane graph* is a drawing of a graph in the plane, representing edges as piecewise linear functions, without edge crossings.

Definition. A graph G is *planar* if it can be drawn in the plane $\mathbb{R}^2$ with no edges crossing, so a graph is planar if it admits a plane graph representation.

Example. K_1, K_2, K_3, K_4 are planar. P_n is planar for $n \in \mathbb{N}$. $K_{n,2}$ is planar, by placing the vertices in the two-vertex set on either side of the other set.

Definition. Let G be a plane graph. One of the finitely many connected components of $\mathbb{R}^2 \setminus G$ is called a *face*. The boundary of a face F is the collection of vertices and edges in ∂f. Therefore, the boundary of any face in G is a subgraph of G.

Remark. The boundary of a face need not be (or contain) a cycle, and need not be connected. Two drawings of a graph can be fundamentally different.

Theorem (Euler)**.** Let G be a connected plane graph with F faces. Then $|V(G)| - |E(G)| + F = 2$.

Remark. The number of faces is uniquely determined by intrinsic properties of a graph, its number of vertices and edges.

Proof. We work by induction on the number of edges $E(G)$. In the case where $E(G) = 0$, we must have $V(G) = 1$ and $F = 1$ by connectedness. Suppose G is acyclic. Then by connectedness, G is a tree, so $V(G) = E(G)+1$ and $F = 1$, satisfying Euler's formula. Now suppose G contains a cycle, and E be an edge in the cycle. Removing this edge, $G-E$ is connected, and has $|V(G)|$ vertices, $|E(G)|-1$ edges, and $F-1$ faces. By induction, Euler's formula holds in this case. $\square$

Corollary. Let G be a planar graph where $|G| \geq 3$. Then $e(G) \leq 3|G| - 6$.

Proof. Consider a planar drawing of G. We may assume G is connected without loss of generality. Let F be a face, and let $\deg F$ be the number of edges that meet at F. Note that the degree of any face is at least 3, since $|G| \geq 3$. Since each edge occurs in at most two faces, $\sum_F \deg F \leq 2e(G)$. Hence, $3f \leq 2e(G)$, where f is the amount of faces. Using Euler's formula, $|G| - e(G) + f = 2 \implies 2(|G| - 2) \geq e(G)$. $\square$

Remark. K_5 is not planar, because $e(K_5) = 10$ and $3|K_5| - 6 = 9$. $K_{3,3}$ does not violate this bound, but is not planar.

Corollary. Let G be a planar graph, $|G| \geq 4$ and there is no cycle of length 3. Then $e(G) \leq 2(|G| - 2)$.

Proof. The minimal degree of a face is 4, because a degree of 3 would imply there is a triangle since there are at least four vertices in the graph. Running the same argument, our bound becomes $e(G) \leq 2(|G| - 2)$, $\square$

This shows that $K_{3,3}$ is not planar.

Definition. A *subdivision* of a graph G is a new graph where some of the edges are replaced by (disjoint) paths.

Remark. A subdivision of a non-planar graph is non-planar. In particular, if G contains a subdivision of $K_{3,3}$ or K_5, G is non-planar.

Theorem (Kuratowski). G is planar if and only if it contains no subdivision of $K_{3,3}$ or K_5.

2 Connectivity and matching

2.1 Matching in bipartite graphs

Definition. Let $G = (X \sqcup Y, E)$ be a bipartite graph. A *matching from X to Y* is a set of edges $E' \subseteq \{xy_x \mid x \in X, y_x \in Y\} = E$ such that the map $x \mapsto y_x$ is injective.

Definition. Let G be a graph, $A \subseteq V(G)$. We define $N_G(A) = \left\{\bigcup_{x \in A} N(x)\right\}$.

Theorem (Hall)**.** Let $G = (X \sqcup Y, E)$ be a bipartite graph. There exists a matching from X to Y if and only if *Hall's criterion* holds: that $|A| \leq |N(A)|$ for all $A \subseteq X$.

Proof. The forward direction is simple, by considering the image of the injective map $x \mapsto y_x : A \to N(A)$ for each subset $A \subseteq X$. Conversely, suppose Hall's criterion is satisfied. We apply induction on $|X|$. If $|X| = 1$, $N(X)$ is nonempty and so the proof is complete.

If there does not exist $\varnothing \neq A \subsetneq X$ such that $|N(A)| = |A|$, we have $|A| < |N(A)|$ for all $\varnothing \neq A \neq X$. Let $xy \in E$, and let $G' = G[X \setminus \{x\} \sqcup Y \setminus \{y\}]$. By induction, it suffices to show Hall's criterion holds for G'. If $B \subseteq X \setminus \{x\}$, we have

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

as required.

However, suppose there exists such a set $\varnothing \neq A \subsetneq X$ with $|A| = |N(A)|$. Let $G_1 = G[A \sqcup N(A)]$ and $G_2 = G[X \setminus A \sqcup Y \setminus N(A)]$. G_1 satisfies Hall's criterion. Indeed, for $B \subseteq A$, $N_{G_1}(B) = N_G(B)$ as required. G_2 also satisfies Hall's criterion. Suppose $B \subseteq X \setminus A$, and consider $N_G(A \cup B)$. We have

$$|A| + |B| \leq |N_G(A \cup B)| = |N_G(A)| + \left|N_{G_2}(B)\right| \implies |B| \leq \left|N_{G_2}(B)\right|$$

Hence Hall's criterion is satisfied.

Then by induction on G_1 and G_2, the proof is complete. $\square$

Definition. A *matching of deficiency d from X to Y* is a matching from $X' \subseteq X$ to Y where $|X'| + d = |X|$.

Theorem (defect Hall)**.** Let $G = (X \sqcup Y, E)$ be a bipartite graph. G contains a matching of deficiency $d \leq |X|$ if and only if $|A| \leq |N(A)| + d$ for all $A \subseteq X$.

Proof. The forward direction is again a simple proof. Let $G = (X \sqcup Y, E)$ be a graph such that $|A| \leq |N(A)| + d$ for all $A \subseteq X$. Let $G' = (X \sqcup (Y \cup \{z_1, \dots, z_d\}), E \cup E')$ where $E' = \{xz_i \mid x \in X, i \in \{1, \dots, d\}\}$. Hall's criterion on G' is satisfied, so there exists a matching. Deleting these new vertices $\{z_1, \dots, z_d\}$ and the edge set E', we construct a matching from X to Y of deficiency at most d. To construct a matching of deficiency precisely d, we can delete extra edges as required. $\square$

Definition. The *maximum degree* $\Delta(G)$ (resp. *minimum degree* $\delta(G)$) of a graph G is the maximum (resp. minimum) degree of a vertex in G.

Definition. A graph is *regular* if all vertices have the same degree, or equivalently, $\delta(G) = \Delta(G)$. A graph is *k-regular* if $\delta(G) = \Delta(G) = k$.

Corollary. Let $G = (X \sqcup Y, E)$ be a k-regular bipartite graph and $k \geq 1$. Then there exists a matching from X to Y.

Proof. It suffices to show Hall's criterion holds. Let $A \subseteq X$. Then

$$e(G[A \cup N(A)]) = \sum_{x \in A} \deg x = k|A|; \quad e(G[A \cup N(A)]) = \sum_{x \in N(A)} \deg v \leq k|N(A)|$$

Hence $|A| \leq |N(A)|$. □

Example. Let Γ be a finite group, and let $H \leq \Gamma$. Let $L_1, \dots, L_n$ be the left cosets, and $R_1, \dots, R_n$ be the right cosets. We want to find $g_1, \dots, g_n$ such that $g_1 H, \dots, g_n H$ are the left cosets and $Hg_1, \dots, Hg_n$ are the right cosets.

Consider the graph $G = (\{L_1, \dots, L_n\} \sqcup \{R_1, \dots, R_n\}, E)$ where an edge lies between L_i and R_j if $L_i \cap R_j \neq \varnothing$. It suffices to find a matching in this graph, because then each edge in the matching implies the existence of a representative for both cosets. Let $A \subseteq \{L_1, \dots, L_n\}$, so $A = \left\{L_{i_1}, \dots, L_{i,k}\right\}$. Consider $\left|\bigcup_{j=1}^k L_{ij}\right| = k|H|$, but since $R_1, \dots, R_n$ partition Γ and have size $|H|$, at least k right cosets of H must intersect $\bigcup R_{ij}$. Hence Hall's criterion is satisfied.

2.2 Connectivity

Let $S \subseteq V(G)$. Then we define $G - S = G[V(G) \setminus S]$.

Definition. Let G be a graph, and $|G| \geq 1$. Then we define the *connectivity parameter* $\varkappa$ of G by

$$\varkappa = \min \{|S| \mid S \subseteq V(G), G - S \text{ is disconnected or a single vertex}\}$$

We say that G is k-connected if $k \leq \varkappa$. Hence G is k-connected if and only if for all sets S of at most $k - 1$ vertices, $G - S$ is connected and not a single vertex.

Example. $\varkappa(\text{Petersen graph}) = 3$, because deleting any two vertices leaves the graph connected, but deleting the neighbourhood of any vertex disconnects the graph. $\varkappa(G) = 1$ if G is a tree. $\varkappa(C_n) = 2$ for $n \geq 3$. $\varkappa(K_n) = n - 1$.

Definition. Let G be a graph, and $a, b \in V(G)$. We say that the a–b paths $P_1, \dots, P_k$ are *disjoint* if $P_i \cap P_j = \{a, b\}$ for $i \neq j$.

Note that $\delta(G) \geq \varkappa(G)$. This follows because removing the neighbours of the vertex of minimum degree disconnects the graph or leaves it a single vertex. Also, we can easily see that $\varkappa(G - x) \geq$

$\kappa(G) - 1$. Note that we can have $\kappa(G - x) > \kappa(G)$ by considering a 2-connected graph with an additional leaf.

Definition. Let G be a graph and $a \neq b \in V(G)$, where $a \nsim b$. We say that $S \subseteq V(G) \setminus \{a, b\}$ is a *a–b separator* if $G - S$ disconnects a and b.

Theorem (Menger, form 1). Let G be a connected graph and $a \neq b \in V(G)$, where $a \nsim b$. The minimum size of an a–b separator is the maximum number of disjoint paths from a to b. Equivalently, if all a–b separators have size at least k, then there exist k disjoint a–b paths.

Proof. We write $\kappa_{a,b}(G)$ for the minimum size of an a–b separator. Note that $\kappa(G - x) \geq \kappa(G) - 1$, and $\kappa(G - xy) \geq \kappa(G) - 1$. We also have the same properties for $\kappa_{a,b}$.

Suppose the theorem does not hold, then there is a nonempty set of counterexamples. Let $\mathcal{G}$ be the set of counterexamples of smallest possible k, and let G be an element of $\mathcal{G}$ with the smallest possible amount of edges. Let S be a minimal a–b separator in G, so $|S| = k$. Note that the theorem is true for $k = 1$, so we may assume $k \geq 2$.

If $S \neq N(a)$ and $S \neq N(b)$, consider $G - S$. Then a, b lie in different connected components. Let A be the component containing a, and B be the component containing b. Define G_a to be the graph $G[A \cup S]$ together with a vertex c with edges to each $s \in S$. Similarly, define G_b to be the graph $G[B \cup S]$ together with a vertex c with edges to each $s \in S$.

Note that $\kappa_{a,c}(G_a) \geq k$, because any a–c separator in G_a is an a–b separator in G, and $\kappa_{b,c}(G_b) \geq k$ by symmetry. Note further that $e(G_a), e(G_b) < e(G)$; because $S \neq N(a)$ and $S \neq N(b)$, the amount of newly added edges is smaller than the amount of edges that must have been removed in each induced graph. Then by minimality of G, the G_a and G_b are not counterexamples to the theorem. Hence there exist disjoint a–c paths $P_1, \dots, P_k$ in G_a and disjoint c–b paths $Q_1, \dots, Q_k$ in G_b. Concatenating P_i with Q_i, we obtain k disjoint a–b paths in G. Then G is not a counterexample.

Now, suppose $S = N(a)$ without loss of generality. We claim that $N(a) \cap N(b) = \varnothing$. If there exists $x \in N(a) \cap N(b)$, then consider the graph $G - x$. We have $\kappa_{a,b}(G - x) \geq k - 1$, so by minimality, there exist disjoint a–b paths $P_1, \dots, P_{k-1}$ in $G - x$. Adding the path a, x, b, which is disjoint from all others, we obtain k disjoint a–b paths, contradicting the assumption.

Let $a, x_1, \dots, x_\ell, b$ be a shortest a–b path. Note that $\ell \geq 2$ since $N(a) \cap N(b) = \varnothing$, and in particular, $x_2 \neq b$. Consider $G - x_1x_2$. We must have that $\kappa_{a,b}(G - x_1x_2) \leq k - 1$, otherwise we have a smaller counterexample. Hence $\kappa_{a,b}(G - x_1x_2) = k - 1$. Therefore there is an a–b separator $\tilde{S}$ with $|\tilde{S}| = k - 1$ in $G - x_1x_2$. We see that either $\tilde{S} \cup \{x_1\}$ or $\tilde{S} \cup \{x_2\}$ is a separator of size k in G, which is not equal to either $N(a)$ or $N(b)$. Then we can use the above construction to find the relevant contradiction. $\square$

Corollary (Menger, form 2). Let G be a connected graph with $|G| \geq 2$. Then G is k-connected if and only if all pairs of distinct vertices a, b admit k disjoint a–b paths.

Proof. Suppose all pairs of vertices a, b have k such paths. Suppose $G - S$ is disconnected, and a, b lie in different components of $G - S$. Note that $a \nsim b$, because there exists a separator for a and b. Then by assumption, there are k disjoint a–b paths, and so S must intersect each path. Therefore, $|S| \geq k$.

Now suppose G is k-connected. Let a, b be vertices in G. If $a \nsim b$, apply the first form of Menger's theorem. Conversely, consider $G - ab$. This graph is $k-1$-connected, so there are $k-1$ disjoint a–b by Menger's theorem. Adding the additional path a, b, we obtain k disjoint paths as required. □

2.3 Edge connectivity

Definition. Let G be a graph. Then $\lambda(G) = \min\{|W| \mid W \subseteq E(G), G - W \text{ disconnected}\}$ is the smallest amount of edges that can be deleted to disconnect G. We say that G is *k-edge connected* if $k \leq \lambda(G)$.

Example. Let C_n be the cycle on n vertices. The vertex connectivity $\varkappa$ and edge connectivity λ of this graph are both two.

Example. Consider a graph with two connected subgraphs K_n, but with one vertex in the intersection between the two. Then $\varkappa = 1$ by deleting the intersection vertex, but $\lambda(G) = n - 1$.

Definition. Paths $P_1, \dots, P_k$ are *edge-disjoint* if the edge sets are disjoint.

Theorem (Menger, edge version, form 1)**.** Let G be a connected graph, and $a \neq b$ be vertices. Then, if every $W \subseteq E(G)$ that separates a from b has size at least k, then there exist k edge-disjoint a–b paths.

Definition. Let G be a graph. The *line graph* of G, denoted $L(G)$, is the graph where $V(L(G)) = E(G)$ and $e, f \in E(G)$ are adjacent if they share an endpoint.

Proof. Let G' be the line graph of G, together with distinguished vertices a', b' that are connected to the edges incident to a and b respectively. Note that there is an a–b path in G if and only if there is an a'–b' path in G'. Thus, $W \subseteq V(G') \setminus \{a', b'\}$ is an a', b' separator if and only if $W \subseteq E(G)$ separates a from b. Therefore, $\varkappa_{a',b'}(G') \geq k$. By the first form of Menger's theorem on G', we can find k disjoint a'–b' paths $P_1, \dots, P_k$ in G'. These paths describe edge-disjoint a–b walks in G, which yield edge-disjoint a–b paths. □

Theorem (Menger, edge version, form 2)**.** Let G be a connected graph. Then $\lambda(G) \geq k$ if and only if all all pairs of vertices $a \neq b$ admit k edge-disjoint a–b paths.

Proof. If there exist k edge-disjoint paths between each pair of vertices, to separate any two vertices we must remove at least one edge from each of these k paths, so we must remove at least k edges. Conversely, if $\lambda(G) \geq k$, apply the above form of Menger's theorem. □

3 Colouring

3.1 Definition

Definition. A function $c : V(G) \to \{1, \dots, k\}$ is a *(proper) k-colouring* of a graph if $x \sim y \implies c(x) \neq c(y)$. The *chromatic number* of G, denoted $\chi(G)$, is the minimum k such that there exists a k-colouring of G.

Example. A path P_n has a 2-colouring. More generally, a graph is bipartite if and only if it has a 2-colouring. An even cycle has chromatic number 2, and an odd cycle has chromatic number 3. A tree has chromatic number 2. The complete graph on n vertices has chromatic number n.

Proposition. Let G be a graph. Then $\chi(G) \leq \Delta(G) + 1$.

Proof. Let $x_1, \dots, x_n$ be an ordering of the vertices of G. We create a colouring of the vertices by induction. Suppose $x_1, \dots, x_i$ have already been coloured, and we want to colour x_{i+1}. Since x_{i+1} has at most $\Delta(G)$ neighbours that have already been coloured, but we have $\Delta(G) + 1$ available colours, there is a free colour that does not match any previous neighbours. Choose the smallest available colour. By induction we can colour the entire graph. □

Remark. This is sometimes known as a *greedy colouring*. The greedy colouring may produce a colouring which is suboptimal for a given graph; consider the path P_4 on the vertex set $\{1, 2, 3, 4\}$ but with the ordering $1, 4, 2, 3$: this gives a 3-colouring. The proposition above is sharp: the chromatic number of the complete graph is n, and its maximum degree is $n - 1$.

3.2 Colouring planar graphs

Proposition. Let G be planar. Then $\delta(G) \leq 5$.

Proof. The average degree of G, given by $n^{-1} \sum_{v \in V(G)} \deg v$, is exactly $2n^{-1}e(G)$. Since $e(G) \leq 3n - 6$, the average degree at most $6 - \frac{12}{n} < 6$, so $\delta(G) \leq 5$. □

Proposition (six-colour theorem)**.** Let G be planar. Then G admits a 6-colouring.

Proof. Apply induction on $|G|$. If $|G| \leq 6$, there admits a trivial 6-colouring. Let G be planar, and let $x \in V(G)$ have degree at most 5. By the inductive hypothesis, $G - x$ admits a 6-colouring. Since x has at most five neighbours, there is a free colour to use for x. □

Theorem (five-colour theorem)**.** Let G be planar. Then G admits a 5-colouring.

Proof. We apply induction on $|G|$. Clearly the theorem holds for $|G| \leq 5$. Suppose $|G| > 5$. Let $x \in V(G)$ be a vertex with degree at most five. Applying induction, there exists a 5-colouring of $G - x$. If the degree is four or lower, we can use the free colour to colour x, so suppose x has degree five. Let $N(x) = \{x_1, x_2, x_3, x_4, x_5\}$ arranged cyclically in the plane, and let the colour of x_i be i. Then without loss of generality, all the x_i must have different colours, since otherwise, we are done.

Suppose there exists no path from x_1 to x_3 in $G - x$ only along vertices coloured 1 and 3. In this case, let C be the component of G of vertices coloured 1 or 3 that contains x_1. This is the connected component of the subgraph of $G - x$ induced by the vertices coloured 1 and 3 that contains x_1. By assumption, x_3 is not in this component. Now, swap the colours 1 and 3 on C; this yields another 5-colouring of $G - x$. We can then extend this 5-colouring to x by colouring x with 1.

Now, suppose there exists no path from x_2 to x_4 in $G - x$ along vertices coloured 2 and 4. If so, we are done as above.

Suppose that there exists an x_1–x_3 path using only colours 1 and 3, and an x_2–x_4 path using only colours 2 and 4. Then since both paths lie in the plane and the vertices are arranged cyclically as above, they must cross. The intersection vertex is coloured either 1 or 3, and also either 2 or 4. This is a contradiction. □

Remark. Any planar graph admits a 4-colouring; this result is known as the *four-colour theorem*. The above method does not work, because when swapping the colours of a component, there is not a free colour to use for the newly added vertex. The four-colour theorem was eventually proven using a computer-aided search after reducing the problem to thousands of specific local configurations. The four-colour theorem is sharp; K_4 is planar.

3.3 Colouring non-planar graphs

Proposition. Let G be a connected graph, and $\delta(G) < \Delta(G)$. Then $\chi(G) \leq \Delta(G)$.

Proof. Order the vertices in G into $x_1, \dots, x_n$ such that $\deg x_n < \Delta(G)$, and x_{n-1} is adjacent to x_n, and also x_{n-2} is adjacent to one of x_n and x_{n-1} and so on. This is always possible since G is connected. This ordering has the property that all vertices have less than $\Delta(G)$ edges facing forward. So the greedy colouring gives a $\Delta(G)$-colouring. □

Theorem (Brooks)**.** Let G be a connected graph. If G is not an odd cycle or complete graph, $\chi(G) \leq \Delta(G)$.

Remark. We have shown above that $\chi(G) \leq \Delta(G) + 1$. This theorem then says that $\chi(G) = \Delta(G) + 1$ if and only if G is an odd cycle or a complete graph.

Proof. We apply induction on $|G|$. We can check that the theorem holds for $|G| \leq 3$. Note that we may assume that $\Delta(G) \geq 3$; otherwise, the graph is bipartite or an odd cycle.

We will show first that if G is 3-connected, the theorem holds. We give an ordering of $V(G)$. Let x_n be a vertex of degree $\Delta(G)$, and let $x_1, x_2 \in N(x)$ be non-adjacent vertices. This is possible; indeed, suppose we could not find such vertices. Then $\{x\} \cup N(x)$ is a complete graph, so $G = K_{\Delta(G)+1}$ by connectedness, contradicting our assumption. Now, consider $G - \{x_1, x_2\}$. Since G is 3-connected, $G - \{x_1, x_2\}$ is connected. We can order the vertices in the same way as above, choosing $x_{n-1} \sim x_n$

and x_{n-2} a neighbour of x_{n-1} or x_n, and so on. Then the greedy algorithm produces the required colouring.

Now, we show that if $\varkappa(G) = 1$, the theorem holds. In this case, we have a separator of size one, so let $\{x\}$ be such a separator (we call x a *cut vertex*). Let $C_1, \dots, C_n$ be the connected components of $G - x$. By induction, we can colour $C_i \cup \{x\}$ for each i; they cannot be complete, by counting the number of edges of x in this graph. We can then permute the colours in each such colouring to make x the same colour. Then we can combine each colouring to produce a colouring of the entire graph.

Finally, we will consider the case when $\varkappa(G) = 2$. Let $S = \{x, y\}$ be a separator for G. Let $C_1, \dots, C_k$ be the components of $G - S$. Define the graphs $G_i = G[C_i \cup S] + xy$ for $i = 1, \dots, k$.

Suppose $\delta(G_i) < \Delta(G)$ for all i. In this case, the G_i can be coloured by induction as they are not complete graphs. Note that x, y get different colours since we have added the edge xy. Therefore, we can permute the colours, such that the colouring agrees on x, y for all G_i. These colourings can be combined into a $\Delta(G)$-colouring of G.

Now suppose without loss of generality that $\delta(G_1) = \Delta(G)$. In this case, $k = 2$, and

$$|N(x) \cap C_1| = \Delta(G) - 1 = |N(x) \cap C_1|; \quad |N(x) \cap C_2| = 1 = |N(y) \cap C_2|$$

Let x', y' be the neighbours of x, y in C_2. Now, note that $\tilde{S} = \{x, y'\}$ is a separator, and now $\delta(G_i) < \Delta(G)$ for all connected components, and we can use the proof from above. □

3.4 Chromatic polynomial

Definition. Let G be a graph. The *chromatic polynomial* of G is $P_G : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ where $P_G(t)$ is the number of t-colourings of G.

Remark. The minimum t for which $P_G(t) > 0$ the chromatic number.

Example. The chromatic polynomial on the empty graph on n vertices is given by $P_G(t) = t^n$.

The chromatic polynomial on the complete graph on n vertices is $P_G(t) = t(t-1)\dots(t-(n-1)) = n!\binom{t}{n}$.

For a path on n vertices, $P_G(t) = t(t-1)^{n-1}$. For any tree, colouring each leaf, removing it, then colouring the remainder inductively, $P_G(t) = t(t-1)^{|G|-1}$.

Definition. Let G be a graph, and $e = xy \in E(G)$. The *contraction* of G along e, denoted G/e, is the graph with vertices $V(G) \setminus \{x, y\} \cup \{a\}$ for a new variable a, and edges $E(G[V \setminus \{x, y\}]) \cup \{az \mid x \sim z\} \cup \{az \mid y \sim z\}$.

Proposition. Let G be a graph and $e \in E(G)$. Then $P_G(t) = P_{G-e}(t) - P_{G/e}(t)$.

Proof. Let $e = xy$. A t-colouring of $G - e$ where x, y are assigned different colours corresponds to a t-colouring of G, by simply adding the edge back. A t-colouring of $G - e$ where x, y are assigned the same colour corresponds to a t-colouring of G/e, by contracting the edge. □

Remark. The above proposition is known as a 'cut-fuse' relation.

Proposition. Let G be a graph. Then $P_G(t)$ is indeed a polynomial with degree $|G|$.

Proof. We apply induction on $e(G)$. If there are no edges in the graph, the graph is empty, and has chromatic polynomial $P_G(t) = t^{|G|}$. Otherwise, let $e \in E(G)$. By induction, $P_{G-e}(t)$ is a polynomial of degree $|G - e| = |G|$, and $P_{G/e}(t)$ is a polynomial of degree $|G/e| = |G| - 1$. Hence $P_G(t) = P_{G-e}(t) - P_{G/e}(t)$ is indeed a polynomial of the required degree. □

Proposition. Let G be a graph with n vertices and m edges. Then $P_G(t) = t^n - mt^{n-1} + p(t)$ where p is a polynomial of degree at most $n - 2$.

Proof. We apply induction on $e(G)$. If there are no edges, we have the empty graph, which has the required form. Otherwise, let $e \in E(G)$. Then

$$P_G(t) = P_{G-e}(t) - P_{G/e}(t) = \left(t^n - (m-1)t^{n-1} + \dots\right) + \left(t^{n-1} + \dots\right) = t^n - mt^{n-1} + \dots$$

as required. □

Remark. Other coefficients of the chromatic polynomial contain other information about the graph. For example, the t^{n-2} coefficient is exactly $\binom{e(G)}{2}$ − number of triangles in G.

If G is planar, $P_G\left(2 + \frac{1+\sqrt{5}}{2}\right) \neq 0$.

A result due to June Huh is that the coefficients $c_0, \dots, c_n$ of P_G are *log-concave*, so $c_i^2 > c_{i-1}c_{i+1}$.

3.5 Edge colouring

Definition. Let G be a graph. A *k-edge colouring* is a function $c : E(G) \to \{1, \dots, k\}$ such that if $c(e) \neq c(f)$ if e, f share an endpoint. The *edge chromatic number*, or the *chromatic index*, denoted $\chi'(G)$, is the minimum k such that there exists a k-edge colouring.

Remark. An edge colouring of G corresponds exactly to a vertex colouring of the line graph of G. In particular, $\chi'(G) = \chi(L(G))$. Note that not every graph can be realised as the line graph of some other graph.

Example. The edge chromatic number of an even cycle is 2. The edge chromatic number of an odd cycle is 3. This is because a cycle is its own line graph.

Example. We have $\Delta(G) \leq \chi'(G)$. If $x \in V(G)$ has degree $\Delta(G)$, all edges indicent to x must be given different colours. We may have $\Delta(G) < \chi'(G)$ for some graphs, such as C_3. The edge chromatic number of the Petersen graph is 4, but it is 3-regular.

We can show that $\chi'(G) \leq 2\Delta(G) - 1$ by the greedy colouring, considering how many vertices each edge can be connected to. χ' and χ can be very different, for instance, consider $\chi(K_{t,1}) = 2$ but $\chi'(K_{t,1}) = t$.

Given an edge colouring $c: E(G) \to \{1, \dots, k\}$, we define the *colour classes* as equivalence classes of colours: $C_i = \{e \in E(G) \mid c(e) = i\}$. Note that $(V(G), C_i \cup C_j)$ is the union of disjoint paths, even cycles, and isolated vertices. We say that the components of this graph are $\{i, j\}$-*components*.

Theorem (Vizing). Let G be a graph. Then $\chi'(G) = \Delta(G)$ or $\chi'(G) = \Delta(G) + 1$.

Proof. We prove this by induction on $|E(G)|$. It suffices to show there is a $\Delta(G) + 1$ colouring of any graph. If there are no edges, the graph can be 0-coloured, so $\chi'(G) = \Delta(G) = 0$ and so there is clearly a 1-colouring. For the inductive step, let G be a graph with $e(G) > 0$, and $xv \in E(G)$. Apply induction to $G - xv$ to obtain a $\Delta(G) + 1$ edge colouring.

Let $y \in V(G)$ and $c \in \{1, \dots, \Delta(G) + 1\}$. We say c is *missing* at y if no edge incident to y are coloured c. Note that there is a colour missing at every vertex since we have $\Delta(G) + 1$ different colours available.

Let c_0 be a colour missing at x. We define a sequence of vertices $v_1, \dots, v_k \in N(x)$ and corresponding colours $c_1, \dots, c_k$ such that c_i is missing at v_i. First, we set $v_1 = v$ and let c_1 be any colour missing at v. Then if v_i and c_i are defined, define v_{i+1} such that $c(xv_{i+1}) = c_i$, and define c_{i+1} to be any colour missing at v_{i+1}. This induction continues until either we find a colour missing at x or we repeat a colour.

Suppose $v_1, \dots, v_k$ are defined and c_k is missing at x. Then we can recolour xv_k with c_k. Now c_{k-1} is missing at x, so inductively, recolour xv_i with c_i. In particular, c_1 is missing at x, so we can colour xv_1 with c_1.

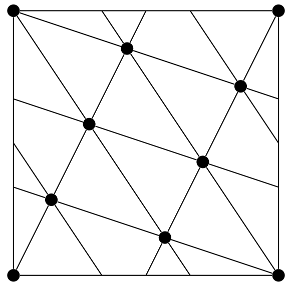
In the other case, suppose $c_k = c_i$ for $i < k$. Note that we may assume $i = 1$: uncolour xv_{i-1} and recolour xv_j with c_j for all $j < i$ as above. So $c_k = c_1$. If v_1 is not in the same $\{c_0, c_1\}$ component as x, we can swap the colours on the $\{c_0, c_1\}$ component containing v_1. Then c_0 is missing at v_1, and the colours of $xv_2, \dots, xv_k$ are unchanged. So we can colour xv_1 with c_0.

Now suppose x, v_1 are in the same $\{c_0, c_1\}$ component. If v_k is not in the same $\{c_0, c_1\}$ component as x, we can similarly swap the colours on the $\{c_0, c_1\}$ component containing v_k. So c_0 is missing at v_k and x, and so we can recolour xv_k to c_0, and inductively xv_i with c_i.

Now finally suppose x, v_1, v_k are all in the same $\{c_0, c_1\}$ component. So one of c_0, c_1 are missing at each of x, v_1, v_k. Since all $\{c_0, c_1\}$-components of the graph are disjoint paths, even cycles, or isolated vertices. So x, v_1, v_k are each endpoints of a path. But since paths only have two endpoints, this is a contradiction. $\square$

3.6 Graphs on surfaces

We have seen that a planar graph has chromatic number $\chi(G) \leq 5$. Drawing graphs on other surfaces give different possible chromatic numbers. For instance, the complete graph on seven vertices K_7 can be drawn on a torus with no edge crossings.



Recall from IB Geometry that for any $g \in \mathbb{N}$, there is a *compact orientable surface of genus* g which is homeomorphic to a sphere with g 'handles' attached. The 2-sphere S^2 is a compact orientable surface of genus 0. The torus T^2 is a compact orientable surface of genus 1.

We have already seen that for a connected planar graph G with f faces, we have $|G| - e(G) + f = 2$. For a disconnected planar graph, we can add edges to make G into a connected graph. Hence, any planar graph with f faces satisfies $|G| - e(G) + f \leq 2$. In general, on the compact orientable surface of genus g, $|G| - e(G) + f \leq E$, where $E = 2 - 2g$ is the Euler characteristic of the surface. Due to results from IB Geometry, the equality holds for connected graphs, and then for any other graph, we can add edges to make it connected.

In particular, if $e(G) \geq 3$, then $3f \leq 2e(G)$ as usual. Therefore, $|G| - e(G) + \frac{2e(G)}{3} \geq E$, and so $e(G) \leq 3(|G| - E)$.

Theorem (Heawood)**.** Let G be a graph drawn on a surface of Euler characteristic $E \leq 0$. Then

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

Remark. Note that $H(2) = 4$, which would prove the four-colour theorem if not for the requirement that $E \leq 0$.

Proof. Let G be a graph drawn on a given surface with Euler characteristic E. Suppose its chromatic number is $\chi(G) = k$. Without loss of generality, we can choose a minimal such graph G with $\chi(G) = k$.

Each vertex has degree at least $k-1$. Indeed, suppose there was a vertex of degree less than $k-1$. Then we could remove this vertex and all associated edges, and we would obtain a strictly smaller graph with chromatic number exactly k, contradicting minimality. Further, we have $|G| \geq k$, otherwise we could colour the graph with only $|G|$ colours contradicting the definition of the chromatic number.

Since $e(G) \leq 3(|G| - E)$, the sum of the degrees of the vertices is $2e(G) \leq 6(|G| - E)$. Hence, $\delta(G) \leq \frac{1}{|G|}6(|G| - E) = 6 - 6\frac{E}{|G|}$. In particular,

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

Note that this step requires the fact that $E \leq 0$. This gives the quadratic equation $k^2 - 7k + 6E \leq 0$. Then,

$$\left(k - \frac{7}{2}\right)^2 - \frac{49}{4} + 6E \leq 0 \implies k \leq \frac{7 + \sqrt{49 - 24E}}{2}$$

$\square$

Remark. The inequality is sharp, since the complete graph $K_{H(E)}$ can be drawn on a surface of characteristic E. An example of this is drawing K_7 on the torus, as demonstrated above. However, this is a very difficult result to prove.

4 Extremal graph theory

4.1 Hamiltonian graphs

Definition. A graph is said to be *Hamiltonian* if it contains a cycle that contains all vertices. Such a cycle is called a *Hamilton cycle*.

Theorem. Let G be a graph on $n \geq 3$ vertices. Then if $\delta(G) \geq \frac{n}{2}$, G is Hamiltonian.

Remark. This theorem is sharp. If n is even, two disjoint $K_{\frac{n}{2}}$ cliques suffices for a counterexample, since $\delta(G) = \frac{n}{2} - 1$. If n is odd, we can take two $K_{\frac{n+1}{2}}$ cliques which intersect in a single vertex, giving $\delta(G) = \frac{n-1}{2}$.

Proof. First, note that G is connected. Indeed, if $x \nsim y$, $|N(x)|, |N(y)| \geq \frac{n}{2}$, but there are only $n-2$ remaining vertices in the graph. So by the pigeonhole principle, there is a path of length 2 between x and y.

Consider a path $x_1, \dots, x_\ell$ of maximum length, and suppose for a contradiction that there is no cycle in G of length ℓ. Observe that $N(x_1) \subseteq \{x_2, \dots, x_{\ell-1}\}$ by maximality, and $N(x_\ell) \subseteq \{x_2, \dots, x_{\ell-1}\}$ by symmetry. Define $N^-(x_1) = \{x_i \mid x_{i+1} \in N(x_1)\}$. Note that $|N^-(x_1) \cup N(x_\ell)| \leq \ell - 1 \leq n - 1$, but $|N^-(x_1)|, |N(x_\ell)| \geq \frac{n}{2}$. So there exists $x_i \in N^-(x_1) \cap N(x_\ell)$. So we can find a cycle $x_i, x_\ell, x_{\ell-1}, \dots, x_{i+1}, x_1, x_2, \dots, x_i$ of length ℓ. $\square$

Remark. Note that there is not an interesting theorem of the form '$e(G) \geq k$ implies G is Hamiltonian', because K_{n-1} adjoined to a single vertex by one edge is not Hamiltonian.

4.2 Paths of a given length

Lemma. Let G be a graph on n vertices, and $n \geq 3$. Let $k < n$. If G is connected and $\delta(G) \geq \frac{k}{2}$, then G contains a path of length k.

Remark. We need the assumption $k < n$, otherwise K_n is a counterexample. We need the assumption that G is connected, otherwise a collection of $\frac{n}{k}$ disjoint graphs give a counterexample if $n \mid k$. The requirement that $\delta(G) \geq \frac{k}{2}$ is sharp, by considering collections of $K_{\frac{k+1}{2}}$ that all intersect in a single vertex.

Proof. Let $x_1, \dots, x_\ell$ be a path of maximum length in G. There is no cycle of length ℓ, because if $\ell = n$ we are done as $k < n$, and if $\ell < n$ we can use a cycle of length ℓ to build a path of length $\ell + 1$ by the same argument from the previous theorem: $N^-(x_1)$ and $N(x_\ell)$ must intersect and so we can build a longer path. $\square$

Theorem. Let G be a graph on n vertices. Then if $e(G) > \frac{n(k-1)}{2}$, G contains a path of length k.

Remark. If $k \mid n$, a collection of $\frac{n}{k}$ disjoint K_k graphs shows that the theorem is sharp.

Proof. Note that if $k = 1$, the theorem clearly holds. Suppose $k \geq 2$, and apply induction on n. The case $n = 2$ holds vacuously. Suppose now we have a graph G on $n \geq 3$ vertices. First note that $\frac{n(k-1)}{2} < e(G) \leq \frac{n(n-1)}{2}$, so $k < n$.

We may assume G is connected without loss of generality, because if it is disconnected, we can apply induction to one of its connected components. Let $C_1, \dots, C_r$ be the components, and $|C_i| = n_i$. Since $\sum_{i=1}^r e(G[C_i]) = e(G) > \frac{n(k-1)}{2}$, we have $\sum_{i=1}^r \left(e(G[C_i]) - \frac{n_i(k-1)}{2}\right) > 0$, so one of the summands is positive. So there exists a connected component C_i such that $e(G[C_i]) > \frac{n_i(k-1)}{2}$, so we can apply induction to this graph to obtain a path of length k as required.

If $\delta(G) \geq \frac{k}{2}$, the proof is complete by the previous lemma. Otherwise, there exists a vertex x of degree less than $\frac{k}{2}$, so $\deg(x) \leq \frac{k-1}{2}$. Note that $e(G - x) > \frac{n(k-1)}{2} - \frac{k-1}{2} = \frac{(n-1)(k-1)}{2}$, so we can apply induction to $G - x$ to obtain a path of length k, completing the proof. $\square$

4.3 Forcing triangles

Proposition (Jensen)**.** Let $a < b$ be real numbers, and $f : [a, b] \to \mathbb{R}$ be a convex function. Let $x_1, \dots, x_n \in [a, b]$. Then, $f\left(\frac{1}{n} \sum_{i=1}^n x_i\right) \leq \frac{1}{n} \sum_{i=1}^n f(x_i)$.

Theorem (Mantel)**.** Let G be a graph on n vertices, and $\frac{n^2}{4} < e(G)$. Then G contains a triangle.

Remark. The bipartite graph $K_{\frac{n}{2}, \frac{n}{2}}$ contains no triangles, and has $\frac{n^2}{4}$ edges, so the above theorem is sharp.

Proof. Suppose the graph contains no triangle. We may assume that $n \geq 3$, otherwise there is nothing to prove. Let $x, y \in V(G)$ such that $x \sim y$. In particular, $\deg x + \deg y \leq n - 2 + 2 = n$. Then, since

$x \mapsto x^2$ is convex,

$$\begin{aligned}
n \cdot e(G) &\geq \sum_{x \sim y} (\deg x + \deg y) \\
&= \frac{1}{2} \sum_x \sum_y (\deg x + \deg y) \mathbb{1}_{x \sim y} \\
&= \sum_x \sum_y \deg x \mathbb{1}_{x \sim y} \\
&= \sum_x \deg x \sum_y \mathbb{1}_{x \sim y} \\
&= \sum_x (\deg x)^2 \\
&= n \left(\frac{1}{n} \sum_x (\deg x)^2 \right) \\
&\geq n \left(\frac{1}{n} \sum_x (\deg x) \right)^2 \\
&= n \left(\frac{2e(G)}{n} \right)^2
\end{aligned}$$

So $e(G) \leq \frac{n^2}{4}$ as required. □

4.4 Forcing cliques

Definition. We say that a graph G is *r-partite* if there is a partition of V into r subsets such that no part contains an edge. Equivalently, G is r-colourable, so $\chi(G) \leq r$.

Definition. Given natural numbers $n_1, \dots, n_r$, define $K_{n_1,\dots,n_r}$ to be the complete r-partite graph with partitions of size $n_1, \dots, n_r$.

Observe that if $r \mid n$, the graph $K_{\frac{n}{r},\dots,\frac{n}{r}}$ is an r-partite graph with $\binom{r}{2}\frac{n^2}{r^2} = \left(1 - \frac{1}{r}\right)\frac{n^2}{2}$ edges.

Theorem (Turán, form 1)**.** Let G be a graph on n vertices, and $\left(1 - \frac{1}{r}\right)\frac{n^2}{2} < e(G)$ for $r \geq 1$. Then G contains a subgraph of the form K_{r+1}, so it has an $(r+1)$-clique.

Proof. Suppose that G has an $(r+1)$-clique. For a given r, we prove the result by induction on n, assuming the theorem holds for all lower values of r, then we can complete the proof by induction.

If $n \leq r$, the result clearly holds. Let G be a graph that contains no $(r+1)$-clique. Suppose $r \geq 2$, otherwise the result is trivial. Then we can find an r-clique by induction on r. Let K be such a clique. Then each vertex in $V(G) \setminus K$ have at most $r-1$ neighbours in K, otherwise, this would be an $(r+1)$-

clique. So

$$e(G) \le \binom{r}{2} + (r-1)(n-r) + e(G \setminus K) \le \binom{r}{2} + (r-1)(n-r) + \left(1 - \frac{1}{r}\right)\frac{(n-1)^2}{2} = \left(1 - \frac{1}{r}\right)\frac{n^2}{2}$$

□

Remark. This is a generalisation of Mantel's theorem. If $r \mid n$, the theorem is sharp by considering the complete r-partite graph.

Definition. The *Turán graph* $T_{r,n}$ is the complete r-partite graph $K_{n_1,\dots,n_r}$ where $\sum_{i=1}^r n_i = n$ and $n_1, \dots, n_r$ differ by at most one.

Proposition. Let G be a r-partite graph on n vertices. Then $e(G) \le e(T_{r,n})$.

Remark. Turán graphs maximise the number of edges among all r-partite graphs on n vertices.

Proof. Let G be an r-partite graph on n vertices with the maximum number of edges. This graph is complete, since if there is a missing edge, there is a graph with more edges. Let $G = K_{n_1,\dots,n_r}$. Suppose that $n_i - n_j \ge 2$, so G is not a Turán graph. Then consider the graph obtained by moving an edge from the part with n_i vertices to the part with n_j vertices. Then we gain a total of $(n_i - 1)$ edges, and remove n_j edges. But this is at least 1, so we have obtained a graph with more edges. □

Theorem (Turán, form 2)**.** Let G be a graph on n vertices and $r \ge 2$. Then if G does not contain an $(r+1)$-clique, $e(G) \le e(T_{r,n})$.

Proof. We will transform a graph G into a complete r-partite graph without decreasing the number of edges. Then, since the Turán graph maximises the amount of edges for such a graph, the result follows.

Let $V(G) = \{1, \dots, n\}$. Let $\alpha_1, \dots, \alpha_r > 0$ be numbers that are linearly independent over $\mathbb{Q}$. For $S \subseteq V(G)$, define $\mu(S) = \sum_{i \in S} \alpha_i$.

If H is a graph on n vertices, we define the transformation of H, denoted $T(H)$, as follows. Let x, y be a pair of vertices maximising $\mu(\{x, y\})$ (to break any ties) such that $N(x) \neq N(y)$ and $x \nsim y$, and also either $\deg x > \deg y$ or both $\deg x = \deg y$ and $\mu(N(x)) > \mu(N(y))$. Now define $T(H)$ to be $H - y$ along with a new vertex x' with $N(x') = N(x)$.

We first show that if H does not contain a K_{r+1}, then $T(H)$ also does not contain a K_{r+1}. Suppose that our new graph H' contains a clique K isomorphic to K_{r+1}. We must have that x' lies inside this clique, because all other vertices remain the same. We know $x \notin K$ since $x \nsim x'$. Then $K \setminus \{x'\} \cup \{x\}$ must be an $(r+1)$-clique in H, which is a contradiction.

Now, consider the sequence $G, T(G), T(T(G)), \dots$, iteratively applying the transformation T. We will now show that this sequence $(T^{(n)}(G))_n$ eventually stabilises. This is because $e(T(H)) \ge e(H)$, so $(e(T^{(n)}(G)))_n$ is an increasing sequence of integers which is bounded above by $\binom{n}{2}$. Note that $\sum_{1 \le x \le n} \mu(N_{T^{(i)}(G)}(x))$ is also an increasing sequence, but since there are only finitely many possible values for this sum, it must also stabilise. Therefore, at some point, the transformation T will do nothing more to our graph. Let G_∞ be the limiting graph in the sequence $(T^{(n)}(G))_n$.

We will show that G_∞ is a complete k-partite graph for some k. Let $k = \chi(G_\infty)$, and c be a k-colouring of G_∞. We write $V(G_\infty) = C_1 \cup \dots \cup C_k$ where C_i is the colour class of vertices with colour i. Note that if $x, y \in C_i$, we have $x \nsim y$, so $N(x) = N(y)$, otherwise the transformation T would have manipulated the neighbourhoods to be equal. Now let $x \in C_i, y \in C_j$ for $i \neq j$. Suppose $x \nsim y$. Then $x' \nsim y'$ for all other $x' \in C_i$ and $y' \in C_j$, so C_i and C_j have no edges between them. But then by merging C_i and C_j, we obtain a more optimal colouring, contradicting our assumption that $k = \chi(G_\infty)$. So G_∞ is a complete k-partite graph for some k.

Since G_∞ does not contain a K_{r+1}, we have $k \leq r$. By the previous proposition, $e(G_\infty) \leq e(T_{r,n})$, and $e(G) \leq e(G_\infty)$ since $e(H) \leq e(T(H))$ for all H. □

4.5 The Zarankiewicz problem

Definition. The *Zarankiewicz number* $Z(n,t)$ is the maximum number of edges in a bipartite graph $G = (X \sqcup Y, E)$ with $|X| = |Y| = n$ such that G does not contain $K_{t,t}$.

Lemma. Let $t \in \mathbb{N}$, and $t \geq 2$. Define the function $f_t(x) = \frac{x(x-1)\dots(x-t+1)}{t!}$. Then $f_t(x)$ is convex for $x \geq t-1$.

Proof. Let $s = x-t+1$, so $f_t(x) = \frac{(s+t-1)(s+t-2)\dots s}{t!}$. This is a polynomial with nonnegative coefficients. Hence it is convex for $s \geq 0$, since $f''(s) \geq 0$. □

Theorem. Let $t \geq 2$. Then $Z(n,t) \leq t^{\frac{1}{t}} n^{2-\frac{1}{t}} + tn$.

Remark. In particular, as n increases, $Z(n,t)$ is eventually lower bounded by $2n^{2-\frac{1}{t}}$.

Proof. Note that we may assume that $\deg y \geq t-1$ for all $y \in Y$. If $\deg y < t-1$, we can add an edge and preserve the property that G contains no $K_{t,t}$.

Let $x_1, \dots, x_t \in X$ be distinct vertices. Then $|N(x_1) \cap \dots \cap N(x_t)| \leq t-1$, otherwise we have a $K_{t,t}$. Now, applying Jensen's inequality,

$$
\begin{aligned}
(t-1)\binom{n}{t} &\geq \sum_{x_1,\dots,x_t \text{ distinct}} |N(x_1) \cap \dots \cap N(x_t)| \\
&= \sum_{x_1,\dots,x_t \text{ distinct}} \sum_y \mathbb{1}_{y\sim x_1} \cdots \mathbb{1}_{y \sim x_t} \\
&= \sum_y \sum_{x_1,\dots,x_t \text{ distinct}} \mathbb{1}_{y\sim x_1} \cdots \mathbb{1}_{y \sim x_t} \\
&= \sum_y \binom{\deg y}{t} \\
&= n\left(\frac{1}{n}\sum_y \binom{\deg y}{t}\right) \\
&\geq n\binom{\overline{d}}{t}
\end{aligned}
$$

where $\overline{d} = \frac{e(G)}{n}$ (since we are in a bipartite graph), using the fact that $\deg y \geq t-1$, so $x \mapsto \binom{x}{t}$ is convex. So

$$
\begin{aligned}
(t-1)\binom{n}{t} &\geq n\binom{\overline{d}}{t} \\
\frac{tn^t}{t!} &\geq \frac{n(\overline{d}-t)^t}{t!} \\
tn^t &\geq n(\overline{d}-t)^t \\
t^{\frac{1}{t}} n^{1-\frac{1}{t}} &\geq \overline{d} - t \\
t^{\frac{1}{t}} n^{1-\frac{1}{t}} &\geq \frac{e(G)}{n} - t \\
e(G) &\leq t^{\frac{1}{t}} n^{2-\frac{1}{t}} + tn
\end{aligned}
$$

□

Remark. If $t = 2$, then it is known that $Z(n,t) \geq cn^{\frac{3}{2}}$ for some contant $c > 0$. If $t = 3$, $Z(n,t) \geq cn^{\frac{5}{3}}$. This is an open problem for $t = 4$.

4.6 Erdős–Stone theorem

Definition. Let H be a fixed graph, and $n \in \mathbb{N}$. Then we define the *extremal number* $\mathrm{ex}(n,H) = \max\{e(G) \mid |G| = n, G \text{ contains no copy of } H\}$.

Example. $\mathrm{ex}(n, K_{r+1}) = e(T_{r,n}) \leq \left(1 - \frac{1}{r}\right)\frac{n^2}{2}$. $\mathrm{ex}(n, P_k) = \frac{n(k-1)}{2}$. $\mathrm{ex}(n, K_{t,t}) \leq 2n^{2-\frac{1}{t}} + tn$.

Theorem. Let H be a fixed nonempty graph. Then

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

Remark. If $\chi(H) \geq 3$, this determines the leading order term in the function $\mathrm{ex}(n,H)$ for large n. If $\chi(H) = 2$, this theorem implies that $\frac{\mathrm{ex}(n,H)}{n^2} \to 0$. But in this case, $H \subseteq K_{t,t}$, and we already know (almost) that $\mathrm{ex}(n,H) \leq cn^{2-\frac{1}{t}}$, which implies the result from the Erdős–Stone theorem. It is easy to see that $\mathrm{ex}(n,H) \geq \left(1 - \frac{1}{\chi(H)-1}\right)\frac{n^2}{2}$, since H is not contained in any $T_{(\chi(H)-1),n}$.

5 Ramsey theory

5.1 Ramsey's theorem

Macroscopically, theorems in Ramsey theory are of the form 'complete disorder in sufficiently large systems is impossible'.

Proposition. Let c be a 2-edge (not proper) colouring of K_6. Then there exists a monochromatic triangle K_3; there exists a subgraph induced on three vertices where all edges have the same colour.

Proof. Suppose our colours are red and blue. Let $x \in V(K_6)$. Without loss of generality, x has three neighbours y_1, y_2, y_3 coloured red. Then the edges between the y_i cannot be coloured red. So they must all be coloured blue, but then this forms a blue triangle. □

Definition. Let $s \geq 2$. Then the *sth Ramsey number*, denoted $R(s)$, is the minimal n such that every 2-edge colouring of K_n contains a monochromatic K_s.

It is not clear *a priori* that such numbers indeed exist.

Definition. Let $s, t \geq 2$. We define $R(s,t)$ be the minimal n such that every 2-edge colouring of K_n contains either a red K_s or a blue K_t.

Remark. $R(s,t)$ is symmetric, and $R(s) = R(s,s)$. Note that $R(2,t)$ is the minimal n that contains a red edge or K_t, so $R(2,t) = t$. We showed above that $R(3,3) = R(3) \leq 6$, and in fact this is an equality by demonstrating a 2-edge colouring of K_5 containing no monochromatic triangle.

Theorem (Ramsey)**.** For all s, t, the Ramsey number $R(s,t)$ exists, and $R(s,t) \leq R(s-1,t) + R(s,t-1)$.

Proof. Apply induction on $s+t$. For $s,t \leq 2$, the result holds. Now suppose $s,t > 2$, and let $a = R(s-1,t), b = R(s,t-1)$. Let $n = a+b = R(s-1,t) + R(s,t-1)$, and consider the complete graph K_n. Let $c : E(K_n) \to \{\text{red, blue}\}$ be a given colouring.

Let $x \in K_n$, and let $N_r(x)$ be the red neighbourhood and $N_b(x)$ be the blue neighbourhood. Suppose that $|N_r(x)| \geq a$. In this case, $N_r(x)$ contains either a red K_{s-1}, in which case $N_r(x) \cup \{x\}$ is a red K_s in K_n; or a blue K_t, in which case we are already done. Now suppose $|N_b(x)| \geq b$. Then $N_b(x)$ contains either a red K_s in which case we are done; or it contains a blue K_{t-1}, in which case $N_b(x) \cup \{x\}$ is a blue K_t in K_n as required. Suppose that neither of these cases occur, so $|N_r(x)| \leq a-1$ and $|N_b(x)| \leq b-1$, so $|N(x)| \leq a+b-2$, which is a contradiction since the graph is complete. □

Corollary. For all s, the Ramsey number $R(s)$ exists.

Definition. $R_k(s_1, \dots, s_k)$ is the minimal n such that every k-edge colouring of K_n contains a K_{s_i} coloured i for some i.

Theorem (multicoloured Ramsey's theorem)**.** For $s_1, \dots, s_k$ for $k \geq 2$, then $R_k(s_1, \dots, s_k)$ exists.

Proof. We will show by induction on k that $R_k(s_1, \dots, s_k) \leq R(s_1, R_{k-1}(s_2, \dots, s_k)) = n$. Let c be a k-colouring of K_n. Apply the two-colour version of Ramsey's theorem to obtain either a K_{s_1} coloured 1, or a $K_{R_{k-1}(s_2,\dots,s_k)}$ coloured in any combination of $2, \dots, k$. If we have a K_{s_1} coloured 1, we are done. Otherwise, apply induction to obtain an edge colouring of $K_{R_{k-1}(s_2,\dots,s_k)}$ to obtain a K_{s_i} coloured i for some $i \geq 2$. □

Remark. We have seen $R(3) = 6$. There are very few known Ramsey numbers. $R(4) = 18$, but $R(5)$ is unknown.

5.2 Infinite graphs and larger sets

Theorem. Let c be a 2-colouring of the countably infinite complete graph, so $c : \mathbb{N}^{(2)} \to \{\text{red}, \text{blue}\}$. Then there exists an infinite set $X \subseteq \mathbb{N}$ which is monochromatic, so $X^{(2)}$ is coloured either entirely red or entirely blue.

Remark. The finite version of Ramsey's theorem cannot be applied here; we can create arbitrarily large cliques, but we do not know if such cliques connect into an infinite set.

Proof. We construct a sequence $x_1, x_2, \dots$ inductively as follows. Let $x_1 \in \mathbb{N}$ be arbitrary. x_1 has either an infinite red neighbourhood or an infinite blue neighbourhood. We define S_1 to be the red neighbourhood of x_1 if it is infinite, or the blue neighbourhood otherwise, so S_1 is infinite. Now let $x_2 \in S_1$. Now, x_2 has either an infinite red neighbourhood in S_1 or an infinite blue neighbourhood in S_1, so we can define S_2 to be one of these that is infinite, and proceed inductively.

For each i, all edges $x_i \sim x_j$ where $i < j$ have the same colour by construction. Label a vertex red if all its forward-facing edges are red, and label an edge blue if all its forward-facing edges are blue. Then there are either infinitely many red vertices or infinitely many blue vertices. Without loss of generality, suppose the set of red vertices X is infinite. Then all edges in X are coloured red, so X is the infinite monochromatic set as required. □

Remark. We can easily construct a version of the above theorem for an arbitrary finite amount of colours, using the same idea as from the multiple-colour version of Ramsey's theorem in the finite case.

Example. It can be difficult to determine which colour has an infinite monochromatic clique. Suppose we colour ij with the maximal n such that $2^n \mid i+j$, modulo 2. The set $\{2^2, 2^4, 2^6, \dots\}$ is an example of an infinite monochromatic clique.

Suppose ij is coloured with the number of distinct prime factors of $i+j$, modulo 2. The colour of the infinite clique is not known.

Remark. It is possible to deduce the existence of $R(s,t)$ from the infinite version.

Theorem. Let c be a 2-colouring of the set of r-sets of $\mathbb{N}$, so $c : \mathbb{N}^{(r)} \to \{\text{red, blue}\}$. Then there exists an infinite set $X \subseteq \mathbb{N}$ such that $X^{(r)}$ is monochromatic.

Proof. Apply induction on r. If $r = 2$, we fall back to the previous theorem. We define a sequence $x_1, x_2, \dots$ and a sequence of infinite sets $S_1, S_2, \dots$ by the following procedure. We start by choosing x_1 arbitrarily. Now, consider the colouring $c_{x_1}(F) = c(\{x_1\} \cup F)$ for $F \in (\mathbb{N} \setminus \{x_1\})^{(r-1)}$. By induction, there exists a set $S_1 \subseteq \mathbb{N} \setminus \{x_1\}$ that is infinite and $S_1^{(r-1)}$ is monochromatic with respect to the colouring c_{x_1}. Now we choose $x_2 \in S_1$, and proceed inductively.

The sequence $x_1, x_2, \dots$ has the property that $F_i = \{\{x_{i_1}, \dots, x_{i_r}\} \mid i_1 < \dots < i_r\}$ are monochromatic for each i. But there are either infinitely many red-coloured x_i or infinitely many blue-coloured x_i. Let X be one of these infinite sets, then $X^{(r)}$ is monochromatic. □

We can produce a similar version of this theorem for the finite case, along with an explicit inductively-defined bound.

Definition. Let $r \in \mathbb{N}$, and $s, t \geq 1$. We define the *r-set Ramsey number* $R^{(r)}(s,t)$ to be the minimal n such that for every 2-colouring of $\{1, \dots, n\}^{(r)}$, it contains either a set S with $|S| = s$ and $S^{(r)}$ are coloured red, or a set T with $|T| = t$ and $T^{(r)}$ are coloured blue.

Remark. $R^{(1)}(s,t) = s + t - 1$. $R^{(2)}(s,t) = R(s,t)$. $R^{(r)}(r,t) = t = R^{(r)}(t,r)$.

Theorem. For all $r, s, t \geq 1$, the number $R^{(r)}(s,t)$ exists.

Proof. Apply induction on r, and then induction on $s + t$. If $s \leq r$ or $t \leq r$, we are done, since $R^{(r)}(r,t) = t$. We claim that $R^{(r)}(s,t) \leq R^{(r-1)}(R^{(r)}(s-1,t) + R^{(r)}(s,t-1)) + 1 = N$.

Consider a 2-coloured set $\{1, \dots, n\}^{(r)}$ where $n \geq N$. Choose a vertex $x \in \{1, \dots, n\}$. Consider the colouring $c_x(F) = c(\{x\} \cup F)$ where $F \in (\{1, \dots, n\} \setminus \{x\})^{(r-1)}$. Applying induction on r, we have a set S_1 such that $|S_1| = R^{(r)}(s-1,t)$ and $S_1^{(r-1)}$ is red, or there is a set S_2 with $|S_2| = R^{(r)}(s,t-1)$ and $S_2^{(r-1)}$ is blue. We consider the first case; the other is similar.

Apply the r-set version of Ramsey's theorem by induction to S_1 to find either a set $A \subseteq S_1$ with $|A| = s-1$ and $A^{(r)}$ is coloured red (with respect to c), or a set $B \subseteq S_2$ with $|B| = t$ and $B^{(r)}$ is coloured blue. If B exists, we are done. If A exists, $A \cup \{x\}$ is coloured red and has size s as required. □

5.3 Upper bounds

Proposition. Let $s, t \geq 2$, we have $R(s,t) \leq \binom{s+t-2}{t-1}$. In particular, $R(s) = R(s,s) \leq 4^s$.

Proof. Apply induction on $s+t$. We know $R(s,2) = s = \binom{s+2-2}{2-1}$ as required. Suppose this holds for $R(s-1,t)$ and $R(s,t-1)$. We have already shown that $R(s,t) \leq R(s-1,t) + R(s,t-1)$. So

$$R(s,t) \leq R(s-1,t) + R(s,t-1) \leq \binom{s+t-2}{s-2} + \binom{s+t-3}{s-1} = \binom{s+t-2}{s-1}$$

□

We are interested in bounding $R^{(r)}(s,t)$. Note that we have the bound $R^{(r)}(s,t) \leq R^{(r-1)}(R^{(r)}(s,t-1), R^{(r)}(s-1,t)) + 1$. Define $f_1(x) = 2x$, and recursively, $f_n(x) = f_{n-1}^x(x)$. Then $f_2(x) \sim 2^x$, and as n increases, f_n increases very rapidly. So our bound on $R^{(r)}(s,t)$ grows asymptotically on the order of $f_r(s+t)$.

5.4 Lower bounds

We can explicitly construct some lower bounds for $R(s)$.

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

Proof. Consider the graph defined by $(s-1)$ disjoint K_{s-1} cliques, all of which are coloured blue, but all lines between cliques are coloured red. This graph has no monochromatic K_s. □

Theorem (Erdős)**.** Let $s \geq 3$. Then $R(s) \geq 2^{\frac{s}{2}}$.

Proof. Consider $G = K_n$ for $n \leq 2^{\frac{s}{2}}$. For each edge e in G, we construct an independent Bernoulli random variable X_e with parameter $\frac{1}{2}$. If $X_e = 0$, we colour e red, and if $X_e = 1$, we colour e blue. Then

$$
\begin{aligned}
\mathbb{P}\left(\text{colouring has a monochromatic } K_s\right) &= \mathbb{P}\left(\bigcup_{K\in\{1,\dots,n\}^{(s)}} \{K \text{ monochromatic}\}\right) \\
&\leq \sum_{K\in\{1,\dots,n\}^{(s)}} \mathbb{P}\left(K \text{ monochromatic}\right) \\
&= \sum_{K\in\{1,\dots,n\}^{(s)}} 2\cdot 2^{-\binom{s}{2}} \\
&= \binom{n}{s} 2\cdot 2^{-\binom{s}{2}} \\
&< \frac{n^s}{s!} 2\cdot 2^{-\frac{s(s-1)}{2}} \\
&= 2\left(\frac{n}{(s!)^{\frac{1}{s}}} 2^{-\frac{s-1}{2}}\right)^s \\
&\leq 2\left(\frac{2^{\frac{1}{2}}}{(s!)^{\frac{1}{s}}}\right)^s
\end{aligned}
$$

Note that $s! \geq 2^{\frac{s}{2}+1}$, so $(s!)^{\frac{1}{s}} \geq 2^{\frac{1}{2}+\frac{1}{s}}$.

$$
\mathbb{P}\left(\text{colouring has a monochromatic } K_s\right) < 2\left(\frac{1}{2^{\frac{1}{s}}}\right)^s \leq 1
$$

Since the probability is less than 1, there must exist a colouring that has no monochromatic K_s. □

Remark. We can think about this proof as follows. Consider the collection of $2^{\binom{n}{2}}$ colourings of K_n. Then for each clique, there are at most $2^{\binom{n}{2}}\cdot 2\cdot 2^{-\binom{s}{2}}$ colourings where that clique is monochromatic. So the collection of all colourings where none of these cliques are monochromatic has at least as many elements as $2^{\binom{n}{2}} - \binom{n}{2}2^{\binom{n}{2}}\cdot 2\cdot 2^{-\binom{s}{2}}$. In general, however, a probabilistic interpretation is more powerful.

Remark. This proof is nonconstructive. It is a major open problem to explicitly construct colourings to show that $R(s) > (1+\varepsilon)^s$.

6 Random graphs

6.1 Lower bounds for Zarankiewicz numbers

Recall the Zarankiewicz numbers $Z(n,t)$, the maximum number of edges between a bipartite graph on (n,n) vertices, before a $K_{t,t}$ is forced. We have shown that $Z(n,t) \leq 2n^{2-\frac{1}{t}}$, but we have found no lower bound.

Theorem. Let $t \geq 2$. Then $Z(n,t) \geq \frac{1}{2} n^{2-\frac{2}{t+1}}$.

Proof excluding the $t+1$ term. Suppose we include each edge in the graph with probability p. Let Z be a random variable that counts the number of $K_{t,t}$ in the bipartite graph G on (n,n) vertices. Then

$$Z = \sum_{A \in X^{(t)}, B \in Y^{(t)}} \mathbb{1}(\text{all edges between } A \text{ and } B \text{ lie in } G)$$

We find

$$\mathbb{E}[Z] = \sum_{A \in X^{(t)}, B \in Y^{(t)}} \mathbb{P}(\text{all edges between } A \text{ and } B \text{ lie in } G) = \binom{n}{t}^2 p^{t^2} \leq \frac{n^{2t}}{4} p^{t^2} = \frac{1}{4}(n^2 p^t)^t$$

So if $p = n^{-\frac{2}{t}}$, then our upper bound is at most $\frac{1}{4}$. Then $\mathbb{P}(X \geq 1) \leq \frac{1}{4}$ by Markov's inequality. Note that $\mathbb{E}[e(G)] = pa^2 = n^{2-\frac{2}{t}}$. So $\mathbb{P}\left(e(G) \leq \frac{pn^2}{2}\right) \leq \frac{1}{2}$. So with probability greater than $\frac{1}{4}$, we have $e(G) > \frac{1}{2}pn^2 = \frac{1}{2}n^{2-\frac{2}{t}}$ and G does not contain a $K_{t,t}$. □

Proof. Let $G = (X \sqcup Y, E)$ be a random bipartite graph with $|X| = |Y| = n$, such that $xy \in E$ with probability $p = n^{-\frac{2}{t+1}}$. Let $\widetilde{G}$ be the graph G with an edge removed from each $K_{t,t}$. By definition, $\widetilde{G}$ has no $K_{t,t}$. Note that $e(\widetilde{G}) \geq e(G) -$(amount of $K_{t,t}$ in G). Taking expectations, $\mathbb{E}\left[e(\widetilde{G})\right] \geq \mathbb{E}[e(G)] - \mathbb{E}\left[\text{amount of } K_{t,t}\right]$. We have $\mathbb{E}[e(G)] = pn^2$, and the expected amount of $K_{t,t}$ subgraphs of G is $\binom{n}{t}^2 p^{t^2}$. Substituting in for p and approximating,

$$\mathbb{E}\left[e(\widetilde{G})\right] \geq n^{2-\frac{2}{t+1}} - \frac{n^{2t}}{2} p^{t^2}$$

Note that

$$n^{2t} p^{t^2} = (n^2 p^t)^t = (n^2 n^{\frac{-2t}{t+1}})^t = (n^{\frac{2(t+1)-2t}{t+1}})^t = n^{\frac{2t}{t+1}} = n^{2-\frac{2}{t+1}}$$

Hence

$$\mathbb{E}\left[e(\widetilde{G})\right] \geq \frac{1}{2} n^{2-\frac{2}{t+1}}$$

So there must exist a graph $\widetilde{G}$ with no $K_{t,t}$ and that has at least $\frac{1}{2}n^{2-\frac{2}{t+1}}$ edges. □

6.2 Girth

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

Proposition (Markov)**.** Let X be a nonnegative random variable. Then for all $t > 0$,

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$$

Proposition. Let G be a graph. Then $\chi(G) \geq \frac{|G|}{\alpha(G)}$, where $\alpha(G)$ is the size of the largest independent set (non-adjacent vertices) in G.

Proof. Let c be a colouring of G with $k = \chi(G)$ colours. Let C_i be the set of vertices coloured i. Then the C_i are each independent sets. We have $|G| = |C_1| + \cdots + |C_k| \leq k\alpha(G) = \chi(G)\alpha(G)$. ☐

Theorem (Erdős)**.** For all $k, g \geq 3$, there exists a graph G with $\chi(G) \geq k$ and girth at least g.

Proof. Let G be a random graph on $\{1, \dots, n\}$ where each edge ij is included with probability $p = n^{-1+\frac{1}{g}}$. Let X_i be the random variable that counts the number of cycles in G of length i. Let $X = X_3 + \cdots + X_{g-1}$. Now, note that $\mathbb{P}\left(X \geq \frac{n}{2}\right) \leq \frac{2}{n}\mathbb{E}\left[X\right]$.

$$\begin{aligned}
\mathbb{E}\left[X\right] &= \sum_{i=3}^{g-1} \mathbb{E}\left[X_i\right] \\
&\leq \sum_{i=3}^{g-1} \frac{n(n-1)\dots(n-i+1)}{i} p^i \\
&\leq \sum_{i=3}^{g-1} (np)^i \\
&= \sum_{i=3}^{g-1} n^{\frac{i}{g}} \\
&\leq cn^{-\frac{1}{g}} < \frac{1}{2}
\end{aligned}$$

for a constant c. Now, let Y be the random variable counting the number of independent sets of $s = \frac{n}{2k}$ vertices (up to rounding).

$$\begin{aligned}
\mathbb{P}\left(Y \geq 1\right) &\leq \mathbb{E}\left[Y\right] \\
&= \binom{n}{s}(1-p)^{\binom{s}{2}} \\
&\leq n^s e^{-p\binom{s}{2}} \\
&= \left(n^2 e^{-p(s-1)}\right)^{\frac{s}{2}} \\
&\leq \left(2n^2 e^{-\frac{n^{\frac{1}{g}}}{2k}}\right)^{\frac{s}{2}} \\
&< \frac{1}{2}
\end{aligned}$$

for n sufficiently large. We have shown that G has at most $\frac{n}{2}$ cycles of length at most $g-1$ with probability at least $\frac{1}{2}$, and G has $\alpha(G) \le \frac{n}{2k}$ with probability at least $\frac{1}{2}$. Hence there is a graph G with both properties. Let $\widetilde{G}$ be G with a vertex deleted from each cycle of length less than g. Then $\widetilde{G}$ has girth at least g. Further,

$$\chi(\widetilde{G}) \ge \frac{|\widetilde{G}|}{\alpha(\widetilde{G})} \ge \frac{\frac{n}{2}}{\alpha(G)} \ge \frac{\frac{n}{2}}{\frac{n}{2k}} = k$$

as required. □

6.3 Binomial random graphs

Definition. The *binomial random graph* on n vertices with parameter $p \in [0,1]$ is the probability space $G(n,p)$ on the graphs on n vertices, where each potential edge is included in the graph independently with probability p.

Let $(a_n), (b_n)$ be sequences of nonnegative numbers, and $b_n \neq 0$ for sufficiently large n. Then we write $a_n \ll b_n$ if $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$. Let X be the random variable that counts the number of triangles K_3 in some random graph $G \sim G(n,p)$. Then $\mathbb{E}[X] = \binom{n}{3}p^3$.

Note that if $p \ll \frac{1}{n}$, so $pn \to 0$, we have $\mathbb{E}[X] \le n^3p^3 \to 0$. By Markov's inequality, $\mathbb{P}(K_3 \subset G) = \mathbb{P}(X \ge 1) \le \mathbb{E}[X] \to 0$.

If $p \gg \frac{1}{n}$, so $pn \to \infty$, then we have $\mathbb{E}[X] \ge \frac{(n-3)^3}{6}p^3 \to \infty$. So asymptotically we have infinitely many triangles. We can also show that $\mathbb{P}(X \ge 1) \to 1$, but this does not follow immediately from the previous result.

Proposition (Chebyshev)**.** Let X be a random variable, and let $t > 0$. Then

$$\mathbb{P}(|X - \mathbb{E}[X]| \ge t) \le \frac{\mathrm{Var}(X)}{t^2}$$

Proposition (second moment method)**.** Let X be a random variable taking values in $\mathbb{N}$. Then

$$\mathbb{P}(X = 0) \le \frac{\mathrm{Var}(X)}{(\mathbb{E}[X])^2}$$

Proof.

$$\mathbb{P}(X = 0) \le \mathbb{P}(|X - \mathbb{E}[X]| \ge \mathbb{E}[X]) \le \frac{\mathrm{Var}(X)}{(\mathbb{E}[X])^2}$$

□

Theorem. Let $G \sim G(n,p)$ be a binomial random graph. Then

$$\lim_{n\to\infty} \mathbb{P}(K_3 \subset G) = \begin{cases} 0 & p \ll \frac{1}{n} \\ 1 & p \gg \frac{1}{n} \end{cases}$$

Proof. Let X be the random variable counting the triangles in G. If $p \ll \frac{1}{n}$, then $\mathbb{E}[X] \to 0$ so $\mathbb{P}(X \geq 1) \to 0$. Now suppose $p \gg \frac{1}{n}$. Let $p \gg \frac{1}{n}$. Now, $\mathbb{P}(X = 0) \leq \frac{\mathrm{Var}(X)}{(\mathbb{E}[X])^2}$. So it suffices to show that $\frac{\mathrm{Var}(X)}{(\mathbb{E}[X])^2} \to 0$. We have

$$\begin{aligned} X &= \sum_{K \in \{0,\dots,n\}^{(3)}} \mathbb{1}(K \text{ is a triangle in } G) \\ X^2 &= \sum_{K \in \{0,\dots,n\}^{(3)}} \sum_{L \in \{0,\dots,n\}^{(3)}} \mathbb{1}(K, L \text{ are triangles in } G) \\ \mathbb{E}\left[X^2\right] &= \sum_{K \in \{0,\dots,n\}^{(3)}} \sum_{L \in \{0,\dots,n\}^{(3)}} \mathbb{P}(K, L \text{ are triangles in } G) \end{aligned}$$

and

$$(\mathbb{E}[X])^2 = \sum_{K \in \{0,\dots,n\}^{(3)}} \sum_{L \in \{0,\dots,n\}^{(3)}} \mathbb{P}(K \text{ is a triangle in } G)\, \mathbb{P}(L \text{ is a triangle in } G)$$

When computing $\mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2$, the only terms that do not cancel are those terms which share edges.

$$\begin{aligned} \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2 \leq & \sum_{K \in \{0,\dots,n\}^{(3)}} \sum_{L \text{ that shares a single edge with } K} \mathbb{P}(K, L \text{ are triangles in } G) \\ &+ \sum_{K \in \{0,\dots,n\}^{(3)}} \mathbb{P}(K \text{ is a triangle in } G) \\ &\leq \sum_{K \in \{0,\dots,n\}^{(3)}} \sum_{L \text{ that shares a single edge with } K} \mathbb{P}(K, L \text{ are triangles in } G) + \mathbb{E}[X] \\ &\leq \underbrace{n^4 p^5}_{four vertices, five edges} + \mathbb{E}[X] \end{aligned}$$

Hence,

$$\frac{\mathrm{Var}(X)}{(\mathbb{E}[X])^2} \leq \frac{n^4 p^5 + \mathbb{E}[X]}{(\mathbb{E}[X])^2} \leq X \frac{n^4 p^5}{(p^3 n^3)^2} + \frac{1}{\mathbb{E}[X]} \leq \frac{1}{pn^2} + \frac{1}{\mathbb{E}[X]} \to 0$$

$\square$

Remark. We see a 'phase transition' from in $\mathbb{P}(K_3 \subset G)$ as p moves from below $\frac{1}{n}$ to above $\frac{1}{n}$. Suppose $p = \frac{\lambda}{n}$ for some fixed $\lambda > 0$. Here, $\lim_{n\to\infty} \mathbb{P}(K_3 \subset G) = 1 - e^{-\frac{\lambda^3}{6}}$, but this result will not be proven.

Remark. We have seen that if the expected number of triangles increases to infinity, then the probability that $G \sim G(n,p)$ contains a triangle converges to 1. However, this is not true in general, replacing 'triangle' with another graph. Consider the graph H defined by a triangle with 1000 extra disjoint vertices. Here, the expected amount of copies of H is $\binom{n}{1003}p^3 \approx \frac{n^{1003}}{1003!}p^3$, which becomes large when $p = n^{-\frac{1003}{3}} < \frac{1}{n}$. If K is the 'densest' subgraph of H, then if the expected amount of copies of K tends to infinity, the probability that G contains a copy of H tends to 1.

6.4 Connectedness

Throughout this section, we will use the inequality $1 - x \leq e^{-x}$.

Proposition. Let $G \sim G(n, p)$. Then, for all $\varepsilon > 0$, we have

$$\lim_{n\to\infty} \mathbb{P}\left(G \text{ has an isolated vertex}\right) = \begin{cases} 0 & p \geq (1+\varepsilon)\frac{\log n}{n} \\ 1 & p \leq (1-\varepsilon)\frac{\log n}{n} \end{cases}$$

where a vertex is *isolated* if its degree is zero.

Proof. Let I be the number of isolated vertices in G. Then,

$$\mathbb{E}\left[I\right] = \sum_{i=1}^{n} \mathbb{P}\left(v_i \text{ is isolated}\right) = \sum_{i=1}^{n} (1-p)^{n-1} = n(1-p)^{n-1}$$

If $p \geq (1+\varepsilon)\frac{\log n}{n}$, then

$$\mathbb{E}\left[I\right] = \frac{n(1-p)^n}{1-p} \leq ne^{-pn} \leq ne^{-(1+\varepsilon)\frac{\log n}{n}n} = ne^{-(1+\varepsilon)\log n} = nn^{-(1+\varepsilon)} = n^{-\varepsilon} \to 0$$

Hence, by Markov's inequality, the probability that G has an isolated vertex is $\mathbb{P}\left(I \geq 1\right) \leq \mathbb{E}\left[I\right] \to 0$. If $p \leq (1-\varepsilon)\frac{\log n}{n}$, then

$$\mathbb{E}\left[I\right] = \frac{n(1-p)^n}{1-p} \geq n(1-p)^n \geq ne^{-\left(1+\frac{\varepsilon}{4}\right)pn}$$

for sufficiently large n, and sufficiently small ε. This statement holds because $1 - p = e^{\log(1-p)}$ and Taylor's theorem implies $\log(1-p) = -p + \frac{p^2}{2} + o(p^2)$. Then

$$\mathbb{E}\left[I\right] \geq ne^{-\left(1+\frac{\varepsilon}{4}\right)(1-\varepsilon)\log n} = nn^{-\left(1+\frac{\varepsilon}{4}\right)(1-\varepsilon)} = nn^{-1+\frac{3\varepsilon}{4}+\frac{\varepsilon^2}{4}} = n^{\frac{3\varepsilon}{4}+\frac{\varepsilon^2}{4}} \to \infty$$

We will apply the second moment method on I. We have $\mathbb{P}(I = 0) \leq \frac{\operatorname{Var}(I)}{(\mathbb{E}[I])^2}$.

$$\begin{aligned}
\operatorname{Var}(I) &= \mathbb{E}\left[I^2\right] - (\mathbb{E}[I])^2 \\
&= \sum_{u,v \in V(G)} \mathbb{P}(d(u) = 0, d(v) = 0) - \sum_{u,v \in V(G)} \mathbb{P}(d(u) = 0)\,\mathbb{P}(d(v) = 0) \\
&\leq \mathbb{E}[I] + \sum_{u \neq v} (\mathbb{P}(d(u) = 0, d(v) = 0) - \mathbb{P}(d(u) = 0)\,\mathbb{P}(d(v) = 0)) \\
&= \mathbb{E}[I] + \sum_{u \neq v} \left((1-p)^{2(n-1)} - (1-p)^{2(n-1)}\right) \\
&\leq \mathbb{E}[I] + n^2(1-p)^{2(n-1)}\left(\frac{1}{1-p} - 1\right) \\
\frac{\operatorname{Var}(I)}{(\mathbb{E}[I])^2} &\leq \frac{1}{\mathbb{E}[I]} \frac{1}{n^2(1-p)^{2(n-1)}} n^2(1-p)^{2(n-1)}\left(\frac{1}{1-p} - 1\right) \\
&\leq \frac{1}{\mathbb{E}[I]} + \frac{1}{1-p} - 1 \to 0
\end{aligned}$$

since $p \to 0$ and $\mathbb{E}[I] \to \infty$ for $p < (1-\varepsilon)\frac{\log n}{n}$, as required. □

Theorem. Let $G \sim G(n, p)$. Then for all $\varepsilon > 0$, we have

$$\lim_{n \to \infty} \mathbb{P}(G \text{ connected}) = \begin{cases} 1 & p \geq (1+\varepsilon)\frac{\log n}{n} \\ 0 & p \leq (1-\varepsilon)\frac{\log n}{n} \end{cases}$$

Remark. This is an example of a *sharp threshold*. Above, we saw the *coarse threshold* $p \gg \frac{1}{n}$ and $p \ll \frac{1}{n}$. Often, sharp thresholds are seen in relation to global properties, and coarse thresholds are seen when analysing local properties.

Proof. Suppose $p \leq (1-\varepsilon)\frac{\log n}{n}$. We want to show that $\lim_{n \to \infty} \mathbb{P}(G \text{ connected})$ converges to zero. This follows from the fact that $\mathbb{P}(G \text{ connected}) \geq \mathbb{P}(G \text{ has no isolated vertex}) \to 0$.

Now suppose $p \geq (1+\varepsilon)\frac{\log n}{n}$. We now want to show that $\lim_{n \to \infty} \mathbb{P}(G \text{ connected})$ converges to one. If G is not connected, we can find $A \subset V(G)$ where $1 \leq |A| \leq \frac{n}{2}$, and there are no edges between A and $V(G) \setminus A$. Consider

$$
\begin{aligned}
\mathbb{P}\left(G \text{ not connected}\right) &= \mathbb{P}\left(\exists A \subset V(G),\, 0 < |A| \leq \frac{n}{2}, e(A, V(G) \setminus A) = 0\right) \\
&= \mathbb{P}\left(\bigcup_{A \subset V(G), 0 < |A| \leq \frac{n}{2}} \{e(A, V(G) \setminus A) = 0\}\right) \\
&\leq \sum_{A \subset V(G),\, 0 < |A| \leq \frac{n}{2}} \mathbb{P}\left(e(A, V(G) \setminus A) = 0\right) \\
&= \sum_{A \subset V(G),\, 0 < |A| \leq \frac{n}{2}} (1-p)^{|A|(n-|A|)} \\
&= \sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} \binom{n}{k} (1-p)^{k(n-k)} \\
&\leq \sum_{k=1}^{\frac{\varepsilon n}{4}} \binom{n}{k} (1-p)^{k(n-k)} + \sum_{k=\frac{\varepsilon n}{4}} \binom{n}{k} (1-p)^{k(n-k)} \\
&\leq \sum_{k=1}^{\frac{\varepsilon n}{4}} n^k e^{-pk(n-k)} + \sum_{k=\frac{\varepsilon n}{4}} \binom{n}{k} (1-p)^{k(n-k)} \\
&\leq \sum_{k=1}^{\frac{\varepsilon n}{4}} \left(n e^{-(1+\varepsilon)\frac{\log n}{n}(n-k)}\right)^k + \sum_{k=\frac{\varepsilon n}{4}} \binom{n}{k} (1-p)^{k(n-k)} \\
&\leq \sum_{k=1}^{\frac{\varepsilon n}{4}} \left(n e^{-(1+\varepsilon)\frac{\log n}{n} n\left(1-\frac{\varepsilon}{4}\right)}\right)^k + \sum_{k=\frac{\varepsilon n}{4}} \binom{n}{k} (1-p)^{k(n-k)} \\
&\leq \sum_{k=1}^{\frac{\varepsilon n}{4}} \left(n e^{-(1+\varepsilon)\log n\left(1-\frac{\varepsilon}{4}\right)}\right)^k + \sum_{k=\frac{\varepsilon n}{4}} \binom{n}{k} (1-p)^{k(n-k)} \\
&\leq \sum_{k=1}^{\frac{\varepsilon n}{4}} \left(n e^{-\log n\left(1+\frac{3\varepsilon}{4}\right)}\right)^k + \sum_{k=\frac{\varepsilon n}{4}} \binom{n}{k} (1-p)^{k(n-k)} \\
&\leq \sum_{k=1}^{\frac{\varepsilon n}{4}} \underbrace{\left(n^{-\left(1+\frac{3\varepsilon}{4}\right)}\right)^k}_{\to 0} + \sum_{k=\frac{\varepsilon n}{4}} \binom{n}{k} (1-p)^{k(n-k)} \\
&\leq \sum_{k=1}^{\frac{\varepsilon n}{4}} \left(n^{-\left(1+\frac{3\varepsilon}{4}\right)}\right)^k + \sum_{k=\frac{\varepsilon n}{4}} 2^n e^{-(1+\varepsilon)\frac{\log n}{n} k(n-k)} \\
&\leq \sum_{k=1}^{\frac{\varepsilon n}{4}} \left(n^{-\left(1+\frac{3\varepsilon}{4}\right)}\right)^k + \sum_{k=\frac{\varepsilon n}{4}} \underbrace{2^n e^{-(1+\varepsilon)\frac{\log n}{n}\frac{\varepsilon n}{4}\frac{n}{2}}}_{\to 0}
\end{aligned}
$$

as required. $\square$

7 Algebraic graph theory

7.1 Graphs of a given diameter

Definition. Let G be a connected graph. The *diameter* of G is

$$\operatorname{diam} G = \max\{d(x, y) \mid x, y \in V(G)\}$$

Remark. The diameter of G is 1 if and only if G is complete, so there are $\binom{n}{2}$ edges.

Proposition. Let G be a graph with diameter at most 2. Then $|G| \leq \Delta(G)^2 + 1$.

Proof. Let $x \in G$. Then $V(G) = \{x\} \cup N(x) \cup N(N(x)) \setminus N(x)$. Hence $|G| \leq 1 + \Delta(G) + \Delta(G)(\Delta(G) - 1) \leq \Delta(G)^2 + 1$. $\square$

Definition. A *Moore graph* is a graph for which $|G| = \Delta(G)^2 + 1$.

Remark. Any Moore graph is regular. Such a graph does not contain a triangle. A graph G is a Moore graph if and only if every distinct $x, y \in V(G)$ have a unique path of length at most 2 between them.

Example. C_5 is a Moore graph with $\Delta(C_5) = 2$. The Petersen graph is a Moore graph with degree 3.

7.2 Adjacency matrices

Definition. The *adjacency matrix* of a graph G on vertex set $\{1, \dots, n\}$ is the $n \times n$ matrix A_G with entries $a_{xy} = \mathbb{1}_{xy \in E(G)}$.

Remark. Adjacency matrices are symmetric and have zero diagonal, hence $\operatorname{tr} A_G = 0$.

Proposition. Let G be a graph, and A_G be its adjacency matrix. Let $k \in \mathbb{N}$. Then $(A_G^k)_{xy}$ is the number of walks of length k from x to y in G.

Proof. If $k = 1$, then the theorem clearly holds. If $k = 2$, then $(A_G^2)_{xy} = \sum_z (A_G)_{xz}(A_G)_{zy} = \sum_z \mathbb{1}_{x \sim z \in E} \mathbb{1}_{z \in y}$ counts the amount of walks of length 2. For $k > 2$, we can proceed by induction. $\square$

A_G acts on $\mathbb{R}^n$ as it is a linear map.

Example. Consider the graph C_4 on vertex set $\{1, 2, 3, 4\}$. This has adjacency matrix

$$A_{C_4} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Let $x = (1, 2, -2, 3)^\intercal$. Then $A_G x = (5, -1, 5, -1)^\intercal$. Note that $(A_G x)_y$ is the sum of x_z for $z \sim y$.

Proposition. Let A be an $n \times n$ symmetric matrix. Then A has real eigenvalues λ_i, and there exists an orthonormal basis u_i where $Au_i = \lambda_i u_i$.

Given a graph G on n vertices, we can now consider its eigenvalues and eigenvectors, which are the eigenvalues and eigenvectors of A_G. Let $\lambda_{\max} = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n = \lambda_{\min}$ without loss of generality. Since $\sum_{i=1}^n \lambda_i = \operatorname{tr} A_G = 0$, if G is a nonempty graph, $\lambda_{\max} > 0$ and $\lambda_{\min} < 0$.

Example. $(1, 1, 1, 1)^\intercal$ is an eigenvector of C_4 with eigenvalue 2. Note that the rank of A_G is 2, so there are two zero eigenvalues. Since the eigenvalues sum to zero, $\lambda_{\min} = -2$. One example of a corresponding eigenvector is $(1, -1, 1, -1)^\intercal$.

Proposition. Let A be a symmetric $n \times n$ matrix. Then

$$\lambda_{\max} = \max_{x \in \mathbb{R}^n \setminus \{0\}} \left(\frac{\langle x, Ax \rangle}{\langle x, x \rangle} \right); \quad \lambda_{\min} = \min_{x \in \mathbb{R}^n \setminus \{0\}} \left(\frac{\langle x, Ax \rangle}{\langle x, x \rangle} \right)$$

Proposition. Let G be a graph.

(i) If λ is an eigenvalue, then $|\lambda| \leq \Delta(G)$.

(ii) If G is connected, then $\Delta(G)$ is an eigenvalue if and only if G is regular. In this case, $\mathbb{1} = (1, \dots, 1)$ is the corresponding eigenvector, and $\Delta(G)$ has multiplicity 1.

(iii) If G is connected, then $-\Delta(G)$ is an eigenvalue if and only if G is regular and bipartite.

(iv) $\lambda_{\max} \geq \delta(G)$.

Proof. Part (i). Let λ be an eigenvalue for G. Let $x = (x_1, \dots, x_n)$ be a corresponding eigenvector. Let x_i be the entry with largest absolute value. We may assume that $x_i = 1$. Then, $\lambda x = Ax$ gives

$$\lambda = \lambda x_i = (\lambda x)_i = (Ax)_i = \sum_{j \sim i} x_j \implies |\lambda| \leq \left| \sum_{j \sim i} x_j \right| \leq \Delta(G)$$

Part (ii). Suppose G is regular. Then observe that $\mathbb{1} = (1, \dots, 1)$ is an eigenvector of G with eigenvalue $\delta(G) = \Delta(G)$. Now suppose $\Delta(G)$ is an eigenvalue. Let $x = (x_1, \dots, x_n)$ be a corresponding eigenvector and let x_i be the entry with largest absolute value. Without loss of generality let $x_i = 1$. We have $\Delta(G) = \Delta(G) x_i = \sum_{j \sim i} x_j$, so $\deg i = \Delta(G)$, and if $j \sim i$, then $x_j = 1$. Proceeding inductively, since the graph is connected, all x_j are equal to 1, and all vertices have degree $\Delta(G)$. So $x = \mathbb{1}$ as required. Since this is the only possible eigenvector with eigenvalue $\Delta(G)$, and A_G is symmetric, the multiplicity of the eigenvalue $\Delta(G)$ is 1.

Part (iii). Suppose G is bipartite and regular. Let $V(G) = X \sqcup Y$, and consider the vector given by $x_i = 1$ if $i \in X$ and $x_i = -1$ if $i \in Y$. Then $Ax = -\Delta(G)x$ as required. Now suppose $-\Delta(G)$ is an eigenvalue. As before, let x be an eigenvector with $x_i = 1$ of maximal absolute value. We have $-\Delta(G) = -\Delta(G)x_i = \sum_{j \sim i} x_j$, hence $\deg i = \Delta(G)$, and if $j \sim i$, we have $x_j = -1$. Since G is connected, we repeat the process to show that G is $\Delta(G)$-regular, and x_j is either $+1$ or -1 giving a natural bipartition of the graph.

Part (iv). Note that

$$\lambda_{\max} = \max_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\langle x, Ax \rangle}{\langle x, x \rangle}$$

Consider $x = \mathbb{1} = (1, \dots, 1)$. Then

$$\lambda_{\max} \geq \frac{\langle \mathbb{1}, A\mathbb{1} \rangle}{\langle \mathbb{1}, \mathbb{1} \rangle} = \frac{1}{n} \sum_{i=1}^{n} \deg(i) \geq \delta(G)$$

□

7.3 Strongly regular graphs

Definition. A graph G is (k, a, b)-*strongly regular* if

(i) G is k-regular;

(ii) for every pair of adjacent vertices $x \sim y$, they have exactly a common neighbours, so $|N(x) \cap N(y)| = a$;

(iii) for every pair of not equal and non-adjacent vertices $x \nsim y$, they have exactly b common neighbours, so $|N(x) \cap N(y)| = b$.

Example. C_4 is $(2, 0, 2)$-strongly regular. C_5 is $(2, 0, 1)$-strongly regular. Any Moore graph is $(\Delta(K), 0, 1)$-strongly regular.

Theorem (strongly regular graphs are rare)**.** Let G be a (k, a, b)-strongly regular graph on n vertices. Then,

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

are integers.

Proof. Let A be the adjacency matrix of G. Then

$$(A^2)_{xy} = \begin{cases} a & x \sim y \\ b & x \neq y, x \nsim y \\ k & x = y \end{cases} \implies A^2 = aA + b(J - I - A) + kI$$

where J is the matrix with $J_{xy} = 1$ for all x, y. Hence, $A^2 + (b-a)A + (b-k)I - bJ = 0$. We know that k is an eigenvalue of A, and the corresponding eigenvector is $\mathbb{1}$. Since G is connected, k has multiplicity 1.

Let λ be an eigenvalue of A such that $\lambda \neq k$. Let x be the corresponding eigenvector. Applying the matrix equation to x, we obtain $\lambda^2 x + (b-a)\lambda x + (b-k)x = 0$ as $Jx = 0$, as x is orthogonal to $\mathbb{1}$. Then $\lambda^2 + (b-a)\lambda + (b-k)$ as $x \neq 0$. Hence,

$$\lambda = \frac{(a-b) \pm \sqrt{(a-b)^2 + 4(k-b)}}{2}$$

In particular, there are only three possible eigenvalues for A, which are k and the two possible solutions to the quadratic equation for λ. Let λ, μ be the solutions to the above equation. Let λ have multiplicity s and μ have multiplicity t. Then,

$$0 = \operatorname{tr} A = \sum_{i=1}^{n} \lambda_i = s\lambda + t\mu + k$$

We also have $s + t + 1 = n$, since there are n eigenvalues. Solving both equations simultaneously, we obtain the result as desired. $\square$

Corollary. Let G be a Moore graph with $\Delta(G) = k$. Then $k \in \{2, 3, 7, 57\}$.

Proof. If G is a Moore graph, it is $(k, 0, 1)$-strongly regular on $k^2 + 1$ vertices. Then, one can check the condition in the previous theorem.

$$\frac{1}{2}\left(k^2 \pm \frac{k^2 - 2k}{\sqrt{4k - 3}}\right) \in \mathbb{Z}$$

$\square$

Remark. It is not known if such a graph G with $k = 57$ exists.