Part II - Graph Theory

Lectured by Dr J. Sahasrabudhe

Artur Avameri

Michaelmas 2022

Contents

0 Introduction

07 Oct 2022, Lecture 1

Notation. We write $[n]$ for $\{1, 2, \ldots, n\}$. For a set X and $k \in \mathbb{N}$, define $X^{(k)} = \{S \subset X \mid |S| = k\}$, i.e. the set of all subsets of size k.

1 Fundamentals

Definition 1.1. A **graph** is an object $G = (V, E)$ where V is a set and $E \subseteq V^{(2)}$.

V is the set of vertices, and E is the set of edges.

$V(G)$ will denote V, $E(G)$ will denote E, and we define $|G| = |V(G)|$ (sometimes called the order) and $e(G) = |E(G)|$ (sometimes called the size).

Example 1.1. The **complete graph** on n vertices is denoted K_n. This is the graph where $V(K_n) = [n]$ and $E(K_n) = [n]^{(2)}$.

Remark. We assume the following:

- We don't order edges;
- We don't allow loops (an edge joining a vertex to itself);
- We don't allow multiple edges;
- Most of the time, $V(G)$ will be finite (we will explicitly say when it's not).

Example 1.2. The **empty graph** on n vertices, denoted $\overline{K_n}$, has $V(\overline{K_n}) = [n]$ and $E(\overline{K_n}) = \emptyset$.

Example 1.3. The path of length n, denoted P_n, is a path: it has $V(P_n) = [n+1]$ and $E(P_n) = \{\{i, i+1\} \mid 1 \leq i \leq n\}$.

Example 1.4. The cycle of length n, denoted C_n, has $V(C_n) = [n]$ and $E(C_n) = \{\{i, i+1\} \mid 1 \leq i \leq n-1\} \cup \{\{1, n\}\}$.

Let G be a graph and $x \in V(G)$. The **neighborhood** of x is $N(x) = \{y \mid xy \in E(G)\}$, i.e. all the vertices connected to x. If $y \in N(x)$, we write $x \sim y$ and say y is a **neighbor** of x or that y is **adjacent** to x.

The **degree** of x is $\deg(x) = |N(x)|$.

Just as a formality, we define graph isomorphism: let G, H be graphs. A graph isomorphism is a bijection $\phi : V(G) \to V(H)$ such that it maps edges to edges, i.e. $uv \in E(G) \iff \phi(u)\phi(v) \in E(H)$.

Definition 1.2 (Subgraph)**.** We say H is a **subgraph** of G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$.

Two subgraph types that are important enough to have their own notation:

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

Definition 1.3 (Path)**.** Let G be a graph, $x, y \in V(G)$. A $x - y$ **path in** G is a sequence $x_1, \ldots, x_k$ where $x_1 = x,\ x_k = y$ and $x_i x_{i+1} \in E(G)\ \forall 1 \leq i \leq k-1$ and all the x_i are distinct.

Definition 1.4. A graph is **connected** if $\forall x \neq y \in V(G)$, there exists an $x - y$ path in G.

Remark. A little annoyingly, if P is a $x - y$ path and P' is a $y - z$ path, then the concatenation PP' may not be a path (since the vertices of the new path might not be unique).

So let an $x - y$ **walk** in a graph G be a sequence $x_1, \ldots, x_k$ where $x_1 = x$, $x_k = y$ and $x_i x_{i+1} \in E(G)\ \forall 1 \leq i \leq k-1$. Then a concatenation of walks is again a walk.

Proposition 1.1. If W is an xy walk, then W contains a xy path.

Proof. Let $W' \subseteq W$ be a minimal xy walk. We claim this is a path. If not, then some vertex x_i must be visited at least twice, say $W' = x_1 x_2 \ldots x_i \ldots x_i x_l \ldots x_k$. Then take $W'' = x_1 x_2 \ldots x_i x_l \ldots x_k$. This contradicts the minimality of W', so we're done. $\square$

10 Oct 2022, Lecture 2

Remark. We may define a **distance** on $V(G)$: for $x, y \in V(G)$, let $d(x, y)$ be the length of the shortest xy path. If G is connected, then this distance defines a metric on $V(G)$.

1.1 Trees

Definition 1.5. A graph G is **acyclic** if it does not contain a cycle as a subgraph.

Definition 1.6. A graph G is a **tree** if it is acyclic and connected.

Proposition 1.2. The following are equivalent:

1. G is a tree;
2. G is minimally connected ($\forall xy \in E(G)$, $G - xy$ is not connected);
3. G is maximally acyclic ($\forall xy \notin E(G)$, $G + xy$ contains a cycle).

Proof. (a) $\implies$ (b): A tree is connected. Assume for contradiction that $\exists xy \in E(G)$ such that $G - xy$ is connected. Let P be a xy path in $G - xy$. But then P defines a cycle in G, contradiction.

(b) $\implies$ (a): Minimally connected implies connected. For acyclicness, assume for contradiction that G contains a cycle C. Let $xy \in E(C)$. We claim that $G - xy$ is connected. Choose $u \neq v \in V(G - xy)$. Let P be a uv path in G. If P does not contain xy, we're done. If P does contain xy, then take paths $u \to x$; $x \to y$ along our cycle without using xy; $y \to v$. The concatenation of these gives a uv walk, which contains a uv path. Hence $G - xy$ is connected, contradiction.

(a) $\implies$ (c): A tree is acyclic. Let $xy \notin E(G), x \neq y$. Let P be a xy path. Then P defines a cycle in $G + xy$.

(c) $\implies$ (a): We have acyclicity. If G is not connected, $\exists x \neq y \in V(G)$ with no xy path. Then $G + xy$ is acyclic. □

Definition 1.7. If T is a tree and $v \in V(T)$ with $\deg(v) = 1$, we call v a **leaf**.

Definition 1.8. Let G be a graph and $X \subseteq V(G)$. Then

$$G[X] = (X, E(G) \cap \{(u, v) \mid u, v \in X\}).$$

We call this **the graph induced on** X.

Definition 1.9. If $x \in V(G)$, define $G - x = G[V(G) \setminus \{x\}]$.

Proposition 1.3. Let T be a tree, $|T| \geq 2$. Then T has a leaf.

Proof. Let $P = x_1 \dots x_k$ be the a longest possible path in T. Note $N(x_k) \subseteq \{x_1, \dots, x_{k-1}\}$. If $x_i \sim x_k$ for some $1 \leq i \leq k - 2$, there is a cycle in T, contradiction. Thus $N(x_k) = \{x_{k-1}\} \implies X_k$ is a leaf. □

Remark. We can show that any T has two leaves, but we can't do any better (consider a path).

Remark. We could have also proved this by taking a non-backtracking walk in G (i.e. go out of a different edge then you came in on). Either we return (hence have a cycle) or get stuck (hence have a leaf).

Proposition 1.4. Let T be a tree on $n \geq 1$ vertices. Then $e(G) = n - 1$.

Proof. By induction. $n = 1$ is trivial. Assume the claim holds for n. Take a tree T with $n + 1$ vertices. Let $x \in V(T)$ be a leaf. Then $T - x$ is connected and acyclic, therefore a tree, thus $e(T - x) = n - 1$. But $e(G) = e(G - x) + 1$ and $|V(G)| = |V(G - x)| + 1$, hence we're done. □

Definition 1.10. Let G be a connected graph. Then a subgraph T of G is a **spanning tree** if T is a tree on $V(G)$.

Proposition 1.5. Every connected graph contains a spanning tree.

Proof. Start with the graph G, then throw away edges of $E(G)$ one by one subject to keeping the graph connected. At some point, the removal of any further edge will disconnect the graph, at which point we have a minimal connected subgraph of G, which by Prop. 1.2 is a tree. $\square$

1.2 Bipartite graphs

Definition 1.11. Let $G = (V, E)$ be a graph. G is **bipartite** if there exists a partition $V = A \cup B$ such that $E(G) \subseteq \{uv \mid u \in A, v \in B\}$.

Definition 1.12. The **complete bipartite graph** $K_{n,m}$ is the graph with vertex set $A \cup B$, $A = \{x_1, \ldots, x_n\}, B = \{y_1, \ldots, y_m\}$ and edge set $E(K_{n,m}) = \{x_i y_j \mid x_i \in A, y_j \in B\}$.

Remark. There obviously exist non-bipartite graphs: odd cycles are not bipartite.

Definition 1.13. A **circuit** is a sequence $x_1, x_2, \ldots x_l x_{l+1}$, where $x_i x_{i+1} \in E(G)$ and $x_{l+1} = x_1$. The length of this circuit is l. We say a circuit is **odd** if its length is odd.

Proposition 1.6. Let C be an odd circuit in a graph G. Then C contains an odd cycle.

Proof. Let $x_1 x_2 \ldots x_i x_{i+1} \ldots x_i x_k \ldots x_l x_1$ be an odd circuit. Consider the circuits $C_1 = x_1 \ldots x_i x_k \ldots x_l x_1$ and $C_2 = x_i x_{i+1} \ldots x_{k-2} x_i$. Then one of C_1, C_2 has odd length and is strictly shorter, so we're done by induction. $\square$

Theorem 1.7. Let G be a graph. Then

$$G \text{ is bipartite} \iff G \text{ does not contain an odd cycle.}$$

Proof. ($\implies$): If G contains an odd cycle, then as odd cycles are not bipartite, G cannot be bipartite.

($\impliedby$): We may assume that G is connected. Let us fix $x_0 \in V(G)$. Let

$$V_0 = \{x \in V(G) \mid d(x, x_0) \equiv 0 \pmod 2\}$$
$$V_1 = \{x \in V(G) \mid d(x, x_0) \equiv 1 \pmod 2\}.$$

We claim this is a bipartition of G. Assume for contradiction that $\exists u, v \in V_0$ s.t. $uv \in E(G)$. But there is an even ux_0 path and and an even vx_0 path, thus putting these three paths together gives an odd circuit in G. By Prop 1.6, G contains an odd cycle, contradiction. (Analogous proof for V_1). $\square$

12 Oct 2022, Lecture 3

1.3 Planar graphs

Definition 1.14. A **planar graph** is a graph that can be drawn in the plane with no edge crossings.

Example 1.5. K_4 is planar. A path P_n is planar.

Definition 1.15. A **plane graph** is one such drawing of a graph in the plane without edge crossings.

Note that this is important, since we can draw K_4 in a way that it does have edges crossing.

Example 1.6. $K_{2,3}$ is planar. $K_{3,3}$ is not planar. K_5 is not planar (we don't prove this right now).

Question. What graphs are planar? Is there a (simple) method to decide if a graph is planar?

Definition 1.16. Let G be a plane graph. Consider $\mathbb{R}^2 \setminus G$. This is broken into finitely many regions. These are called the **faces** of the plane graph.

Definition 1.17. The **boundary** of a face F is the collection of vertices and edges on the topological boundary.

Remark. The boundary of a face need not be a cycle. It need also not be connected. In fact, the boundary need not contain a cycle at all (consider a tree).

Remark. We also note that two different drawings of a graph in the plane can be fundamentally different.

Theorem 1.8 (Euler)**.** Let G be a connected plane graph with n vertices, m edges and f faces. Then $n - m + f = 2$.

Proof. We induct on m. $m = 1$ is clear. If G is acyclic, then G is a tree, so $m = n - 1$, $f = 1$ and we're done.

So assume G contains a cycle and let e be an edge on this cycle. Delete e. Then n stays fixed, m decreases by 1, and f decreases by 1, so by induction, $n - (m - 1) + (f - 1) = 2$ and we're done. $\square$

Remark. We really do need the graph to be connected, consider t triangles in the plane as a counterexample.

Corollary 1.9. Let G be a planar graph, $|G| \geq 3$. Then $e(G) \leq 3|G| - 6$.

Proof. Draw G in the plane. We may assume that G is connected. Let F be a face, let $\deg(F)$ = the number of edges in G that touch F. Note $\deg(F) \geq 3$. Now note that since every edge touches at most two faces, we get

$$3f \leq \sum_{F \text{ a face}} \deg(F) \leq 2e(G) \implies f \leq \frac{2}{3}e(G).$$

Put this into Euler's formula to get

$$n - \frac{1}{3}e(G) \geq n - e(G) + f = 2 \implies 3(n-2) \geq e(G).$$

□

Remarks. (i): This is a statement about planar graphs only.
(ii): This is quite restrictive. K_n has $\binom{n}{2} \approx n^2/2$ edges, while our above corollary says the number of edges of a planar graph is linear in n.

Corollary 1.10. K_5 is not planar.

Proof. We have $e(K_5) = 10, n = 5$, so $10e(G) \not\leq 3|G| - 6 = 9$, so we're done by the above corollary. □

But $K_{3,3}$ does not fail this test. So we need to improve our argument:

Corollary 1.11. Let G be a planar graph, $|G| \geq 4$ and G has no cycles of length 3. Then $e(G) \leq 2|G| - 4$.

Proof. Repeat the proof of Corollary 1.9, but use $\deg(F) \geq 4$ for every face. □

Now we can see that $K_{3,3}$ is not planar. $K_{3,3}$ has no cycle of length 3 by definition, $n = 6$, $e(G) = 9$, so $9 = e(G) \not\leq 2 \cdot (6-2) = 8$.

14 Oct 2022, Lecture 4

Definition 1.18. A **subdivision** of a graph G is a subgraph where we replace some of the edges of G with disjoint paths.

Observation. A subdivision of a **non-planar** graph is non-planar.
Observation. If G contains a $K_{3,3}$ or K_5 subdivision as a subgraph, then G is non-planar.

Theorem 1.12 (Kuratowski's theorem)**.** G is planar $\iff$ G does not contain a subdivided $K_{3,3}$ or K_5.

We do not prove this, but the proof is actually not too hard.

2 Connectivity & matching

2.1 Matching in bipartite graphs

Let $G = (X \sqcup Y, E)$ be bipartite with bipartition X, Y.

Definition 2.1. A **matching from** X **to** Y is a set of edges $\{xy_x \mid x \in X, y_x \in Y\}$ and $x \to y_x$ is an injection.

Question. When does a bipartite graph have a X to Y matching?

We can first think about examples where we do not have a matching. For example, we clearly have no matching if $|X| > |Y|$.

Definition 2.2. Let G be a graph, $A \subseteq V(G)$. Define $N_G(A) = \bigcup_{x \in A} N(x)$.

Then we clearly also don't have a matching if we have $A \subset X$ such that $|N(A)| < |A|$. But this is actually the only obstruction:

Theorem 2.1 (Hall's Marriage Theorem)**.** Let G be a bipartite graph $G = (X \sqcup Y, E)$. Then

$$G \text{ has a matching from } X \text{ to } Y \iff \forall A \subseteq X,\ |N(A)| \geq A.$$

The right-hand side is called Hall's criterion.

Proof. ($\implies$) is the easy direction.

Now let M be a matching and let $A \subseteq X$. Then if $\{y_1, \ldots, y_{|A|}\}$ are matched to A, we show $|N(A)| \geq |\{y_1, \ldots, y_{|A|}\}| \geq |A|$.

($\impliedby$): Apply induction on $|X|$. If $|X| = 1$, we're done. For the induction step, consider the following question: is there $\emptyset \neq A \subsetneq X$ such that $|N(A)| = |A|$?

If the answer is no, then $\forall A \subsetneq X$ we have $|N(A)| \geq |A| + 1$. Let $xy \in E(G)$ and let $G' = G[X \setminus \{x\} \cup Y \setminus \{y\}]$. We now check Hall's criterion for G'. If $B \subseteq X \setminus \{x\}$, then $|N_{G'}(B)| \geq |N_G(B)| - 1 \geq |B|$, so done by induction.

If the answer is yes, then let $G_1 = G[A \cup N(A)]$ and $G_2 = G[X \setminus A \cup Y \setminus N(A)]$.

Claim 1: G_1 satisfies Hall's criterion. Let $B \subseteq A$, then

$$|N_{G_1}(B)| = |N_G(B)| \geq B.$$

Claim 2: G_2 satisfies Hall's criterion. Let $B \subset X \setminus A$. Consider $N_G(A \cup B)$. One the one hand, $|N_G(A \cup B)| \geq |A| + |B|$. On the other hand, $|N_G(A)| + |N_{G_2}(B)| = |N_G(A \cup B)|$. As $|N(A)| = |A|$, we get $|N_{G_2}(B)| \geq |B|$.

From claims 1 and 2 we can apply induction in G_1, G_2 to get a matching in these graphs, and then put them together to get a matching in G. □

Definition 2.3. A matching of deficiency of d from X to Y is a matching from X' to Y where $X' \subseteq X$, $|X| - d = |X'|$.

Theorem 2.2 (Defect Hall's Theorem)**.**

$$G \text{ contains a matching of deficiency } d \iff \forall A \subseteq X,\ |N(A)| \geq |A| - d.$$

Proof. ($\implies$) : easy.

($\impliedby$) : Add d phantom vertices to Y, which we join to all vertices in X, so we now satisfy Hall's condition. Apply Hall to get a matching, and then remove the d vertices we added, which removes at most d elements of X. $\square$

Definition 2.4. Let G be a graph. The **minimum degree** in G is $\delta(G) = \min_{x \in V(G)} d(x)$, and the **maximal degree** in G is $\Delta(G) = \max_{x \in V(G)} d(x)$.

Definition 2.5. A graph is **regular** if $\delta(G) = \Delta(G)$. It is k**-regular** if $k = \delta(G) = \Delta(G)$.

17 Oct 2022, Lecture 5

Corollary 2.3. For $k \geq 1$, if $G = (X \sqcup Y, E)$ is a k-regular bipartite graph, then there exists a matching from X to Y.

Proof. We check Hall's criterion. Let $A \subseteq X$. On the one hand,

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

On the other hand,

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

Hence $|N(A)| \geq |A|$ and we're done. $\square$

Let Γ be a finite group and let H be a subgroup of Γ. Let $L_1, \dots, L_n$ be the set of left cosets and $R_1, \dots, R_n$ be the right cosets (of the forms gH and Hg respectively).

Question. Is there $g_1, \dots, g_n \in \Gamma$ such that $g_1H, \dots, g_nH$ are the left cosets and $Hg_1, \dots, Hg_n$ are the right cosets?

Corollary 2.4. There exist $g_1, \dots, g_n \in \Gamma \geq H$ such that $g_1H, \dots, g_nH$ are the left cosets and $Hg_1, \dots, Hg_n$ are the right cosets.

Proof. It is enough to find a pairing $L_i \leftrightarrow R_{\sigma(i)}$ such that $L_i \cap R_{\sigma(i)} \neq \emptyset\ \forall i$. Then choose $g_i \in L_i \cap R_{\sigma(i)}$ and we have $g_iH = L_i$, $Hg_i = R_{\sigma(i)}$.

Define $X = \{R_1, \ldots, R_n\}$ and $Y = \{L_1, \ldots, L_n\}$, and define $R_i \sim L_j$ when $R_i \cap L_j \neq \emptyset \; \forall i, j$. Let $A = \{r_{i_1}, \ldots, R_{i_k}\}$. Note

$$\left| \bigcup_{j=1}^{k} R_{i_j} \right| = k|H|.$$

But $L_1, \ldots, L_n$ partition Γ and $|L_i| = |H|$, so at least k left cosets must intersect $\bigcup R_{i_j}$. Thus Hall's criterion is satisfied and we're done. $\square$

2.2 Connectivity

For a tree, $G - x$ (where x is any non-leaf) is disconnected. On the other hand, remove any 2 vertices from the Petersen graph and it stays connected (but if you remove any 3, you disconnect it).

Notation. Let $S \subseteq V(G)$, and let $G - S = G[V(G) \setminus S]$.

Definition 2.6. Let G be a graph, $|G| \geq 1$. Define

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

We say a graph G is k**-connected** if $\kappa(G) \geq k$.

In other words, G is k-connected if and only if $G - S$ is connected for all $S \subseteq V(G), |S| \leq k - 1$.

Example 2.1.

- $\kappa(\text{Tree}) = 1$.
- $\kappa(\text{Petersen graph}) = 3$, so we can say the Petersen graph is 3-connected.
- $\kappa(\text{Cycle}) = 2$.
- $\kappa(K_n) = n - 1$.

We have another natural definition of connectivity.

Definition 2.7. Let G be a graph and let $a, b \in V(G)$. Say that ab paths $P_1, \ldots, P_k$ are **disjoint** if $V(P_i) \cap V(P_j) = \{a, b\} \; \forall i \neq j$.

Amazingly, we have Menger's theorem: These two notions of connectivity (# of disjoint paths and $\kappa(G)$) are equivalent.

Remarks:

- We have $\delta(G) \geq \kappa(G)$. To see this, delete $N(x)$ for $x \in V(G)$ of minimal degree, then $G - N(x)$ is disconnected (or a single vertex).
- We have $\kappa(G - x) \geq \kappa(G) - 1$. This is clear: if $S \subset V(G - x)$ disconnects $G - x$ with $|S| \leq \kappa(G) - 2$, then $S \cup \{x\}$ disconnects G, contradiction.

- We can have $\kappa(G-x) > \kappa(G)$. For example, a cycle is 2-connected, but a cycle with one protruding edge is 1-connected.

Definition 2.8. A **component** in G is a maximal connected subgraph.

Definition 2.9. Let G be a graph, let $a, b \in V(G), a \neq b, a \not\sim b$. Say $S \subseteq V(G) \setminus \{a, b\}$is a $a-b$ **separator** if $G-S$ disconnects a from b (i.e. a, b are in different components of $G-S$).

Theorem 2.5 (Menger's theorem, form 1)**.** Let G be a connected graph and fix $a, b \in V(G), a \neq b, a \not\sim b$. Then the minimum size of an $a-b$ separator is equal to the maximal number of disjoint paths from a to b.

In other words, if all $a-b$ separators have size $\geq k$, then there exist $P_1, \ldots, P_k$, disjoint paths between a and b.

19 Oct 2022, Lecture 6

Note. Define $\kappa_{a,b}(G)$ be the size of the minimal $a-b$ separator.

Note. Recall $\kappa(G-x) \geq \kappa(G)-1$, and also $\kappa(G-xy) \geq \kappa(G)-1$. We also have $\kappa_{a,b}(G-x) \geq \kappa_{a,b}(G)-1$ and $\kappa_{a,b}(G-xy) \geq \kappa_{a,b}(G)-1$ (exercise, not hard).

Proof. Assume for contradiction that the statement of the theorem is false. Let G be a minimal counterexample to the theorem that

(a) minimizes k;

(b) subject to (a), choose G to minimize $e(G)$.

Now let S be a minimal a, b separator in G. We have $|S| = k$. Note that the theorem is true for $k = 1$, so assume $k \geq 2$.

If $S \neq N(A)$ and $S \neq N(B)$, consider $G - S$ and let A be the component containing a and B be the component containing B.

Define $G_a = G[A \cup S]$ along with a vertex c joined to each vertex in S, and $G_b = G[B \cup S]$ along with a vertex c joined to each vertex in S. Note that $\kappa_{a,c}(G_a) \geq k$, since any $a-c$ separator in G_a is a a, b separator in G. Likewise, $\kappa_{b,c}(G_b) \geq k$.

Note that $e(G_a) < e(G), e(G_b) < e(G)$ since $N(a) \not\subset S, N(b) \not\subset S$. So there exists a neighbor x of b in B with $\deg(x) \geq 2$, else we can remove x and apply minimality.

So by minimality of G, we can find k disjoint a, c paths, say $P_1, \ldots, P_k$ in G_a, and likewise we can find k b, c paths $Q_1, \ldots, Q_k$ in G_b. We can put these paths together to get paths $P_1 Q_{\sigma(1)}, \ldots, P_k Q_{\sigma(k)}$, which are k disjoint a, b paths, contradiction, done.

Let us now assume WLOG that $S = N(a)$.

Claim: $N(a) \cap N(b) = \emptyset$.

Indeed, if $\exists x \in N(a) \cap N(b)$, consider $G - x$. We have $\kappa_{a,b}(G - x) \geq k - 1$. Thus, by minimality, we can find $k - 1$ disjoint ab paths in $G - x$, so all of these, plus axb, gives us k disjoint ab paths in G, contradiction.

Let $ax_1 \dots x_l b$ be a shortest ab path. Note that $l \geq 2$ and $x_2 \neq b$. Consider $G - x_1x_2$. We must have $\kappa_{a,b}(G - x_1x_2) \leq k - 1$ by minimality, so $\kappa_{a,b}(G - x_1x_2) = k - 1$. So there is a a, b separator $\tilde{S}$, $|\tilde{S}| = k - 1$ in $G - x_1x_2$. We see that $\tilde{S} \cup \{x_1\}$ and $\tilde{S} \cup \{x_2\}$ are a, b separators in G of size at most k. Now either $\tilde{S} \cup \{x_1\} \neq N(a), N(b)$ or $\tilde{S} \cup \{x_2\} \neq N(a), N(b)$, so we're done. $\square$

Corollary 2.6 (Menger's theorem, form 2)**.** Let G be a connected graph, $|G| \geq 2$. Then G is k-connected $\iff$ $\forall a, b \in V(G), a \neq b$, there exist k disjoint ab paths in G.

Proof. $\Longleftarrow$ is the easy direction. Say $G - S$ is disconnected and let a, b be in different components of $G - S$. Note $a \not\sim b$. Then $\exists$ k disjoint $a - b$ paths and S must intersect each of these, so $|S| \geq k$.

$\Longrightarrow$. Let $a, b \in V(G), a \neq b$. If $a \not\sim b$, then just apply Menger form 1 and we're done. If $a \sim b$, then consider $G - ab$. We have $\kappa_{a,b}(G - ab) \geq k - 1$, so apply Menger form 1 to get $k - 1$ disjoint paths and add back ab as a k^{th} path. $\square$

2.2.1 Edge connectivity

Let G be a graph. Let $\lambda(G) = \min\{|W| \mid W \subseteq E(G), G - W \text{ is disconnected}\}$. We say that G is k**-edge connected** if $\lambda(G) \geq k$.

21 Oct 2022, Lecture 7

Example 2.2.

- A cycle has $\kappa(C_n) = 2, \lambda(C_n) = 2$.
- A "bowtie graph" has $\kappa(G) = 1, \lambda(G) = 2$. We can generalize this and take two copies of K_n which intersect in one vertex, then $\kappa(G) = 1$ and $\lambda(G) = n - 1$.

Definition 2.10. We say paths $P_1, \dots, P_k$ are **edge disjoint** if

$$E(P_i) \cap E(P_j) = \emptyset \; \forall i \neq j.$$

Theorem 2.7 (Menger, edge version)**.** Let G be a connected graph and $a, b \in V(G), a \neq b$. Then, every $W \subseteq E(G)$ that separates a from b having size $\geq k \implies \exists k$ edge disjoint $a - b$ paths $P_1, \dots, P_k$.

Definition 2.11. Let G be a graph. The **line graph** of G, denoted $L(G)$, is defined to be the graph

$$V(L(G)) = E(G);$$
$$\text{If } e, f \in E(G), \text{ then } e \sim f \text{ if they share a vertex.}$$

Proof of Thorem 2.7. Given G, define a new graph G' by taking the line graph of G and adding a vertex a', which we join to all edges incident to $a \in G$, and similarly adding a vertex b', which we join to all edges incident to $b \in G$.

Note that there is a ab path in G if and only if there is an $a'b'$ path in G'. Thus $W \subseteq V(G') \setminus \{a, b\}$ is a $a'b'$ separator if and only if $W \subseteq E(G)$ is an ab separator. Hence $\kappa_{a,b}(G') \geq k$.

Now apply Menger (form 1) to find k disjoint ab paths $P_1, \dots, P_k$ in G'. These describe edge disjoint walks in G from a to b. Thus there are disjoint paths $\tilde{P}_1 \subseteq P_1, \dots, \tilde{P}_k \subseteq P_k$ and we're done. ☐

Theorem 2.8 (Menger, edge version 2)**.** Let G be a connected graph. Then $\lambda(G) \geq k \iff \forall a, b \in V(G)$ with $a \neq b, \exists k$ edge disjoint ab paths.

Proof. $\Longleftarrow$ is the easy direction. To separate any two vertices, say a, b, we must remove an edge from each of the ab paths, so $\lambda(G) \geq k$. $\Longrightarrow$ follows from Menger, edge version 1. ☐

3 Graph coloring

Definition 3.1. We say that $c : V(G) \to \{1, \dots, k\}$ is a k**-coloring** (or a proper k-coloring) if $c(x) \neq c(y)\ \forall x \sim y$.

Definition 3.2. The **chromatic number of** G is

$$\chi(G) = \min\{k \mid \exists \text{ a } k\text{-coloring of } G\}.$$

Example 3.1.

- A path has chromatic number 2. $\chi(P_n) = 2$.
- For a cycle, $\chi(G) = \begin{cases} 2 & \text{if } n \text{ is even.} \\ 3 & \text{if } n \text{ is odd.} \end{cases}$
- A tree has chromatic number 2 by induction.
- A complete graph has $\chi(K_n) = n$.
- A bipartite graph has $\chi(K_{m,n}) = 2$. In fact, a graph G is bipartite if and only if $\chi(G) = 2$.

Proposition 3.1. Let G be a graph. Then $\chi(G) \leq \Delta(G) + 1$.

Proof. Let $x_1, \dots, x_n$ be an ordering of $V(G)$. We color the (x_i) one at a time in this order. When we come to vertex x_i, at most Δ colors have been used in $N(x_i)$, so there is a free color for x_i. ☐

Remark. This proposition is sharp (e.g. on K_n).

Remark. This is sometimes called a greedy coloring. But a greedy coloring may produce a coloring that is not optimal!

3.1 Coloring planar graphs

Observation. Let G be a planar graph. Then $\delta(G) \leq 5$.

Proof. The average degree of G is

$$\frac{1}{n} \sum_{v \in V(G)} \deg(v) = \frac{2e(G)}{n} \leq \frac{2(3n-6)}{n} = 6 - \frac{12}{n} < 6.$$

But all degrees of vertices are integers, so the minimal degree is ≤ 5. $\square$

Proposition 3.2. If G is planar, then $\chi(G) \leq 6$.

Proof. We induct. Base step: if $|G| \leq 6$, then we're clearly done.

Induction step: Given a graph G, let x be a vertex with $\deg(x) \leq 5$. Apply induction to $G - x$, which gives a coloring of $G - x$ with 6 colors. But x has degree ≤ 5, thus there is a free color to color x with. $\square$

24 Oct 2022, Lecture 8

Theorem 3.3. If G is planar, then $\chi(G) \leq 5$.

Proof. We apply induction on $|G|$. If $|G| \leq 5$, we're done.

Now let G be a planar graph and $x \in V(G)$ with $\deg(x) \leq 5$. Apply induction to $G - x$. Let neighbor x_i of x have color i, ordered clockwise around x.

Question: Can we get from x_1 to x_3 only walking along vertices colored 1 and 3?

If no, let C be the component of G of vertices colored 1 or 3 that contains x_1, so $x_3 \notin C$. "Swap" the colours 1 and 3 on C. This is a proper coloring of $G - x$, so we can color x with color 1 and we're done.

If yes, ask the same question for x_2 and x_4 - if the answer is no, swap colors on the 2–4 component containing x_2, and we can color x with color 2.

If the answer is again yes, then we get a contradiction to planarity since the 1–3 path and the 2–4 path have to intersect somewhere and we're done. $\square$

Theorem 3.4 (Four color theorem, non-examinable)**.** If G is planar, then $\chi(G) \leq 4$.

To see this is equivalent to the map version, take the dual of our graph (place a vertex inside each face (and one for the infinite face) and connect two vertices by an edge if the two faces have any common boundary).

Kempe "proved" the four colour theorem in 1879, but his proof had a mistake. It was then proved in 1976 by reducing the problem to about 2000 configurations and checking them by computer.

Also, this is the best we can do, since K_4 is planar - there is no "three color theorem".

3.2 Coloring graphs

Proposition 3.5. Let G be connected and $\delta(G) < \Delta(G)$. Then $\chi(G) \leq \Delta(G)$.

Proof. Let x_n have $\deg(x_n) < \Delta(G)$. Then choose x_{n-1} to be adjacent to x_n, x_{n-2} to be adjacent to one of $\{x_n, x_{n-1}\}$, etc. Since G is connected, we eventually order everything and the ordering has the property that all vertices have less than $\Delta(G)$ edges going to vertices with smaller index, so color greedily and we're done. □

Theorem 3.6 (Brooks)**.** Let G be a connected graph which is not an odd cycle or a complete graph. Then $\chi(G) \leq \Delta(G)$.

Proof. Apply induction on $|G|$. The claim is clearly is true for $|G| \leq 3$.

Note that we may assume that $\Delta \geq 3$ (where $\Delta = \Delta(G)$).

Claim 1. If G is 3-connected, then we're done.

Proof. Define an ordering of G as in the proposition above. Let x_n have $\deg(x_n) = \Delta$ and choose $x_1, x_2 \in N(x_n)$ with $x_1 \neq x_2, x_1 \not\sim x_2$. This is possible, since G is not $K_{\Delta+1}$.

Now consider $G \setminus \{x_1, x_2\}$. We order this in the same way as above: connect x_{n-1} to x_n, x_{n-2} to x_{n-1} or x_n, etc. Since $G \setminus \{x_1, x_2\}$ is connected, we eventually order every vertex. Hence we're done by coloring greedily.

Claim 2. If $\kappa(G) = 1$, we're done.

Proof. Let x be a cut vertex and let $C_1, \ldots, C_k$ be the components of $G - x$. By induction, we can color each $G[C_i \cup x]$. Then permute colors to have x always be the same color, and put everything back together to get a coloring of G. 26 Oct 2022, Lecture 9

Claim 3. If $\kappa(G) = 2$, we're done.

Proof. Let $S \subset V(G)$, $S = \{x, y\}, x \neq y$ be our separator. Let $C_1, \ldots, C_k$ be the components of $G - S$ and let $G_i = G[C_i \cup S] + xy$ for each i.

- Case 1: $\delta(G_i) < \Delta(G)\ \forall i$. In this case we can always color $C_1, \ldots, C_k$ by Proposition 3.5. Note that in each of these colorings, x, y have different colors, so we can permute colors in a way that x, y always have the same color, and put everything back together to get a coloring of G.

- Case 2: $\delta(G_1) = \Delta(G)$ (wlog assume $i = 1$). In this case, $k = 2$ and $|N(x) \cap C_1| = \Delta - 1$. Also, $|N(x) \cap C_2| = 1$ and $|N(y) \cap C_2| = 1$. Thus let $x', y' \in C_2$ be such that $x' \sim x, y' \sim y$. Now observe that $\tilde{S} = \{x, y'\}$ is not of this bad form, so we're done by Case 1.

□

3.3 Chromatic polynomials

Definition 3.3. For a graph G, define the **chromatic polynomial** of G to be the function $P_G : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ by $P_G(t) =$ the number of t–colorings of G.

In particular, the minimal value of t for which $P_G(t) > 0$ is $t = \chi(G)$.

Example 3.2.

- $P_{\overline{K_n}}(t) = t^n$.
- $P_{K_n}(t) = t(t-1)\dots(t-(n-1))$.
- $P_{P_{n-1}}(t) = t(t-1)^{n-1}$.
- $P_T(t) = t(t-1)^{n-1}$ where T is a tree with n vertices (proof: take a leaf, induct).

Definition 3.4. Let G be a graph and $e \in E(G)$. Define G/e to be the **contraction** of G along e, which is the graph

$$\begin{aligned} V(G/e) &= V(G) \setminus \{x,y\} \cup \{xy\}, \\ E(G) &= E(G[V \setminus \{x,y\}]) \cup \{ez \mid x \sim z\} \cup \{ez \mid y \sim z\}. \end{aligned}$$

(In other words, you just "squish" the edge e and its endpoints become a single vertex.)

Proposition 3.7. Let G be a graph. We have $P_G(t) = P_{G-e}(t) - P_{G/e}(t)$.

Proof. Let $e = xy$. A t–coloring of $G - e$ where x, y get different colors corresponds exactly to a t–coloring of G. On the other hand, a t–coloring of $G - e$ where x, y get the same color corresponds exactly to a t–coloring of G/e. □

Proposition 3.8. For a graph G, P_G is a polynomial.

Proof. Apply induction on $e(G)$. For the base case $e(G) = 0$, $P_{\overline{K_n}}(t) = t^n$. Induction step: For $e \in E(G)$, $P_G(t) = P_{G-e}(t) - P_{G/e}(t)$ is a polynomial (as it is the difference of two polynomials), so we're done by induction. □

Proposition 3.9. Let G be a graph with n vertices and m edges. Then

$$P_G(t) = t^n - mt^{n-1} + \dots.$$

Proof. Induction on the number of edges. If $e(G) = 0$, then we're done.

Induction step: Let $e \in E(G)$ and apply the deletion-contraction relation, $P_G(t) = P_{G-e}(t) - P_{G/e}(t)$. By induction, the leading terms on the RHS are

$$(t^n - (m-1)t^{n-1} + \dots) - (t^{n-1} - e(G/e)t^{n-2} + \dots) = t^n - mt^{n-1} + \dots.$$

□

Remarks. We're going to stop here, but chromatic polynomials have a lot of other properties. For example:

- Other coefficients of P_G contain other information about G, e.g.

$$P_G = t^n - mt^{n-1} + \left(\binom{m}{2} - \text{the number of triangles in } G\right) t^{n-2} + \dots.$$

- If G is planar, then $P_G(2 + \frac{1+\sqrt{5}}{2}) \neq 0$.
- The coefficients of $c_0, \dots, c_n$ of $P_G(t)$ are log-concave, i.e. $c_i^2 \geq c_{i-1}c_{i+1}$. June Huh won a Fields medal for this in 2021.

28 Oct 2022, Lecture 10

3.4 Edge coloring

Definition 3.5. For a graph G, a k–edge coloring is a function $c : E(G) \to [k]$ such that $c(e) \neq c(f)$ whenever $e, f \in E(G)$ share an endpoint.

Define the **edge-chromatic number** (sometimes called the **chromatic index**) to be

$$\chi'(G) = \min\{k \mid \exists k\text{-edge coloring of } G\}.$$

Note that an edge coloring of G corresponds exactly to a vertex coloring of the line graph of G, and so $\chi'(G) = \chi(L(G))$.

Remarks.

- $\chi'(G) = \begin{cases} 2 & \text{if } n \text{ even.} \\ 3 & \text{if } n \text{ odd.} \end{cases}$
- $\Delta(G) \leq \chi'(G)$.
- We can have $\Delta(G) < \chi'(G)$. For example, $\chi'(\text{Petersen}) = 4$.
- $\chi'(G) \leq 2\Delta(G) - 1$ by coloring greedily.
- χ and χ' can be very different. For example, $\chi'(K_{t,1}) = t, \chi(K_{t,1}) = 2$.

We now prove Vizing's theorem. First, some setup:

Definition 3.6. Given an edge-coloring $c : E(G) \to [k]$, define the **color classes** of c to be

$$C_i = \{e : c(e) = i\}.$$

Question. What does $(V(G), C_i \cup C_j)$ look like? It is the union of disjoint paths, isolated vertices and even cycles.

We say the components of this graph are $\{i, j\}$–components.

Definition 3.7. Let $y \in V(G), c \in [\Delta + 1]$. We say that c is **missing** at y if none of the edges incident at y have color c.

Theorem 3.10 (Vizing's theorem)**.** Let G be a graph. Then

$$\chi'(G) \in \{\Delta(G), \Delta(G) + 1\}.$$

Proof. Induction on $e(G)$. If $e(G) = 0$, we're done.

Induction step: given G, $e(G) > 0$, let $xv \in e(G)$. Apply induction to the graph $G - xv$ to get a $\Delta + 1$ edge coloring. We want to extend this coloring to G.

Note that there is a color missing at every vertex and let c_0 be the color missing at x. Define a sequence of $v_1, \ldots, v_k \in N(x)$ and corresponding colors $c_1, \ldots, c_k$ such that c_i is missing at v_i as follows:

First set $v_1 = v$ and let c_1 be any color missing at v_1. Now say we have defined v_i and c_i. Define v_{i+1} such that xv_{i+1} is colored c_i. Continue this way until one of two things happens:

- Case 1: Say we have defined v_k, c_k and c_k is missing at x. Then recolor xv_k with color c_k, so now c_{k-1} is missing at x. Hence recolor xv_{k-1} with color c_{k-1}, so now c_{k-2} is missing at x, etc, until we color xv_i with color c_i for all $i \leq k$. This includes $i = 1$, so we're done.
- Case 2: Say $c_k = c_i$ for some $i < k$. Note that we may assume $i = 1$ by recoloring as in Case 1 (i.e. uncolor xv_i and color $xv_j \to c_j$ for all $j < i$). The question we ask is whether v_1 is in the same $\{c_0, c_1\}$–component as x.
  - If not, swap colors on the $\{c_0, c_1\}$–component containing v_1. Then c_0 is missing at both x and v_1, and so we can color xv_1 with c_0 and we're done.
  - If yes, we need another question: Is x in the same $\{c_0, c_1\}$–component as v_k?
    - If the answer is no, then swap colors in the $\{c_0, c_1\}$–component containing v_k. So c_0 is missing at v_k and x, so we are done by Case 1.
    - If yes, then x, v_1, v_k are all in the same $\{c_0, c_1\}$–component. But one of c_0, c_1 is missing at each of x, v_1, v_k. Thus x, v_1, v_k all have to be endpoints of a path in the same $\{c_0, c_1\}$–component, which is a contradiction.

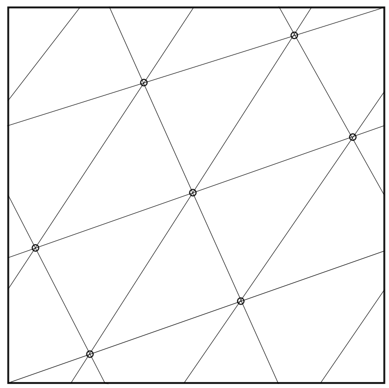
$\square$

3.5 Coloring graphs on surfaces

31 Oct 2022, Lecture 11

If G is planar, we know that$\chi(G) \leq 5$. (And really $\chi(G) \leq 4$). If G is drawn on a surface S, can we say anything about $\chi(G)$?

K_7 can be drawn on a torus without edge crossings, so we definitely don't have a 4 color theorem for the torus:



Question. Is there a "7 color theorem" for the torus?

Definition 3.8. We define the **surface of genus** g (more formally, a compact orientable surface of genus g) to be the sphere with g handles attached (or equivalently, the sphere with g holes).

Do we have an Euler-type theorem in this setting? Different cycles on the torus give different values for $V - E + F$, we get all of 2, 1, 0. *picture*

Proposition 3.11. Let G be a graph drawn on a surface of genus g with n vertices, m edges, and f faces. Then

$$n - m + f \geq 2 - 2g.$$

Remarks.

- $E = 2 - 2g$ is the **Euler characteristic** of the surface of genus g.
- This is the direction of Euler's theorem ($\geq$, not $\leq$) that we need to bound the number of edges of G.
- We do not assume that the graph is connected here.

Corollary 3.12. Let G be a graph with $|G| \geq 3$, drawn on a surface of Euler characteristic E. Then

$$e(G) \leq 3(|G| - E).$$

Proof. If $e(G) \leq 2$, we can check the claim holds. When $e(G) \geq 3$, then every face in G has length at least 3, so

$$3f \leq \sum_{F \text{ a face}} \deg(F) \leq 2e(G).$$

Thus $f \leq \frac{2}{3}e(G)$, so $e(G) \leq n + f - E \implies e(G) \leq 3(n - E)$. □

Theorem 3.13 (Heawood's theorem)**.** Let G be a graph drawn on a surface of Euler characteristic $E \leq 0$. Then

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

Proof. Let G be a graph drawn on a surface of Euler characteristic E and let $k = \chi(G)$. We may assume that G is minimal, i.e. all proper subgraphs of G have $\chi \leq k - 1$.

We esimate $\delta(G)$ in two ways. Note $\delta(G) \geq k - 1$: if not, let x be a vertex with $\deg(x) \leq k - 2$. We have $\chi(G - x) \leq k - 1$ by minimality, so we can put x back and color it $\implies \chi(G) \leq k - 1$, contradiction.

Note

$$\delta(G) \leq \text{the average degree of the graph} = \frac{1}{n} \sum_{x \in V(G)} \deg(x) =$$
$$\frac{2e(G)}{n} \leq \frac{6(n - E)}{n} = 6 - \frac{6E}{n} \leq 6 - \frac{6E}{k}$$

because $n \geq k$ and $E \leq 0$. So

$$k - 1 \leq 6 - \frac{6E}{k} \implies k^2 - 7k + 6E \leq 0 \implies k \leq \frac{7 + \sqrt{49 - 24E}}{2}.$$

As k is an integer, we're done. □

Remarks.

- Define $H(E) = \left\lfloor \frac{7+\sqrt{49-24E}}{2} \right\rfloor$.
- Did we just prove the four color theorem, since the above gives $H(2) = 4$? No, since we assumed $E \leq 0$.
- Amazingly, Heawood's theorem is sharp. In particular, $K_{H(E)}$ can be drawn in the surface of Euler characteristic E (but this is very hard!).

4 Extremal graph theory

Question. What is the minimum number of edges a graph must have before it is forced to have a triangle? In other words, what is the value of k such that if $e(G) > k$, then $K_3 \subset G$?

Question. How large does $\delta(G)$ have to be before G is forced to have a cycle of length $|G|$?

02 Nov 2022, Lecture 12

Definition 4.1. 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 4.1 (Dirac)**.** Let G be a graph on n vertices, $n \geq 3$. If $\delta(G) \geq \frac{n}{2}$, then G is Hamiltonian.

Remark. This theorem is sharp. If n is even, take two disjoint $K_{n/2}$'s. If n is odd, take two copies of $K_{(n+1)/2}$ that overlap at exactly one vertex.

Proof. First note that G is connected. If $x \not\sim y$, then $N(x), N(y) \subset G - x - y$, but $|N(x)|, |N(y)| \geq \frac{n}{2}$, so $N(x) \cap N(y) \neq \emptyset$ by the pigeonhole principle, so x and y have a common neighbor.

Now take a path of maximal length $x_1 \ldots x_l$. The claim is that there is no cycle in G of length l. Indeed, if such a cycle did exist, then either $l = n$ and we're done, or $l < n$ and we can use this cycle to find a path of length $l + 1$ in G, contradicting maximality.

Observe that $N(x_1) \subset \{x_2, \ldots, x_{l-1}\}$ and $N(x_l) \subset \{x_2, \ldots, x_{l-1}\}$. Now define $N^-(x_1) = \{x_i \mid x_{i+1} \in N(x_1)\}$. Note

$$|N^-(x_1) \cup N^-(x_l)| \leq \{x_1, x_2, \ldots, x_{l-1}\} \leq l - 1 \leq n - 1,$$

but $|N^-(x_1)|, |N(x_2)| \geq \frac{n}{2}$. So there must exist x_i in their intersection.

So we have $x_1 \ldots x_i x_l x_{l-1} \ldots x_{i+1} x_1$, a cycle of length l, a contradiction and we're done. □

Remark. Note that there's no interesting theorem of the form "if $e(G) > k$, then G is Hamiltonian", since $K_{n-1} + x$ where x has degree 1 is not Hamiltonian.

Theorem 4.2. Let G be a graph on n vertices. Then $e(G) > \frac{n(k-1)}{2} \implies G$ contains a path of length k.

Remark. If $k \mid n$, then this is sharp: take the disjoint union of $\frac{n}{k}$ copies of K_k.

Lemma 4.3. Let G be a graph on n vertices, $n \geq 3$, and let $k < n$. If G is connected and $\delta(G) \geq \frac{k}{2}$, then G contains a path of length k.

Remarks.

- We need the assumption that $k < n$, otherwise take K_k.
- We need the assumption that G is connected, otherwise take a bunch of disjoint K_k.
- The $\frac{k}{2}$ is sharp. Take a bunch of K_k's that overlap at a single vertex.

Proof. Let $x_1 \dots x_l$ be a path of maximum length. We have that there is no cycle of length l: If $l = n$, then we're done since $k < n$, and if $l < n$, then use a cycle of length l to build a path of length $l + 1$ (since G is connected), a contradiction.

So we get $N(x_1) \subset \{x_2, \dots, x_{l-1}\}, N(x_l) \subset \{x_2, \dots, x_{l-1}\}$, and $N^-(x_1) = \{x_i \mid x_{i+1} \in N(x_1)\} \subset \{x_1, \dots, x_{l-2}\}$. Thus

$$|N^-(x_1) \cup N(x_l)| \le |\{x_1, \dots, x_{l-1}\}| \le l - 1 \overset{(\star)}{\le} k - 1$$

(where $(\star)$ follows since else we're immediately done), but $|N^-(x_1)|, |N(x_l)| \ge \frac{k}{2}$, so $N^-(x_1) \cap N(x_l) \neq \emptyset$. So as before, we can build a cycle of length l, contradiction, done. □

Proof of Theorem 4.2. If $k = 1$, we're done. Now assume $k \ge 2$ and apply induction on n, $n = 2$ being clear.

For the induction step, given a graph G on $n \ge 3$ vertices, first note that $\frac{n(k-1)}{2} < e(G) \le \frac{n(n-1)}{2} \implies k \le n$.

We may assume that G is connected: say $C_1, \dots, C_r$ are the components of G and $|C_i| = n_i$. Then $\sum_{i=1}^r e(G[C_i]) = e(G) > \frac{n(k-1)}{2}$, so since $\sum_{i=1}^r n_i = n$,

$$\sum_{i=1}^r \left(e(G[C_i]) - \frac{n_i(k-1)}{2} \right) > 0.$$

So there exists some component C_i that has $e(G[C_i]) > \frac{n_i(k-1)}{2}$, so apply induction to this component and we're done.

If $\delta(G) \ge \frac{k}{2}$, then apply Lemma 4.3 to see that G contains a path of length k. Otherwise, there exists a vertex $x \in V(G)$ with $\deg(x) \le \frac{k-1}{2}$. Note that we have

$$e(G - x) > \frac{n(k-1)}{2} - \frac{k-1}{2} = \frac{(n-1)(k-1)}{2},$$

hence G contains a path of length k and we're done. □

04 Nov 2022, Lecture 13

Question. How many edges in a graph with n vertices do we need before we are forced to have a triangle?

Note. For n even, $K_{n/2,n/2}$ contains no triangle and has $\frac{n^2}{4}$ edges.

Theorem 4.4. Let G be a graph on n vertices. If $e(G) > \frac{n^2}{4}$, then G contains K_3.

Fact (Jensen). Let $a < b$, $f : [a, b] \to \mathbb{R}$ convex, $x_1, \ldots, x_n \in [a, b]$. Then

$$\frac{1}{n}\sum_{i=1}^{n} f(x_i) \geq f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right).$$

Proof. Assume for contradiction that $K_3 \not\subset G$. We may assume that $n \geq 3$. Consider $x \sim y$. We have $d(x) + d(y) \leq n$, else we have a triangle (since x, y have a common neighbor by pigeonhole). Thus

$$ne(G) = \sum_{x \sim y}(d(x) + d(y)) = \frac{1}{2}\sum_{x}\sum_{y}(d(x) + d(y))\mathbb{1}(x \sim y) =$$

$$\sum_{x}\sum_{y} d(x)\mathbb{1}(x \sim y) = \sum_{x} d(x)\sum_{y}\mathbb{1}(x \sim y) = \sum_{x} d(x)^2.$$

Now by Jensen,

$$ne(G) = n\left(\frac{1}{n}\sum_{x} d(x)^2\right) \geq n\left(\frac{1}{n}\sum_{x} d(x)\right)^2 = n\left(\frac{2e(G)}{n}\right)^2$$

$$\implies e(G) \leq \frac{n^2}{4}.$$

□

Definition 4.2. We say that a graph $G = (V, E)$ is r–**partite** if there is a partition of $V = V_1 \sqcup V_2 \sqcup \ldots \sqcup V_r$ such that $\forall xy \in E$, $\exists V_i, V_j, i \neq j$ s.t. $x \in V_1$ and $y \in V_j$. This is the same as saying $\chi(G) \leq r$.

Definition 4.3. Given $n_1, \ldots, n_r \in \mathbb{Z}_{\geq 0}$, define $K_{n_1,\ldots,n_r}$ to be the **complete** r–**partite graph** where $V = V_1 \sqcup V_2 \sqcup \ldots \sqcup V_r$, $|V_i| = n_i$ $\forall i$, and $\forall x \in V_i, y \in V_j, i \neq j, xy \in E(K_{n_1,\ldots,n_r})$.

Note that if $r \mid n$, then

$$e(K_{\frac{n}{r},\ldots,\frac{n}{r}}) = \frac{n^2}{r^2}\binom{r}{2} = \left(1 - \frac{1}{r}\right)\frac{n^2}{2}.$$

Theorem 4.5 (Turan)**.** Let G be a graph on n vertices, $r \geq 1$. Then

$$e(G) > \left(1 - \frac{1}{r}\right)\frac{n^2}{2} \implies K_{r+1} \subset G.$$

Proof. We prove the contrapositive. Apply induction first on $r \geq 1$ and then on n. If $n \leq r$, we're done. Note that the theorem is trivial if $r = 1$, so assume $r \geq 2$. Now let G be a graph, $K_{r+1} \not\subset G$.

If there is no K_r in G, then apply induction (on r) to get an even stronger bound $e(G) \leq (1 - \frac{1}{r-1})\frac{n^2}{2}$.

Hence let K be a K_r in G. Each vertex in $V(G) \setminus K$ has at most $r - 1$ neighbors in K (else we'd have $K_{r+1} \subset G$), So

$$\begin{aligned} e(G) &\leq \binom{r}{2} + (r-1)(n-r) + e(G \setminus K) \leq \\ &\leq \binom{r}{2} + (r-1)(n-r) + \left(1 - \frac{1}{r}\right)\frac{(n-r)^2}{2} = \left(1 - \frac{1}{r}\right)\frac{n^2}{2}. \end{aligned}$$

□

Definition 4.4. The **Turan graph** $T_{r,n}$ is the graph $K_{n_1,\ldots,n_r}$ where $\sum n_i = n$ and the $\{n_i\}$ are as equal as possible, i.e. $|n_i - n_j| \leq 1 \ \forall i, j$.

Lemma 4.6. Let G be an r–partite graph on n vertices. Then $e(G) \leq e(T_{r,n})$.

Proof. Let G be a r–partite graph with maximal number of edges. First note G is a complete r–partite graph $K_{n_1,\ldots,n_r}$ for some $n_1, \ldots, n_r$. If $\exists i, j$ such that $n_i \geq n_j + 2$, move a vertex from part i to part j. Then we gain $n_i - 1$ edges and lose n_j edges, so we get a graph with more edges, contradiction. Hence $|n_i - n_j| \leq 1 \ \forall i, j$, so $G = T_{r,n}$. □

Theorem 4.7 (Turan, version 2)**.** Let G be a graph on n vertices and $r \geq 2$. If G does not contain a K_{r+1}, then $e(G) \leq e(T_{r,n})$.

Idea for proof. Given G, we "transform" G into a complete r–partite graph without decreasing the number of edges or creating any K_{r+1}'s. □

Proof. Say $V(G) = [n]$. Let $\alpha_1, \ldots, \alpha_n > 0$ be linearly independent over $\mathbb{Q}$ and define $\mu(S) = \sum_{i \in S} \alpha_i$ for $S \subseteq [n]$.

07 Nov 2022, Lecture 14

If H is a graph on n vertices, define $T(H)$ as follows: Let x, y be a pair maximizing $\mu(\{x, y\})$ such that $N(x) \neq N(y)$ and $x \not\sim y$. If $\deg(x) > \deg(y)$ or $\deg(x) = \deg(y)$ but $\mu(N(x)) > \mu(N(y))$, define $T(H)$ to be $H - y$ along with a vertex x' with $N(x') = N(x)$.

Claim. If $K_{r+1} \not\subset H$, then $K_{r+1} \not\subset T(H)$.

Proof. Assume for contradiction that $T(H)$ contains K_{r+1}, call it K. We must have $x' \in K$, because all other vertices stayed the same. But $x \notin K$ since $x \not\sim x'$. But then $K \setminus \{x'\} \cup x$ is a K_{r+1} in H, contradiction. □

Now consider the sequence, $G, T(G), T(T(G)), \ldots$.

Claim. This sequence eventually stabilizes.

Proof. Note $e(T(H)) \geq e(H)$, so $e(T^{(i)}(H))$ is an increasing sequence. It is also bounded above, so eventually we reach a maximum. Now note

$$\sum_{x \in [n]} \mu(N_{T^{(i)}(G)}(x))$$

is always increasing past that point, so eventually it stabilizes. □

So at some point, our transformation does nothing anymore. Let G_∞ be the graph where we stop.

Claim. G_∞ is a complete k–partite graph for some $k \in \mathbb{N}$.

Proof. Let $k = \chi(G_\infty)$ and let c be a k–coloring of G_∞ with $V(G_\infty) = C_1 \cup \ldots \cup C_k$, where $C_i = \{x \mid c(x) = i\}$.

Note that if $x, y \in c(i)$, then $N(x) = N(y)$, otherwise $T(G_\infty) \neq G_\infty$.

Now let $x \in C_i$ and $y \in C_j$ for $i \neq j$. If $x \not\sim y$, then $x' \not\sim y\ \forall x' \in C_i$. It follows that $\forall x' \in C_i$ and $\forall y' \in C_j$, $x' \not\sim y'$. We can now recolor i, j to be the same color to get a $k-1$–coloring, contradiction. So $x \sim y\ \forall x \in C_i,\ \forall y \in C_j,\ \forall i \neq j$. □

So G_∞ is complete k–partite for some k. Since G_∞ does not contain K_{r+1} by our first claim, we have $k \leq r$. By Lemma 4.6, $e(G_\infty) \leq e(T_{r,n})$ and also $e(G) \leq e(G_\infty)$, so we're done. □

The Zarankiewicz problem.

Let $Z(n,t)$ be the maximal number of edges in a bipartite graph $G = (X \cup Y, E)$, with $|X| = |Y| = n$ and $K_{t,t} \not\subset G$.

Theorem 4.8. Let $t \geq 2$. Then

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

In particular, if n is large compared to t, we have $Z(n,t) \leq 2n^{2-\frac{1}{t}}$.

Lemma 4.9. Let $t \geq 2$ be an integer. The function

$$f_t(x) = \frac{x(x-1)\ldots(x-t+1)}{t!}$$

is convex for $x \geq t-1$.

Proof. Let $s = x - t + 1$, so $f_t(s) = \frac{(s+t-1)(s+t-2)\dots(s)}{t!}$ which is a polynomial with non-negative coefficients. So $f''(s) \geq 0$ if $s \geq 0 \iff x \geq t-1$. $\square$

Proof of Theorem 4.8. We may assume that $d(y) \geq t-1\ \forall y \in Y$, since if $d(y) < t-1$, we could add an edge incident to y without creating a $K_{t,t}$. Note

$$|N(x_1) \cap \ldots \cap N(x_t)| \leq t - 1.$$

Hence

$$(t-1)\binom{n}{t} \geq \sum_{x_1,\dots,x_t \in X \text{ distinct}} |N(x_1) \cap \ldots \cap N(x_t)|.$$

Rewrite the RHS as

$$\sum_{x_1,\dots,x_t \text{ distinct}} \sum_y \mathbb{1}(y \sim x_1)\mathbb{1}(y \sim x_2)\dots\mathbb{1}(Y \sim x_t) =$$

$$= \sum_y \sum_{x_1,\dots,x_t} \mathbb{1}(y \sim x_1)\dots\mathbb{1}(y \sim x_t) = \sum_y \binom{d(y)}{t}.$$

Let $\overline{d} = \frac{e(G)}{n}$. As $d(y) \geq t-1$, our function is convex by Lemma 4.9, so by Jensen's inequality,

$$\sum_y \binom{d(y)}{t} \geq n\binom{\overline{d}}{t}.$$

So

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

$\square$

Remarks.

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

Non-examinable aside:

Theorem 4.10 (Non-examinable aside)**.** $\exists c > 0$ such that for infinitely many n, $Z(n, 2) \geq cn^{\frac{3}{2}}$.

Proof. Think of X as lines in a finite geometry and Y as points, and we have an edge in our graph if a line contains a given point.

For p a prime, let $\mathcal{L} = \mathbb{F}_p^2 \setminus \{(0,0)\}$ and $\mathcal{P} = \mathbb{F}_p^2 \setminus \{(0,0)\}$. For $(a,b) \in \mathcal{L}$, define $N((a,b)) = \{(x, ax+b) \mid x \in \mathbb{F}_p\} \setminus \{(0,0)\}$. 09 Nov 2022, Lecture 15

Note that if we have two lines $(a,b),(a',b') \in \mathcal{L}$, then $|N(a,b) \cap N(a',b')| \leq 1$, since if (x,y) is in the intersection, then $y = ax+b, y = a'x+b'$. If $a' = a$, then $b = b'$ and we have the same line. Otherwise $x = \frac{-(b-b')}{a-a'} \bmod p$, so x (and hence y) is uniquely determined.

Note $\forall (a,b) \in \mathcal{L}$, $|N(a,b)| \geq p-1$, so $V(G) = 2(p^2-1) = 2n$ and $e(G) \geq (p-1)(p^2-1) \geq cp^3 = cn^{\frac{3}{2}}$ for an absolute constant c. □

End of aside.

The Erdős–Stone theorem.

Definition 4.5. Let H be a (fixed) graph and $n \in \mathbb{N}$. Define

$$\mathrm{ex}(n, H) = \max\{e(G) \mid V(G) = n \text{ and } H \not\subset G\}.$$

Example 4.1.

- We have $\mathrm{ex}(n, K_{r+1}) = e(T_{r,n}) \leq \left(1 - \frac{1}{r}\right)\frac{n^2}{2}$.
- We have $\mathrm{ex}(P_k, n) = \frac{n(k-1)}{2}$.
- We've basically shown $\mathrm{ex}(K_{t,t}, n) \leq 2n^{2-\frac{1}{t}} + tn$.

Theorem 4.11 (Erdős–Stone)**.** Let H be a fixed graph, H nonempty. Then

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

Remarks.

- If $\chi(H) \geq 3$, then this determines the leading order term in $\mathrm{ex}(n,t)$ for large n.
- If $\chi(H) = 2$, then this just says $\frac{\mathrm{ex}(n,H)}{n^2} \to 0$. But we have a stronger result already: $H \subseteq K_{t,t}$ for $t = |H|$, so we (almost) know $\mathrm{ex}(H,n) \leq cn^{2-\frac{1}{t}}$.
- It is easy to see that $\mathrm{ex}(n,H) \geq (1 - \frac{1}{\chi(H)-1})\frac{n^2}{2}$, since $H \not\subset T_{(\chi(H)-1),n}$.

We sketch the proof of this. It is non-examinable.

Sketch of proof, non-examinable. We only consider $\chi(H) = 3$. Let $K_3(t)$ be the complete 3–partite graph with t vertices in each part (i.e. $T_{3,3t}$). Note that it is enough to prove our theorem for $H = K_3(t)\ \forall t$, since $H \subseteq K_3(|H|)$.

We prove the following:

Theorem 4.12. Let $\epsilon > 0$, $t \in \mathbb{N}$, G a graph with $|G| = n$ and

$$e(G) \geq \left(\frac{1}{2} + \epsilon\right)\frac{n^2}{2}$$

where $n \geq n_0(\epsilon, t)$. Then $G \supset K_3(t)$.

Lemma 4.13. Let G be a graph on n vertices with average degree $\geq (\frac{1}{2} + \epsilon)n$ (in other words, $e(G) \geq (\frac{1}{2} + \epsilon)\frac{n^2}{2}$). Then $\exists \tilde{G} \subset G$ with minimum degree at least $(\frac{1}{2} + \frac{\epsilon}{2})m$, where $m = |\tilde{G}| \geq \frac{\epsilon^{\frac{1}{2}} n}{8}$.

Proof. Omitted, but the idea is to remove edges with degrees smaller than the number stated above and eventually we're left with a subgraph that works. $\square$

Proof of theorem. Let $\tilde{G} \subset G$ be a subgraph with $\delta(\tilde{G}) \geq (\frac{1+\epsilon}{2})m$, $|\tilde{G}| = m$. Find a $K_{s,s}$ in $\tilde{G}$ with $s \approx \frac{10t}{\epsilon}$, call this graph K. Consider

$$e(Y, K) = \sum_{x \in K} |N(x) \cap Y| = \sum_{x \in K} (\deg(x) - s) \geq \sum_{x \in K} \left(\frac{1}{2} + \frac{\epsilon}{4}\right) m = \left(\frac{1}{2} + \frac{\epsilon}{4}\right) m|K|.$$

But also $e(Y, K) = \sum_{y \in Y} |N(y) \cap K|$, so

$$\left(\frac{1}{2} + \frac{\epsilon}{4}\right)|K| < \frac{1}{m}\sum_{y \in Y} |N(y) \cap K| = (\star).$$

Let $\mathcal{B} = \{y \in Y \mid |N(y) \cap K| \geq (\frac{1}{2} + \frac{\epsilon}{8})|K|\}$. We claim $|\mathcal{B}| \geq \frac{\epsilon m}{8}$. Indeed, suppose not. Then

$$(\star) \leq \frac{1}{m}\frac{\epsilon m}{8}|K| + \left(\frac{1}{2} + \frac{\epsilon}{8}\right)|K| = \left(\frac{1}{2} + \frac{\epsilon}{4}\right)|K|,$$

a contradiction. $\square$

and now finish, but we're out of time. $\square$

5 Ramsey Theory

11 Nov 2022, Lecture 16

Fact. Let c be a 2–edge coloring of K_6. Then there exists a monochromatic K_3 (i.e. a triangle T) for which c is constant on $E(T)$.

This is just saying "if we color the edges of a K_6, we get a monochromatic triangle".

Proof. Let $x \in V(K_6)$. x has 5 neighbors, so WLOG it has 3 red neighbors y_1, y_2, y_3. If there is no red triangle, then y_1y_2, y_2y_3, y_3y_1 are all blue, but then $y_1y_2y_3$ is a blue triangle. □

Definition 5.1. For $s \geq 2$, define

$$R(s) = \min\{n \mid \text{every 2–edge coloring of } K_n \text{ contains a monochromatic } K_s\}.$$

Definition 5.2. Let $s, t \geq 2$. Define

$$R(s,t) = \min\{n \mid \text{every 2–coloring of } K_n \text{ contains either a red } K_s \text{ or a blue } K_t\}.$$

Note that $R(s,t) = R(t,s)$ and $R(2,t) = t, R(s,2) = s$. We just showed that $R(3,3) = R(3) \leq 6$ (in fact, $R(3) = 6$).

Theorem 5.1 (Ramsey)**.** For all s, $R(s)$ exists.

This is implied by:

Theorem 5.2 (Ramsey, version 2)**.** For all s, t, $R(s,t)$ exists, and $\forall s, t \geq 2$,

$$R(s,t) \leq R(s-1,t) + R(s,t-1).$$

Proof. We induct on $s + t$. If $s, t \leq 2$, we're done.

Given s, t, fix $n = \underbrace{R(s-1,t)}_{=a} + \underbrace{R(s,t-1)}_{=b}$. Let $c : E(K_n) \to \{\text{Red}, \text{Blue}\}$ be given. Pick $x \in V(K_n)$ and let $N_{\text{red}}(x), N_{\text{blue}}(x)$ be the red and blue neighborhoods of x.

If $|N_{\text{red}}(x)| \geq a$, then $N_{\text{red}}(x)$ contains either a red K_{s-1}, so this K_{s-1} along with x is a red K_s in K_n, or a blue K_t.

Similarly, if $|N_{\text{blue}}(x)| \geq b$, then we get either a blue K_{t-1}, which extends to a blue K_t, or a red K_s.

If $|N_{\text{red}}(x)| < a$ and $|N_{\text{blue}}(x)| < b$, then $n \leq (a-1) + (b-1) + 1 < a + b$, contradiction. □

Definition 5.3. For $k \in \mathbb{N}, s_1, \ldots, s_k \geq 2$, define

$$R_k(s_1, \ldots, s_k) =$$
$$\min\{n \mid \text{every } k\text{–coloring of } K_n \text{ contains a monochromatic } K_{s_i} \text{ in color } i \text{ for some } i\}$$

Theorem 5.3 (Multicolor Ramsey theorem)**.** For $s_1, \ldots, s_k$ and $k \geq 2$, the number $R_k(s_1, \ldots, s_k)$ exists.

Proof. Induct on k. We will show that

$$R_k(s_1, \ldots, s_k) \leq R_2(s_1, R_{k-1}(s_2, \ldots, s_k)) = (\star).$$

Say $(\star) = n$. Let c be a k–coloring of $E(K_n)$. Now apply the 2–color Ramsey theorem to get either a subgraph K_{s_1} which is monochromatic in color 1, or we get a $K_{R_{k-1}(s_2,\ldots,s_k)}$ on which color 1 does not appear.

In the second case, induct on this complete graph to get a K_{s_i} which is monochromatic in color i. □

Theorem 5.4 (Infinite Ramsey)**.** Let $c : \mathbb{N}^{(2)} \to \{\text{Red}, \text{Blue}\}$. Then there exists an infinite monochromatic complete subgraph of $\mathbb{N}$.

Remark. This does not follow from the previous theorems, which only give proofs for larger and larger finite cliques.

Proof. Given $c : \mathbb{N}^{(2)} \to \{\text{Red}, \text{Blue}\}$, we construct a sequence $x_1, x_2, x_3, \ldots$ inductively as follows.

Let $x_1 \in \mathbb{N}$. x_1 has either an infinite blue neighborhood or an infinite red neighborhood. Define S_1 to be either $N_{\text{red}}(x_1)$ or $N_{\text{blue}}(x_1)$, depending on which one is infinite. Choose $x_2 \in S_1$ and let $S_2 = S_1 \cap N_{\text{red}}(x_2)$ or $S_2 = S_1 \cap N_{\text{blue}}(x_2)$ depending on which one is infinite. Continue this indefinitely.

Note that the sets $F_i = \{x_i x_j \mid j > i\}$ are all monochromatic (but possibly some are blue and some are red). Call x_i red if F_i is monochromatic in red and call x_i blue otherwise.

As every x_i is either red or blue, we must have infinitely many red or infinitely many blue vertices, WLOG assume we have infinitely many red vertices. Then we're done, as $X = \{x_i \mid x_i \text{ is red}\}$ works. □

14 Nov 2022, Lecture 17

Remarks.

- We can prove a k–color version of infinite Ramsey using an analogous argument to before.
- It can be hard to determine which of the two colors (or both) have the infinite monochromatic complete graph.

  As an (easy) example, color ij with $\max\{n : 2^n \mid i + j\} \pmod 2$. Then we can find an infinite subgraph $\{2^2, 2^4, 2^6, \ldots\}$.

  As a (hard) example, color ij with red if $i + j$ has as even number of distinct prime factors and with blue otherwise. In this case, we don't know how to construct such a set explicitly.

- It is possible to deduce the existence of $R(s,t)$ from infinite Ramsey, but it requires a trick (but apparently it's not too hard, example sheet difficulty).

Theorem 5.5 (Ramsey for r–sets)**.** Let c be a 2–coloring of $\mathbb{N}^{(r)}$ for some $r \geq 2$. Then there exists $X \subseteq \mathbb{N}, |X| = \infty$ for which $X^{(r)}$ is monochromatic.

Proof. We induct on r. $r = 2$ we've done already.

We define a sequence $x_1, x_2, x_3, \ldots$ and a sequence of infinite sets $S_1 \supset S_2 \supset S_3 \ldots$. We start by choosing $x_1 \in \mathbb{N}$. Now consider the coloring

$$c_{x_1}(F) = c(x_1 \cup F),$$

where $F \in (\mathbb{N} - x_1)^{(r-1)}$. By induction, there exists a set $S_1 \subset \mathbb{N} - x_1$ such that $|S_1| = \infty$ and $S_1^{(r-1)}$ is monochromatic with respect to the coloring c_{x_1}. Now choose $x_2 \in S_1$ and only consider $x_2, x_3, \ldots$ inside of S_1.

Iterating this, we get a sequence of points $x_1, x_2, x_3, \ldots$ which has the property that $F_i = \{\{x_i, x_{i_2}, \ldots, x_{i_r}\} \mid i < i_2 < \ldots < i_r\}$ is monochromatic.

Call an x_i red if F_i is monochromatic in red or blue if it is monochromatic in blue. WLOG we may assume that there are infinitely many red x_i, so $X = \{x_i \mid x_i \text{ is red}\}$ works. $\square$

Definition 5.4. For $r \in \mathbb{N}$, $s, t \geq 1$, define $R^{(r)}(s,t)$ to be the minimal value for n for which every 2–coloring of $[n]^{(r)}$ contains either

(a) $S \subseteq [n]$, $|S| = s$, $S^{(r)}$ is monochromatic red.

(b) $T \subseteq [n]$, $|T| = t$, $T^{(r)}$ is monochromatic blue.

Remark. We get by pigeonhole that $R^{(1)}(s,t) = s + t - 1$, and we also find $R^{(2)}(s,t) = R(s,t)$, $R^{(r)}(r,t) = t, R^{(r)}(s,r) = s$.

Lemma 5.6.

$$R^{(r)}(s,t) \leq R^{(r-1)}(R^{(r)}(s-1,t), R^{(r)}(s,t-1)) + 1 = (\star).$$

Proof. Apply induction on r, and then on $s + t$. We will be done if $s \leq r$ or $t \leq r$. Choose a vertex in a 2–colored set $[n]^{(r)}$ where $n \geq (\star)$. Let $x \in [n]$. Consider $c_x(F) = c(\{x\} \cup F)$, where $F = ([n] - x)^{(r-1)}$. We have either a set S_1 with $|S_1| = R^{(r)}(s-1,t)$ and $S_1^{(r-1)}$ is monochromatic red or a set S_2 with $|S_2| = R^{(r)}(s,t-1)$ and $S_2^{(r-1)}$ is monochromatic blue. We consider the first case, the second case being similar.

Apply Ramsey (the r–set version) to find either a set $A \subseteq S_1$ with $|A| = s-1$ and $A^{(r)}$ is monochromatic red (with respect to c), or $B \subseteq S_1$ with $|B| = t$ and $B^{(r)}$ is monochromatic blue.

In the latter case, we're done. In the former case, consider $A \cup \{x\}$, which is monochromatic red and has size s. $\square$

Quantitative approach to Ramsey theory

Question. How large is $R(s,t)$?

Proposition 5.7. For $s, t \geq 2$, we have

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

and in particular,

$$R(s,s) \leq 4^s.$$

Proof. Apply induction on $s+t$. We know $R(s,2) = s = \binom{s}{1}$ and simiarly for $R(2,t)$.

Now assume the claim is true for $(s-1,t)$ and $(t,s-1)$. We proved before that $R(s,t) \leq R(s-1,t) + R(s,t-1)$, so

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

□

16 Nov 2022, Lecture 18

Question. How large are $R^{(r)}(s,t)$? We had

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

So to get $R^{(r)}(s,t)$, we need to iterate $R^{(r-1)}$ $s+t$ times. Let's abstract this for a second. Let $f_1(x) = 2x$, and define

$$f_n(x) = \underbrace{f_{n-1}(f_{n-1}(\dots f_{n-1}}_{x \text{ times}}(x))).$$

Then $f_2(x) = x \cdot 2^x$, $f_3(x) = 2^{2^{\cdot^{\cdot^{\cdot^x}}}}$ with the 2 being iterated x times. So $f_4(1) = f_3(1) = 2$ and $f_4(2) = f_3(f_3(2)) = f_3(2^2) = 2^{2^{2^2}} = 65536$, and $f_4(3) = f_3(f_3(f_3(2))) = 2^{2^{2^{2^2}}} = 2^{65536}$, which is a pretty big number.

So our bound on $R^{(r)}(s,t)$ gives that it grows at least as fast as $f_r(s,t)$.

Exact Ramsey numbers

We know $R(3) = R(3,3) = 6$. In fact, there are very few known Ramsey numbers. We know $R(4) = 18$, but already $R(5)$ is unknown.

5.1 Lower bounds on $R(s)$

.

First idea: take $s-1$ disjoint copies of graphs of size $s-1$ and color everything inside these graphs blue and every edge between different copies red. This gives us a graph with no monochromatic K_s and tells us

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

This bound was thought to be close to the truth in the 1940s. However:

Theorem 5.8 (Erdös, 1940s)**.** For $s \geq 3$,

$$R(s) \geq s^{s/2}.$$

Idea: Take a K_n ($n = 2^{s/2}$) and color each edge of our K_n by flipping a fair independent coin. Then we show

$$\mathbb{P}(K_n \text{ has a monochromatic } K_s) < 1,$$

which means there is some coloring without a monochromatic K_s, so we get our desired lower bound.

Proof. Let $n \leq 2^{s/2}$ be an integer. Color the edges as described above.

$$\begin{aligned}
&\mathbb{P}(\text{the coloring has a monochromatic } K_s) = \\
&\mathbb{P}\Big(\bigcup_{k\in[n]^{(s)}} \{K \text{ is a monochromatic } K_s\}\Big) \leq \\
&\sum_{k\in[n]^{(s)}} \mathbb{P}(K \text{ is a monochromatic } K_s) = \\
&\binom{n}{s}\cdot 2 \cdot 2^{-\binom{n}{2}} < \frac{n^s}{s!}\cdot 2\cdot 2^{-\frac{s(s-1)}{2}} = \\
&2\left(\frac{n}{(s!)^{\frac{1}{s}}}2^{-\frac{s-1}{2}}\right)^s \leq 2\left(\frac{s^{\frac{1}{2}}}{(s!)^{\frac{1}{s}}}\right)^s = (\star).
\end{aligned}$$

But we know $s! \geq 2^{\frac{s}{2}+1}$ for $s \geq 3$, hence $(s!)^{\frac{1}{s}} \geq s^{\frac{1}{2}+\frac{1}{s}}$, so the above gives

$$(\star) \leq 2\left(\frac{1}{2^{\frac{1}{s}}}\right)^s \leq 1.$$

So $\mathbb{P}(\text{the coloring has a monochromatic } K_s) < 1$, so there exists a coloring with no monochromatic K_s. $\square$

Remarks.

- We can think about this proof as follows: start with every possible coloring, and then go through every clique and remove every monochromatic K_s.

Hence we want

$$2^{\binom{n}{2}} - 2^{\binom{n}{2}} \cdot 2 \cdot 2^{-\binom{s}{2}} \binom{n}{s} \geq 1.$$

This is the same calculation as we just did, but this is not as powerful of an approach.

- This proof does not help us at all in constructing such a coloring. In fact, it is a major open problem to construct a coloring to show $R(s) > (1+\epsilon)^s$.
- This is an example of a "random graph" argument.

6 Random graphs

Recall the Zarankiewicz number (n, t): we have a bipartite graph G with $V(G) = X \sqcup Y$, $|X| = |Y| = n$, how many edges can we add between X and Y before we get a $K_{t,t}$?

We proved $Z(n,t) \leq 2n^{2-\frac{1}{t}}$ for n much larger than t.

Theorem 6.1. Let $t \geq 2$, then $Z(n,t) \geq \frac{1}{2} n^{2-\frac{2}{t+1}}$.

Idea. Take X, Y, a bipartition with $|X| = |Y| = n$ and include each edge xy with $x \in X, y \in Y$ independently with probability p. Let $\mathcal{X}$ denote the number of $K_{t,t}$'s in G. Then

$$\begin{aligned}
\mathcal{X} &= \sum_{A \in X^{(t)}, B \in Y^{(t)}} \mathbb{1}(\text{all edges between } A, B \text{ are in } G) \\
\mathbb{E}[\mathcal{X}] &= \sum_{A \in X^{(t)}, B \in Y^{(t)}} \mathbb{P}(\text{all edges between } A, B \text{ are in } G) \\
&= \binom{n}{t}^2 p^{t^2} \leq \frac{n^{2t}}{4} p^{t^2} = \frac{1}{4}\left(n^2 p^t\right)^t = (\star).
\end{aligned}$$

So if $p = n^{-\frac{2}{t}}$, then $(\star) \leq \frac{1}{4}$. We have $\mathbb{P}(\mathcal{X} \geq 1) \leq \frac{1}{4}$. Note

$$\mathbb{E}[e(G)] = pn^2 = n^{2-\frac{2}{t}}.$$

So (by Markov's inequality, which we state on the next page)

$$\mathbb{P}\left(e(G) < \frac{pn^2}{2}\right) < \frac{1}{2}.$$

So with probability $> \frac{1}{4}$ we have that $e(G) > \frac{1}{2}pn^2 = \frac{1}{2}n^{2-\frac{2}{t}}$ and $G \not\supset K_{t,t}$. 18 Nov 2022, Lecture 19

Proof. Let $G = (X \cup Y, E)$ with X, Y disjoint, $|X| = |Y| = n$, and include each pair $xy \in E$ ($x \in X, y \in Y$) with probability $p = n^{-\frac{2}{t+1}}$. We let $\tilde{G}$ be the graph G with an edge deleted from each $K_{t,t}$. So $\tilde{G} \not\supset K_{t,t}$ by definition. But

$$e(\tilde{G}) \geq e(G) - (\# \text{ of } K_{t,t}\text{'s})$$
$$\mathbb{E}[e(\tilde{G})] \geq \mathbb{E}[e(G)] - \mathbb{E}[(\# \text{ of } K_{t,t}\text{'s})] = pn^2 - \binom{n}{t}^2 p^{t^2} \geq n^{2-\frac{2}{t+1}} - \frac{n^{2t}}{2} p^{t^2} = (\star).$$

As we have

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

we conclude

$$(\star) \geq \frac{1}{2} n^{2-\frac{2}{t+1}},$$

so there exists a graph $\tilde{G}$ which is $K_{t,t}$–free and has $e(\tilde{G}) \geq \frac{1}{2} n^{2-\frac{2}{t+1}}$. □

Question. Let G have large $\chi(G)$. Can we say anything about the structure of G?

First idea: If $\chi(G) > k$, does G contain a K_t?

This is totally wrong - by an example sheet question, we have graphs with arbitrarily large $\chi(G)$ and no triangles.

Second idea: If $\chi(G)$ is large, can we bound the girth of G (i.e. the length of the shortest cycle in G).

This is also wrong:

Theorem 6.2 (Erdös)**.** For all $k, g \geq 3$, there exists a graph G with $\chi(G) \geq k$ and girth $\geq g$.

Fact. (Markov) If X is a random variable which takes only non–negative values, then $\forall t > 0$ we have $\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$.

Fact. Let G be a graph. Then $\chi(G) \geq \frac{|G|}{\alpha(G)}$, where $\alpha(G)$ is the size of the largest independent set.

Proof. Let c be a coloring with $k = \chi(G)$ and let C_i be the set of all vertices that got colored with color i. Then

$$|G| = |C_1| + \ldots + |C_k| \leq k\alpha(G).$$

□

Proof of Theorem 6.2. Let G be a random graph on $[n]$, where each edge ij is included in $E(G)$ independently with probability $p = n^{-1+\frac{1}{g}}$. Let

$$X_i = \# \text{ of cycles in } G \text{ of length } i$$

and let $X = X_3 + \ldots + X_{g-1}$. By Markov,

$$\mathbb{P}(X \geq \frac{n}{2}) \leq \frac{2}{n}\mathbb{E}[X] = \frac{2}{n}\sum_{i=3}^{g-1}\mathbb{E}[X_i] = \frac{2}{n}\sum_{i=3}^{g-1}\frac{n(n-1)\ldots(n-i+1)}{i!}p^i \leq$$
$$\leq \frac{2}{n}\sum_{i=3}^{g-1}(np)^i = \frac{2}{n}\sum_{i=3}^{g-1}n^{\frac{i}{g}} \leq 10n^{-\frac{1}{g}} < \frac{1}{2}$$

for n large enough. Let

$$Y = \# \text{ of independent sets of size} = \frac{n}{2k}.$$

Let $s = \frac{n}{2k}$. We have (using $1 - x \leq e^{-x}$, which holds $\forall x \in (0,1)$)

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

for n sufficiently large.

We conclude that G has at most $\frac{n}{2}$ cycles of length $\leq g-1$ with probability $> \frac{1}{2}$, and G has $\alpha(G) \leq \frac{n}{2k}$ with probability $> \frac{1}{2}$.

So there exists a graph that satisfies both conditions. Let G be such a graph. Let $\tilde{G}$ be G with a vertex deleted from each cycle of length $\leq g-1$. So the girth of $\tilde{G}$ is $\geq g$, and

$$\chi(\tilde{G}) \geq \frac{|\tilde{G}|}{\alpha(\tilde{G})} \geq \frac{\frac{n}{2}}{\alpha(G)} \geq \frac{\frac{n}{2}}{\frac{n}{2k}} = k.$$

□

Definition 6.1 (Binomial random graph)**.** For $n \in \mathbb{N}, p \in [0,1]$, define the **binomial random graph** to be the probability space $G(n,p)$ on all graphs on n vertices, where each potential edge is included in the graph with probability p independently.

21 Nov 2022, Lecture 20

Question. For what values of $p \in (0,1)$ is $G \sim G(n,p)$ (a graph G sampled from the probability space $G(n,p)$) likely to contain a triangle? Note that here we're thinking of $p = p(n)$ as a function of n, and we mostly only care about large values of n.

Notation. If $(a_n), (b_n)$ are sequences of nonnegative numbers and $b_n \neq 0$ for all sufficiently large n, then we write $a_n \ll b_n$ if $\lim_{n\to\infty} \frac{a_n}{b_n} \to 0$.

Now let X be the random variable that counts the number of triangles in $G \sim G(n,p)$, so
$$\mathbb{E}[X] = \binom{n}{3} p^3.$$
Note that if $p \ll \frac{1}{n}$ (i.e. $pn \to 0$), then
$$\binom{n}{3} p^3 \le n^3 p^3 = (np)^3 \to 0.$$
By Markov,
$$\mathbb{P}(G \supset K_3) = \mathbb{P}(X \ge 1) \le \mathbb{E}[X] \to 0$$
for $p \ll \frac{1}{n}$.

Note that if $p \gg \frac{1}{n}$ ($pn \to \infty$), then
$$\binom{n}{3} p^3 \ge \frac{(n-3)^3 p^3}{6} \overset{\text{for } n \text{ large}}{\ge} \frac{1}{10}(np)^3 \to \infty.$$

Question. If $\mathbb{E}[X] \to \infty$, do we have $\mathbb{P}(G \supset K_3) \to 1$?

Fact. (Chebyshev) Let X be a random variable and $t > 0$. Then
$$\mathbb{P}(|X - \mathbb{E}[X]| \ge t) \le \frac{\mathrm{Var}(X)}{t^2}.$$

Fact. (Second moment method) Let X be a random variable. 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 6.3. Let $G \sim G(n,p)$ and X the number of triangles in G. Then
$$\lim_{n\to\infty} \mathbb{P}(G \supset K_3) = \begin{cases} 0 & \text{if } p \ll \frac{1}{n} \\ 1 & \text{if } p \gg \frac{1}{n}. \end{cases}$$

Proof. If $p \ll \frac{1}{n}$, then $\mathbb{E}[X] \to 0$ by above, so $\mathbb{P}(G \supset K_3) = 0$ by above.

If $p \gg \frac{1}{n}$, then we have

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

so we need to find $\mathrm{Var}(X)$. Now

$$\begin{aligned} X &= \sum_{\Delta \in [n]^3} \mathbb{1}(\Delta \text{ is a } K_3 \text{ in } G) \\ X^2 &= \sum_{\Delta_1, \Delta_2 \in [n]^3} \mathbb{1}(\Delta_1, \Delta_2 \text{ are } K_3\text{'s in } G) \\ \mathbb{E}[X^2] &= \sum_{\Delta_1, \Delta_2} \mathbb{P}(\Delta_1, \Delta_2 \text{ are } K_3\text{'s in } G). \end{aligned}$$

On the other hand,

$$\mathbb{E}[X]^2 = \sum_{\Delta_1, \Delta_2} \mathbb{P}(\Delta_1 \text{ is a } K_3 \text{ in } G)\, \mathbb{P}(\Delta_2 \text{ is a } K_3 \text{ in } G).$$

Hence, as triangles that don't share any edges cancel out,

$$\begin{aligned} &\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \leq \\ &\leq \sum_{\Delta_1, \Delta_2 \text{ share an edge}} \mathbb{P}(\Delta_1, \Delta_2 \text{ are } K_4\text{'s in } G) + \sum_{\Delta_1} \mathbb{P}(\Delta_1 \text{ is a } K_3 \text{ in } G) \leq \\ &\leq n^4 p^5 + \mathbb{E}[X]. \end{aligned}$$

Thus

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

$\square$

Remarks.

- We see a "phase transition" in $\mathbb{P}(G \supset K_3)$ – it goes from 0 to 1 in a "small" interval.
- What if $p = \frac{\lambda}{n}$ for $\lambda > 0$ fixed? We then have

  $$\lim_{n \to \infty} \mathbb{P}(G \supset K_3) = 1 - e^{-\lambda^2/6},$$

  but we will not prove this.

- Question: is it true that

$$\mathbb{E}[\# \text{ of copies of } H] \to \infty \implies \mathbb{P}(G \supset H) \to 1?$$

The answer to this is no in general (i.e. consider H to be 1000 disjoint vertices and one triangle and think what happens[1]). But if $K \subset H$ is the "densest" subgraph, then this is true.

Question. For which p is $G \sim G(n,p)$ likely to be connected?

23 Nov 2022, Lecture 21

Theorem 6.4. Let $G \sim G(n,p)$, then $\forall \epsilon > 0$ we have

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

Remark. This is an example of a "sharp threshold", which has a different qualitative behavior than the property "$G \supset K_3$, which is what we saw earlier.

Proposition 6.5. Let $G \sim G(n,p)$, then $\forall \epsilon > 0$ we have

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

Proof. Let I be the number of isolated vertices in $G \sim G(n,p)$. Then

$$\mathbb{E}[I] = n(1-p)^{n-1}.$$

Note that if $p \geq (1+\epsilon)\frac{\log n}{n}$, then (using $1 - x \leq e^{-x}$)

$$(1-p)\mathbb{E}[I] = n(1-p)^n \leq ne^{-pn} \leq ne^{-(1+\epsilon)\frac{\log n}{n}n} = n \cdot n^{-(1+\epsilon)} = n^{-\epsilon} \to 0.$$

So by Markov, $\mathbb{P}(G \text{ has an isolated vertex}) = \mathbb{P}(I \geq 1) \leq \mathbb{E}[I] \to 0$.

If $p \leq (1-\epsilon)\frac{\log n}{n}$, then

$$\mathbb{E}[I] = \frac{n(1-p)^n}{1-p} \geq ne^{-(1+\epsilon/4)pn} = (\star)$$

for n sufficiently large and ϵ sufficiently small. Now plug in p:

$$(\star) \geq ne^{-(1+\epsilon/4)(1-\epsilon)\log n} = ne^{-\log n + \frac{3\epsilon}{4}\log n + \epsilon^2 \log n/4} = n^{\frac{3\epsilon}{4} + \frac{\epsilon^2}{4}} \to \infty.$$

[1] $\mathbb{E}[\#H] = \binom{n}{1003}p^3 \approx n^{1003}p^3$, which goes to infinity when $p = n^{-\frac{1003}{3}} \ll n^{-1}$.

We now want to use the second moment method.

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

Hence

$$
\mathbb{P}(I=0) \leq \frac{\operatorname{Var}(I)}{\mathbb{E}[I]^2} \leq \frac{1}{\mathbb{E}[I]} + \left(\frac{1}{1-p} - 1\right) \to 0
$$

(since $p \to 0$ and $\mathbb{E}[I] \to \infty$). □

Proof of Theorem 6.4. (0–statement): We want to show that if $p \leq (1-\epsilon)\frac{\log n}{n}$, then $\lim_{n\to\infty} \mathbb{P}(G \text{ connected}) = 0$. But this follows from the fact that

$$
\mathbb{P}(G \text{ is connected}) \leq \mathbb{P}(G \text{ has no isolated vertices}) \to 0
$$

for this p by above.

(1–statement): Let $p \geq (1+\epsilon)\frac{\log n}{n}$. We need to show that

$$
\mathbb{P}(G \text{ is connected}) \to 1 \text{ as } n \to \infty.
$$

But

$$
\begin{aligned}
&\mathbb{P}(G \text{ is not connected}) = \\
&= \mathbb{P}(\exists A \subseteq V(G), 1 \leq |A| \leq \frac{n}{2}, e(A, V(G)\setminus A) = 0) \\
&= \mathbb{P}\left(\bigcup_{A\subseteq V(G), 0\leq |A| \leq \frac{n}{2}} \{(e(A), V(G)\setminus A) = 0\}\right) \\
&\leq \sum_{A\in V(G), 1\leq |A|\leq \frac{n}{2}} \mathbb{P}(e(A, V(G)\setminus A) = 0) \\
&= \sum_{k=1}^{n/2} \binom{n}{k} (1-p)^{k(n-k)} = (\star).
\end{aligned}
$$

Now we just need to show the last expression goes to zero, but we have

$$
\begin{aligned}
(\star) &\leq \sum_{k=1}^{\epsilon n/4} n^k e^{-pk(n-k)} + \sum_{k=\epsilon n/4}^{n/2} 2^n e^{-(1+\epsilon)\frac{\log n}{n}k(n-k)} \\
&\leq \sum_{k=1}^{\epsilon n/4} \left(ne^{-(1+\epsilon)\frac{\log n}{n}(n-k)}\right)^k + \sum_{k=\epsilon n/4}^{n/2} 2^n e^{-(1+\epsilon)\frac{\log n}{n}\frac{\epsilon n}{4}\frac{n}{2}} \\
&\leq \sum_{k=1}^{\epsilon n/4} \left(ne^{-(1+\epsilon)\frac{\log n}{n}n(1-\epsilon/4)}\right)^k + \sum_{k=\epsilon n/4}^{n/2} 2^n e^{-(1+\epsilon)\epsilon n\log n/8} \\
&\leq \sum_{k=1}^{\epsilon n/4} \left(ne^{-\log n-\frac{3\epsilon}{4}\log n+\frac{\epsilon^2}{4}\log n}\right)^k + \sum_{k=\epsilon n/4}^{n/2} 2^n e^{-(1+\epsilon)\epsilon n\log n/8} \to 0.
\end{aligned}
$$

□

28 Nov 2022, Lecture 23 - lecture 22 was cancelled due to strike action

7 Algebraic graph theory

Definition 7.1. Let G be a graph. The **diameter of** G is defined as

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

We have $\operatorname{diam}(G) = 1 \iff G$ is complete $\iff \binom{n}{2}$ edges.

Question. Now let diam(2). What is the minimum number of edges in a diameter 2 graph?

Question. What is the minimum $\Delta(G)$ over all graphs G with $\operatorname{diam}(G) = 2$ on n vertices?

Fact. Let G be a graph with $\operatorname{diam}(G) \geq 2$. Then $|G| \leq \Delta^2 + 1$.

Proof. Let $x \in V(G)$, then $V(G) \subseteq \{x\} \cup N(x) \cup N(N(x)) \setminus N(x)$, so

$$|V(G)| \leq 1 + \Delta + \Delta(\Delta - 1) = \Delta^2 + 1.$$

□

Definition 7.2. A **Moore graph** is a graph for which $|G| = \Delta^2 + 1$.

Remarks.

- If G is a Moore graph, then it is regular.
- A Moore graph is triangle–free.

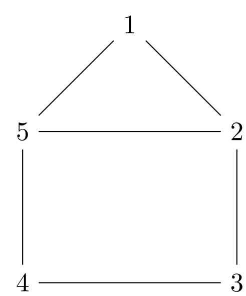
- More generally, a graph G is a Moore graph $\iff$ every distinct $x, y \in V(G)$ have a unique path of length ≤ 2 between them.

All of these follow from considering the equality case in the proof above.

Example 7.1.
- If $\Delta(G) = 2$, then C_5 is a Moore graph.
- If $\Delta(G) = 3$, then the Petersen graph is a Moore graph.
- If $\Delta(G) = 4$, we can draw our graph: draw the root vertex, 4 of its children, and then three leaves for each of the children. Now the question is the following: can we add the rest of the edges between the leaves to end up with a unique path of length ≤ 2 between each pair of vertices?

Definition 7.3. Let G be a graph on vertex set $[n]$. We define the **adjacency matrix** of G to be the $n \times n$ matrix $(A_G)_{xy} = \begin{cases} 1 & \text{if } xy \in E(G) \\ 0 & \text{otherwise} \end{cases}$ $\forall x, y \in [n]$.

Example 7.2. Let G be the following graph.



Then

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

Fact. A_G is symmetric and $(A_G)_{xx} = 0$ $\forall x \in V(G)$. So $\mathrm{tr}(A) = 0$.

Fact. Let G be a graph and A_G its adjacency matrix. Let $k \in \mathbb{N}$. Then

$$(A_G^k)_{xy} = \text{number of walks of length } k \text{ from } x \text{ to } y \text{ in } G.$$

Proof. If $k = 1$, then $xy \in E(G) \iff$ a path of length 1. If $k = 2$, then

$$(A_G^2)_{xy} = \sum_z (A_G)_{xz}(A)_{zy} = \sum_z \mathbb{1}(x \sim z)\mathbb{1}(z \sim y) = \text{the number of paths of length 2.}$$

For $k \geq 3$, proceed by induction. $\square$

We have that A_G acts on $\mathbb{R}^n \to \mathbb{R}^n$ as a linear map. What does this say about G?

Example 7.3. Let $G = C_4$, so $A_G = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$. Take some vector, e.g. $x = \begin{pmatrix} 1 & 2 & -2 & 3 \end{pmatrix}^\top$. We can think of this as labeling our vertices of the graph. Then $A_G x$ is just summing these labels for neighboring vertices, which gives us a new graph with labels $5, -1, 5, -1$, i.e. $A_G x = \begin{pmatrix} 5 & -1 & 5 & -1 \end{pmatrix}^\top$.

Fact. If A is a $n \times n$ symmetric matrix, then A has real eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$. Moreover, there exist eigenvectors $u_1, \ldots, u_n$ which form an orthonormal basis of $\mathbb{R}^n$.

Thus, given a graph on n vertices, we can talk about its eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$, which are the eigenvalues of A_G. We write $\lambda_{\max} = \lambda_1$ and $\lambda_{\min} = \lambda_n$.

Note that $\sum_{i=1}^n \lambda_i = \operatorname{tr}(A) = 0$, so if G is a nonempty graph, then $\lambda_{\max} > 0$ and $\lambda_{\min} < 0$.

Let's compute the eigenvalues of C_4. We can check that $\begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}^\top$ is an eigenvector with eigenvalue 2, preferably by thinking about it by summing neighboring labels. We also see that A_G has rank 2, so we have 0 as an eigenvalue and two zero–eigenvectors. Since the eigenvalues sum to zero, the last eigenvalue must be -2 (with eigenvector $\begin{pmatrix} 1 & -1 & 1 & -1 \end{pmatrix}^\top$).

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

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

Proposition 7.1. Let G be a graph.

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

(ii) Let G be connected. Then Δ is an eigenvalue $\iff$ G is regular; $\mathbb{1} = \begin{pmatrix} 1, \ldots, 1 \end{pmatrix}^\top$ is the corresponding eigenvector, and Δ has multiplicity 1.

(iii) Let G be connected. Then $-\Delta$ is an eigenvalue $\iff$ G is regular and bipartite.

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

30 Nov 2022, Lecture 24

Proof. (i) Let λ be an eigenvalue with an eigenvector $x = (x_1, \dots, x_n)^\top$. Let x_i be such that $|x_i|$ is maximal. We may assume that $x_i = 1$ by scaling. We have

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

(ii) ($\Longleftarrow$): If G is regular, then $\mathbb{1} = (1, \dots, 1)^\top$ is an eigenvector of G with eigenvalue Δ.

($\Longrightarrow$): For Δ an eigenvalue of G, let $x = (x_1, \dots, x_n)^\top$ be a corresponding eigenvector with $|x_i|$ maximal, and assume $x_i = 1$. We have

$$\Delta = \Delta x_i = \sum_{j \sim i} x_j \implies \deg(i) = \Delta \text{ and } x_j = 1 \ \forall j \sim i.$$

Now repeat this argument on each of the neighbors of i and continue until we reach every vertex in G (as G is connected) to see that G is Δ–regular and $x = (1, \dots, 1)^\top$. Since we've shown that $\mathbb{1}$ is the only eigenvector possible for eigenvalue Δ, we conclude that Δ has multiplicity 1.

(iii) ($\Longleftarrow$): If G is bipartite and regular with bipartition $V(G) = X \sqcup Y$, then $x = (\underbrace{1, \dots, 1}_{|X|}, \underbrace{-1, \dots, -1}_{|Y|})^\top$ is an eigenvector with eigenvalue $-\Delta$.

($\Longrightarrow$): Let $-\Delta$ be an eigenvalue of G with corresponding eigenvector $x = (x_1, \dots, x_n)$. Let x_i be such that $|x_i| = 1$ and assume that $x_i = 1$. We have

$$-\Delta = -\Delta x_i = \sum_{j \sim i} x_j \implies \deg(i) = \Delta \text{ and } x_j = -1 \ \forall j \sim i.$$

Now repeat this for all $j \in N(i)$ to see that $\deg(j) = \Delta$ and $x_k = 1 \ \forall k \sim j \sim i$. Analogously keep repeating until we reach all vertices in G and see that G is Δ–regular and $x = (1, \dots, 1, -1, \dots, -1)$ up to a permutation of the coordinates. Moreover, $V(G) = \{j \mid x_j = 1\} \cup \{j \mid x_j = -1\}$ defines a bipartition of G.

(iv) We know that

$$\lambda_{\max} = \max_{x \in \mathbb{R}^n \setminus 0} \frac{\langle x, Ax \rangle}{\langle x, x \rangle} \geq \frac{\langle \mathbb{1}, A\mathbb{1} \rangle}{\langle \mathbb{1}, \mathbb{1} \rangle} = \frac{1}{n} \sum_{i=1}^{n} \deg(i) \geq \delta(G).$$

□

Definition 7.4. We say that a graph G is (k, a, b)–**strongly regular** if:

- G is k–regular.
- For every $xy \in E(G)$, $|N(x) \cap N(y)| = a$.
- For every $x \neq y, x \not\sim y$ in G, $|N(x) \cap N(y)| = b$.

Example 7.4.

- C_4 is $(2, 0, 2)$–strongly regular.
- C_5 is $(2, 0, 1)$–strongly regular.
- In fact, any Moore graph is $(k, 0, 1)$–strongly regular.

Theorem 7.2 (Strongly regular graphs are rare)**.** If there exists a (k, a, b)–strongly regular graph on n vertices, then

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

Proof. Let A be the adjacency matrix of a (k, a, b)–strongly regular graph on n vertices. Note that

$$(A^2)_{xy} = \begin{cases} a & x \sim y. \\ b & x \not\sim y, x \neq y. \\ k & x = y. \end{cases}$$

This means that

$$A^2 = aA + b(J - I - A) + kI,$$

where J is the matrix of all ones. Hence

$$A^2 + (b-a)A + (b-k)I - bJ = 0 \ (\dagger).$$

We know by Proposition 7.1 that k is an eigenvalue of A with corresponding eigenvector $\mathbb{1} = (1, \ldots, 1)^\top$ and k has multiplicity 1. Let λ be an eigenvalue with $\lambda \neq k$, and let x be a corresponding eigenvector. Multiply $(\dagger)$ by x to get that

$$\lambda^2 x + \lambda(b-a)x + (b-k)x = 0$$

since $Jx = 0$, since $x \perp \mathbb{1}$ (we choose an orthonormal basis). Hence

$$\begin{aligned} &\lambda^2 + \lambda(b-a) + (b-k) = 0 \\ \implies\ &\lambda = \frac{1}{2}\left((a-b) \pm \sqrt{(b-a)^2 + 4(k-b)}\right). \end{aligned}$$

Let λ, μ be the solutions to this quadratic equation and assume that λ has multiplicity s and μ has multiplicity t. We have

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

hence

$$\begin{cases} s\lambda + t\mu = -k \\ s + t = n - 1. \end{cases}$$

Solve these for $s, t \in \mathbb{Z}$ to get the number in our theorem as desired. $\square$

Corollary 7.3. Let G be a Moore graph with $\Delta(G) = k$. Then $k \in \{2, 3, 7, 57\}$.

Proof. The idea is that if G is a Moore graph, then it is $(k, 0, 1)$–strongly regular on $k^2 + 1$ vertices. Hence we have

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

which gives $k \in \{2, 3, 7, 57\}$. $\square$