Combinatorics I Lecture Notes

Guilherme Zeus Dantas e Moura
zeusdanmou@gmail.com

IMPA

January – February 2021
Last update: February 26, 2021

This is IMPA's master class Combinatorics 1, instructed by Robert Morris, with the help of Letícia Mattos. All errors are my responsability.

Use these notes only as a guide. There is a non-trivial chance that some things here are wrong or incomplete (especially proofs).

This class is being taught remotely via Gooogle Meet and YouTube videos. The recommended material can be found here.

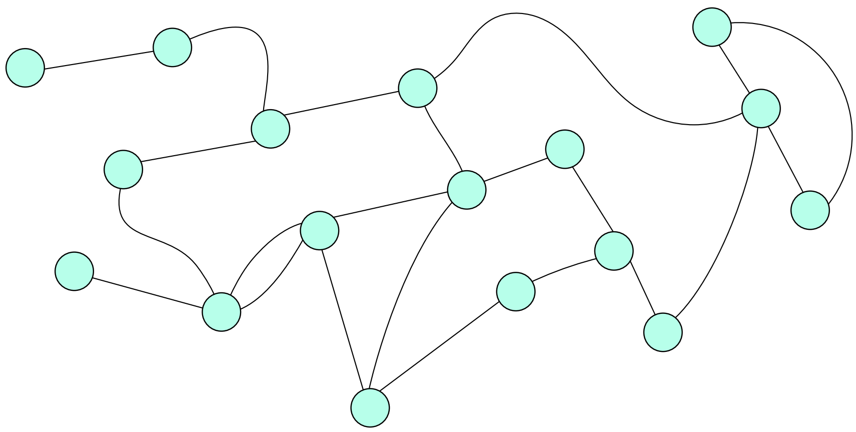

Figure 1: This is a graph.

Contents

1 Which problems we'll study?

YouTube, Lec. 1
January 04, 2021

In this summer course, we'll study extremal, counting and probabilistic problems. Here are some examples:

Problem 1.1

Let A be a subset of $\{1, 2, \ldots, 2n\}$ such that $a \nmid b$, for all $a \neq b \in A$.

How large can $|A|$ be?

Solution. $A = \{n+1, \ldots, 2n\}$ is a good example. This yields $|A| = n$.

Consider the partition of $\{1, 2, \ldots, 2n\}$ given by the following sets:

- $\{2^t\}$
- $\{3 \cdot 2^t\}$
- $\{5 \cdot 2^t\}$

$\vdots$

- $\{(2n-1) \cdot 2^t\}$

There can't be two elements in the same set of the partition, so $|A| \leq n$.

Problem 1.2

Let A be a subset of $\{1, 2, \ldots, 2n\}$ such that $a + b \neq c$, for all $a, b, c \in A$. We'll call such set *sum-free*.

How large can $|A|$ be?

Solution. $A = \{n+1, \ldots, 2n\}$ is a good example. Another good example are the odd numbers. Both yield $|A| = n$.

Suppose $|A| \geq n+1$. Let $a = \max A$.

Consider the following partition with $\left\lfloor \frac{a}{2} \right\rfloor$ sets:

- $\{1, a-1\}$
- $\{2, a-2\}$

$\vdots$

- $\{\left\lfloor \frac{a}{2} \right\rfloor, \left\lceil \frac{a}{2} \right\rceil\}$

There can't be two elements in the same set of the partition.

If $a \leq 2n-1$, then there are at most $n-1$ sets listed above, which implies $|A| \leq n$.

If $a = 2n$, then $n \notin A$, and then the $n-1$ first sets listed above cover A, thus $|A| \leq n$.

Theorem 1.1 (Schur, 1916)

Given $c\colon \mathbb{Z}_{>0} \to \{1, \ldots, r\}$, the there are x, y, z such that:

- $x + y = z$
- $c(x) = c(y) = c(z)$

Problem 1.3

How many sum-free sets are in $[n]$?

Conjecture 1.2 (Cameron and Erdős)

The number of sum-free sets in $[n]$ is $\leq C \cdot 2^{n/2}$.

2 Ramsey's Theory

Theorem 2.1 (Ramsey's Theorem)

If $c : \binom{\mathbb{N}}{2} \to \{1, \ldots, r\}$, then there exists $A \subset \mathbb{N}$ infinite and monochromactic, i.e, such that $c(ab) = c$, for all $a, b \in A$.

Proof of Theorem 2.1. Let $S_0 = \mathbb{N}$.

For each i, do the following: Pick $v_i \in S_{i-1}$. Look at the colors of $\{v_i, u\}$, for u in S_{i-1}. Since S_{i-1} is infinite and there are finitely many colors, there is some color that appears infinitely many times; we'll call this color c_i, and define $S_i = \{u \in S_{i-1} : c(\{v_i, u\}) = c_i\}$.

Now, we have an infinite sequence $v_1, v_2, \ldots$, such that $c(\{v_i, v_j\}) = c_i$, for $i < j$. Since there are finitely many colors, there is some color that appears in infinitely many c_i's; call this color c, and define $A = \{v_i : c_i = c\}$.

The set A satisfies our condition.

Proof of Theorem 1.1. Given a coloring $c : \mathbb{N} \to \{1, \ldots, r\}$, we define $c' : \binom{\mathbb{N}}{2} \to \{1, \ldots, r\}$ by $c'(\{a, b\}) = c(b - a)$, for $b > a$.

By Theorem 2.1, there is A infinite and monochromactic. Pick $x < y < z \in A$, then we have $c(y - x) = c(z - y) = c(z - x)$, and $(y - x) + (z - y) = z - y$, so we're done!

Definition 2.2 (Ramsey Number)

Let $R(k)$ denote the smallest n such that, for every coloring with two colors of the edges of the complete graph K_n, i.e., for every $c : E(K_n) \to \{R, B\}$, there exists a monochromatic copy of K_k.

Let $R(s, t)$ denote the smallest n such that, for every coloring with two colors of the edges of the complete graph K_n, i.e., for every $c : E(K_n) \to \{R, B\}$, there exists a red copy of K_s or a blue copy of K_t.

Clearly, $R(k) = R(k, k)$.

Theorem 2.3

$$R(k) \lesssim 2^{2k}.$$

Sketch. Let $n = 2^{2k}$, and pick any coloring c of K_n. Let $S_0 = [n]$.

For each $i < 2k$ do the following: Pick $v_i \in S_{i-1}$. Look at the colors of $\{v_i, u\}$, for u in S_{i-1}. There is some color that appears more times; we'll call this color c_i, and define $S_i = \{u \in S_{i-1} : c(\{v_i, u\}) = c_i\}$.

~~Note that, the size of S_m is at least $\frac{n}{2^m} \geq 1$.~~ This is not quite correct. At each step, we're taking one vertice away, and then dividing by two. We'll properly prove a better bound later.

Now, we have an sequence $v_1, v_2, \ldots, v_{2k-1}$, such that $c(\{v_i, v_j\}) = c_i$, for $i < j$. Since there are two colors, there is some color that appears at least k times; call this color c, and define $A = \{v_i : c_i = c\}$. The size of A is at least k. Pick any subset B of A that has exactly k elements.

The subgraph of K given by deleting all vertices but those in B is a monochromatic copy of K_k.

Lemma 2.4

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

Proof. Let $n = R(s,t) - 1$. By definition, there exists a coloring $c\colon E(K_n) \to \{R, B\}$ without a red K_s or a blue K_t.

Pick any vertex v. v it connected to some of the vertices through a red edge, which we'll put in the set S_R; the others are connected to v through a blue edge, those we'll put in the set S_B.

Since there are no red K_s or blue K_t, there can't be any red K_{s-1} or blue K_t in S_R; thus, $|S_R| \leq R(s-1,t)$. Analougously, $|S_B| \leq R(s,t-1)$.

Therefore,

$$\begin{aligned} R(s,t) - 1 &\leq R(s-1,t) - 1 + R(s,t-1) - 1 + 1 \\ R(s,t) &\leq R(s-1,t) + R(s,t-1). \end{aligned}$$

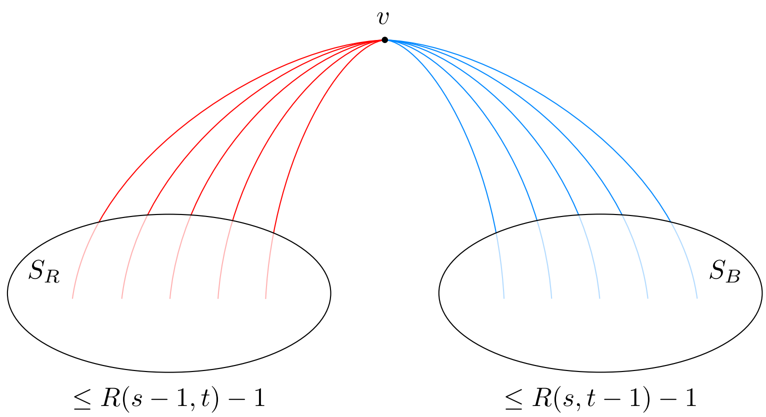

Figure 2: S_R and S_B.

Theorem 2.5

$$R(s,t) \leq \binom{s+t}{s}.$$

Proof. Follows from Lemma 2.4.

Theorem 2.6 (Erdős-Szekeres, 1935)

$$R(k) \leq \binom{2k}{k} \approx \frac{1}{\sqrt{k}} 4^k.$$

YouTube, Lec. 2
January 05, 2021

Let's now try to find a lower bound. It is very difficult to show a good construction. Luckly, we are not going to do that.

Theorem 2.7 (Erdős, 1947)

$$\sqrt{2}^k \leq R(k)$$

Proof. Let $n \leq \sqrt{2}^k$. Let's choose colors randomly. Let

$$\mathbb{P}(c(e) = R) = \frac{1}{2},$$

for every edge e in K_n, independently.

We want to show that

$$\mathbb{P}(\text{there is not a monochromatic copy of } K_k) > 0.$$

Let X be the number of monochromactic copies of K_k in c. Then,

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}\left[\sum_{\substack{S \subset K_n \\ S \text{ is a copy of } K_k}} \mathbb{1}[S \text{ is monochromatic}]\right] \\
&= \sum_{\substack{S \subset K_n \\ S \text{ is a copy of } K_k}} \mathbb{E}\left[\mathbb{1}[S \text{ is monochromatic}]\right] \\
&= \sum_{\substack{S \subset K_n \\ S \text{ is a copy of } K_k}} \mathbb{P}(S \text{ is monochromatic})) \\
&= \sum_{\substack{S \subset K_n \\ S \text{ is a copy of } K_k}} \left(\frac{1}{2}\right)^{\binom{k}{2}-1} \\
&= \binom{k}{n}\left(\frac{1}{2}\right)^{\binom{k}{2}-1} \\
&\leq 2\left(\frac{en}{k}\right)^k \left(\frac{1}{2}\right)^{\frac{k(k-1)}{2}} \\
&\leq 2\left(\frac{e\sqrt{2}}{k}\right)^k \\
&< 1, \text{ for } k \geq 5.
\end{aligned}$$

Therefore, since $\mathbb{E}[X] < 1$, we have $\mathbb{P}(X = 0) > 0$.

The bounds have not improved much since then

Theorem 2.8 (Conlon, 2009)

$$R(k) \leq \frac{4^k}{k^{\sqrt{\log k}}}$$

3 Extremal Graph Theory

3.1 Complete Graphs

Definition 3.1

Let $\text{ex}(n, H)$ be the maximum number of edges a graph $G \subset K_n$ can have such that there are no copies of H in G.

Theorem 3.2 (Mantel, 1907)

$$\text{ex}(n, K_3) = \left\lfloor \frac{n^2}{4} \right\rfloor.$$

Proof. The example is the bipartite graph with $\lfloor \frac{n}{2} \rfloor$ and $\lceil \frac{n}{2} \rceil$ vertices.

Let's prove by indution on n.

Now, suppose G does not have a triangle. Pick an edge uv. Let G' be the graph G deleting u and v. The subgraph G' also does not contain triangles, so $e(G') \geq \left\lfloor \frac{n^2}{4} \right\rfloor$.

Notice that cannot exist $w \in G'$ such that uw and vw are edges of G, because G does not have triangles. Therefore, there can be at most $n-2$ edges from u or v to vertices on G'. Including the edge uv, we conclude that

$$\begin{aligned} e(G) &\leq e(G') + n - 1 \\ &\leq \left\lfloor \frac{(n-2)^2}{4} \right\rfloor + n - 1 \\ &\leq \left\lfloor \frac{n^2}{4} \right\rfloor. \end{aligned}$$

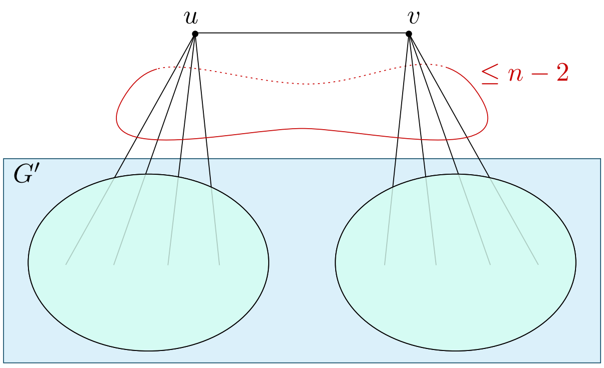

Figure 3: Edge uv on a triangle-free graph.

Definition 3.3 (Turán's Graph)

The graph $T_r(n)$ consists of r sets with roughly n/r elements each (some rounded up, some rounded down).; we create an edge uv if, and only if, u and v are on different sets.

We'll denote by $t_r(n)$ the number of edges in $T_r(n)$.

Theorem 3.4 (Turán, 1941)

$$\text{ex}(n, K_{r+1}) = t_r(n) \approx \left(1 - \frac{1}{r}\right)\binom{n}{2}.$$

Proof. We'll use induction on n. For $n \leq r$, we're good.

Pick a maximal graph G that doesn't have a copy of K_{r+1}. Pick a copy of K_r, let's call it H. Define $G' = G - H$. Of course, G' has no copies of K_r; thus $e(G') \leq t_r(n-r)$, by induction.

Futhermore, if $v \in G'$, there can be at most $r-1$ edges connecting v to some vertex in H.

Wrapping everything up, we have

$$\begin{aligned} e(G) &\leq e(G') + (n-r)(r-1) + \binom{r}{2} \\ &\leq t_r(n-r) + (n-r)(r-1) + \binom{r}{2} \\ &\leq t_r(n). \end{aligned}$$

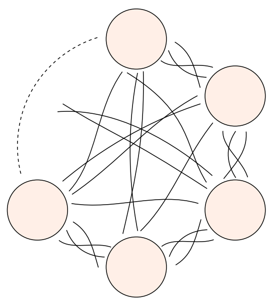

Figure 4: Turán's Graph

3.2 Bipartite Graphs

Theorem 3.5 (Erdős, 1935)

$$\text{ex}(n, C_4) \leq \frac{n^{3/2}}{2}.$$

Proof. Let's count cherries! A *cherry* is a pair $(v, \{u, w\})$, in which vu and vw are edges of the graph.

Since there is no C_4, there is at most one cherry for each pair $\{u, w\}$. This implies that:

$$\begin{aligned} \binom{n}{2} \geq \#(\text{cherries}) &= \sum_{v \in G} \binom{d(v)}{2} \\ &\geq n \binom{\frac{2e(G)}{n}}{2}. \end{aligned}$$

Solving this quadractic inequation on $e(G)$ yields to

$$e(g) \geq \frac{n^{3/2}}{2}.$$

Question 3.1

For which graphs we have

$$\text{ex}(n, H) = \Theta(n^2)?$$

Proposition 3.6

For every non-bipartite graph H, we have

$$\mathrm{ex}(n, H) \geq \frac{n^2}{4}.$$

Proof. Take G as the complete bipartite graph with n vertices. It has roughly $\frac{n^2}{4}$ vertices and it cannot contain a non-bipartite graph.

Theorem 3.7 (Kővári–Sós–Turán, 1954)

Let H be a bipartite graph. Then,

$$\mathrm{ex}(n, H) = o(n^2).$$

Proof. Since H is bipartite, there is some $K_{s,t}$ such that $H \subset K_{s,t}$. Then,

$$\mathrm{ex}(n, H) \leq \mathrm{ex}(n, K_{s,t}).$$

Let's bound $\mathrm{ex}(n, K_{s,t})$.

We'll count s-cherries: (v, S), in which S has size s and $vx \in E(G)$ for all $x \in S$.

There are at most $t-1$ s-cherries for each subset S with size s. This implies that:

$$\begin{aligned}(t-1)\binom{n}{s} \geq \#(\text{cherries}) &= \sum_{v \in G} \binom{d(v)}{s} \\ &\geq n\binom{\frac{2e(G)}{n}}{s} \geq \frac{e(G)^s}{s^s \cdot n^{s-1}}.\end{aligned}$$

This implies that, for some constant C,

$$e(G) \leq C \cdot n^{2-\frac{1}{s}}$$

Question 3.2

For which H it holds that

$$\mathrm{ex}(n, H) = O(n)?$$

3.3 Trees

YouTube, Lec. 3
January 07, 2021

Definition 3.8 (Tree)

A tree is a connected graph that has no cycles.

Proposition 3.9

Given a graph G, the following are equivalent:

- (i) G is a tree;
- (ii) G is a maximal graph without cycles, i.e., G does not have cycles and there is no graph $H \supset G$ such that H does not have cycles;
- (iii) G is a minimal connected graph, i.e., G is connected and there is no graph $H \subset G$ such that H is connected.

Theorem 3.10

Let T be a graph with k vertices. Then,

$$\frac{(k-2)}{2}n \leq \text{ex}(n,T) \leq (k-1)\cdot n.$$

Proof of the lower bound. Pick $\frac{n}{k-1}$ disjoint $k-1$-cliques. There cannot be a copy of a connected graph with k vertices inside this graph. This graph has roughly

$$\binom{k-1}{2}\frac{n}{k-1} = \frac{k-2}{2}n$$

edges.

Proof of the upper bound. Let's start with a lemma.

Lemma 3.11

Let G be a graph with mean degree d, then, there exists a subgraph $G' \subset G$ with minimum degree at least $\frac{d}{2}$.

Proof. While there are vertices with degree smaller than $\frac{d}{2}$, throw them away.

If we stopped before throwing away all vertices, we're done. Suppose we threw away all vertices. At each step, we threw away at most $\frac{d}{2}$ edges. Since we threw away all edges, this means $n \cdot \frac{d}{2} < e(G) = n\frac{d}{2}$; a contradiction.

Lemma 3.12

Let G be a graph with $\delta(G) \geq k-1$. Then, there is a copy of T in G for every tree T with k vertices.

Proof. We'll use induction on k. If $k=1$, we're done!

Pick a leaf v of T. Its unique edge connects it to u. Let T' be the tree without v. By induction, there is a copy $C_{T'}$ of T' in G. Let c_u be the copy of u in $C_{T'}$. Since $\deg(c_u) \leq k-2$ in $C_{T'}$, but $\deg(c_u) \geq k-1$ in G, there is some vertex that is connected to u outside of $C_{T'}$, say c_v. Thus, let C_T be $C_{T'}$, adding c_v. C_T is a copy of T inside G.

Finally, $e(G) = (k-1)n \implies \bar{d}(G) = 2(k-1) \implies$ there exists a subgraph $G' \subset G$ such that $\delta(G') \geq k-1 \implies T \subset G'$.

Conjecture 3.13 (Erdős-Sós, 1960's)

$$\text{ex}(n,T) \leq \frac{(k-2)n}{2}$$

Definition 3.14 (Random graph of Erdős-Rónyi)

We define $G(n,p)$ as a random distribution of graphs with n vertices, with

$$\mathbb{P}(e \in E(G(n,p))) = p,$$

chosen independently.

Lemma 3.15 (Markov's inequality)

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

Proof. Left to the reader. Use the definition of $\mathbb{E}[X]$.

Theorem 3.16

$$\text{ex}(n, C_t) \geq O\left(n^{1+\frac{1}{2k-1}}\right) \gg n.$$

Proof. Let $t = 2k$. We want to choose $p = p(n)$ such that:

- $e(G(n,p)) \gg n$;
- $C_{2k} \not\subset G(n,p)$.

$$\mathbb{E}[e(G(n,p))] = p\binom{n}{2}.$$

Moreover, $e(G(n,p))$ is a binomial distribution, therefore, $e(G(n,p)) \approx np^2$ with high probability. Thus, we should pick $p \gg 1/n$, i.e., $pn \to \infty$.

Define X as the number of copies of C_{2k} in $G(n,p)$.

$$\begin{aligned}\mathbb{E}[X] &= \sum_{\substack{\text{copies } S \text{ of} \\ C_{2k} \text{ in } K_n}} \mathbb{P}(S \subset G(n,p)) \\ &\approx n^{2k}p^{2k} = (pn)^{2k}.\end{aligned}$$

Let $0 < \varepsilon < \frac{1}{2k-1}$, and define $p = p(n) = n^{-1+\varepsilon}$. Then, we have $pn \gg n^{-1}$ and $(pn)^{2k} \ll pn^2$. Therefore, each of the following happen with high probability:

- $e(G(n,p)) \approx pn^2$;
- The number of copies of C_{2k} in $G(n,p) \approx (pn)^{2k}$.

Therefore, the intersection also occours with high probability. Pick a graph G in the intersection.

For each of the $(pn)^{2k}$ cycles in G delete an edge in it; call this new graph G'. Thus $e(G') \approx pn^2 - (pn)^{2k} \approx n^{1+\epsilon}$, and G' has no C_{2k}.

Theorem 3.17

$$\text{ex}(n, H) = O(n) \iff H \text{ does not have cycles.}$$

Proof. All the work has been done. The proof, which is simply a jigsaw puzzle, is left to the reader.

4 Planar graphs

Definition 4.1 (Planar Graph)
A planar graph is a graph that can be drawn on the plane without having crossing edges. Edges may not be straight.

Lemma 4.2 ($V + F - E = 2$)
Let G be a planar connected graph, and $v(G) \geq 1$. For any planar drawing of G, we have

$$v(G) + f(G) - e(G) = 2.$$

Sketch. Induction on $e(G)$.

(i) **If there is a leaf,** then we can take it away.

$$\begin{aligned} v(G') &= v(G) - 1, \\ e(G') &= e(G) - 1, \\ f(G') &= f(G). \end{aligned}$$

(ii) **If there is no leaf,** there is a cycle, take away an edge from the cycle.

$$\begin{aligned} v(G') &= v(G), \\ e(G') &= e(G) - 1, \\ f(G') &= f(G) - 1. \end{aligned}$$

Watch an animated version of this classic demonstration at 3Blue1Brown.

Theorem 4.3
Let G be a planar graph with $n \geq 3$ vertices. Then,

$$e(G) \leq 3n - 6$$

Proof. Without loss of generalitym G is maximal.

Maximal and $n \geq 3$ implies all regions are triangles. Double counting implies

$$3f(G) = 2e(G).$$

Also,

$$v(G) + f(G) - e(G) = 2.$$

It follows that

$$e(G) = 3n - 6.$$

Theorem 4.4
K_5 is not planar.

Proof.

$$e(K_5) = 10 > 3 \cdot 5 - 6 = 3v(K_5) - 6.$$

Theorem 4.5
Let G be a triangle-free planar graph with $n \geq 4$ vertices. Then,

$$e(G) \leq 3n - 6$$

Proof. Without loss of generalitym G is maximal.

Maximal and $n \geq 4$ implies all regions have at least 4 sides. Double counting implies

$$4f(G) \leq 2e(G).$$

Also,

$$v(G) + f(G) - e(G) = 2.$$

It follows that

$$e(G) = 2n - 4.$$

Theorem 4.6

$K_{3,3}$ is not planar.

Proof. $K_{3,3}$ is triangle-free.

$$e(K_{3,3}) = 9 > 2 \cdot 6 - 4 = 2v(K_{3,3}) - 4$$

Theorem 4.7

G is planar if, and only if, G does not have a topological copy of K_5 or $K_{3,3}$ if, and only if, G does not have a K_5-minor or a $K_{3,3}$-minor.

5 More colors

Definition 5.1 (Chromatic Number of a Graph)

The chromatic number of G, denoted by $\chi(G)$, is the smallest r such that there is a coloring $c\colon V(G) \to [r]$ such that $c(u) \neq c(v)$ whenever $uv \in E(G)$.

YouTube, Lec. 4
January 11, 2021

Definition 5.2

Let $r(G, H)$ denote the minimum n such that, for every coloration $c\colon E(K_n) \to \{R, B\}$, there must exist a red G or a blue H.

Proposition 5.3

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

Sketch. Greedy algorithm.

Theorem 5.4 (4-color Theorem, 1970's)

If G is planar, then $\chi(G) \leq 4$.

Proposition 5.5

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

Proof. Induction on n.

Since G is planar, $e(G) \leq 3n - 6$, thus $\delta(G) \leq 5$. Pick v with degree at most 5. Define G' as G without v, then G' has a proper coloring. Now, v has at most five neighbors, thus we can pick one color for v out of six such that no neighbor of v has this color.

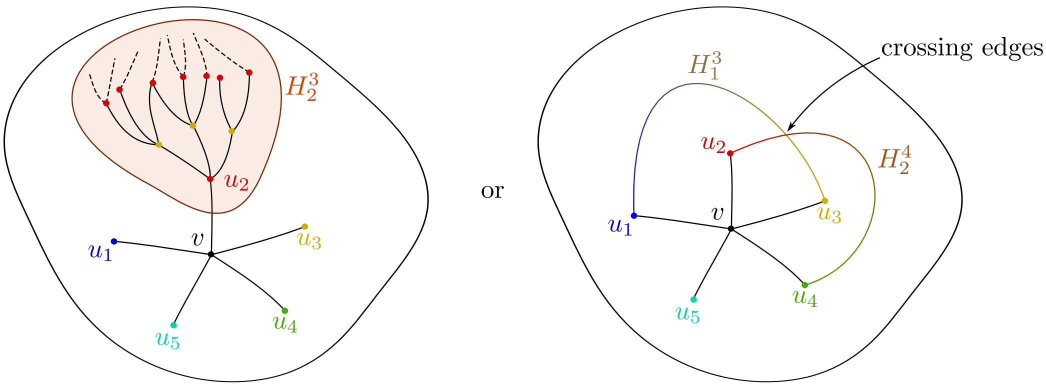

Figure 5: Five color theorem

Theorem 5.6

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

Proof. Induction on n.

Since G is planar, $e(G) \leq 3n - 6$, thus $\delta(G) \leq 5$. Pick v with degree at most 5. Define G' as G without v, then G' has a proper coloring. Now, v has at most five neighbors. If there at most four colors are used in the neighbors of v, we can paint v with a distinct color.

Suppose all neighbors of v have different colors. Let's call the neightbors u_1, u_2, u_3, u_4, u_5, in clockwise order, with colors $1, 2, 3, 4, 5$.

Define ${G'}_a^b$ as the subgraph of G' that only contains vertices with colors a and b. Let H_a^b be the connected component of ${G'}_a^b$ that contains u_a.

- **If there exists a, b such that $u_b \notin H_a^b$,** then we flip the colors a and b inside H_a^b and define $c(v) := a$.
- **If, for all a, b, $u_b \in H_a^b$,** $H_{1,3}$ and $H_{2,4}$ are vertex disjoint, but have to go through each other; a contradiction. See fig. 5.

Lemma 5.7

Se T é uma árvore, então $\chi(T) \leq 2$.

Sketch 1. Induction on number of vertices. Paint a leaf by the oposite color of its neighbor.

Theorem 5.8 (Erdős-Stone, 1946)

$$\operatorname{ex}(n, H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right)\binom{n}{2}.$$

Sketch. The example is the Turán's Graph $T_{\chi(H)-1}(n)$.

Let $r = \chi(H)$. We'll show it by induction on r.

If $r \leq 2$, then the theorem says $\operatorname{ex}(n, H) = o(n^2)$, which is true by Kővári-Sós-Turán, 1954.

Lemma 5.9

Let $\varepsilon > 0$ such that $\epsilon\binom{n}{2} > \binom{m_0}{2}$ and G be a graph with density β. Then, there exists $G^* \subset G$ with $m \geq m_0$ vertices and

$$\delta(G^*) \geq (\beta - \varepsilon)m.$$

Sketch. Throw away vertices with small degree. The first one we threw away had degree at most $< (\beta - \varepsilon)n$, the second one had degree at most $< (\beta - \varepsilon)(n - 1)$, and so on.

If we threw $n - m_0$ vertices away, then

$$\begin{aligned} e(G) &< (\beta - \varepsilon)(n + (n-1) + \cdots + m_0) + \binom{m_0}{2} \\ &< (\beta - \varepsilon)\binom{n}{2} + \binom{m_0}{2} \\ &< \beta\binom{n}{2}. \end{aligned}$$

The graph H is contained in $K_r(t)$, the complete r-partite with t vertices on each part, with $t = t(H)$.

Suppose $e(G) \geq \left(1 - \frac{1}{r-1} + \alpha\right)\binom{n}{2}$. Applying Lemma 5.9 with $\varepsilon = \frac{\alpha}{2}$, $m_0 = \frac{\alpha n}{2}$, and $\beta = 1 - \frac{1}{r-1} + \alpha$, we conclude that there exists $G^* \subset G$, with $m \geq \frac{\alpha n}{2}$ vertices, and $\delta(G^*) \geq \left(1 - \frac{1}{r-1} + \frac{\alpha}{2}\right)m$.

Induction hypothesis implies that, for large n, G^* has a copy F of $K_{r-1}(q)$, the complete $(r-1)$-partide graph with q vertices on each part, for $q > \frac{2(t-1)}{(r-1)\alpha}$.

Let $X = V(G^*) \backslash V(F)$. Let Y be the set of vertices in X that have at least $(r-2)q + t$ neighbors in $V(F)$.

Let's call $F_1, F_2, \ldots, F_{r-1}$ the parts of F, a complete $(r-1)$-partite graph. Let's count the number of *hyper-cherries* $(v, S_1, S_2, \ldots, S_{r-1})$, in which $v \in X$, $S_1 \subset F_1$, ..., $S_{r-1} \subset F_{r-1}$, and $v \sim u$, for all u in some S_i. See fig. 6.

For each vertex v in Y (of $|Y|$), there are $\prod_i \binom{\deg_i(v)}{t} \geq \binom{q}{t}^{r-2}$ hyper-cherries. On the other hand, for each possible subsets $S_1, \dots, S_{r-1}$ (of $\binom{q}{t}^{r-1}$), there are at most $t-1$ hyper-cherries. This implies

$$|Y| \leq (t-1)\binom{q}{t}.$$

Thereore, the number of edges between X and $V(F)$ is at most

$$\left(m - (r-1)q - \binom{q}{t}(t-1)\right)((r-2)q + t - 1) + \binom{q}{t}(t-1)(r-1)q,$$

which simplifies to

$$m((r-2)q + t - 1) + \text{constant}.$$

On the other hand, since every vertex of $V(F)$ has degree at least $\left(1 - \frac{1}{r-1} + \frac{\alpha}{2}\right)m$ the number of vertices between X and $V(F)$ is at least

$$\left(\left(1 - \frac{1}{r-1} + \frac{\alpha}{2}\right)m - (r-2)q\right)(r-1)q,$$

which simplifies to

$$m\left((r-2)q + \frac{(r-1)q\alpha}{2}\right) + \text{constant},$$

which yeilds to a contradiction to large n (i.e. large m).

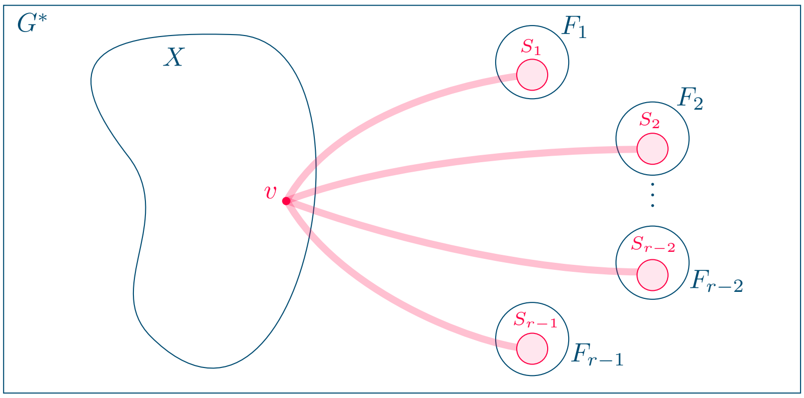

Figure 6: Hyper-cherry

6 Ramsey's Theory again

YouTube, Lec. 5
January 18, 2021

Definition 6.1

Let $R_r^{(k)}(m)$ is the minimal n such that, for all colorings $c\colon \binom{[n]}{k} \to [r]$, there exists a monochromatic copy of $K_m^{(k)}$.

We'll consider $r = 2$ and $k = 2$, if not otherwise stated.

Remark. $K_m^{(k)}$ is the k-uniform complete hypergraph with n vertices. $E\left(K_m^{(k)}\right) = \binom{V(K_n^{(k)})}{k}$. See Wikipedia.

Theorem 6.2 (Ramsey, 1930)

$$R_r^{(k)}(m) < \infty.$$

Sketch. Induction on k.

Pick $v_1 \in G$. Given $c\colon \binom{V(G)}{k} \to [r]$, define $c_1 : \binom{v(G)\setminus\{v\}}{k-1}$. Induction hypothesis implies that there exists a monochromatic copy of $K_{m_1}^{(k-1)}$, for $n \geq R_r^{(k-1)}(m_1)$.

Repeat the process inside this copy of $K_{m-1}^{(k-1)}$.

Similarly to the proof of Theorem 2.1, we'll have a sequence $v_1, v_2, \ldots, v_\ell$ (that gets larger as n gets larger), for which $c(\{v_{a_1}, v_{a_2}, \ldots, v_{a_k}\}) = f(a_1)$, if $a_1 < a_2 < \cdots < a_r$.

Pick large n such that $\ell \geq (r-1)m + 1$, for which there exists a subsequence $a_{b_1}, \ldots, a_{b_r}$ such that $f(a_{b_i})$ is the same for all i.

Theorem 6.3 (Erdős–Hajnal)

$$R^{(k)}(m) \leq 2^{\binom{R^{(k-1)}(m)}{k-1}}$$

Sketch for $k = 3$. Suppose $e(G) \gtrsim 2^{\binom{R(m)}{2}}$

Pick a edge $v_1v_2 \in E(G)$. Given $c\colon \binom{V(G)}{3} \to \{1,2\}$, define $c'\colon \binom{V(G)\setminus\{v_1,v_2\}}{2} \to \{1,2\}$ by $c'(v) := c(v_1v_2v)$. The coloring c' naturally partitions $V(G)\setminus\{v_1, v_2\}$ into two parts, one for each color — denote the largest part by A_3, this has $\gtrsim n/2$ vertices. This implies that $c(v_1v_2v)$ is constant for all $v \in A_3$ — denote this constant by $f(v_1v_2)$.

Now, pick a vertex in $v_3 \in A_3$. Create similar colorings for the edges v_1v_3 and v_2v_3. There is a subset $A_4 \subset A_3$, with $\gtrsim n/8$ vertices, such that $c(v_1v_3v)$ and $c(v_2v_3v)$ are constant for all $v \in A_3$ — denote those constants by $f(v_1v_3)$ and $f(v_2v_3)$.

Repeat this process $R(m)$ times, which we can because $n \geq 2^{\binom{R(m)}{2}}$. Now, we have vertices $v_1, \ldots, v_{\binom{R(m)}{2}}$, with a coloring f of each 2-edge, in which $f(v_{a_1}v_{a_2}) = c(v_{a_1}v_{a_2}v_{a_3})$, for all $a_1 < a_2 < a_3$. By definition, there is a monochromatic K_m over the coloring f, which implies that there exists a monochromatic $K_m^{(3)}$ over the coloring c.

6.1 Happy Ending Problem

Problem 6.1

Given 5 points on the plane, prove that there are 4 of them that form a convex polygon.

Solution. If the convex hull has size 5 or 4, we're ok. If it has size 3, then draw a line through the 2 points inside the convex hull, it meets two of the three sides of the convex hull. The two points inside and the two points in the side not crossed form a convex polygon.

Definition 6.4

Let $f(k)$ be the minimal n such that, for any set of n points in $\mathbb{R}^2$ in general position, there are k points that form a convex polygon.

Theorem 6.5 (Erdős-Szekeres, 1935)

$$f(k) \leq R^{(4)}(k) \leq 2^{2^{2^{ck}}}.$$

Proof. Suppose $n > R^{(4)}(k)$.

Define $c\colon \binom{[n]}{4} \to R, B$ by $c(\{A, B, C, D\}) = R$ if, and only if, $\{A, B, C, D\}$ does form a convex polygon.

By definition, there exists a monochoromatic $K_k^{(4)}$. For $k \geq 5$, it cannot be blue. Therefore, it's red, which would not be possible if those k vertices didn't form a convex polygon.

6.2 Monochromatic Arithmetic Progression

Definition 6.6

Let $W(r, k)$ be the minimal n such that for all $c\colon [n] \to [r]$, there exists a monochromatic arithmetic progression of size k.

Theorem 6.7 (Van der Waerden, 1927)

Let $c\colon \mathbb{N} \to [r]$. There is a monochromatic arithmetic progression of size k, for all positive integers k.

Equivalently,

$$W(r, k) < \infty.$$

Definition 6.8

Denote $\{a, a + d, a + 2d, \ldots, a + (k-1)d\}$ by $PA_k(a, d)$.

The arithmetic progressions $PA_k(a_1, d_1), PA_k(a_2, d_2) \ldots, PA_k(a_s, d_s)$ are color-focused if:

- (i) They are monochromatic with diferent colors.
- (ii) They have the same "focus" f, i.e.,

$$a_1 + kd_1 = \cdots = a_s + kd_s = f$$

Proof of Van der Waerden, 1927. We will use induction on k. Note that $W(r, 1) = 1$.

We shall find r color-focused $(k-1)$-arithmetic progressions.

Lemma 6.9

There exists $n = n(s, r)$ such that, for every coloring $c\colon [n] \to [r]$, there exists a monochromatic k-arithmetic progression or s color-focused $(k-1)$-arithmetic progressions.

Proof. Induction on s. $n(1, r) = W(r, k-1) < \infty$.

Let $N = 2n(s-1, r)$. Consider $W(r^N, k-1) < \infty$ blocks of size N. There is an arithmetic progression of equally-colored blocks of size $k-1$, let D be the distance of consecutive blocks in the arithmetic progression of blocks. Since the first half of the block has $n(s-1, r)$ elements, there exists a monochromatic k-arithmetic progression (which means we're done), or $s-1$ color-focused $(k-1)$-arithmetic progressions – their focus f surely lies inside the block of size N.

Let the $s-1$ color-focused $(k-1)$-arithmetic progressions in the first block be $PA_{k-1}(a_1, d_1), \ldots, PA_{k-1}(a_{s-1}, d_{s-1})$, with focus f_1. The proposed s color-focused $(k-1)$-arithmetic progressions are $PA_{k-1}(a_1, d_1 + d), \ldots, PA_{k-1}(a_{s-1}, d_{s-1} + d), PA_{k-1}(f_1, d)$.

Therefore,

$$n(s, r) \leq 2 \cdot W(r^{2n(s-1,r)}, k-1) \cdot 2n(s-1, r).$$

Therefore, for suitable large n, there must exist a large k-arithmetic progression.

7 Extremal olympiad-like problems

YouTube, Lec. 7
January 19, 2021

Definition 7.1

$\mathcal{A} \subset \mathcal{P}([n])$ is an *anti-chain* if $A \not\subset B$, for all $A, B \in \mathcal{A}$, $A \neq B$.

Theorem 7.2 (Sperner, 1910's)

If $\mathcal{A} \subset \mathcal{P}([n])$ is an anti-chain, then $|\mathcal{A}| \leq \binom{n}{n/2}$.

The example is $\binom{[n]}{n/2}$.

Proof of Sperner, 1910's. We know that $\binom{n}{k} \leq \binom{n}{n/2}$. Thus, by LYMB, 1960's,

$$1 \geq \sum_{A \in \mathcal{A}} \frac{1}{\binom{n}{|A|}} \geq \sum_{A \in \mathcal{A}} \frac{1}{\binom{n}{n/2}} = \frac{|\mathcal{A}|}{\binom{n}{n/2}}$$

Lemma 7.3 (LYMB, 1960's)

If $\mathcal{A} \subset \mathcal{P}([n])$ is an anti-chain, then $\sum_{A \in \mathcal{A}} \frac{1}{\binom{n}{|A|}} \leq 1$.

Proof of LYMB, 1960's. Let's count the pairs (π, A) such that π is a permutation of $[n]$, $A \in \mathcal{A}$, and $\{\pi(1), \pi(2), \ldots, \pi(|A|)\} = A$.

For each $A \in \mathcal{A}$, the number of π such that $\{\pi(1), \pi(2), \ldots, \pi(|A|)\}$ is equal to $|A|!(n - |A|)!$.

For each π, the number of $A \in \mathcal{A}$ such that $\{\pi(1), \ldots, \pi(|A|)\}$ is at most 1, since $\mathcal{A}$ is an anti-chain.

Therefore,

$$\sum_{A \in \mathcal{A}} |A|!(n - |A|)! \leq n! \implies \sum_{A \in \mathcal{A}} \frac{1}{\binom{n}{|A|}} \leq 1.$$

Definition 7.4

$\mathcal{A}$ is *intersecting* if $A \cap B \neq \varnothing$, for all $A, B \in \mathcal{A}$.

Proposition 7.5

$\mathcal{A} \subset \mathcal{P}([n])$ is intersecting $\implies |\mathcal{A}| \leq 2^{n-1}$.

Sketch. At most one of $(S, \overline{S})$ can be on $\mathcal{A}$.

Theorem 7.6 (Erdős-Ko-Rado, 1961)

$\mathcal{A} \subset \binom{[n]}{k}$ is intersecting $\implies |\mathcal{A}| \leq \binom{n-1}{k-1}$, for $k < \frac{n+1}{2}$.

Proof. Let's count the number of pairs (π, A) such that π is a circular permutation and $A \in \mathcal{A}$ is an interval in π.

For each $A \in \mathcal{A}$, the number of permutations such that A is an interval in π is $k!(n-k)!$

For each circular permutation π, the number of $A \in \mathcal{A}$ such that A is an interval in π is at most k.

Therefore,

$$|A|k!(n-k)! \leq (n-1)!k \implies |A| \leq \binom{n-1}{k-1}.$$

8 Supersaturation and Stability

Youtube, Lec. 10
January 19, 2021

Definition 8.1

G is *t-close to bipartite* if there exists $T \subset E(G)$, $|T| \leq t$ such that $G - T$ is bipartite.

Otherwise, G is *t-far from bipartite.*

Theorem 8.2 (Füredi)

If G is t-far from bipartite, then

$$\#\ K_3 \text{ in } G \geq \frac{n}{6}\left(e(G) + t - \frac{n^2}{4}\right).$$

Proof. Let $N(v)$ be the neighborhood of v. Then,

$$\#K_3 \text{ in } G = \frac{1}{3}\sum_{v \in G} e(N(v)).$$

Also, since G is t-far from bipartite,

$$e(N(v)) + e(\overline{N(v)}) > t$$

Lastly,

$$\begin{aligned}\sum_{u \in \overline{N(v)}} d(u) &= e(\overline{N(v)}, N(v)) + 2e(\overline{N(v)}) \\ &= e(G) + e(\overline{N(v)}) - e(N) \\ &> e(G) + t - 2e(N(v)).\end{aligned}$$

Therefore,

$$\begin{aligned}\#K_3 \text{ in } G &= \frac{1}{3}\sum_{v \in G} e(N(v)) \\ &> \frac{1}{6}\sum_{v \in G}\left(e(G) + t - \sum_{u \in \overline{N(v)}} d(u)\right) \\ &> \frac{1}{6}\sum_{v \in G}(e(G) + t) - \frac{1}{6}\left(\sum_{v \in G}\sum_{u \in \overline{N(v)}} d(u)\right) \\ &> \frac{n}{6}(e(G) + t) - \frac{1}{6}\sum_{\substack{v \in G \\ u \in G \\ u \not\sim v}} d(u) \\ &> \frac{n}{6}(e(G) + t) - \frac{1}{6}\sum_{u \in G} d(u)(n - d(u)) \\ &> \frac{n}{6}(e(G) + t) - \frac{1}{6}\frac{n^3}{4} \\ &> \frac{n}{6}\left(e(G) + t - \frac{n^2}{4}\right)\end{aligned}$$

Corollary 8.3

$e(G) \geq \frac{n^2}{4} + t \implies \#K_3 \text{ in } G \geq \frac{tn}{3}$.

Corollary 8.4

$e(G) > \frac{n^2}{4} - t$, $K_3 \not\subset G \implies G$ is t-close to bipartite.

Theorem 8.5 (Generalization of Füredi)

If G is t-far from r-partite, then

$$\#K_{r+1} \text{ in } G \geq c(r)n^{r-2}\left(e(G) + t - \left(1 - \frac{1}{r}\right)\binom{n}{2}\right).$$

Proof. Left as an exercise. Same idea; use induction; use Hölder.

Theorem 8.6 (Stability Theorem of Erdős–Simonovits, 1970's)

Let H be a graph. For all $\varepsilon > 0$, there is $\delta > 0$ such that the following property holds.

If $H \not\subset G$ and

$$e(G) \geq \left(1 - \frac{1}{\chi(H) - 1} - \delta\right)\binom{n}{2},$$

then G is εn^2-close to $(\chi(H) - 1)$-partite.

Sketch. For simplicity, let $\chi(H) = 3$. Thus, we shall prove $e(G) \geq \frac{n^2}{4} - \delta n^2 \implies G$ is εn^2-close to bipartite.

Let's have $H \subset K_3(s)$, for some s. G is εn^2-far from bipartite, and $e(G) \geq \frac{n^2}{4} - \delta n^2$. We shall prove that $K_3(s) \subset G$.

By Füredi, for $t = \varepsilon n^2$,

$$\begin{aligned} K_3(G) &\geq \frac{n}{6}\left(e(G) + \varepsilon n^2 - \frac{n^2}{4}\right) \\ &\geq \frac{\varepsilon n^3}{12}. \end{aligned}$$

By Theorem 8.7, we're done!

Theorem 8.7 (Erdős)

Let H be a k-uniform hypergraph. if $e(H) \geq \alpha\binom{n}{k}$, then there exists a copy of $K_k^{(k)}(t)$, the complete k-partite hypergraph, inside H.

Sketch.

9 Random Graphs and Thresholds

YouTube, Lec. 6
January 27, 2021

In this section, we'll recall some things we've seen before. Namely, the Random graph of Erdős-Rónyi and Markov's inequality. Another inequality that will be useful is Chebychev's inequality.

Lemma 9.1 (Chebychev's inequality)

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

Definition 9.2 (Variance)

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \\ &= \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2\end{aligned}$$

Sketch of Chebychev's inequality. Apply Markov's inequality with $(X - \mathbb{E}(X))^2$ and t^2.

9.1 Triangle-free

Proposition 9.3
If $p \ll \frac{1}{n}$, i.e., $pn \to 0$, then $\mathbb{P}(\text{copy of } K_3 \subset G(n,p)) \to 0$, in other words, $K_3 \not\subset G(n,p)$ with high probability.

Proof. Define X as the number of copies of K_3 in $G(n,p)$.

$$\mathbb{E}[X] = \binom{n}{3} p^3 \leq n^3 p^3 \to 0.$$

Therefore,

$$\mathbb{P}(X \geq 1) \to 0 \implies \mathbb{P}(X = 0) \to 1.$$

Proposition 9.4
If $p \gg \frac{1}{n}$, i.e., $pn \to \infty$, then $\mathbb{P}(\text{copy of } K_3 \subset G(n,p)) \to 1$.

Proof. Again, define X as the number of copies of K_3 in $G(n,p)$. First, $\mathbb{E}[X]^2 = \binom{n}{3}^2 p^6 \to \infty$. Second,

$$\begin{aligned}\mathbb{E}\left[X^2\right] &= \sum_{\substack{S,T \text{ copies} \\ \text{of } K_3}} \mathbb{P}(S \subset G(n,p) \text{ and } T \subset G(n,p)) \\ &\leq \binom{n}{3}^2 p^6 + n^4 p^5 + n^3 p^3.\end{aligned}$$

Therefore, $\mathrm{Var}(X) \leq n^4 p^5 + n^3 p^3 \ll \mathbb{E}[X]^2$.

Applying Chebychev's inequality with $t = \frac{\mathbb{E}[X]}{2}$, we have

$$\mathbb{P}\left(X \leq \frac{\mathbb{E}[X]}{2}\right) \leq \frac{4\,\mathrm{Var}(X)}{\mathbb{E}[X]^2} \to 0.$$

This means that $\frac{1}{k}$ is a threshold. If p is much smaller than $\frac{1}{n}$, then there is no triangle with high probability. If p is much bigger than $\frac{1}{n}$, then there is a triangle with high probability.

Can we do the same for K_r? Let's try.

Proposition 9.5

If $p \ll n^{-\frac{2}{r-1}}$, then $\mathbb{P}(\text{copy of } K_r \subset G(n,p)) \to 0$.

Proof. Define X as the number of copies of K_r in $G(n,p)$.

$$\mathbb{E}[X] = \binom{n}{r} p^{\binom{r}{2}} \to 0.$$

Therefore,

$$\mathbb{P}(X \geq 1) \to 0.$$

Proposition 9.6

If $p \gg n^{-\frac{2}{r-1}}$, then $\mathbb{P}(\text{copy of } K_r \subset G(n,p)) \to 1$.

Proof. Again, define X as the number of copies of K_r in $G(n,p)$. First, $\mathbb{E}[X]^2 = \left(\binom{n}{r} p^{\binom{r}{2}}\right)^2 \to \infty$. Second,

$$\begin{aligned}\mathbb{E}\left[X^2\right] &= \sum_{\substack{S,T \text{ copies} \\ \text{of } K_r}} \mathbb{P}(S \subset G(n,p) \text{ and } T \subset G(n,p)) \\ &\leq \left(\binom{n}{r} p^{\binom{r}{2}}\right)^2 + \sum_{k=2}^{r} n^{2r-k} p^{2\binom{r}{2} - \binom{k}{2}}.\end{aligned}$$

Note that, for $2 \leq k \leq r$, $n^{2r-k} p^{2\binom{r}{2}-\binom{k}{2}} \ll n^{2r} p^{2\binom{r}{2}} \approx \mathbb{E}[X]^2$.

Therefore, $\mathrm{Var}(X) \leq \sum_{k=2}^{r} n^{2r-k} p^{2\binom{r}{2}-\binom{k}{2}} \ll \mathbb{E}[X]^2$.

Applying Chebychev's inequality with $t = \frac{\mathbb{E}[X]}{2}$, we have

$$\mathbb{P}\left(X \leq \frac{\mathbb{E}[X]}{2}\right) \leq \frac{4\,\mathrm{Var}(X)}{\mathbb{E}[X]^2} \to 0.$$

9.2 Mathings

YouTuhe, Lec. 8
January 28, 2021

Definition 9.7 (Matching)

A *matching* is a graph in which each vertex has degree at most 1.

A *perfect matching* is a graph in which each vertex has degree 1.

Theorem 9.8 (Dilworth)

Let P be a partially ordered set with n vertices. Then, there exists a chain of size k (i.e., a sequence $v_1, v_2, \ldots, v_k$ such that $v_i < v_j$ whenever $i < j$) or an anti-chain of size n/k (i.e., a set of vertices such that for any pair u, v, $u \not< v$).

Theorem 9.9

Let G be a bipartite graph with n vertices on each part. Let's call the parts A and B.

There exists a perfect matching inside G if, and only if, for all subsets $S \subset A$,

$$N(S) := |\cup_{u \in S} N(u)| \geq |S|.$$

Sketch. It is clear that it is a necessary condition. We shall prove that it is a sufficient condition. We'll use induction on n.

Suppose that there is a set $A_1 \subset A$, $A_1 \neq A, \varnothing$ such that $|N(A_1)| = |A_1|$. Consider the graphs G_1 and G_2 by restraining the vertices to $A_1 \cup N(A_1)$ and $\overline{A_1 \cup N(A_1)}$, respectively. Show that G_1 and G_2 satisfy the hypothesis. Imply that there is a matching inside G.

Suppose that $|N(S)| > |S|$ for all $S \subset A$, $S \neq A, \varnothing$. Pick pick any edge uv and fix it. Consider $G' = G - u - v$. Show that G' satisfy the hypothesis. Imply that there is a matching inside G.

Proposition 9.10

If $p < (1-\varepsilon)\frac{\log n}{n}$, then there exists an isolated vertex in $G(n,p)$ with high probability.

Proof. Let X denote the number of isolated vertices in $G(n,p)$.

$$\begin{aligned}
\mathbb{E}[X] &= n(1-p)^{n-1} \\
&\gtrsim ne^{-(1-\varepsilon)\log n} \\
&\gtrsim n^{\varepsilon}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\mathbb{E}[X^2] &= \sum_{u,v} \mathbb{P}(u, v \text{ are isolated}) \\
&= n(n-1)(1-p)^{2n-3} + n(1-p)^{n-1} \\
&= \mathbb{E}[X]^2 \frac{n-1}{n}(1-p)^{-1} + \mathbb{E}[X].
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&\leq 2p\mathbb{E}[X]^2 + \mathbb{E}[X] \\
&\ll \mathbb{E}[X]^2.
\end{aligned}$$

Thus, by Chebychev's inequality, we're done.

Theorem 9.11

Suppose n is even.

If $p < (1-\varepsilon)\frac{\log n}{n}$, then there is no perfect matching in $G(n,p)$ with high probability.

If $p > (2+\varepsilon)\frac{\log n}{n}$, then there is a perfect matching in $G(n,p)$ with high probability.

9.3 Connectivity

YouTube, Lec. 9
January 28, 2021

Theorem 9.12

If $p < (1-\varepsilon)\frac{\log n}{n}$, then $G(n,p)$ is not connected with high probability.

If $p > (1+\varepsilon)\frac{\log n}{n}$, then $G(n,p)$ is connected with high probability.

Proof of the first part. Directly from Proposition 9.10.

Proof of the second part. A graph G is disconnected if, and only if, there exits a complete bipartite graph which is a subgraph of $\overline{G}$.

For $k \in \{1, \ldots, n/2\}$, let X_k be the number of copies of $K_{k,n-k}$ in $\overline{G(n,p)}$.

$$\begin{aligned} \mathbb{E}[X_k] &= \binom{n}{k}(1-p)^{k(n-k)} \\ &\leq \left(\frac{en}{k}e^{-p(n-k)}\right)^k \\ &\leq \left(\frac{en}{k}n^{-(1+\varepsilon)\left(1-\frac{k}{n}\right)}\right)^k \\ &\leq n^{-\varepsilon k/2} \to 0 \end{aligned}$$

Since $X_k = 0$, for $k \in \{1, \ldots, n/2\}$ with high probability, then $G(n,p)$ is connected with high probability.

9.4 Thresholds for General Properties

Definition 9.13 (Sharp threshold)

An event $\mathcal{A} = \mathcal{A}(n)$ has a *sharp threshold* if there exists p_c such that:

- $p \geq (1+\varepsilon)p_c \implies \mathbb{P}(\mathcal{A}) \to 1$, as $n \to \infty$;
- $p \geq (1-\varepsilon)p_c \implies \mathbb{P}(\mathcal{A}) \to 0$, as $n \to \infty$.

Definition 9.14 (Coarse threshold)

An event $\mathcal{A} = \mathcal{A}(n)$ has a *coarse threshold* if there exists p_c such that:

- $p \gg p_c \implies \mathbb{P}(\mathcal{A}) \to 1$, as $n \to \infty$;
- $p \ll p_c \implies \mathbb{P}(\mathcal{A}) \to 0$, as $n \to \infty$.

Theorem 9.15 (Bollobás–Thomason, 1980s)

Every increasing property (in the sense of adding edges) has a coarse threshold.

Sketch. Define $p_\varepsilon = \inf\{p : \mathbb{P}(G(n,p) \in \mathcal{A}) \geq \varepsilon\}$.

Proposition 9.16

There exists $C = C(\varepsilon)$ such that $\mathbb{P}(G(n, Cp_\varepsilon) \in \mathcal{A}) \geq 1 - \varepsilon$.

Sketch. We shall use sprinkling.

$G_1, \ldots, G_C$ indepentent copies of $G(n, p_\varepsilon)$.

$$G_1 \cup \cdots \cup G_c \sim G\left(n, 1 - (1-p_\varepsilon)^c\right) \subset G(n, Cp_\varepsilon).$$

Then,

$$\begin{aligned} \mathbb{P}(G(n, Cp_\epsilon) \in \mathcal{A}) &\geq \mathbb{P}\left(\bigcup_{i=1}^{C} G_i \in \mathcal{A}\right) \\ &\geq 1 - (1-\epsilon)^C \\ &\geq 1 - e^{-C\epsilon} \\ &\geq 1 - \epsilon, \end{aligned}$$

if $C \geq \frac{-\log(\epsilon)}{\epsilon}$.

Recall that $p \gg p_c \implies \forall \varepsilon > 0, p > C(\varepsilon)p_c$; and $p \ll p_c \implies \forall \varepsilon > 0, p < \frac{p_c}{C(\varepsilon)}$.

9.5 Hamiltonian Cycles

Theorem 9.17 (Dirac)

If $\delta(G) \geq n/2$, then there exists a Hamiltonian cycle.

Sketch. Take the longest path $v_1, v_2, \ldots, v_k$. Clearly, $k > n/2$. Use pidgeonhole principle to find $v_1 \sim v_{x+1}$ and $v_k \sim v_x$. We have a cycle $v_1, v_2, \ldots, v_x, v_k, v_{k-1}, \ldots, v_{x+1}$.

If there is a vertex outside of this cycle, it must connect to some vertex inside the cycle. If such thing happens, we can find a larger path than the one we started with; a contradiction.

Therefore, this cycle goes through all vertices.

Theorem 9.18

In the random graph process G_m,

$$\min\{m : \delta(G_m) \geq 2\} = \min\{m : G_m \text{ has a Hamiltonian cycle}\}$$

with high probability.

10 Janson's Inequality and Aplications

YouTube, Lec. 11
February 09, 2021

Lemma 10.1 (Harris, 1960; FKG, 1970s)
If E and F are increasing events (or both decreasing), then

$$\mathbb{P}(E \cap F) \geq \mathbb{P}(E)\mathbb{P}(F).$$

If E is increasing and F is decreasing, then

$$\mathbb{P}(E \cap F) \leq \mathbb{P}(E)\mathbb{P}(F).$$

Proof. Left to the reader.

Theorem 10.2 (Janson, 1987)
Let $A_1, \ldots, A_m \subset [N]$. Define the events $B_j := \{A_j \subset R\}$, in which $\mathbb{P}(i \in R) = p$, $\forall i \in [N]$ independently.

Let $X = \sum_{j=1}^{m} \mathbb{1}[B_j]$, $\mu = \mathbb{E}[X]$, $\Delta = \sum_{i \sim j} \mathbb{P}(B_i \cap B_j)$, in which $i \sim j$ means $A_i \cap A_j \neq \varnothing$.

Then,

$$\mathbb{P}\left(\bigcap_{j=1}^{m} \overline{B_j}\right) \leq \begin{cases} \exp\left(-\mu + \frac{\Delta}{2}\right), \text{ if } \Delta \leq \mu, \\ \exp\left(-\frac{\mu^2}{2\Delta}\right), \text{ if } \Delta \geq \mu. \end{cases}$$

Proof of the first bound, by Boppara-Spencer.

$$\begin{aligned} \mathbb{P}\left(\bigcap_{j=1}^{m} \overline{B_j}\right) &= \prod_{j=1}^{m} \left(1 - \mathbb{P}\left(B_j \mid \overline{B_1} \cap \cdots \cap \overline{B_{i-1}}\right)\right) \\ &\leq \exp\left(-\sum_{j=1}^{m} \mathbb{P}(B_j \mid \overline{B_1} \cap \cdots \cap \overline{B_{i-1}})\right) \\ &\underset{\substack{\uparrow \\ \text{Proposition 10.3}}}{\leq} \exp\left(-\mu + \frac{\Delta}{2}\right). \end{aligned}$$

It suffices to prove the following claim:

Proposition 10.3

$$\mathbb{P}\left(B_j \mid \overline{B_1} \cap \cdots \cap \overline{B_{i-1}}\right) \geq \mathbb{P}(B_j) - \sum_{\substack{i \sim j \\ i < j}} \mathbb{P}(B_i \cap B_j)$$

Proof. Let $E = \bigcap_{\substack{i\sim j \\ i<j}} \overline{B_j}$, and $F = \bigcap_{\substack{i\not\sim j \\ i<j}} \overline{B_i}$. Then,

$$\begin{aligned}
\mathbb{P}\left(B_j \mid E \cap F\right) &\geq \mathbb{P}(B_j \cap E \mid F) \\
&= \mathbb{P}(B_j \mid F)\mathbb{P}(E \mid B_j \cap F) \\
&= \mathbb{P}(B_j)\mathbb{P}(E \mid B_j \cap F) \\
&\underset{\substack{\uparrow \\ \text{Harris}}}{\geq} \mathbb{P}(B_j)\mathbb{P}(E \mid B_j) \\
&\underset{\substack{\uparrow \\ \text{union bound}}}{\geq} \mathbb{P}(B_j)\left(1 - \sum_{\substack{i\sim j \\ i<j}} \mathbb{P}(B_i \mid B_j)\right) \\
&\geq \mathbb{P}(B_j) - \sum_{\substack{i\sim j \\ i<j}} \mathbb{P}(B_i \cap B_j).
\end{aligned}$$

Proof of the second bound. Suppose that $\Delta \geq \mu$.

Let S be a random subset of $[m]$, with $\mathbb{P}(j \in S) = q$, $\forall j$, independently.

Let's calculate the expectated values of $\mu(S)$ and $\Delta(S)$.

$$\mathbb{E}[\mu(S)] = \sum \mathbb{P}(j \in S)\mathbb{P}(B_j) = q\mu.$$
$$\mathbb{E}[\Delta(S)] = \sum_{i\sim j} \mathbb{P}(i, j \in S)\mathbb{P}(B_i \cap B_j) = q^2\Delta.$$

Thus,

$$\begin{aligned}
\mathbb{E}\left[\mu(S) - \frac{\Delta}{2}\right] &= qu - \frac{q^2\Delta}{2} \\
&= \frac{\mu^2}{2\Delta},
\end{aligned}$$

for $q := \frac{\mu}{\Delta} \leq 1$.

Thus, there exists S such that $\mu(S) - \frac{\Delta(S)}{2} \geq \frac{\mu}{2\Delta}$.

Then,

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{j=1}^{m} \overline{B_j}\right) &\leq \mathbb{P}\left(\bigcap_{j\in S} \overline{B_j}\right) \\
&\leq \exp\left(-\mu(S) + \frac{\Delta(S)}{2}\right)
\end{aligned}$$

10.1 Triangle-freeness is not sharp

We saw before that

$$\mathbb{P}(K_3 \subset G(n,p)) \to \begin{cases} 0, \text{ if } p \ll \frac{1}{n}, \\ 1, \text{ if } p \gg \frac{1}{n}. \end{cases}$$

What happens if $p = \frac{c}{n}$?

Lemma 10.4

If $p = \frac{c}{n}$, then, as $n \to \infty$,

$$\mathbb{P}(K_3 \not\subset G(n,p)) \to \exp\left(-\frac{c^3}{6}\right)$$

Proof. Let $T_1, T_2, \ldots, T_{\binom{n}{3}}$ be the triangles in K_n. Define the events $B_j := \{T_j \in G(n,p)\}$.

On one hand,

$$\begin{aligned}\mathbb{P}\left(K_3 \not\subset G(n,p)\right) &= \mathbb{P}\left(\bigcap_{j=1}^{\binom{n}{3}} \overline{B_j}\right) \\ &= \prod_{i=1}^{\binom{n}{3}} \mathbb{P}\left(\overline{B_i} \mid \overline{B_1} \cap \cdots \cap \overline{B_{i-1}}\right) \\ &\geq \prod_{i=1}^{\binom{n}{3}} \mathbb{P}\left(\overline{B_i}\right) \\ &= (1-p^3)^{\binom{n}{3}} \to \exp\left(-\frac{c^3}{6}\right).\end{aligned}$$

On the other hand, $\mu = p^3\binom{n}{3}$, and $\Delta \approx p^2 n\mu \ll \mu$. Thus,

$$\begin{aligned}\mathbb{P}\left(K_3 \not\subset G(n,p)\right) &= \mathbb{P}\left(\bigcap_{j=1}^{\binom{n}{3}} \overline{B_j}\right) \\ &\underset{\substack{\uparrow \\ \text{Janson}}}{\leq} \exp\left(-(1+o(1))p^3\binom{n}{3}\right) \\ &\to \exp\left(-\frac{c^3}{6}\right).\end{aligned}$$

10.2 Chromatic number of a random graph

Theorem 10.5 (Bollobás, 1980s)

$$\chi\left(G\left(n, \tfrac{1}{2}\right)\right) = \left(1+o(1)\right)\frac{n}{2\log_2 n},$$

with high probability.

Sketch for upper bound. We will start the proof with the following lemma:

Lemma 10.6

For all subsets $S \subset V\left(G\left(n, \frac{1}{2}\right)\right)$, $|S| \geq \frac{n}{(\log n)^2}$, there exists an independent set $I \subset S$, with size $\geq (2-\varepsilon)\log_2 n$.

Sketch. Let $N = \binom{|S|}{2}$; $m = \binom{|S|}{k}$; $A_1, \dots, A_m$ be sets of edges of the complete bipartite graphs inside with vertices on S; $B_1, \dots, B_m$ be the events $\{A_j \subset R\}$; and $R = \overline{G(n, \frac{1}{2})}[S]$.

Let's set things up for applying Janson, 1987.

$$\mu = \binom{|S|}{k} 2^{-\binom{k}{2}} \geq \left(\frac{|S|}{k} 2^{-k/2} \right)^k \geq \left(\frac{n^{\varepsilon/2}}{(\log n)^3} \right)^k \geq n^{(\varepsilon \log_2 n)/2}$$

$$\delta \underset{\substack{\uparrow \\ \text{worst cases are } \ell = 2 \text{ or } \ell = k-1}}{\lesssim} \mu^2 \frac{(\log n)^8}{n^2} + \mu \frac{1}{n^{1-\varepsilon}}.$$

From Lemma 10.6, we some have disjoint independent sets with size at least $(2 - \varepsilon) \log_2 n$ elements, and at most $\frac{n}{(\log n)^2}$ outside a independent set. Thus,

$$\chi\left(G(n, \tfrac{1}{2})\right) \leq \frac{n}{(2 - \varepsilon) \log_2 n} + \frac{n}{(\log n)^2}.$$