Advanced Graph Theory Notes

Sam Spiro

May 2, 2026

Contents

0 Overview and Basic Definitions

The following are my personal notes for a graduate level course in graph theory. The content of these notes differs a fair amount compared to that of the standard texts in the field, namely because I wanted to design a course which focuses primarily on the topics that I am interested in and study as a researcher. For example, I put a very heavy emphasis on extremal graph theory topics such as Turán problems since these are the problems that I tend to personally work on, while other important topics like connectivity are (currently) barely discussed for the sole reason that I tend to not work on such problems.

One place where these notes differ significantly from most texts which we feel the need to highlight is in the emphasis on the more general object of hypergraphs rather than talking strictly about graph theory alone. We do this because we believe that any modern graph theorist ought to know at least the basics of hypergraphs due to their significant importance in the field and because there often is not a dedicated course where students are ever formally taught the fundamentals of these objects. We emphasize though that the notes are structured in such a way that any mention of hypergraphs typically occurs only in the final subsection of a given chapter, so any graph-purist can safely read through these notes without fear of having to think beyond their familiar worldview of vertices and edges.

The notes are structured as follows:

- In the first chapter we establish the basic concepts we need throughout the book. Outside on some of the definitions around hypergraphs, this chapter is meant to serve primarily as a refresher on the standard graph theoretic definitions we use throughout the text, as well as a convenient place where one can refer to standard notions we use throughout the text such as with asymptotic notation and basic analytic inequalities. We will typically assign the bulk of this chapter to be read as the first homework assignment of the course since most of this content is intended to be review and we will often restate the relevant definitions of this chapter around when they are first used in the main part of the text.

- The part on "Fundamental Concepts" is the real heart of the notes and is intended to cover the most important definitions and results in (extremal) graph theory.

- The part on "Methods" covers a variety of tools which can be used to solve a wide range of graph theoretic problems.

- The next two parts on "Bonus Topics" and "Advanced Topics" are intended to be more supplementary in nature, covering topics which we think are interesting but not necessarily essential for every graph theorist to have an intimate familiarity with.

Maybe talk about other things: mantras, exercises.

With all this out of the way, let's go ahead and get started.

0.1 Graphs and Hypergraphs

Formally, a *graph* G is a pair of sets (V,E) with E a set of 2-element subsets of V, i.e. $E \subseteq \{\{x,y\} : x,y \in V,\ x \neq y\}$. The set V is called the *vertex set* of G and its elements are called *vertices*, while the set E is called the *edge set* of G and its elements are called *edges*. We will typically denote edges $\{x,y\}$ by the simpler notation xy. For example, $(\{1,2,3,4\},\{12,23,13,14\})$ is a graph. Often it's easier to depict graphs by pictures (and how exactly we draw the picture doesn't matter). Insert two drawings of the graph stated above. We emphasize that our definition of graphs does not allow for loops (which would correspond to an edge having a single vertex) nor repeated edges (which would correspond to E being a multiset).

While almost all of the content of these notes focuses on graphs, we will sometimes look at other related objects, and in particular we will take note of the more general notion of a hypergraph. Formally, a *hypergraph* H is a pair of sets (V,E) with E a collection of subsets of V. Again V is called the vertex set and E the edge set. For example, insert example is a hypergraph. When drawing hypergraphs, we will typically represent its edges as blobs which contain the given set of vertices see example. We say a hypergraph H is r*-uniform* if every edge has size exactly r, and we say H is *uniform* if it is r-uniform for some r. For convenience we will sometimes refer to r-uniform hypergraphs simply as r-graphs.

We emphasize that 2-graphs are simply graphs, and as such the notion of r-graphs and more generally that of arbitrary hypergraphs gives a significant generalization of graphs, and this in turn gives rise to the following general idea.

Mantra 1. If you see a definition, result, or problem from graph theory, you should ask if it has a reasonable generalization to r-graphs or even to hypergraphs in general.

Note that since hypergraphs are a more general object than that of r-graphs, it is less likely that somethng generalizes to arbitrary hypergraphs compared to simply graphs.

While we will not get into it in such detail here, we also emphasize that there are often many non0equivalent ways one can try to generalize a given graph theoretic definition. For example, one possible notion of an independent set in a hypergraph is a set of vertices which does not contain any edge. Another non-equivalent generalization is a set of vertices such that every edge contains at most one vertex from the set. Both these notions recover the usual notion of an independent set from graph theory, and while the former definition is the more standard definition in the literature, there can be contexts where the latter definition is the right notion to consider.

0.2 Basic Definitions

In this subsection we establish a number of definitions which are valid for both graphs and hypergraphs. Throughout we use G to denote an arbitrary hypergraph in order so that anyone who does not want to think about hypergraphs can think of G simply as an arbitrary graph.

We will often write $V(G)$ and $E(G)$ to denote the vertex set and edge set of G, respectively, and we write $v(G) = |V(G)|$ and $e(G) = |E(G)|$. In many cases we will assume G has n vertices,, i.e. $v(G) = n$. Given a vertex $x \in V(G)$, we define its *degree* $\deg_G(x)$ to be the number of edges of G which contain x as an element, and whenever G is clear from context we will simply denote this quantity by $\deg(x)$.

We say a graph or hypergraph $G' = (V', E')$ is a *subgraph* of another graph or hypergraph $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$. In this case we write $G' \subseteq G$. Given a set of vertices $S \subseteq V(G)$, we define $G - S$ to be the graph or hypergraph obtained by deleting S and all edges incident to it. That is, $V(G - S) = V(G) \setminus S$ and $E(G - S) = E(G) \setminus \{e : e \cap S \neq \emptyset\}$. If $S = \{x\}$ then we will denote this simply by $G - x$. Similarly if e is an edge of G we define $G - e$ to be the graph obtained by deleting the edge e (but keeping all other vertices and edges). A subgraph $G' \subseteq G$ is said to be *induced* if it is of the form $G - S$ for some set of vertices S. Given a set of vertices V we will sometimes write $G[V]$ to be the induced subgraph with vertex set V, i.e. $G[V] = G - V(G) \setminus V$. A subgraph $G' \subseteq G$ is called *spanning* if $V(G') = V(G)$.

For graphs or hypergraphs G, H, we say a map $\phi : V(G) \to V(H)$ is a *homomorphism* if for every edge $e \in E(G)$, we have $\phi(e) := \{\phi(x) : x \in e\} \in E(H)$, and we say ϕ is an isomorphism if ϕ is bijective and if ϕ^{-1} is a homomorphism. We say G, H are isomorphic if there exists an isomorphism $\phi : V(G) \to V(H)$.

We say G is *r-partite* for an integer $r \geq 1$ if there exists a partition $V_1 \cup \cdots \cup V_r$ of G such that every edge of G intersects each V_i set in at most one vertex. We will typically refer to 2-partite graphs as *bipartite* graphs.

Other things: independent set, independence number, chromatic number

0.3 Graph Specific Definitions

We now turn to standard definitions and notation for graphs. All of these definitions can in fact be extended to hypergraphs, but there is not a "standard" way to do this. As before we let G denote an arbitrary graph.

We say two vertices x, y are *adjacent* or *neighbors* if $xy \in E(G)$, and we sometimes denote this be writing $x \sim y$. Given a vertex x we define the *neighborhood* of x by $N_G(x) = \{$vertices that are adjacent to x in $G\}$. Again if G is clear from context we will denote this set simply by $N(x)$. Note that $\deg(x) = |N(x)|$ for graphs.

I'll make the formatting nicer at some point but for now:

Paths and Connectivity:

- A *path* in a graph G is sequence of distinct adjacent vertices $(x_1, x_2, \ldots, x_t)$, and we say such a path is a path from x_1 to x_t and that it has *length* $t - 1$ (i.e. the length of the path is the number of edges it has).
- A graph is *connected* if for any two pair of vertices there exists a path from x to y.
- The *distance* between two vertices x, y, denoted $\mathrm{dist}(x, y)$, is the length of the shortest path from x to y (with $\mathrm{dist}(x, y) = \infty$ if no such path exists).

Graph Operations and Subgraphs

- Given a set S and an integer k, we let $\binom{S}{k}$ denote the set of all subsets of S of size k. For example, our definition of a graph is equivalent to saying that $E \subseteq \binom{V}{2}$.

- Given a graph G we define its *complement* $\overline{G}$ to be the graph obtained by replacing all edges with non-edges and vice versa. That is, $\overline{G}$ is the graph with vertex set $V(G)$ and edge set $\binom{V(G)}{2} \setminus E(G)$.

Independent Sets and Colorings

- A set of vertices I is *independent* if no two vertices $x, y \in I$ are adjacent to each other.

- Given a graph G and an integer k, a *proper k-coloring* is a map $\phi : V(G) \to [k]$ with the property that adjacent vertices $x, y \in V(G)$ have $\phi(x) \neq \phi(y)$. The smallest k for which G has a proper k-coloring is called the *chromatic number* of G and is denoted $\chi(G)$.

Forests and Trees

- A graph is a *forest* if it contains no cycles (i.e. no subgraph isomorphic to a cycle graph C_ℓ). A *tree* is a forest which is connected.

- A vertex of degree 0 is called an *isolated vertex*. A vertex of degree 1 (especially in the context of trees and forests) is called a *leaf*.

We record notation for graphs that will appear throughout the text.

- K_n denotes the n-vertex complete graph, i.e. the unique n-vertex graph with all $\binom{n}{2}$ edges.

- $K_{s,t}$ denotes the complete bipartite graph which has s vertices in one part and t vertices in the other.

- C_ℓ denotes the cycle graph of length ℓ.

- P_r denotes the path graph with r vertices (NOTE: some authors would denote this by P_{r-1}).

We record notation for graph parameters that will appear throughout the text, where here G denotes an arbitrary graph.

- $\delta(G)$ is the minimum degree of G, i.e. $\delta(G) = \min_{x \in V(G)} \deg(x)$.

- $\Delta(G)$ is the maximum degree of G, i.e. $\Delta(G) = \max_{x \in V(G)} \deg(x)$.

- $\alpha(G)$ is the independence number of G, which is the largest size of an independent set of G.
- $\chi(G)$ is the chromatic number of G, which is the smallest integer k such that G has a proper k-coloring.

0.4 Hypergraph Specific Definitions

The following definitions and notations for hypergraphs H will be restated later on in the text, but for ease of reference we record some of this notation here. Again this will be better organized at some point:

- For a set of vertices $S \subseteq V(H)$, we define $\deg(S)$ to be the number of edges of H which contain S as a subset. This quantity will sometimes be referred to either as the *degree* or *codegree* of the set.
- For an integer k, we let $\Delta_k(H) = \max_{S:|S|=k} \deg(S)$ and we call this the *maximum k-degree* of H.
- $K_n^{(r)}$ denotes the complete r-uniform hypergraph on n vertices, i.e. the hypergraph with vertex set $[n]$ and edge set $\binom{[n]}{r}$.
- Probably more things.

0.5 Asymptotic Notation

Eventually in the text it will be convenient for us to make use of the following asymptotic notation which we record here for ease of reference. We emphasize that this notation will be redefined when it first appears in the text, so there is no need to memorize this right now.

Let $f(n), g(n)$ be two functions.

- We write $f(n) = O(g(n))$ if there exists a constant $C > 0$ such that $f(n) \leq Cg(n)$ for all n.
- We write $f(n) = \Omega(g(n))$ if there exists a constant $c > 0$ such that $f(n) \leq cg(n)$ for all n.
- We write $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$. In this case we say that f, g have the same *order of magnitude*.
- We write $f(n) \sim g(n)$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$. In this case we say that f, g are *asymptotic* to each other.

- We write $f(n) = o(g(n))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$. In particular, writing $f(n) = o(1)$ means $\lim_{n\to\infty} f(n) = 0$.

0.6 Inequalities

Many proofs in extremal combinatorics rely on basic inequalities from analysis. Here we record the most important of these that we will use.

Theorem (Cauchy-Schwarz Inequality)**.** *If $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ are real numbers, then*

$$\sum_{i=1}^{n} x_i y_i \le \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \left(\sum_{i=1}^{n} y_i^2\right)^{1/2}$$

Personally, we like to remember the statement of Cauchy-Schwarz by noting that it follows from the vector equality $\langle \mathbf{x}, \mathbf{y} \rangle = \cos\theta \, \|\mathbf{x}\| \, \|\mathbf{y}\|$ where θ is the angle between the vectors $\mathbf{x}, \mathbf{y}$.

For the next inequality, recall that a function $\phi : \mathbb{R} \to \mathbb{R}$ is *convex* if for all $0 \le t \le 1$ and $x, y \in \mathbb{R}$ we have $\phi(tx + (1-t)y) \le t\phi(x) + (1-t)\phi(y)$.

Theorem (Jensen's Inequality)**.** *If ϕ is a convex function and $x_1, \ldots, x_n \in \mathbb{R}$, then*

$$\sum_{i=1}^{n} \phi(x_i) \ge n\phi\left(n^{-1} \sum_{i=1}^{n} x_i\right).$$

That is, this sum is minimized when each x_i is equal to their average $n^{-1} \sum x_{i'}$. We note for later that for any integer $t \ge 1$, the functions x^t and $\binom{x}{t} := \frac{x(x-1)\cdots(x-t+1)}{t!}$ are convex.

Maybe also include: Markov, AMGM.

0.7 Exercises

Each chapter will end with a set of exercises. Following the notation of Stanley, we will add numbers after each exercise to indicate the problem's rough level of difficulty as follows:

- [1] problems are elementary and routine requiring little to no thought,
- [2] problems have simple solutions (though that does not necessarily mean it is easy to find such a solution!),
- [3] problems tend to have involved solutions,
- [4] problems have extremely difficult solutions (to the extent that such questions should never be used in a classroom setting),
- [5] problems are unsolved open problems.

Additionally, plus and minus symbols may be used to indicate higher or lower levels of difficulty for the problem. For example, a [2+] problem might have a simple solution that's pretty challenging to find, while a [3-] problem might have an involved solution that's actually not too hard to work out. Ultimately, all of the ratings that I give are only rough estimates and the reader may find a given [3] problem easier to solve than a [2-] depending on the circumstances.

With that preamble out of the way, we begin with some "elementary" (though not necessarily easy) graph theory problems.

Add some hypergraph exercises here, eg ask to prove many of these and other classical results for the setting of hypergraphs.

1. (Handshaking Lemma) Prove that every graph G has $\sum_{x\in V(G)} \deg(x) = 2e(G)$ [2-].

2. Prove that every graph G with $v(G) \geq 2$ contains two vertices with the same degree [2-].

3. Prove that for every graph G, either G or its complement $\overline{G}$ is connected [2-].

4. Prove that a graph is bipartite if and only if it contains no odd cycles [2-].

5. Prove that for every graph G, the set of edges $E(G)$ can be partitioned into cycles if and only if every very vertex of G has even degree [2+].

* * *

6. Recall that a graph is d-regular if $\deg(u) = d$ for every vertex u. Prove for all integers $0 \leq d < n$ that there exists an n-vertex d-regular graph if and only if at least one of d or n is even [2].

7. A graph is said to have girth g if it contains a cycle of length g and no cycles of shorter length.

   (a) Prove that for all integers $d, g \geq 2$, there exists a d-regular graph of girth g [2+].

   (b) Prove that if G is a d-regular graph of girth g, then
   $$v(G) \leq ???.$$
   [2]

   (c) Show that the bound above is tight for $d = 3, g = 5$ [1+].

* * *

8. Prove that $\chi(G)\alpha(G) \leq v(G)$ for all graphs G [2-].

9. Prove that $\alpha(G) \geq \frac{v(G)}{\Delta(G)+1}$ for all graphs G [2].

10. Prove that if a graph G is triangle-free (i.e. if G contains no subgraph isomorphic to K_3) then $\alpha(G) \geq \sqrt{v(G)}$ [2-].

$*\ *\ *$

11. Prove that every tree T with $v(T) \geq 2$ has at least two leaves.

12. Prove that for every tree T, there exists an ordering of its vertices $v_1, \ldots, v_n$ such that for all $2 \leq i \leq n$, there exists an integer j_i such that $N(v_i) \cap \{v_1, \ldots, v_{i-1}\} = \{v_{j_i}\}$ [1+].

13. Prove various characterizations of trees

14. (Helly Theorem for Trees) Let T be a tree and $\mathcal{T}$ a set of subtrees of T (i.e. a set of subgraphs of T which are themselves trees). Prove that if $V(T') \cap V(T'') \neq \emptyset$ for all $T', T'' \in \mathcal{T}$, then there exists a vertex $v \in \bigcap_{T' \in \mathcal{T}} V(T')$ [2+].

Part I

Fundamental Concepts

This pitch needs to be adjusted now that at least in principle the text isn't just about extremal combinatorics, and/or it should be moved either to the preamble or start of chapter 1.

As mentioned in the introduction, extremal graph theory broadly speaking asks questions of the form: how "large" can a graph be if it satisfies a certain property?

What exactly "large" means depends on the type of problem one is considering, with some popular choices being the number of edges, the number of vertices, and the minimum degree of the graph in question. Each of these choices (together with an appropriate choice of "property") gives rise to three of the main topics of extremal graph theory: Turán problems, Ramsey problems, and Dirac problems; see the table below for a brief outline. Each of these types of problems will be the main topic of focus for the forthcoming chapters.

Measurement		Property		Type of Problem
Number of edges	+	Triangle-free	=	Turán Problems: Section 1
Number of vertices	+	G and $\overline{G}$ are triangle-free	=	Ramsey Problems: Section 5
Minimum degree	+	non-Hamiltonian	=	Dirac Problems: Section 2

Figure 1: A table of measures of "largeness", properties that one can consider, and the problems that these produce. Note that in each case, the given property is harder to fulfill the "larger" G is with respect to its measurement, which is a hallmark of a good extremal problem.

1 Forbidden Subgraphs and Turán Problems

Turán Problems broadly ask: how many edges can an n-vertex graph have if it does not contain a copy of a given graph F? Specifically, the we will work with the following throughout this chapter.

Definition 1. Given two graphs F, G, we say that G is *F-free* if G does not contain a subgraph which is isomorphic to F. Given an integer $n \geq 1$, we define the *Turán number* or *extremal number* $\mathrm{ex}(n, F)$ to be the maximum number of edges that an n-vertex F-free graph can have.

The name of the game now is to try and either determine or bound $\mathrm{ex}(n, F)$ for various choices of F.

1.1 Forbidding C_4 and Complete Bipartite Graphs

Perhaps the first question we need to answer is: why should we care about Turán problems in the first place? There are many possible answers to this question, here are a few of my own personal reasons:

- They are natural extremal problem to consider.
- They have applications to various areas of mathematics.
- Solutions to Turán problems often use cool and deep results from other areas of mathematics in interesting ways.
- They're fun!

To try and illustrate these points above, we will begin by studying the Turán problem for $F = C_4$. Historically, this is the second Turán problem to be considered (we will look at the first problem in the following section) and was largely solved by Erdős in year due to its connection to a certain problem in number theory.

The Upper Bound. We begin by establishing an upper bound for this Turán number.

Theorem 1.1. *We have*

$$\mathrm{ex}(n, C_4) \leq \frac{n\sqrt{4n-3}+n}{4}.$$

That is, every n-vertex C_4-free graph has at most this many edges.

We emphasize that this is not a very pretty looking upper bound; we will address this further shortly after the proof.

Proof. In order to prove any upper bound for this problem, we need to get some understanding of what it means for a graph to be C_4-free graph. After thinking about it for long enough, one might come up with the following observation: a graph is C_4-free if and only if every pair of distinct vertices u, v has at most one common neighbor, i.e. there is at most one vertex in $N(u) \cap N(v)$. Indeed, the existence of two vertices in this set together with u, v would exactly define a C_4 in our graph.

Now, a priori, it is not immediate how to use the fact that pairs of vertices have at most one common neighbor to bound the number of edges in our graph. However, one can use it to bound the number of some other object which is "almost" an edge. Namely, let

$$\mathcal{P} = \{(\{u, v\}, x) \in V(G)^3 : u \sim x \sim v,\ u \neq v\},$$

which essentially just encodes the set of P_3's in G. Note that each element of $\mathcal{P}$ can be uniquely identified by picking two distinct vertices to play the roles of u, v together with a common neighbor of these vertices to play the role of x. As such, we have

$$|\mathcal{P}| = \sum_{u \neq v} |N(u) \cap N(v)| \leq \sum_{u,v} 1 = \binom{n}{2},$$

with the inequality using that our graph is C_4-free. Now, we got the first equality above by identifying each element of $\mathcal{P}$ by its first and last vertices u, v and then picking some common neighbor x. Alternatively, we could identify each element of $\mathcal{P}$ by specifying its middle vertex x together with two distinct neighbors u, v of x. As such, we also have

$$|\mathcal{P}| = \sum_{x \in V(G)} \binom{\deg(x)}{2} \geq n \binom{n^{-1} \sum_x \deg(x)}{2} = n \binom{n^{-1} \cdot 2e(G)}{2},$$

where this inequality used Jensen's inequality together with the fact that $\binom{a}{2}$ is a convex function, and the last equality used that $\sum_x \deg(x) = 2e(G)$. Comparing this to the upper bound for $|\mathcal{P}|$ we found above gives

$$n \binom{n^{-1} \cdot 2e(G)}{2} \leq \binom{n}{2}, \tag{1}$$

or equivalently

$$(2e(G))(2n^{-1}e(G) - 1) \leq n(n-1).$$

This in turn is equivalent to having

$$4e(G)^2 - 2ne(G) - n^2(n-1) \leq 0,$$

and solving this exactly gives the desired bound on $e(G)$.

Somewhere in the text I should call this a double counting argument and maybe mention the word cherries/P_2.

□

While the bound of Theorem 1.1 is truly the best we can do using our approach, it is often not a good idea in extremal combinatorics to do things so precisely.

Mantra 2. It is often better to use (slightly) "wastefull" bounds in extremal combinatorics to have cleaner proofs and theorem statements.

Knowing when exactly and how to derive such "crude" bounds is an important skill to have in extremal combinatorics, since in practice we do not know a priori if the approach we are currently playing around with is going to give something useful in the end, and until that point it is a bad idea to harp over minute details in the argument.

For example, let us consider the point in the proof where we reached (1). Here an expert might simplify their lives by observing that simple inequalities for binomial coefficients yield

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

and rearranging this gives

$$n^{-1}2e(G) - 1 \leq n^{1/2},$$

and hence

$$e(G) \leq \frac{1}{2}n^{3/2} + n.$$

Note that this is extremely close to the optimal bound we get in Theorem 1.1. In particular, one can show that both bounds are ultimately of the form $\mathrm{ex}(n, C_4) \leq \frac{1}{2}n^{3/2} + Cn$ for some sufficiently large constant C. This means that our weakening above captures the "main part" of the bound from Theorem 1.1, in the sense that for n very large the two numbers are very close to each other.

It will be useful going forward to develop notation to measure more precisely what exactly we mean by "very close to each other".

Definition 2. Let $f(n), g(n)$ be two functions.

- We write $f(n) = O(g(n))$ if there exists a constant $C > 0$ such that $f(n) \leq Cg(n)$ for all n. In particular, our remark in the paragraph above is equivalent to saying that our two bounds give[1] $\mathrm{ex}(n, C_4) \leq \frac{1}{2}n^{3/2} + O(n)$.

- We write $f(n) = \Omega(g(n))$ if there exists a constant $c > 0$ such that $f(n) \leq cg(n)$ for all n. Whenever we write this, we will often implicitly assume that we consider n large enough so that $f(n) > 0$. For example, if we write $\mathrm{ex}(n, F) = \Omega(1)$ we will implicitly be assuming $n \geq 2$.s

- We write $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$. In this case we say that f, g have the same *order of magnitude*.

- We write $f(n) \sim g(n)$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$. In this case we say that f, g are *asymptotic* to each other.

[1]A very persnickety reader might object that actually this doesn't exactly agree with the definition given: the real thing that should be written is $\mathrm{ex}(n, C_4) - \frac{1}{2}n^{3/2} = O(n)$ and the "algebra" of moving $\frac{1}{2}n^{3/2}$ to the other side is not actually valid. It is, however, common practice in the field to use these somewhat imprecise notational implementations in order to make statements easier to read and write, which is the ultimate goal of introducing this in the first place.

- We write $f(n) = o(g(n))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$. In particular, writing $f(n) = o(1)$ means $\lim_{n\to\infty} f(n) = 0$.

Applications. Theorem 1.1 has a number of applications to other areas of mathematics. We will consider one quick example from discrete geometry.

Let $\mathcal{P}$ be a set of points of $\mathbb{R}^2$ and let $\mathcal{L}$ be a set of lines in $\mathbb{R}^2$. We say that a point $p \in \mathcal{P}$ and a line $\ell \in \mathcal{L}$ are *incident* if p lies on the line ℓ. We let $I(\mathcal{P}, \mathcal{L})$ to denote the number of pairs $(p, \ell) \in \mathcal{P} \times \mathcal{L}$ with p and ℓ incident. A natural extremal question to ask is: what is the maximum number of incidences that a given number of points and line can obtain? Trivially one can do no better than n^2, but it is not so immediate how to improve this. We will be able to achieve such an improvement using our Turán result Theorem 1.1.

Corollary 1.2. *If $\mathcal{P}$ is a set of n points in $\mathbb{R}^2$ and if $\mathcal{L}$ is a set of n lines in $\mathbb{R}^2$, then*

$$I(\mathcal{P}, \mathcal{L}) = O(n^{3/2}).$$

Proof. As is often the case for applications, we begin by defining an auxiliary graph related to our problem at hand. To this end, define a bipartite graph G whose vertex set is $\mathcal{P} \cup \mathcal{L}$ where we have $p \sim \ell$ if and only if p and ℓ are incident. Observe that $I(\mathcal{P}, \mathcal{L}) = e(G)$, so bounding the number of incidences is exactly the same thing as bounding the number of edges of G.

Now, for arbitrary bipartite graphs G we could of course have $e(G)$ as large as n^2, but we have some additional structure to work with because G is coming from a set of points and lines. In particular, because every pair of lines intersect in at most one point, G can not contain a C_4 (since such a subgraph would consist of vertices p_1, p_2, ℓ_1, ℓ_2 with p_1, p_2 points common to both ℓ_1 and ℓ_2). This together with the fact that $v(G) = |\mathcal{P}| + |\mathcal{L}| = 2n$ immediately implies that

$$I(\mathcal{P}, \mathcal{L}) = e(G) \le \mathrm{ex}(2n, C_4) = O((2n)^{3/2}) = O(n^{3/2}),$$

with this last step using that this "big oh" notation is not affected by multiplying by a fixed constant. □

While it is neat that we could obtain this purely geometric result using graph theory, we should note that the bound of Theorem 1.2 is not tight, and in fact the true bound is $I(\mathcal{P}, \mathcal{L}) = O(n^{4/3})$. The fact that we obtained a subpar bound should perhaps not come as a surprise, as we used almost no information about the geometry of the Euclidean plane $\mathbb{R}^2$ in our argument. It is, however, possible to derive this optimal bound of $O(n^{4/3})$ if one uses Theorem 1.1 together with some appropriate geometric tools (such as real polynomial partitiontionings). We will not go into this here, but see eg the book by Sheffer for a lot more on this problem and more.

The Lower Bound. Theorem 1.1 shows that $\mathrm{ex}(n, C_4) = O(n^{3/2})$. The immediate question is: is this tight? This is an important question for us to figure out, since e.g. any improvement to Theorem 1.1 would give an improvement to our bound in Theorem 1.2 as well as to any other application we can come up with for $\mathrm{ex}(n, C_4)$.

To see whether our bound is tight, we need to prove a lower bound for $\mathrm{ex}(n, C_4)$, i.e. to construct n-vertex graphs with many edges and no C_4's. This, as the reader is welcome to try for themselves, is not so easy to do. To make some headway on this, we usethe following mantra.

Mantra 3. To find a lower bound construction for extremal problems, we should ask ourselves what would need to happen for our extremal upper bound to be (exactly) sharp.

In our case we ask: what would need to happen for us to have $\mathrm{ex}(n, C_4) = \frac{n\sqrt{4n-3}+n}{4}$? Well, this would happen precisely if every inequality throughout our proof of Theorem 1.1 were in fact an *equality*. In particular, our very first inequality $\sum_{u \neq v} |N(u) \cap N(v)| \leq \binom{n}{2}$ must be an equality, and this would imply that *every* pair of distinct vertices in G has exactly 1 common neighbor. Now we have to ask...is this ever possible?

Well, if you think about it for long enough, you might have the wild idea that "every two vertices has exactly 1 common neighbor" is kind of analogous to the statement "every two points in $\mathbb{R}^2$ lie on exactly one common line." Riffing off of this as well as what we did for our application in Theorem 1.2, what if we defined a bipartite graph G by taking a set of points $\mathcal{P}$ and a set of lines $\mathcal{L}$ and making a point p adjacent to a line ℓ if and only if they are incident? Such a graph will automatically be C_4-free due to the geometry of the situation, so we will win if we can find some points and lines with many incidences.

As hinted at just after Theorem 1.2, it is possible to find n points and lines in $\mathbb{R}^2$ such that $I(\mathcal{P}, \mathcal{L}) = \Omega(n^{4/3})$, giving a corresponding lower bound to $\mathrm{ex}(n, C_4)$, but this is as good as we can hope to do in Euclidean space. However, another wild thought based on what we said around Theorem 1.2 is that our idea of using points and lines does not fundamentally rely on the full geometry of Euclidean space: we only needed the very basic property that two points line on at most one line, and such a property holds for many different types of geometries. In particular, since we're working with finite graphs...why not try and do something with geometries over finite fields?

Recall from algebra[2] that for every prime power q there exists a field $\mathbb{F}_q$ of order q. Again going off what we did in Euclidean space, we want to consider a set of points and lines from the plane $\mathbb{F}_q^2 = \{(x, y) : x, y \in \mathbb{F}_q\}$. There might be some particularly clever choices of points and lines that we could make here, but since we are just playing around, why don't we go ahead and just take all of them. That is, we will take $\mathcal{P} = \mathbb{F}_q^2$ and $\mathcal{L}$ all of the lines in $\mathbb{F}_q^2$. To be clear, lines in $\mathbb{F}_q^2$ are just sets of points in $\mathbb{F}_q^2$ taking on one of two forms: for $a, b \in \mathbb{F}_q$ we define the line $\ell_{a,b} = \{(x, ax + b) : x \in \mathbb{F}_q\}$, and for $c \in \mathbb{F}_q$ we define the vertical lines $\ell_c = \{(c, y) : y \in \mathbb{F}_q\}$. Now define a bipartite graph G_q on $\mathcal{P} \cup \mathcal{L}$ where $p \sim \ell$ if and only if $p \in \ell$. We leave it as an exercise to the reader to verify that G_q is indeed C_4-free. To count $e(G_q)$, we observe that the total number of lines is $q^2 + q$ and that each line is incident to exactly q points, and as such $e(G_q) = q^3 + q^2$. Because the total number of vertices in G_q is exactly $2q^2 + q$, we in total conclude for any prime power q that

$$\mathrm{ex}(2q^2 + q, C_4) \geq q^3 + q^2.$$

By considering $n = 2q^2 + q \approx 2q^2$ or equivalent $q \approx (n/2)^{1/2}$, we find that for infinitely many integers n that $\mathrm{ex}(n, C_4)$ is at least $q^3 \approx (n/2)^{3/2} = 2^{-3/2} n^{3/2}$. As such, the upper bound of $\mathrm{ex}(n, C_4) = O(n^{3/2})$ really is the best we can do for general n! In fact, some basic number theory facts let us prove the following.

Theorem 1.3. *We have* $\mathrm{ex}(n, C_4) = \Theta(n^{3/2})$.

[2] Any reader scared of algebra should be reassured that this is the only fact you need to recall from algebra.

Proof. By Theorem 1.1 we have for n large enough that, say, $\mathrm{ex}(n, C_4) \le n^{3/2}$, proving $\mathrm{ex}(n, C_4) = O(n^{3/2})$.

Now consider any integer $n \ge 12$. By Bertrend's postulate, there exists a prime number p with $\frac{1}{2}\sqrt{n/3} \le p \le \sqrt{n/3}$. This in particular implies $n \ge 3p^2 \ge 2p^2 + p$, which together with our discussion above implies

$$\mathrm{ex}(n, C_4) \ge \mathrm{ex}(2p^2 + p, C_4) \ge p^3 \ge (12)^{-3/2} n^{3/2},$$

proving that $\mathrm{ex}(n, C_4) = \Omega(n^{3/2})$ and hence that $\mathrm{ex}(n, C_4) = \Theta(n^{3/2})$ as desired. □

We personally find it fascinating that one can use ideas from algebra and geometry to solve the purely combinatorial problem of determining $\mathrm{ex}(n, C_4)$. This is in fact a very common phenomenon.

Mantra 4. To solve a combinatorics problem, one often needs ideas and tools from other areas of math. As such, any extra knowledge you have outside of combinatorics is always useful to keep in the back of your mind!

This mantra is intended to be inspirational rather than intimidating. In particular, even if you don't have hardly any knowledge in areas outside of combinatorics (such as myself), you can still make it very far, its just that some problems in particular may elude your grasps until you figure out the right tool needed to crack it.

Even Better Lower Bounds. We've done pretty good so far with our lower bounds for $\mathrm{ex}(n, C_4)$, but we can go even farther.

Mantra 5. Once you prove something, see if you can prove something even better.

In particular, given that we have determined the order of magnitude $\mathrm{ex}(n, C_4) = \Theta(n^{3/2})$, we should next ask ourselves if we can prove that $\mathrm{ex}(n, C_4) \sim cn^{3/2}$ for some constant c. We emphasize that doing this will require a bit more algebra/geometry than before, and as such the reader may wish to skip over this part of the text if they're already overwhelmed.

Returning back to the problem at hand, we know up to this point (at least for certain values of n) that

$$2^{-3/2} n^{3/2} + o(n^{3/2}) \le \mathrm{ex}(n, C_4) \le 2^{-1} n^{3/2} + o(n^{3/2}),$$

and we need to figure out if we can sharpen either of these bounds. For this, it is useful to analyze "why" our lower bound proof does not match the bound we got in the upper bound. After all, in our construction every pair of points really does have exactly one common neighbor. However, if we look back at what motivated our construction in the first place, we recall that for the upper bound for Theorem 1.1 to be exactly sharp that we need every pair of *vertices* to have a common neighbor, and there is no hope of that happening for our current graph because G_q is bipartite (meaning a given point and a given line will never have any common neighbors in G_q).

It is not so immediate how to fix this problem, as the underlying motivation for our construction relied on working with both points and lines which intrinsically are different objects from each other. But, if we stare at things long enough, we might realize that our lines $\ell_{a,b}$ are indexed by points in $\mathbb{F}_q^2$, and as such, one might possibly have the idea where we could consider a graph G where its vertex set is just $\mathbb{F}_q^2$ but where a point (x,y) corresponds to both the point itself and the line $\ell_{x,y}$. That is, we want to define a graph on $\mathbb{F}_q^2$ where $(x,y) \sim (a,b)$ if and only if $(x,y) \sim \ell_{a,b}$. While this is a noble idea, an immediate issue in this definition is that this edge relation is not symmetric. That is, having $(x,y) \in \ell_{a,b}$ does not imply $(a,b) \in \ell_{x,y}$ (i.e. $y = ax+b$ does not mean $b = xa + y$). At a very high level the issue here with the idea of identifying points with a corresponding line is that points and lines are not truly "dual" to each other in $\mathbb{F}_q^2$. However, this can be fixed by going to yet another type of geometry, namely projective geometry.

Insert better intuition on projective geometries at some point.

To define things, consider the set of triples $T = \{(x,y,z) : x,y,z \in \mathbb{F}_q^3\} \setminus \{(0,0,0)\}$ and define an equivalence relation (not to be confused with an edge relation) by having $(x,y,z) \equiv (\alpha x, \alpha y, \alpha z)$ for all $\alpha \in \mathbb{F}_q \setminus \{0\}$. Let $[x,y,z]$ denote the equivalence class containing (x,y,z), and define our set of "points" $\mathcal{P}$ to be the set of all such equivalence classes. For each $[a,b,c] \in \mathcal{P}$ we define the line $\ell_{[a,b,c]} = \{[x,y,z] : ax+by+cz = 0\}$. Note that this definition is well-defined (i.e. it does not matter whether we write $[x,y,z]$ or $[\alpha x, \alpha y, \alpha z]$) since having $ax+by+cz = 0$ implies $\alpha ax + \alpha by + \alpha cz = 0$ for all $\alpha \neq 0$. Also note that this definition is truly "dual" in points and lines, in that $[x,y,z] \in \ell_{[a,b,c]}$ if and only if $[a,b,c] \in \ell_{[x,y,z]}$. Motivated by this and our ideas from above, we define a graph G_q^* on $\mathcal{P}$ where $[x,y,z] \sim [a,b,c]$ if and only if $[x,y,z] \in \ell_{[a,b,c]}$. We leave it as an exercise to verify that G_q^* is C_4-free, that $v(G_q^*) = q^2+q+1$, and that $e(G_q^*) = \frac{1}{2}(q+1)(q^2+q+1)$.

Similar to before, if we take $n = q^2+q+1 \approx q^2$, then we see that this shows $\mathrm{ex}(n, C_4)$ is at least $\frac{1}{2}q^3 \approx \frac{1}{2}n^{3/2}$, exactly matching the asymptotic bound from Theorem 1.1! Actually, even more is true: one can check that the upper bound $\frac{n\sqrt{4n-3}+n}{4}$ is actually *exactly* tight in this case. That is, for all prime powers q, we have

$$\mathrm{ex}(q^2+q+1, C_4) = \frac{1}{2}(q+1)(q^2+q+1).$$

Generalizations. Given our success with studying the Turán problem for C_4, we should go on and ask to what extent can the ideas here be used to prove bounds for other graphs F. Naively one might first consider the problem for other cycles C_ℓ, but this turns out to be pretty difficult. Instead, the "correct" generalization of the ideas we have here are for complete bipartite graphs $K_{s,t}$ in general beyond just that of $K_{2,2} = C_4$. For example, we leave it as an exercise to generalize the upper bound in Theorem 1.1 to prove the following general upper bound.

Theorem 1.4 (Kővári-Sós-Turán Theorem). *For all integers $s,t \geq 1$, we have*

$$\mathrm{ex}(n, K_{s,t}) = O_{s,t}(n^{2-1/s}).$$

Here we add the s,t subscript to the big-oh notation to emphasize that the implicit constant depends on s,t. This is not entirely necessary since we fix s,t at the start of the theorem, but it is sometimes nice to emphasize this for clarity.

This gives an upper bound, what about a corresponding lower bound? Our lower bound $\mathrm{ex}(n, C_4) = \Omega(n^{3/2})$ immediately implies $\mathrm{ex}(n, K_{2,t}) = \Omega(n^{3/2})$ for all $t \geq 2$, giving the correct order of magnitude. In fact, Füredi**REF**improved the lower bound for $\mathrm{ex}(n, K_{2,t})$ even further, giving a tight asymptotic bound. With some effort, one can generalize the geometric intuition we had for C_4 to prove $\mathrm{ex}(n, K_{3,t}) = \Theta(n^{5/3})$ for all $t \geq 3$, roughly be replacing the intuition of "two lines intersect in at most one point" with "three spheres intersect in at most two points." Despite this success, the next case of this problem remains open.

Open Problem 1.5. *Determine the order of magnitude of* $\mathrm{ex}(n, K_{4,4})$.

Similarly $\mathrm{ex}(n, K_{s,s})$ remains open for all $s \geq 4$. However, it turns out that we can solve this problem for $K_{s,t}$ whenever t is sufficiently large in terms of s.

Theorem 1.6. *For all $s \geq 2$, there exists an integer t_0 such that* $\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s})$ *for all $t \geq t_0$.*

The first result of this form was proven by Authors who showed one can take t_0 =Something by using an explicit algebraic construction like we had for G_q^*. The best current bound is due to Bukh who recently showed one can take $t_0 = 9^{(1+o(1))s}$ by using a *random* algebraic construction.

1.2 Forbidding Cliques

Now that we've all been convinced that studying $\mathrm{ex}(n, F)$ is an interesting problem, we need to figure out some graphs F for which we can effectively bound (or even determine) $\mathrm{ex}(n, F)$. As a starting step, we can think about this problem for small graphs F. A moment's thought shows that it is quite easy to determine $\mathrm{ex}(n, F)$ for every graph F with $v(F) \leq 3$ *except* for the graph $F = K_3$, which is the smallest non-trivial instance of this problem. The full solution to this problem is a classical result of Mantel from 1907.

Theorem 1.7 (Mantel's Theorem)**.** *We have* $\mathrm{ex}(n, K_3) = \lfloor n^2/4 \rfloor$ *for all $n \geq 1$. Moreover, the only n-vertex K_3-free graphs with $\lfloor n^2/4 \rfloor$ edges are those which are isomorphic to the balanced complete bipartite graph* $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

There are many proofs for Mantel's Theorem (the textbook "Proofs from the Book" contains 7 proofs, and there are many more than just these!). We will content ourselves with only a single proof here, though we sketch out a few more in the exercises.

Proof. One reasonable approach to consider when given a problem like this is to try and prove things by induction on n, which is indeed what we shall ultimately do, though we will have to be a little careful with the details.

Indeed, consider the following naive approach using induction: let G be an n-vertex K_3-free graph and v an arbitrary vertex of G. Inductively we know that $e(G - v) \leq \lfloor (n-1)^2/4 \rfloor$, and hence $e(G) \leq \lfloor (n-1)^2/4 \rfloor + \deg(v)$. Unfortunately this bound is not good enough: if, say $G = K_{1,n-1}$ and v were the center of the star then this would give a bound of $\lfloor (n-1)^2/4 \rfloor + n - 1$, which is too large. One can try and be smarter by picking v to be a vertex of minimum degree, but we do not know if this is enough to prove the result. To deal with this, we will prove the result by removing *two* vertices at a time from G rather than just one.

To this end, observe that the result is true for $n = 1, 2$. Assume we have proven the result up to some value $n \geq 3$ and let G be an n-vertex triangle-free graph. If $e(G) = 0$ then we are done, so we can assume G has an edge xy. By induction, we know that $e(G-x-y) \le \lfloor (n-2)^2/4 \rfloor = \lfloor n^2/4 \rfloor - n + 1$, and hence that

$$e(G) = e(G-x-y) + \deg(x) + \deg(y) - 1 \le \lfloor n^2/4 \rfloor + \deg(x) + \deg(y) - n,$$

where the -1 in the first equality comes from the fact that $xy \in E(G)$ and hence is counted by both $\deg(x)$ and $\deg(y)$. Finally, because G is triangle-free (which is a fact we must use somewhere in our argument), we must have $N(x) \cap N(y) = \emptyset$, as any common neighbor z would form a triangle with the edge xy. We conclude then that

$$\deg(x) + \deg(y) = |N(x)| + |N(y)| = |N(x) \cup N(y)| \le n,$$

which combined with the bound above give the desired bound.

To prove the equality case, again one can show this holds for $n = 1, 2$. Inductively then, the only way for the bound $e(G-x-y) \le \lfloor (n-2)^2/4 \rfloor$ to be tight is if $G - x - y = K_{\lfloor n/2 \rfloor - 1, \lceil n/2 \rceil - 1}$, and similarly the only way the bound $|N(x) \cup N(y)| \le n$ can be tight is if every vertex of $G - x - y$ is adjacent to exactly one of x, y, which is only possible if G is $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. □

Similar to how the "correct" way to generalize our bound for C_4 in Theorem 1.1 was to consider complete bipartite graphs, it turns out that the "correct" way to generalize Mantel's Theorem is to consider larger cliques K_r. And indeed, just like the case of triangles, the Turán number for cliques in general can be solved exactly and has a unique extremal construction which is defined as follows.

Definition 3. Given integers $r, n \ge 1$, we define the *Turán graph* $T_{r-1}(n)$ to be the $(r-1)$-partite graph whose part sizes are as equal as possible, i.e. such that each part either has size $\lfloor n/(r-1) \rfloor$ or size $\lceil n/(r-1) \rceil$

Theorem 1.8 (Turán's Theorem)**.** *For all integers $r \ge 2$ and $n \ge 1$, we have $\mathrm{ex}(n, K_r) = e(T_{r-1}(n))$. Moreover, the only n-vertex K_r-free graph with $e(T_{r-1}(n))$ edges are those which are isomorphic to $T_{r-1}(n)$.*

Again there are many different proofs of Turán's Theorem, and again we limit ourselves to just a single one here based on the following idea.

Mantra 6. If you think an extremal problem has a unique optimal construction, then try and prove this by "shifting" an arbitrary construction to look like the optimal construction.

For example, in the setting of Turán's Theorem we might want to try shifting an arbitrary K_r-free graph into a graph that, like the Turán graph $T_{r-1}(n)$, is complete $(r-1)$-partite. And indeed this is always possible to do.

Lemma 1.9 (Zykov Symmeterization)**.** *For every K_r-free graph G, there exists a graph G' satisfying the following:*

- $V(G') = V(G)$,
- $\deg_{G'}(x) \ge \deg_G(x)$ *for all* $x \in V(G)$, *and*

- *G' is complete $(r-1)$-partite.*

In particular, these last two conditions imply that G' is a K_r-free graph with at least as many edges as G.

Emphasize somewhere how the high-level idea of the proof is to duplicate vertices of high degree while deleting certain vertices of low degree.

Proof. We prove the result by induction on r, the case $r = 2$ being trivial. Let $x \in V(G)$ be a vertex of maximum degree. Observe that $H := G[N(x)]$ must be K_{r-1}-free, as any K_{r-1} in H together with x would form a K_r. By induction we can find a complete $(r-2)$-partite graph H' satisfying the conditions of the lemma for H. Now define G' to be the graph formed by starting with H' and then adding every edge from $V(H') = N_G(x)$ to the remaining vertices $x \cup (V(G) \setminus N_G(x))$.

Observe that $V(G') = V(G)$ and that G' is complete $(r-1)$-partite (namely by considering the $r-2$ parts from H' together with the part $x \cup (V(G) \setminus N_G(x))$, so it remains to check the degree condition. If $y \notin V(H') = N_G(x)$ then

$$\deg_{G'}(y) = v(H') = \deg_G(x) \geq \deg_G(y),$$

with this last inequality using that x was chosen to be a vertex of maximum degree. If instead $y \in V(H')$ then

$$\deg_{G'}(y) = \deg_{H'}(y) + |V(G) \setminus N_G(x)| \geq \deg_H(y) + |N_G(y) \setminus N_G(x)| = \deg_G(y),$$

where the inequality used $\deg_{H'}(y) \geq \deg_H(y)$ by definition of H. □

We now use this result to prove Turán's Theorem, though for simplicity we omit the proof of uniqueness.

Proof of Turán's Theorem. Let G be an n-vertex K_r-free graph. By Zykov symmeterization, we know that there exists an n-vertex complete $(r-1)$-partite graph G' with at least as many edges as G, and it is a simple exercise to show that any such graph has at most as many edges as $T_{r-1}(n)$, proving the result. □

As a historical aside, Turán proved this result without being aware of Mantel's Theorem, and in this paper he went on to introduce the general problem of determining $\text{ex}(n, F)$ for various graphs F, which is why the "Turán number" bears his name.

1.3 Forbidding Trees

We have now solved the Turán problem for the "densest" graphs K_r. We now turn to solving the problem for the "sparsest" graphs, namely that of forests and trees. The simplest case of this problem is that of stars, which is easy to solve exactly.

Proposition 1.10. *For all $r \geq 2$, we have $\text{ex}(n, K_{1,r-1}) \leq \frac{r-2}{2}n$ with equality if and only if at least one of r or n is even.*

Proof. A graph G being $K_{1,r-1}$-free is the same as saying that G has maximum degree at most $r-2$. Thus, any n-vertex $K_{1,r-1}$-free graph satisfies

$$e(G) = \frac{1}{2}\sum \deg(x) \leq \frac{1}{2}\sum r-2 = \frac{r-2}{2}n,$$

proving the upper bound. This upper bound is tight whenever there exists an n-vertex $(r-2)$-regular graph, which holds precisely if at least one of r or n is even. □

Note that in this example there are infinitely many extremal constructions, which is a significantly different phenomenon compared to what we saw when forbidding cliques.

We next turn to the problem of avoiding an arbitrary tree T, for which we might ideally like to generalize our argument for stars. Unfortunately unlike in this case we can not say that an arbitrary T-free graph has small maximum degree, but we can prove the slightly weaker statement that such a graph has small minimum degree.

Lemma 1.11. *If T is a tree with r vertices and if G is a graph with minimum degree at least $r-1$, then G contains a copy of T.*

Note that the bound of $r-1$ is best possible, as can be seen by considering graphs G which are disjoint unions of copies of K_{r-1}. We present two essentially equivalent proofs of this result, the first of which is a little vaguer but requires less knowledge of trees while the second is a bit more explicit/algorithmic.

First Proof. We prove the result by induction on r, the case $r=2$ being trivial. Assume we have the proven the result up to some $r \geq 3$ and let T be an arbitrary r-vertex tree.

Because T is a tree, there exists some leaf x with some vertex y its unique neighbor. Because G has minimum degree at least $r-1 \geq r-2$, we inductively can assume that G has a copy of $T' = T - x$. Now the vertex playing the role of y in this copy of T' has at least $r-1$ neighbors, of which at most $r-2$ of them lie in this copy of T'. In particular, there exists at least one neighbor which is not in T', and taking this together with the copy of T' gives a copy of T gives the desired result. □

Second Proof. We build up our copy of T algorithmically "vertex by vertex." To do this we require the fact that for every r-vertex tree, there exists an ordering of the vertices $v_1, \ldots, v_r$ such that for all $2 \leq i \leq r$ there exists an integer $j_i < i$ such that $N_T(v_i) \cap \{v_1, \ldots, v_{i-1}\} = \{v_{j_i}\}$.

Let y_1 be an arbitrary vertex of G. Iteratively given that we have chosen vertices $y_1, \ldots, y_{i-1}$ in G for some $i \leq r$, we choose y_i to be an arbitrary vertex in $N_G(y_{j_i})$ which is not in the set $\{y_1, \ldots, y_{i-1}\} \setminus \{y_{j_i}\}$. Note that the number of such vertices is at least $r-1-(i-2) \geq 1$, so there does indeed exist a valid choice for y_i, and as such this algorithm will successfully terminate. With this, it is not difficult to see that the y_i vertices form a copy of T, giving the result. □

The result above gives a tight bound on the minimum degree needed to contain a copy of T, but we ultimately want a bound on $\mathrm{ex}(n, T)$, i.e. on the *average* degree needed to find a copy of T. Fortunately, there is a general result which allows us to translate between the concept of minimum degrees and average degrees.

Proposition 1.12. *If G is a graph of average degree at least d, then there exists a non-empty subgraph $G' \subseteq G$ with minimum degree at least $d/2$ and average degree at least d.*

For most applications of this result we will only need the conclusion that G' has large minimum degree, but sometimes it is useful to also have this additional average degree condition (see for example Theorem 2.9). Again we offer two essentially equivalent proofs of this result, both of which implicitly use that the average degree by definition is

$$v(G)^{-1} \sum \deg(x) = \frac{2e(G)}{v(G)}.$$

First Proof. Assume the result is false for a given d and graph G, and choose such a counterexample with $v(G)$ as small as possible. If $\delta(G) \geq d/2$ then taking $G' = G$ gives the desired subgraph, a contradiction. As such, we can assume that G contains a vertex x with $\deg(x) < d/2$. In this case, the graph $G-x$ has a smaller number of vertices and average degree

$$\frac{2e(G-x)}{v(G-x)} = \frac{2e(G) - 2\deg(x)}{v(G)-1} \geq \frac{2e(G)-d}{v(G)-1} \geq d,$$

with this last step using that $2e(G) \geq dv(G)$ by hypothesis. Since $G-x$ is a graph with fewer vertices than G and with average degree d, our choice of G having $v(G)$ as small as possible implies that there exists $G' \subseteq G - x \subseteq G$ satisfying the properties of the statement, giving another contradiction. ☐

Second Proof. The key idea of the argument is to start with $G' = G$ and then iteratively remove vertices of low degree, i.e. as long as G' contains a vertex of degree less than $d/2$ then we remove this vertex and we repeat this until no such vertices exist. Note that the total number of edges that we remove in this process is certainly less than

$$(d/2) \cdot v(G) \leq e(G),$$

with this inequality being equivalent to saying that G has average degree at least $d/2$. As such, the resulting graph G' has at least one edge and has minimum degree at least $d/2$ by construction. One can similarly check that it has average degree at least d, proving the result. ☐

This in total lets us prove the following.

Theorem 1.13. *For any r-vertex tree T, we have*

$$\frac{r-2}{2}n - O_r(1) \leq \mathrm{ex}(n,T) \leq (r-2)n$$

Proof. For the lower bound we take the disjoint union of copies of K_{r-1}, which is certainly T-free and which has the stated number of edges.

For the lower bound, assume that there exists an n-vertex T-free graph G with $e(G) > (r-2)n$, i.e. with average degree more than $2(r-2)$. By Theorem 1.12 there exists a subgraph G' of G with minimum degree more than $r-2$, i.e. with $\delta(G') \geq r-1$. By Theorem 1.11 $G' \subseteq G$ contains a copy of T, a contradiction. ☐

While Theorem 1.13 solves the Turán problem for trees up to a factor of 2, one can ask if one can give an even more precise answer. In particular, given that the lower bound of Theorem 1.13 is the truth for the case of stars, it is natural to believe this should be the answer in general.

Conjecture 1.14 (Erdős-Sós)**.** *Every r-vertex tree T satisfies* $\mathrm{ex}(n,T) \le \frac{r-2}{2}n$.

There are a number of special cases for which the Erdős-Sós Conjecture is known to be true (such as for paths; see Theorem 2.9), but overall the problem of improving the small gap from Theorem 1.13 for all T seems difficult to do

1.4 Forbidding Hypergraphs

In this subsection, we very briefly explore Turán problems for hypergraphs. This area is a massive field of study, and in order to narrow our scope, the results we consider here are not the most fundamental results in the area, but rather a few simple examples which have the benefit of highlighting key concepts in the studies of hypergraphs, such as designs, codegrees, and shadows.

To begin our study, we will need to define an appropriate generalization of the Turán problem to hypergraphs. To this end, given two hypergraphs H, F, we say that H is F-free if it contains no subgraph isomorphic to F. Given an r-graph F, we then define $\mathrm{ex}(n,F)$ to be the maximum number of edges in an n-vertex r-uniform hypergraph which is F-free.

Before going further, let us briefly comment that this definition fo $\mathrm{ex}(n,F)$ is significantly less interesting if we allow F and our F-free hypergraph to be arbitrary (i.e. not necessarily uniform) hypergraphs. Indeed, for any $e \in E(F)$ we can find n-vertex F-free graphs with at least $2^n - \binom{n}{|e|}$ edges, namely by keeping every edge except those of size $|e|$ which certainly can't contain $F \ni e$ as a subhypergrpah. Thus the Turán problem (as well as many others) is most natural only in the uniform setting.

Back to the problem of exploring $\mathrm{ex}(n,F)$, it perhaps makes sense as a starting point to consider hypergraphs F which are small. For example, if F has a single edge then $\mathrm{ex}(n,F) = 0$ for all n. However, already when we go up to F containing two edges we reach a problem which is not entirely solved. More precisely, let $F_t^{(r)}$ denote the r-uniform hypergraph which consists of two edges intersecting in exactly t vertices. Determining $\mathrm{ex}(n, F_t^{(r)})$ is known as the Erdős-Sós problem and is still open in full generality, though there are a number of cases where it is solved. For example, in the extreme case of $t = 0$ we see that $\mathrm{ex}(n, F_0^{(r)})$ is the maximum number of edges in an intersecting r-uniform hypergraph, i.e. a hypergraph where every two edges have a non-trivial intersection. While this problem is nearly trivial in the setting of graphs, the determination of $\mathrm{ex}(n, F_0^{(r)})$ is a famous result known as the Erdős-Ko-Rado Theorem which is the fundamental result in the area known as extremal set theory.

Later in these notes we will give a linear algebraic proof of the Erdős-Ko-Rado Theorem. For now, we consider the other extremal case of $t = r-1$ where we observe the following.

Proposition 1.15. *For all $r \ge 2$ we have* $\mathrm{ex}(n, F_{r-1}^{(r)}) \le r^{-1}\binom{n}{r-1}$.

Proof. Our proof will use a hypergraph generalization of our upper bound proof for the Turán number of stars, and for this we will need to prove a generalization of the handshaking lemma.

Given a set of vertices $S \subseteq V(H)$, we define $\deg(S)$ to be the number of edges of H which contain the set S. We claim that for every hypergraph H and every integer k we have

$$\sum_{S \in \binom{V(H)}{k}} \deg(S) = \sum_{e \in E(H)} \binom{|e|}{k}.$$

Indeed, consider the set of all pairs (S, e) where $e \in E(H)$ and $S \subseteq e$ is a set of size k. On the one hand, the number of pairs can be counted by first specifying $S \in \binom{V(H)}{k}$ and then specifying one of the $\deg(S)$ edges containing it to play the role of e. On the other hand we could first pick $e \in E(H)$ and then choose one of the $\binom{|e|}{k}$ subsets of size k to play the role of S. These two methods of counting give the two expressions above and hence they must equal each other.

With this claim in mind, let H be an n-vertex hypergraph with no $F_{r-1}^{(r)}$. Equivalently this says $\deg(S) \le 1$ for all sets S of size $r-1$, and hence

$$\sum_{e \in E(H)} \binom{|e|}{r-1} = \binom{r}{r-1} e(H) = \sum_{S \in \binom{V(H)}{r}} \deg(S) \le \binom{n}{r-1},$$

and rearranging gives the desired bound on $e(H)$. □

The natural followup question from here is to ask when this upper bound for $\mathrm{ex}(n, F_{r-1}^{(r)})$ is tight. Examining the proof, we see that this can happen if and only if there exists an n-vertex r-uniform hypergrpah H such that every $(r-1)$-set is contained in exactly 1 edge. Such hypergraphs are somewhat analogous to the notion of regular graphs (where we demand every 1-set becontained in some d edges) and more generally is an example of something known as a combinatorial design.

More generally, a hypergraph H is said to be[3] an (n, r, k, d)-Steiner system if H is an n-vertex r-uniform hypergrpah with the property that every set of k elements is contained in exactly d edges. Designs such as Steiner systems originally arose in problems around designing statistical experiments and since then they have been recognized as interesting and useful objects within combinatorics.

Our proof above shows that the Turán problem for $F_{r-1}(r)$ is a hypergraph analog of the Turán problem for the graph $K_{1,2}$, and it is natural to ask whether we can find hypergraph analogs of our other graph Turán results. Notably, we might consider hypergraph generalizations of 4-cycles, which was the first Turán problem we considered.

There are several different notions of what it could mean to be a "hypergraph 4-cycle". For simplicity, we will consider only the so-called *loose 4-cycle* $C_4^{(3)}$ which is the 3-uniform hypergraph on 8 vertices with 4 edges $\{x_1, y_{1,2}, x_2\}, \{x_2, y_{2,3}, x_3\}, \{x_3, y_{3,4}, x_4\}, \{x_4, y_{4,1}, x_1\}$. One can also think of this as the 3-graph obtained by starting with a graph C_4 and then inserting a new vertex inside each edge. A simple argument for this gives the following.

Proposition 1.16. *We have* $\mathrm{ex}(n, C_4^{(3)}) = \Theta(n^2)$.

[3] The exact letters (n, r, k, d) are not standard and different text will use different notation.

Proof. The lower bound comes from considering an n-vertex hypergraph star, i.e. a 3-graph on $[n]$ which contains all $\binom{n-1}{2}$ edges containing the element 1. Such a hypergraph is indeed $C_4^{(3)}$-free since any copy of $C_4^{(3)}$ contains two edges which are disjoint from each other which is something the hypergraph star does not have.

For the upper bound, we use a variant of the average degree to minimum degree argument except we will apply this not to minimum degrees of vertices but rather minimum degrees of pairs of vertices. To this end, given a pair of vertices x, y in H we define[4] its *codegree* $\deg(x, y)$ to be the number of edges of H containing both x and y.

Now let H be an n-vertex 3-graph with, say, $e(H) \geq 7\binom{n}{2}$ and assume for contradiction that H has no $C_4^{(3)}$. Iteratively, if H contains a pair of vertices x, y with $1 \leq \deg(x, y) < 7$ then we remove from H all edges which contain both x and y. Letting $H' \subseteq H$ denote the hypergraph at the end of this process, we see that

$$e(H') > e(H) - 7\binom{n}{2} \geq 0.$$

Thus there exists some edge $\{x_1, x_2, x_3\} \in E(H')$. By definition, this implies that the pair x_1, x_3 must be contained in at least 7 edges of H', and in particular there must exist an edge of the form $\{x_1, x_3, x_4\}$ for some $x_4 \neq x_2$. Iteratively now for each $1 \leq i \leq 4$, we observe that because x_i, x_{i+1} belong to an edge of H' that there must exist at least 7 edges containing this pair, and in particular there exists some edge of the form $\{x_i, y_{i,i+1}, x_{i+1}\}$ where $y_{i,i+1}$ does not equal any of the x_j vertices or any other vertices of the form $y_{j,j+1}$. This yields a copy of $C_4^{(3)}$, giving a contradiction. □

We emphasize that this proof is not a generalization of our argument for $\mathrm{ex}(n, C_4)$, and instead it is much closer to our general argument for Turán numbers of trees. Indeed, it turns out one can view $C_4^{(3)}$ as the subgraph of a certain "hypergraph tree" maybe which I'll discuss in the exercises, which in total implies that the somewhat boring construction of a hypergraph star is close to extremal for $C_4^{(3)}$.

This in turn leads to the somewhat boring fact that $\mathrm{ex}(n, C_4^{(3)}) = \Theta(n^2)$, in the sense that many 3-graphs F have $\mathrm{ex}(n, C_4^{(3)}) = \Omega(n^2)$ via considering H a hypergraph star, and in particular the extremal construction for $C_4^{(3)}$ is a lot less interesting than what we saw for the graph C_4. We could thus ask, similar to our work in Dirac's problem I guess I need to move the mantra from that chapter to this one and/or motivate things by us wanting to be closer to the graph case, if there's a way to change our setting to forbid us from taking a hypergraph star as a construction. One way to do this is to further restrict the class of hypergraphs we work with to resemble the graph case even more.

Definition 4. We say that a hypergraph H is *linear* if any two distinct edges $e, f \in H$ has $|e \cap f| \leq 1$. Given a linear r-uniform hypergraph F, we define the linear Turán number $\mathrm{ex}_{lin}(n, F)$ to be the maximum number of edges in an n-vertex F-free r-uniform linear hypergraph.

[4] Here we emphasize that the "degree" of a pair of vertices (or more generally any set of size larger than 1) is often but not always referred to as "codegrees" to distinguish it from the study of degrees of single vertices.

In particular, every (simple) graph is linear, and restricting to this setting indeed produces Turán behavior that more closely resembles graphs.

Proposition 1.17. *We have* $\mathrm{ex}_{lin}(n, C_4^{(3)}) = O(n^{3/2})$.

Proof. Our proof utilizes a useful object known as the shadow of a hypergraph. In general, given an r-uniform hypergraph H we define its k-shadow $\partial^k H$ to be the k-uniform hypergraph on the same vertex set of H where a k-set K is an edge of $\partial^k H$ if and only if there exists an edge $e \in H$ with $K \subseteq e$. In particular, $\partial^2 H$ is simply the graph where two vertices are adjacent to each other if and only if there's some edge in H containing the vertices.

The rough idea for our proof will be to try and find something like a graph C_4 in $\partial^2 H$, say on $x_1, \ldots, x_4$. For each edge $x_i x_{i+1}$ in this C_4, we have by definition of $\partial^2 H$ that there exists an edge $\{x_i, x_{i+1}, y_i\} \in H$ and we would like to say these 4 hyperedges give a copy of $C_4^{(3)}$ in H. The problem with this idea is that some of these y_i may equal each other, or even equal some x_j vertex, which means this will not actually give rise to a $C_4^{(3)}$ in H. However, because H is linear, the number of times we can have y_i equaling some other vertex is relatively small, allowing us to prove the following analog of this approach.

Claim 1.18. *If H is a linear 3-uniform hypergraph which is $C_4^{(3)}$-free, then $\partial^2 H$ is $K_{2,7}$-free.*

Proof. Assume for contradiction that there was a $K_{2,7}$, say with u, v its part of size 2 and $x_1, \ldots, x_7$ its parts of size 7. Because H is linear and because $ux_i, vx_i \in E(\partial^2 H)$ for all i, there exist unique vertices u_i, v_i for each i such that $\{u, u_i, x_i\}, \{v, v_i, x_i\} \in E(H)$. Motivated by our strategy above, we wish to pick some x_i, x_j such that all 4 of these associated hyperedges of H define a $C_4^{(3)}$.

We begin by observing that $u_i \neq v_i$ for all i since otherwise $\{u, u_i, x_i\}, \{v, v_i, x_i\}$ would be two distinct edges intersecting in two vertices, a contradiction to H being linear. Also, $u_i \neq u_{i'}$ and $v_i \neq v_{i'}$ for $i \neq i'$ as otherwise e.g. $\{u, u_i, x_i\}, \{u, u_{i'}, x_{i'}\}$ would be two edges containing the pair $u, u_i = u_{i'}$, and an analogous argument shows $u_i \neq v$ and $v_i \neq u$ for all i. In a similar spirit, there is at most one i such that $\{u, v, x_i\} \in E(H)$, and without loss of generality we can assume that this edge does not exist for any $1 \leq i \leq 6$. We will now aim to find a $C_4^{(3)}$ which includes the edges $\{u, u_1, x_1\}, \{v, v_1, x_1\} \in E(H)$. Note that there is at most one index $2 \leq i \leq 6$ such that $u_1 = v_i$ since we already noted that $v_i \neq v_{i'}$ for distinct i, i'. Similarly there is at most one index $2 \leq i \leq 6$ such that $v_1 = v_i$. Finally, there is at most one index $2 \leq i \leq 6$ such that $\{u, x_1, x_i\} \in E(H)$ and also at most one such index such that $\{v, x_1, x_i\} \in E(H)$. Taking any $j \in \{2, 3, \ldots, 6\}$ not equal to any of these four potential bad indices i, we find that $\{u, u_1, x_1\}, \{v, v_1, x_1\}, \{u, u_j, x_j\}, \{v, v_j, x_j\}$ forms the edges of a $C_4^{(3)}$ (as can be checked by verifying that each of u_1, v_1, u_j, v_j belong to exactly one of these edges). ☐

Observe now that for any 3-uniform linear hypergraph H that $e(\partial^2 H) = 3e(H)$ since each edge of H contains three pairs and each pair is contained in exactly one edge because H is linear. We thus have that if H is an n-vertex $C_4^{(3)}$-free 3-uniform linear hypergraph then $e(H) \leq \frac{1}{3}\mathrm{ex}(n, K_{2,7}) = O(n^{3/2})$, proving the result. ☐

We note that this bound of $O(n^{3/2})$ is in fact tight: if we take the polarity graph G_q^* and then define a hypergraph H with $V(H) = V(G_q^*)$ whose edges are the triangles of G_q^*, then one can check that this hypergraph has about $n^{3/2}$ edges and no $C_4^{(3)}$ (since in particular $\partial^2 H$ does not even contain a C_4).

The previous example used graph theoretic ideas to solve a hypergraph problem, and we'll close with one more example of this form for another variant of hypergraph 4-cycles. To this end, we define the wheel hypergraph W to be the 3-graph with $V(W) = \{v_1, v_2, v_3, v_4, w\}$ and with edges of the form $\{v_i, v_{i+1}, w\}$ for all $1 \le i \le 4$ with these indices written mod 4. Note that like the loose 4-cycle, the wheel is defined by taking a graph C_4 and adding a new vertex into each edge, the only difference being that in the wheel the same vertex is added to each edge while the loose 4-cycle had distinct vertices added.

Proposition 1.19. *The wheel hypergraph W satisfies* $\mathrm{ex}(n, W) = O(n^{5/2})$.

Proof. For a 3-graph H and a vertex v, we define the *link graph* $L_H(v)$ to be the graph on $V(H) \setminus \{v\}$ where $u \sim u'$ if and only if $\{u, u', v\} \in E(H)$. That is, edges of $L_H(v)$ are exactly the pairs such that v together with this pair forms an edge of H.

Let H be an n-vertex wheel-free 3-graph. Crucially, we observe that being wheel-free means $L_H(v)$ is C_4-free for all v, as otherwise this C_4 on $v_1', \ldots, v_4'$ together with v playing the role of w would give a wheel. Thus $e(L_H(v)) \le \mathrm{ex}(n-1, C_4) = O(n^{3/2})$. On the other hand, $e(L_H(v)) = \deg(v)$. By the handshaking lemma,

$$3e(H) = \sum_v \deg(v) = \sum_v O(n^{3/2}) = O(n^{5/2}),$$

proving the result. □

Need to remember what the current bounds are: I think $n^{3/2}$ lower by taking C_4 and adding an apex vertex, best upper bound I think is for $K_{1,s,t}$ in general and is some little oh term initially from removal arguments.

1.5 The Landscape of Turán Problems

At this point we've studied $\mathrm{ex}(n, F)$ for many particular choices of F, but we've said almost nothing about graphs in general. Part of the issue is that the Turán number behaves in very different ways depending on the structure of the graph F, in the following sense.

Proposition 1.20. *Let F be a graph.*

- *If F is non-bipartite, then* $\mathrm{ex}(n, F) = \Theta(n^2)$.
- *If F is bipartite, then* $\mathrm{ex}(n, F) = O(n^{2-1/v(F)})$.

Proof. For any graph we have $\mathrm{ex}(n, F) \le e(K_n) = \binom{n}{2} = O(n^2)$. If F is also non-bipartite, then the balanced complete bipartite graph $K_{\lfloor n/2\rfloor, \lceil n/2\rceil}$ is F-free and shows that $\mathrm{ex}(n, F) \ge \lfloor n^2/4 \rfloor = \Omega(n^2)$, proving the first part. For the second part, because F is bipartite, we have $F \subseteq K_{v(F), v(F)}$, and hence by Kővári-Sós-Turán,

$$\mathrm{ex}(n, F) \le \mathrm{ex}(n, K_{v(F), v(F)}) = O(n^{2-1/v(T)}).$$

□

This observation divides the study of Turán number of graphs into two distinct cases: the *non-degenerate* case which studies non-bipartite F (i.e. those graphs with $\mathrm{ex}(n, F) = \Theta(n^2)$), and the *degenerate* case which studies bipartite F (i.e. those graphs with $\mathrm{ex}(n, F) = o(n^2)$). In what follows we very briefly survey results for these cases. Some of these results require non-trivial machinery to prove, and as such will be deferred until much later in the text.

The Non-Degenerate Case. For non-bipartite graphs F, the most important theorem is the following.

Theorem 1.21 (Erdős-Stone-Simonovits)**.** *For any graph F with at least one edge, we have*

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

In particular, this result determines the asymptotic value of $\mathrm{ex}(n, F)$ for *any* non-bipartite[5] graph F. The lower bound for this is rather easy: the Turán graph $T_{\chi(F)-1}(n)$ has chromatic number $\chi(F) - 1$ and hence is F-free and has about $\frac{\chi(F)-2}{\chi(F)-1}\binom{n}{2}$ edges. The upper bound is somewhat difficult to prove, and we will defer this until we have the regularity lemma at our disposal.

Because of the Erdős-Stone-Simonovits Theorem, the non-degenerate case of the Turán problem is often considered to be a solved problem. That being said, for any given non-bipartite graph F one can still ask for sharper (or even exact) bounds on $\mathrm{ex}(n, F)$, as well as to determine the full set of optimal extremal constructions. The most useful result in this direction is the following due to Simonovits.

Theorem 1.22. *Let F be a graph which has a critical edge, i.e. edge e with $\chi(F - e) < \chi(F)$. Then $\mathrm{ex}(n, F) = e(T_{\chi(F)-1}(n))$ for all n sufficiently large, and moreover the unique extremal construction for n sufficiently large is $T_{\chi(F)-1}(n)$.*

We emphasize that the assumption of n being sufficiently large is necessary in general. Indeed, we always have $\mathrm{ex}(n, F) = \binom{n}{2}$ whenever $n < v(F)$, and for small n this will typically be better than the bound given in Theorem 1.22. It is also an easy exercise to show that if F does not have a critical edge, then $\mathrm{ex}(n, F) > e(T_{\chi(F)-1}(n))$ for all n sufficiently large. As such, Theorem 1.22 exactly classifies which graphs F have Turán graphs as their extremal constructions.

The Degenerate Case. While non-degenerate Turán problems for graphs are largely solved, very little is known about degenerate Turán problems. Indeed, even determining the order of magnitude of relatively simple bipartite graphs remain open despite decades of study. We already mentioned that for complete bipartite graphs that $\mathrm{ex}(n, K_{s,s})$ remains open for all $s \geq 4$. Similarly, for even cycles (which are perhaps the next most natural class of bipartite graphs to study) our knowledge can largely be summarized as follows.

Theorem 1.23. *For all $\ell \geq 2$, we have $\mathrm{ex}(n, C_{2\ell}) = O(n^{1+1/\ell})$. Moreover, this is best possible whenever $\ell = 2, 3,$ or 5.*

[5] If F is bipartite the theorem simply says $\mathrm{ex}(n, F) = o(n^2)$, which follows from Kővári-Sós-Turán.

That is, we know the Turán number for C_4, C_6, and C_{10}, but frustratingly not for C_8! This is roughly because there exists a class of very particular algebraic objects which just so happen to solve these three cases and no others. Another frustrating problem is that of the 3-dimensional hypercube graph Q_3, which can be viewed as the "skeleton" of a usual cube. Determining $\mathrm{ex}(n, Q_3)$ was one of the original problems that Turán raised back in his 1941 paper on the topic, but to date only the following bounds are known.

Theorem 1.24. *We have* $\mathrm{ex}(n, Q_3) = O(n^{8/3})$ *and* $\mathrm{ex}(n, Q_8) = \Omega(n^{3/2})$.

The lower bound comes simply by considering an extremal C_4-free graph. The upper bound is based on a "supersaturation" argument of Erdős and Simonovits from 1969.

While nothing as strong as the Erdős-Stone-Simonovitis Theorem exist for bipartite graphs, there are a few nice general bounds. For lower bounds, essentially the best we know is the following.

Theorem 1.25. *If F is a graph with v vertices and e edges with $e \geq v$, then*

$$\mathrm{ex}(n, F) = \Omega(n^{2-\frac{v-2}{e-1}}).$$

This bound comes from a probabilistic argument that we will see in **REF**. A general upper bound which can be proven by a more sophisticated argument is as follows.

Theorem 1.26 (Füredi)**.** *If F is a bipartite graph where every vertex on one side of the bipartition has degree at most r, then*

$$\mathrm{ex}(n, F) = O(n^{2-1/r}).$$

Much more can be said about what we do not know about Turán numbers of bipartite graphs, see Survey.

Hypergraphs. As in the graph case, the study of $\mathrm{ex}(n, F)$ for r-graphs can be divided into non-degenerate and degenerate cases, where here what distinguishes these two cases is whether or not F is r-partite or not.

Theorem 1.27. *Let F be an r-graph.*

- *If F is not r-partite, then there exists some real number $\pi(F) > 0$ such that $\mathrm{ex}(n, F) \sim \pi(F)\binom{n}{r}$.*
- *If F is r-partite, then $\mathrm{ex}(n, F) = O(n^{r-\varepsilon})$ for some $\varepsilon > 0$ depending on F.*

Just like in the case of graph, the degenerate case for Turán problems of r-partite r-graphs is largely unsolved. To make matters worse, the non-degenerate case is also largely unsolved: not only is there no analog of the Erdős-Stone-Simonovits Theorem, even for very simple F the quantity $\pi(F)$ is unknown.

For example, let $K_t^{(r)}$ denote the complete r-uniform hypergrpah on t vertices, i.e. the t-vertex r-uniform hypergraph which has every possible edge. A natural hypergraph generalization of Mantel's Theorem would be to try and determine $\mathrm{ex}(n, K_4^{(3)})$, i.e. the Turán number for the smallest non-trivial clique. However, despite years of hard work, even determining $\pi(K_4^{(3)})$ remains a major open problem in the field. Part of the difficulty with this problem is that unlike for graphs where there was a unique extremal example for K_3, there in fact exist exponentially many non-isomorphic constructions which are all believed to be extremal for $K_4^{(3)}$, making it significantly harder to solve. Also talk about how the possible values of $\pi(F)$ are also more complicated with there being non-jumps and accumulation points and irrational values for simple hypergraphs, and so on.

Other weird things can also happen for hypergraphs, eg exponents not being simple and compactness conjecture not being true.

Maybe talk about other variants of Turan problems, eg general and random

1.6 Exercises

1. Verify that the graphs G_q, G_q^* defined in the first subsection are C_4-free and that $v(G_q^*) = q^2 + q + 1$ and $e(G_q^*) = \frac{1}{2}(q+1)(q^2+q+1)$ [1+].

2. Prove the Kővári-Sós-Turán Theorem, Theorem 1.4 [1+].

3. Given integers $m, n, s, t \geq 1$, define the *Zarankiewicz number*[6] $z(m, n; s, t)$ to be the maximum number of edges in a bipartite graph G with parts U, V satisfying $|U| = m, |V| = n$, and that G no copy of $K_{s,t}$ with the part of size s in U and the part of size t in V.

   (a) Prove that
   $$z(m, n; s, t) \leq (t-1)^{1/s} m n^{1-1/s} + (s-1)n.$$
   (Hint: if you're struggling with this, try solving the previous problem first) [2].

   (b) Prove that if G is an n-vertex bipartite C_4-free graph then $e(G) \leq 2^{-3/2} n^{3/2} + o(n^{3/2})$, i.e. the lower bound we got for $\mathrm{ex}(n, C_4)$ using G_q was best possible in the setting of bipartite graphs [2].

   (c) Prove that for all s, t there exists a constant $C > 0$ such that if G is an n-vertex $K_{s,t}$-free graph, then the number of edges $xy \in E(G)$ with $\deg(x) \geq C n^{1-1/s}$ is at most $O(n)$. Find an example of a graph which has $\Theta(n)$ edges of this form (Hint: the intended proof I have in mind works with $C \approx (s+t-1)^{1/s}$) [2].

   (d) Use (a) with $s = t = 2$ to give a generalization of Theorem 1.2 [1].

4. The Turán problem involves graphs with 0 copies of a given graph F (where here by a *copy* we mean a subgraph isomorphic to F). What about graphs with more copies?

[6] Some texts define $z(m, n; s, t)$ with respect to G which are $K_{s,t}$-free rather than simply avoiding things on one side like we have here.

(a) Prove that if G is an n-vertex graph then G contains at least $e(G) - \mathrm{ex}(n, F)$ copies of F for any graph F with at least one edge [1].

(b) Prove that if G is an n-vertex graph with $e(G) \geq 100n^{3/2}$ then G contains at least $\Omega(n^{-4}e(G)^4)$ copies of C_4 (the number 100 does not matter in case you'd rather prove this result with a different constant). [2].

Note that the number of copies guaranteed in (b) is far more than the naive bound given by (a). This sort of phenomenon of graphs with $e(G)$ just above $\mathrm{ex}(n, F)$ having a surprisingly large jump in the number of copies of F is known as *supersaturation*.

(c) Prove that for all m with $100n^{3/2} \leq m \leq \binom{n}{2}$ that there exists an n-vertex graph G with $e(G) = \Theta(m)$ and with $\Theta(n^{-4}m^4)$ copies of C_4 (Hint: consider something random) [2+].

5. Prove that $\mathrm{ex}(n, K_{3,3}) = \Omega(n^{5/3})$ [3].

6. Prove that $\mathrm{ex}(n, K_{s,t}) = \Omega(n^{2-1/s})$ for all t sufficiently large in terms of s [3+].

* * *

7. Determine $\mathrm{ex}(n, F)$ for all graphs F with $2 \leq v(F) \leq 3$ other than $F = K_3$. Why did I leave out the case $v(F) = 1$? [1].

8. Verify that if G' is an n-vertex complete $(r-1)$-partite graph then $e(G') \leq e(T_{r-1}(n))$ [1+].

9. Here we sketch a few alternative proofs of Mantel's Theorem and Turán's Theorem.

   (a) Observe that if G is a triangle-free graph, then $\deg(x) + \deg(y) \leq v(G)$ for all $xy \in E(G)$. Use this to prove Mantel's Theorem (which is in fact the original way Mantel proved his result) [2].

   (b) Generalize our inductive proof of Mantel's Theorem to give an alternative proof of Turán's Theorem (which is in fact the original way that Turán proved his result). For simplicity you can choose to prove only that

$$\mathrm{ex}(n, K_r) \leq \left(1 - \frac{1}{r-1}\right)\frac{n^2}{2},$$

   which one can check is equivalent to proving the upper bound of Turán's Theorem [2].

10. Let F denote the unique 4-vertex graph with 5 edges (i.e. the graph consisting of two triangles sharing an edge). Prove (without using Theorem 1.22) that $\mathrm{ex}(n, F) = \lfloor n^2/4 \rfloor$ for all $n \geq 4$ [2].

11. If F denotes the "bowtie" graph consisting of two triangles sharing a vertex, show that $\text{ex}(n, F) = \lfloor n^2/4 \rfloor + 1$ for all $n \geq 6$ [3-].

* * *

12. Determine $\text{ex}(n, P_4)$ exactly for all n (Hint: characterize all connected P_4-free graphs) [2].

13. Prove that for every integer $s \geq 1$ and real $\varepsilon > 0$, there exists a graph with average degree at least $2s - \varepsilon$ which contains no non-empty subgraph with minimum degree greater than $s + 1$; that is, the $d/2$ in Theorem 1.12 is essentially best possible [2-].

* * *

14. One can consider Turán problems which avoids more than just a single graph at a time. To this end, given a set of graphs $\mathcal{F}$, we say that a graph G is $\mathcal{F}$-free if G is F-free for all $F \in \mathcal{F}$.

    Prove (without using Theorem 1.23) that for all $\ell \geq 2$ we have $\text{ex}(n, \{C_3, C_4, \ldots, C_{2\ell}\}) = O(n^{1+1/\ell})$ (Hint: first prove the result under the additional assumption that every vertex of G has degree at least $n^{1/\ell} + 1$) [2].

15. Prove that if F is a graph with $\text{ex}(n, F) = \Omega(n)$ and if F' is a graph obtained from F by adding a new vertex x and making it adjacent to a vertex $y \in V(F)$, then $\text{ex}(n, F') = \Theta(\text{ex}(n, F))$. In other words, to determine the order of magnitude of $\text{ex}(n, F)$ for all graphs F, it suffices to do so for all graphs with minimum degree at least 2 [2].

* * *

16. Prove for every r-uniform hypergraph F that the limit

$$\lim_{n \to \infty} \frac{\text{ex}(n, F)}{\binom{n}{r}}$$

    exists [1+].

17. Prove that $\text{ex}(n, F_1^{(3)}) \leq n$ where $F_1^{(3)}$ is the 3-uniform hypergraph consisting of two edges intersecting in a single vertex. Show that this bound is tight for infinitely many n [2+].

18. Prove that if an (n, r, k, d)-Steiner system exists then $\binom{r-i}{k-i}$ divides $d\binom{n-i}{k-i}$ for all $0 \leq i \leq k$. As an aside, a very deep result obtained independently by Keevash and by Glock, Kühn, Lo, and Osthus implies that this necessary condition is sufficient whenever n is sufficiently large in terms of the parameters [2].

19. For all $\ell, r \geq 3$, define the r-uniform loose ℓ-cycle $C_\ell^{(r)}$ to be the r-uniform hypergraph obtained by inserting $r - 2$ new vertices inside each edge. More precisely, this is the hypergraph with edges $\{x_i, x_{i+1}, y_{i,i+1,1}, \ldots, y_{i,i+1,r-2}\}$ for all $1 \leq i \leq \ell$ with indices written modulo ℓ.

(a) Prove that $\mathrm{ex}(n, C_\ell^{(3)}) = \Theta(n^2)$ for all $\ell \geq 3$ [2].

(b) Prove that $\mathrm{ex}(n, C_\ell^{(r)}) = \Theta(n^{r-1})$ for all $\ell, r \geq 3$. We emphasize that this is not fundamentally more difficult than the previous problem but the notation does becomes a bit more cumbersome to work with [2].

20. For an arbitrary graph F let $F^{+(r-2)}$ denote the r-uniform hypergraph obtained by inserting $r-2$ new vertices into each edge of F. For example, $C_\ell^{(r)} = C_\ell^{+(r-2)}$. Recall also for graphs K, F that $\mathrm{ex}(n, K, F)$ denotes the maximum number of copies of K in an n-vertex F-free graph. Prove for every graph F which is not a subgraph of a star that

$$\mathrm{ex}(n, F^{+(r-2)}) = \mathrm{ex}(n, K_r, F) + \Theta(n^{r-1}).$$

Equivalently, prove that $\mathrm{ex}(n, F^{+(r-2)}) \geq \mathrm{ex}(n, K_r, F)$, that $\mathrm{ex}(n, F^{+(r-2)}) = \Omega(n^{r-1})$, and that $\mathrm{ex}(n, F^{+(r-2)}) \leq \mathrm{ex}(n, K_r, F) + O(n^{r-1})$ [2+].

2 Spanning Subgraphs: Dirac Problems

Up to this point we have considered the Turán number $\mathrm{ex}(n, F)$ where we think of F as a fixed graph and n as tending towards infinity, but this is not the only regime that could be considered. For example, $\mathrm{ex}(n, C_n)$ asks for the maximum number of edges that an n-vertex graph can have without containing a Hamiltonian cycle. More generally, we might consider $\mathrm{ex}(n, F_n)$ where F_n is some sequence of spanning subgraphs of K_n.

Unfortunately the Turán problem for spanning subgraph tends not to be very interesting. For example, we have $\mathrm{ex}(n, C_n) \ge \binom{n-1}{2} + 1$ by taking G to be a clique on $n-1$ vertices together with a single vertex of degree 1, and one can show that this somewhat silly construction is best possible. More generally, $\mathrm{ex}(n, F_n)$ tends to be ludicrously large for a number of natural choices of F_n simply by considering graphs G which have a single vertex of small degree. This leads us to another mantra.

Mantra 7. If an extremal problem has a known or boring optimal construction, try modifying or adding extra restrictions to the problem in such a way that any solution to this new problem must be "far" from the known/boring construction.

In particular, our current construction for $\mathrm{ex}(n, C_n)$ is boring because we can trivially make constructions by using vertices of very small degrees. So what if we instead forced our constructions to have large minimum degree? This leads to the following broad type of problem.

Open Problem 2.1 (The Dirac Problem)**.** *Given a graph F, determine the smallest number δ such that any $v(F)$-vertex graph G with $\delta(G) \ge \delta$ has a copy of F as a spanning subgraph.*

Note that we have already problems similar to this when we were working on Turán numbers for trees via Theorem 1.11. We will see another application of min degree results to Turán problems with Theorem 2.9.

2.1 Hamiltonian Cycles

Recall that a graph G is Hamiltonian if it contains a cycle passing through all of its vertices. Historically, the first study of Dirac problems came from Dirac who studied the case when $F = C_n$, i.e. in determining the smallest minimum degree of an n-vertex graph G which guarantees that G is Hamiltonian.

To start our investigation, let us try to think of some graphs with large minimum degree which do not have a Hamiltonian cycle. One immediate way to tell that a graph does not have a Hamilotonian cycle is if the graph is disconnected. In particular, if we consider G to be the n-vertex graph which is the disjoint union of $K_{\lceil n/2 \rceil}$ and $K_{\lfloor n/2 \rfloor}$, then this is a graph with no Hamiltonian cycle and with minimum degree $\lfloor n/2 \rfloor - 1$, showing that we must have $\delta(G) \ge \lfloor n/2 \rfloor$ to force a Hamiltonian cycle. While perhaps not as obvious, there exists another construction that gives a very similar bound which one might discover by looking at the cases of small n, for example. Specifically, any graph of the form $K_{m,n-m}$ with $m < \lceil n/2 \rceil$ will fail to be Hamiltonian. Indeed, if n is odd this is immediate because $K_{m,n-m}$ is bipartite and hence can not contain C_n. If n is even then any Hamilton cycle in such a graph must have exactly

$n/2 = \lceil n/2 \rceil$ of its vertices lying in each part of $K_{m,n-m}$, which is impossible to do under the condition $m < \lceil n/2 \rceil$. This construction thus implies that we need $\delta(G) \geq \lceil n/2 \rceil$ to force a Hamiltonian cycle, which matches the bound in the previous construction if n is even and does a little better if n is odd. In total it turns out that this bound is indeed the correct one.

Theorem 2.2 (Dirac's Theorem)**.** *Every n-vertex graph G with $\delta(G) \geq n/2$ contains a Hamiltonian cycle.*

Equivalently this says $\delta(G) \geq \lceil n/2 \rceil$ is enough to guarantee a Hamiltonian cycle, which is best possible by the constructions given above. Before we get on with the proof, let us make the meta-observation that for n even there are two extremal constructions for Dirac's Theorem (the disjoint union of two equally sized cliques, and a slightly unbalanced complete bipartite graph). This is non-ideal due to the following

Mantra 8. Extremal problems tend to be harder if they have more than one extremal constructions, especially if these constructions look very different from each other.

Indeed, part of the ease of proving Turán's Theorem is that there is only one possible extremal construction, which means we can hope to do arguments like Zykov symmeterization which move us closer to this unique extremal example. However, this approach as well as many others fail when there are multiple different looking extremal examples because whatever argument we make must simultaneously be optimal for all of our possible constructions.

To partially deal with this issue, we will utilize another mantra.

Mantra 9. If during a proof you assume that there exists some counterexample to your statement, it is sometimes useful to assume this counterexample is "extremal" in some sense.

We will see a concrete example of this in our following proof of Dirac's Theorem, which is originally due to Posa maybe.

Proof of Dirac's Theorem. Assume for some integer n that there exists a counterexample G and, crucially, choose such a counterexample with as many edges as possible. Intuitively by choosing a graph with more edges should make it easier for us to construct a Hamiltonian cycle, giving the desired contradiction. In particular, this assumption gives us the following key fact.

Claim 2.3. *The graph G contains a Hamiltonian path $x_1 \cdots x_n$.*

Proof. This is trivial if $G = K_n$, so assume this is not the case, i.e. that there exists some non-edge $xy \notin E(G)$. Because $G + xy$ is an n-vertex graph with $\delta(G + xy) \geq \delta(G) \geq n/2$ and with strictly more edges than G, it must be that $G + xy$ contains a Hamiltonian cycle C by assumption of G being a counterexample with the maximum number of edges. The subgraph $C - xy$ then must be a Hamiltonian path. □

The other key observation we will need is the following.

Claim 2.4. *If there exists an integer $2 \leq i \leq n$ such that $x_i \sim x_1$ and $x_{i-1} \sim x_n$, then G is Hamiltonian.*

Proof. Consider the following sequence of vertices:

$$P = (x_1, x_i, x_{i+1}, \ldots, x_{n-1}, x_n, x_{i-1}, x_{i-2}, \ldots, x_2).$$

It is not difficult to see that P is a Hamiltonian path (i.e. every vertex appears exactly once and consecutive vertices are adjacent) with its first and last vertices being adjacent to each other. Therefore this defines a Hamiltonian cycle in G, proving the claim. □

As an aside, the idea in this claim of "rotating" the Hamiltonian path we started with into a new one P is a common idea known as a Pósa rotation.

Back to our problem at hand, we want to show that an index i as in the claim exists. To this end, define

$$X_1 = \{i : x_i \sim x_1\},$$
$$X_n = \{i : x_{i-1} \sim x_1\}.$$

By the claim above and our assumption that G is not Hamiltonian, we can assume that X_1, X_n are disjoint subsets of $\{2, \ldots, n\}$. This implies that

$$n - 1 \geq |X_1 \cup X_n| = |X_1| + |X_n| = \deg(x_1) + \deg(x_n) \geq n,$$

a contradiction. □

Even though Dirac's Theorem is tight, it is possible to ask for further strengthenings as follows.

Mantra 10. After proving a theorem, check to see where you used the hypothesis of your theorem and if this can be relaxed in any way.

For example, the only place we used $\delta(G) \geq n/2$ in our proof was to show that $\deg(x_1) + \deg(x_n) \geq n$. A moment's thought then shows that our proof actually implies the following stronger result.

Theorem 2.5 (Ore's Theorem)**.** *If G is an n-vertex graph such that every non-edge $xy \notin E(G)$ has $\deg(x) + \deg(y) \geq n$, then G is Hamiltonian.*

In fact, our proof has even more flexibility that can be exploited to prove other extensions. We state one such extension here and leave its proof as an exercise to the reader.

Theorem 2.6 (Pósa's Theorem)**.** *If G is an n-vertex graph such that for all integers $k < n/2$,*

$$|\{x \in V(G) : \deg(x) \leq k\}| < k,$$

then G is Hamiltonian.

These extensions of Dirac's Theorem, in addition to being nice on their own, have various applications such as the following.

Theorem 2.7. *If G is an n-vertex graph with $\delta(G) \geq \frac{n+1}{2}$, then for every edge $xy \in E(G)$ there exists a Hamiltonian cycle in G which uses the edge xy.*

Proof. Let $xy \in E(G)$ be an arbitrary edge, and consider a new graph G' obtained by adding a new vertex v which is adjacent to only x, y. This $(n+1)$-vertex graph G' satisfies the conditions of Pósa's Theorem (it has only 1 vertex of degree at most 2, and every other vertex has degree at least $v(G')/2$), so G' contains a Hamiltonian cycle C. Note that this Hamiltonian cycle must contain the edges xv, vy since these are the only two neighbors of v. As such, the graph $C - v + xy$ is a Hamiltonian cycle in G using the edge xy, proving the result. □

2.2 Applications to Paths

Having just determined the optimal minimum degree needed to guarantee a graph contains a Hamiltonian cycle, it is natural to ask what conditions guarantee a Hamiltonian path. In fact, this turns out to be a consequence of Dirac's Theorem.

Theorem 2.8. *Every n-vertex graph G with $\delta(G) \geq \frac{n-1}{2}$ contains a Hamiltonian path.*

Note that this result is best possible by considering G to be the disjoint union of two cliques of sizes $\lfloor n/2 \rfloor, \lceil n/2 \rceil$.

Proof. Let G be an n-vertex graph with $\delta(G) \geq \frac{n-1}{2}$ and consider a new graph G' obtained by adding a vertex v which is adjacent to every vertex of G. Then $\delta(G') \geq (n+1)/2 = v(G')/2$, so by Dirac's Theorem G' contains a Hamiltonian cycle C, and thus $C - v$ is a Hamiltonian path in G. □

The trick we used in the proof above lets us easily translate many of the results that we have for Hamiltonian cycles to that of Hamiltonian paths; see the exercises for more.

We can also use Dirac's Theorem to prove good bounds for Turán numbers of (small) paths.

Theorem 2.9 (Erdős-Gallai)**.** *For all $r \geq 2$, we have $\mathrm{ex}(n, P_r) \leq \frac{r-2}{2} n$.*

Note that this bound is tight whenever $r - 1 | n$, as can be seen by considering G to be the disjoint union of K_{r-1}'s.

Proof. By prove the result by double induction on r and n. The result for all n is trivial when $r = 2$, so assume we have proven the result for all n up to some value r. This result in turn is trivial if $n \leq r - 1$, so we assume we have the proven the result up to some value $n \geq r$. With this in mind, let G be an extremal n-vertex P_r-free graph and assume for contradiction that $e(G) > \frac{r-2}{2} n$.

Because our extremal example looks like a disjoint union of K_{r-1}'s, a perhaps reasonable thing to try and prove is the following.

Claim 2.10. *The graph G contains a cycle C with $r - 1$ vertices.*

Proof. By Theorem 1.12, there exists a subgraph $G' \subseteq G$ with minimum degree at least $\frac{r-1}{2}$ (i.e. strictly more than $\frac{r-2}{2}$) and average degree strictly more than $r - 2$. By induction on r and the fact that G' has average degree more than $r - 2$, we conclude that G' must contain a path $x_1 \cdots x_{r-1}$.

Now all of the neighbors for x_1, x_{r-1} must lie within $\{x_1, \ldots, x_{r-1}\}$, as otherwise $G' \subseteq G$ would contain a path on r vertices. Because $\deg_{G'}(x_1), \deg_{G'}(x_{r-1}) \geq \frac{r-1}{2}$, the exact same argument that we used in the proof of Dirac's Theorem implies that there exists a cycle C using all of the vertices in $\{x_1, \ldots, x_{r-1}\}$. □

Observe that every vertex in C can only be adjacent to other vertices of C, as one could use any additional neighbor together with C to construct a P_r in G. As such, the number of edges incident to the vertices of C is at most $\binom{r-1}{2}$, and as such the graph $G - V(C)$ is a smaller order graph which has

$$e(G - V(C)) > \frac{r-2}{2}n - \binom{r-1}{2} = \frac{r-2}{2}(n - r + 1),$$

and since $G - V(C)$ has $n - r + 1$ vertices, we conclude by induction on n that $G - V(C)$ has a P_r, giving the result. □

2.3 Clique Factors

Perhaps after Hamiltonian cycles and paths, the next most natural spanning structure to consider is that of a perfect matching, i.e. a disjoint union of K_2's which cover every vertex of the graph exactly once. Note that perfect matchings can only exist if the number of vertices in our graph is even.

While a natural problem to consider, perfect matchings will turn out to not be very interesting to study for two reasons. First, any graph with an even number of vertices and a Hamiltonian cycle (or path) contains a perfect matching, so by Dirac's Theorem we know that $\delta(G) \geq n/2$ is enough to guarantee a perfect matching, and this is best possible by considering $K_{n/2-1,n/2+1}$. Second, one can in fact characterize *exactly* when a given graph has a perfect matching as we shall see in later section, so just proving a sufficient condition is not so interesting.

While the exact problem of determining minimum degree conditions for perfect matchings is not exciting, there are generalizations of perfect matchings which are more interesting. To this end, we say that a K_r*-matching* in a graph G is a subgraph of G which is the disjoint union of copies of K_r, and we say that G has a K_r*-factor* if G has a K_r-matching which contains every vertex of G exactly once. Note that G can only hope to have a K_r-factor if $r|n$.

Theorem 2.11 (Hajnal-Szemerédi Theorem Version I)**.** *If G is an n-vertex graph with $r|n$ and $\delta(G) \geq (r-1)n/r$, then G contains a K_r-factor.*

The Hajnal-Szemerédi Theorem is a deep result with a number of applications, see for example coloring chapter. The original proof of this result was very difficult. There does exist a quite short proof due to Kierstead and Kostochka, but it is a little too dense to present here I think that's the case; double check. Rather than spending time on proving this in full, we will instead sketch how to prove a somewhat weaker result.

Proposition 2.12. *If G is an n-vertex graph with $r|n$ and $\delta(G) \geq (r-1)n/r$, then G has a K_r-matching which contains all but at most $(r-1)^2 r$ vertices of G.*

Sketch of Proof. The rough idea is to consider a largest K_r-matching in G and argue that it has at least this size. However, to make the argument work we need to assume something slightly stronger about our matching.

To this end, let $S_1, \ldots, S_{n/r}$ be a partition of $V(G)$ into sets of size r such that $G[S_i]$ contains K_r for as many i as possible, and conditional on this, we choose this partition so that $G[S_i]$ contains a K_{r-1} for as many i as possible, and so on. Let $C_i \subseteq S_i$ denote a largest clique in $G[S_i]$ and assume for contradiction that $G[C_i] \neq K_r$ for at least $(r-1)^2 + 1$ values of i. By the Pigeonhole principle, this implies there is some $\ell \in [r-1]$ such that $|C_i| = \ell$ for at least r values of i, say for all $i \in [r]$ without loss of generality. Let $N(C_i)$ denote the set of common neighbors of C_i, i.e. the vertices adjacent to every vertex of C_i.

Claim 2.13. *We have $|N(C_i)| \geq (r-\ell)n/r$ and $N(C_i) \cap C_j = \emptyset$ for all $i, j \in [r]$.*

Proof. The lower bound $|N(C_i)| \geq (r-\ell)n/r$ follows from the fact that each of the ℓ vertices of C_i have minimum degree at least $(r-1)n/r$, i.e. are non-adjacent to at most n/r vertices. For the second part, assume for contradiction that there exists some $v \in N(C_i) \cap C_j$ and let $w \in S_j \setminus C_j$ be arbitrary (which exists since $|C_j| < r = |S_j|$). In this case, we could change our partition by replacing S_i, S_j with $S_i \cup \{v\} \setminus \{w\}$ and $S_j \setminus \{v\} \cup \{w\}$, which would increase the number of sets in the partition which contain a $K_{\ell+1}$ while not decreasing the number of sets containing any larger clique, contradicting how we chose our partition. We conclude that no such v exists. □

In total this claim implies $\sum_{i=1}^{r} |N(C_i) \cap \bigcup_{j>r} C_j| \geq (r-\ell)n$, which by the Pigeonhole principle implies there is some $j > r$ such that

$$\sum_{i=1}^{r} |N(C_i) \cap C_j| \geq \left\lceil \frac{(r-\ell)n}{n/r - r} \right\rceil \geq r(r-\ell) + 1.$$

Claim 2.14. *There exists some distinct $i', i'' \in [r]$ and disjoint $C_j', C_j'' \subseteq C_j$ of sizes 1 and $r-\ell$ such that $C_j' \subseteq N(C_{i'}) \cap C_j$ and $C_j'' \subseteq N(C_{i''}) \cap C_j$.*

Proof. By the inequality above and the Pigeonhole principle, there exists $i' \in [r]$ such that $|N(C_{i'}) \cap C_j| \geq r - \ell + 1$, and since $|N(C_{i'}) \cap C_j| \leq r$ we have

$$\sum_{i \in [r] \setminus \{i'\}} |N(C_i) \cap C_j| \geq r(r-\ell-1) + 1,$$

so again by the Pigeonhole principle there exists $i'' \neq i'$ such that $|N(C_{i''}) \cap C_j| \geq r - \ell$. Let $C_j'' \subseteq N(C_{i''}) \cap C_j$ be an arbitrary subset of size $r-\ell$ and let $C_j' \subseteq N(C_{i'}) \cap C_j$ be an arbitrary vertex disjoint from C_j', giving the result. □

Let $w \in S_{i'} \setminus C_{i'}$ be arbitrary. If we consider modifying the partition by replacing $S_{i'}, S_{i''}, S_j$ (whose largest cliques have sizes ℓ, ℓ, r) with the r-sets $S_{i'} \cup C_j' \setminus \{w\}$, $C_{i''} \cup C_j''$, and $S_j \cup \{w\} \setminus (C_j' \cup C_j'')$ (whose largest cliques have sizes at least $\ell+1, r, 1$), we see that this strictly increases the number of sets in our partition containing a $K_{\ell+1}$ while maintaining the sizes of all larger cliques, a contradiction to how we chose our partition. □

We emphasize that for many Dirac-type problems it is relatively easy to find an "almost spanning" subgraph like we did here, but finding a genuinely spanning structure is often difficult. One general tool for doing this is the absorption method which we probably won't talk about, but we'll see what happens.

2.4 Matchings

In Section 2 we saw sufficient conditions for G to contain a Hamiltonian cycle, and it is natural to ask if there exist nice necessary and sufficient conditions for Hamiltonicity. This turns out to be essentially hopeless. Indeed, it is known that the computational problem of determining whether or not a given graph is Hamiltonian is NP-complete, which means that if a "simple" necessary and sufficient condition existed then a large number of seemingly intractable problems for computer science would all have efficient algorithms. Similarly determining whether a graph has a Hamiltonian path is also NP-complete. However, there does exist a nice characterization for when a graph has a perfect matching.

Definition 5. Given a graph G, a *matching* M is a subgraph of G such that every vertex has degree 1, i.e. M is the disjoint union of some number of edges. The *matching number* $\nu(G)$ is the maximum number of edges in a matching of G. A *perfect matching* is a matching which is incident to every vertex of G. Note that a perfect matching can only exist if $v(G)$ is even.

I have been told you learned most of this in 6420 so I'll just give a very quick recap and prove Tutte unless I am told otherwise.

Matchings are particularly nice in bipartite graphs where we have the following two fundamental (and ultimately equivalent) theorems.

Theorem 2.15 (Kőnig's Theorem)**.** *Given a graph G, let $\tau(G)$ denote the smallest size of a set of vertices S which are incident to every edge of G. If G is bipartite, then*

$$\nu(G) = \tau(G).$$

The inequality $\nu(G) \le \tau(G)$ trivially holds for every graph G, so the difficult part is in proving $\nu(G) \ge \tau(G)$ for bipartite graphs. For this next result, we define for a set of vertices S its neighborhood $N(S) = \{x : \exists y \in S,\ xy \in E(G)\}$.

Theorem 2.16 (Hall's Theorem)**.** *Let G be a bipartite graph with bipartition $U \cup V$ with $|U| = |V|$. Then G has a perfect matching if and only if $|N(S)| \ge |S|$ for all $S \subseteq U$.*

The fact that $|N(S)| \ge |S|$ is necessary for a perfect matching is immediate, so again the difficulty lies in proving this is sufficient for bipartite graphs.

We now move on to prove a slightly less well-known result characterizing when *arbitrary* graphs G have a perfect matching by showing that an "obvious" necessary condition is also sufficient. To this end, given a graph G we let $\mathrm{odd}(G)$ denote the number of connected components of G which have an odd number of vertices.

Theorem 2.17 (Tutte's Theorem)**.** *A graph G has a perfect matching iff $\mathrm{odd}(G - S) \le |S|$ for all $S \subseteq V(G)$.*

Proof. The statement of this theorem as well as both directions of this proof will be motivated by the following observation.

Claim 2.18. *Let G be a graph and $S \subseteq V(G)$. If G has a perfect matching M, then for every connected component C in $G - S$ of odd order, there must exist an edge in M which is incident to a vertex of C and a vertex of S.*

Proof. Each vertex v of C must be contained in an edge $e_v \in M$ by definition of the matching being perfect. Since C has odd order, there must exist some vertex $v \in V(C)$ such that $|e_v \cap V(C)| = 1$. Because C is a component of $G - S$, any edge of G which is incident to exactly one vertex of C must also be incident to a vertex of S, proving the claim. □

For the first direction of the theorem, let G be a graph such that $\text{odd}(G - S) > |S|$ for some S and assume for contradiction that G contained a perfect matching M. By assumption there must exist some connected component C of odd order in $G - S$ which is not incident to any of the at most $|S| < \text{odd}(G - S)$ edges of M incident to S, a contradiction to the claim.

For the second (and harder) direction, assume for contradiction that there exists a graph G with $\text{odd}(G - S) \leq |S|$ for all $S \subseteq V(G)$ which does not have a perfect matching. From now on we fix such a graph with $v(G)$ as small as possible. We begin with a key observation which will guide us on how to construct a perfect matching in G.

Claim 2.19. *There exists some non-empty set $S \subseteq V(G)$ such that* $\text{odd}(G - S) = |S|$.

Proof. We will show that this holds for $S = \{v\}$ for any vertex v. Indeed, if G has an even number of vertices then $G - \{v\}$ necessarily has at least one connected component of odd order, proving $\text{odd}(G - \{v\}) \geq |\{v\}|$, and equality must hold by our hypothesis on G. If G has an odd number of vertices then we would have $\text{odd}(G) \geq 1 > |\emptyset|$, a contradiction to our choice of G. □

Crucially, we observe that for any set S as in Theorem 2.19, any perfect matching M that we wish to construct must have the property that each edge incident to S must also be incident to a distinct odd connected component of $G - S$ by Theorem 2.18. As such, when constructing M we have to somehow take into account all of the sets S of this form. Motivated by this, from now on we fix some $S \subseteq V(G)$ with $\text{odd}(G - S) = |S|$ and we choose such a set S with $|S|$ as large as possible. We begin by showing that the even components of $G - S$ are easy to deal with.

Claim 2.20. *If C is a connected component of $G - S$ with an even number of vertices, then C has a perfect matching.*

Proof. Because $v(C) \leq v(G) - |S| < v(G)$, we have by our choice of G being a minimal counterexample that C either has a perfect matching or there exists some $T \subseteq V(C)$ satisfying $\text{odd}(C - T) > |T|$. In this latter case we have

$$\text{odd}(G - (S \cup T)) = \text{odd}(G - S) + \text{odd}(C - T) > |S| + |T| = |S \cup T|,$$

a contradiction to our assumption on G. Thus it must be that C contains a perfect matching. □

Similarly the odd components are almost as easy to deal with.

Claim 2.21. *If C is a connected component of $G-S$ with an odd number of vertices, then for every $v \in V(C)$ the graph $C-v$ has a perfect matching.*

Proof. Again, if this failed to be true then there must exist some $T \subseteq V(C) \setminus \{v\}$ such that $\mathrm{odd}(C-T-v) \geq |T|+1$, but this means

$$\mathrm{odd}(G-(S \cup T \cup \{v\})) = \mathrm{odd}(G-S) + \mathrm{odd}(C-T-v) \geq |S|+|T|+1 = |S \cup T \cup \{v\}|.$$

By hypothesis on G this is only possible if $\mathrm{odd}(G-(S \cup T \cup \{v\})) = |S \cup T \cup \{v\}|$, but this contradicts are choice of S being the largest subset of G such that equality holds, a contradiction. □

From these observations, we can determine precisely what we need to show to prove that G has a perfect matching.

Claim 2.22. *Define an auxiliary bipartite graph B with one part being vertices of S, the other part the odd connected components of $G-S$, and where $v \sim C$ if and only if there exists an edge of G incident to both v and a vertex of C. If B contains a perfect matching then so does G.*

Proof. Say there existed such a perfect matching M'. For each edge $e' = \{v, C\} \in M'$, let $\tilde{e}$ be an edge in G which contains v and a vertex from C, which exists by definition of G, and let $\tilde{M} = \{\tilde{e} : e' \in M'\}$. For each even connected component C of $G-S$ we let M_C denote a perfect matching of C (which exists by Theorem 2.20), and for each odd connected component C of $G-C$ we let M_C denote a perfect matching of $C-\tilde{e}$ with $\tilde{e} \in \tilde{M}$ the unique edge incident to a vertex of C (and again such a perfect matching exists by Theorem 2.21). It is not difficult to check that $\tilde{M} \bigcup_C M_C$ is a perfect matching of G, proving the result. □

With this claim, we have reduced the problem of finding a perfect matching in an arbitrary graph to finding one in a *bipartite* graph, and as such it suffices for us to verify that the conditions of Hall's Theorem are satisfied[7]. And indeed, for any set $\mathcal{C}$ of odd connected components of $G-S$, if we had $N_B(\mathcal{C}) = T$ with $|T| < |\mathcal{C}|$, then this would imply that $\mathrm{odd}(G-T) = |\mathcal{C}| > |T|$, a contradiction to our condition on G. We conclude that B satisfies the condition of Hall's Theorem and hence has a perfect matching. This implies G has a perfect matching, a contradiction to us assuming no such matching existed. □

One can derive an analog of Kőnig's Theorem from Tutte's Theorem for the setting of arbitrary graphs as follows. This was originally done by Berge as a followup to Tutte's Theorem, hence the name of this result.

Theorem 2.23 (Tutte-Berge Formula)**.** *For any graph G we have*

$$\nu(G) = \min_{U \subseteq V(G)} \frac{1}{2}(|V(G)| - |U| + \mathrm{odd}(G-U))$$

[7]In fact, any proof of Tutte's Theorem must either use or give an alternative proof of Hall's Theorem since the statement of Tutte's Theorem is strictly stronger.

2.5 Exercises

1. Let's look at Turán numbers of spanning subgraphs.

   (a) Prove that $\text{ex}(n, C_n) = \binom{n-1}{2} + 1$ [2].

   (b) Prove that $\text{ex}(n, P_n) = \binom{n-1}{2}$ [2-].

2. We've seen that a minimum degree of about $n/2$ is the threshold for guaranteeing both a perfect matching and a Hamiltonian cycle. In the next few exercises we show that the behaviors for matchings and cycles differ greatly from one another when other sorts of degree conditions are imposed.

   (a) Prove that if G is an n-vertex graph with minimum degree $d \geq 2$, then G contains a cycle on at least $d+1$ vertices. Moreover, prove that for infinitely many n there exists an n-vertex graph which are $(d-1)$-regular and which have no cycle on at least $d+1$ vertices [1+].

   (b) Prove that if G is an n-vertex graph with minimum degree $d \geq 2$ and $d \leq n/2$, then G contains a matching on at least $2d$ vertices. Moreover, prove that this is best possible, i.e. that the result is false if we do not impose the condition $d \leq n/2$ and that there exist infinitely many $n \geq 2d$ with minimum degree $d-1$ with no matching on at least $2d$ vertices (Hint: the argument you use here can't be a direct analog of a proof of Dirac's Theorem since the previous part shows such an approach will fail for cycles) [2].

   (c) Prove that if G is an n-vertex graph with $d \geq 1$ and maximum degree Δ, then every maximal matching of G (i.e. every matching which is not a subset of any larger matching) has at least $\frac{d}{2\Delta} n$ vertices. Moreover, prove for all integers $1 \leq d \leq \Delta$ with d even that there exists a graph G with minimum degree d and maximum degree Δ which contains a matching on at most $\frac{d}{\Delta+1} v(G)$ vertices [2].

   (d) Prove that if G is an n-vertex graph with minimum degree $d \geq 1$ and maximum degree Δ, then G contains a matching on at least $\frac{d}{d+\Delta} n$ vertices. Moreover, prove for all integers $1 \leq d \leq \Delta$ that there exist graphs G with minimum degree d and maximum degree Δ such that no matching has size larger than $\frac{2d}{d+\Delta} v(G)$ (Hint: what would you need to assume about G for the same argument from (c) to give you the desired bound? Can you make this assumption here?) [2+].

3. The original proof of Dirac's theorem went as follows:

   (a) Define a *lollipop* to be a graph which consists of a cycle on vertices $v_1, \ldots, v_\ell$ together with a path on vertices $u_1, \ldots, u_t$ with $u_1 = v_1$. Given a graph G, consider its "largest" lollipop, i.e. the one which has ℓ as large as possible and conditional on this has t as large as possible.

Prove that if such a largest lollipop has $\ell \geq 3$ and $t \geq 2$, then u_t is not adjacent to any two consecutive vertices in $v_1, \ldots, v_\ell$. Similarly prove that if $\ell \geq 3$ then u_t is not adjacent to any v_i vertex which is "close" to v_1. In particular, prove this is true for v_ℓ, v_2, then generalize this as much as you can (Hint: use the previous exercise) [2].

(b) Conclude Dirac's Theorem [2]. .

4. Prove Pósa's Theorem [2].

5. Prove that if G is an n-vertex graph with $\delta(G) \geq (n+k)/2$ for some integer $k \geq 0$, then for any path $P \subseteq G$ on k edges there exists a Hamiltonian cycle of G which contains P as a subgraph (Hint: the trick we did before for $k = 1$ using Pósa's Theorem no longer works here, so you'll have to go back and modify our proof of Dirac's Theorem instead) [2].

6. Prove that if G is an n-vertex graph and $\delta(G) \geq n/2$, then for every edge of G there exists a Hamiltonian path of G containing this edge [1+].

* * *

7. Prove that every graph G has $\nu(G) \leq \tau(G)$ where $\nu(G)$ is the size of a largest matching in G and $\tau(G)$ is the smallest size of a set of vertices S which are incident to every edge of G [1+].

8. Let G be a bipartite graph with bipartition $U \cup V$ and $|U| = |V| = n$. Prove that G contains a matching with $n - d$ edges if and only if every $X \subseteq U$ satisfies

$$|N(X)| \geq |X| - d$$

[2] This is slightly wrong: I want the matching to just cover all of eg X..

9. Use Tutte's Theorem to prove that any n-vertex graph G with n even and $\delta(G) \geq n/2$ has a perfect matching; note that your proof must somewhere use that n is even [2].

10. Let G be a bipartite graph with bipartition $U \cup V$ and $|U| = n$ and $|V| = 2n$. We say that G has a $K_{1,2}$-factor if there exists a subgraph $G' \subseteq G$ which consists of the disjoint union of n copies of $K_{1,2}$ (so that every vertex of G is in exactly one of these copies). Prove that G has a $K_{1,2}$-factor if and only if $|N(X)| \geq 2|X|$ for all $X \subseteq U$ [2].

11. A d-regular graph G is said to have a 1-factorization if it contains d edge disjoint perfect matchings $M_1, \ldots, M_d$ (which means every edge of G is contained in exactly one such matching). Prove that every d-regular bipartite graph has a 1-factorization [2].

12. Prove that every bipartite graph has $\chi'(G) = \Delta(G)$, i.e. prove that for every bipartite graph of maximum degree Δ one can partition $E(G)$ into Δ edge-disjoint matchings [2+].

13. If G is a graph and $f : V(G) \to \mathbb{Z}_{\geq 0}$ is a function, we say that G has an f-factor if it contains a subgraph $G' \subseteq G$ with $\deg_{G'}(x) = f(x)$ for all $x \in V(G)$.

    Let G be a bipartite graph with bipartition $U \cup V$ and $f : V(G) \to \mathbb{Z}_{\geq 0}$. Prove that G has an f-factor if and only if $\sum_{x \in U} f(x) = \sum_{y \in V} f(y)$ and if for all $X \subseteq U$ and $Y \subseteq V$ we have

    $$\sum_{x \in X} f(x) \leq e(X, Y) + \sum_{y \in V \setminus Y} f(y),$$

    where to be clear $e(X, Y)$ denotes the number of edges with one vertex incident to X and the other incident to Y (Hint: if f is identically 1, how do your recover Hall's condition from this?) [2+].

3 Connectivity

TODO. Likely topics: k-connectivity, blocks, Menger's Theorem, ear decompositions

Possibly combine with flows as more examples of min-max theorems a la Konig

I might skip most of this since I was told you covered a lot of this in 6420.

4 Coloring Vertices and Chromatic Numbers

Recall that a *proper k-coloring* of a graph G is a map $c : V(G) \to [k]$ such that $c(u) \neq c(v)$ whenever $uv \in E(G)$. That is, we color each vertex using one of k colors such that no edge is monochromatic. We say that G is *k-colorable* if there exists a proper k-coloring of G, and we define the *chromatic number* $\chi(G)$ to be the smallest k such that G is k-colorable.

Colorings arise in various applied and theoretical contexts, and many problems in graph theory center around determining $\chi(G)$ for various graphs G. However, it is well known that determining whether $\chi(G) = k$ is an NP-hard problem for all $k \geq 3$, meaning one can not hope to find some "simple" way of determining if a graph has a given chromatic number. As such, the best one can realistically hope for in general is to establish reasonable bounds on $\chi(G)$ based on easy to compute parameters of G. We discuss two of the most fundamental bounds in the following sections.

4.1 Upper Bounds

Here and throughout this chapter we let $\Delta(G)$ denote the maximum degree of G, and whenever G is clear from context we will denote this quantity simply by Δ. Perhaps the most important bound for the chromatic number of a graph is the following.

Theorem 4.1. *If G is a graph then*

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

Proof. We define a "greedy" coloring $c : V(G) \to [\Delta + 1]$ as follows. Let $v_1, \ldots, v_n$ be an arbitrary ordering of the vertices of G. Iteratively given that we have defined $c(v_1), \ldots, c(v_{i-1})$ we choose $c(v_i)$ to be any element in $[\Delta + 1] \setminus \{c(v_j) : v_j \in N(v_i),\ j < i\}$; note that such an element must exist since $|N(v_i)| < \Delta + 1$.

We claim that c is a proper $(\Delta + 1)$-coloring. Indeed, if $v_i v_j \in E(G)$ with, say, $i > j$ then we chose $c(v_i)$ to be disjoint from $c(v_j)$. Thus c is a proper $(\Delta+1)$-coloring, proving the result. □

Theorem 4.1 is important in the field of coloring because it and its proof serves as the starting point for a number of other foundational results, several of which we discuss now.

An immediate question to ask upon seeing Theorem 4.1 is if this bound is tight. And indeed, one quickly sees that it is for $K_{\Delta+1}$ for all $\Delta \geq 1$, and for $\Delta = 2$ it is tight if and only if G contains a connected component which is an odd cycle. This turns out to exactly describe the cases of equality for Theorem 4.1.

Theorem 4.2 (Brooks's Theorem). *If G is a connected graph of maximum degree Δ and if G is not an odd cycle or $K_{\Delta+1}$ then $\chi(G) \leq \Delta$.*

Sketch of Proof. Essentially one can show that if G is as in the hypothesis, then there exists an ordering $v_1, \ldots, v_n$ of $V(G)$ such that (1) $v_1, v_2 \in N(v_n)$, (2) $v_1 \not\sim v_2$, and (3) $|\{v_j \in N(v_i) : j < i\}| < \Delta$ for all $i < n$. We now consider a greedy coloring $c : V(G) \to [\Delta]$ as we did before except (crucially) we set $c(v_1) = c(v_2)$ which will not create an improper coloring by (2). By (3), every vertex $v_i < n$ will have at least 1 choice when it is time to be colored, and by (1) the set $\{c(v_j) : v_j \in N(v_n)\}$ has at most $\Delta - 1$ used colors since $c(v_1) = c(v_2)$, meaning that we can also color v_n successfully. This gives a proper Δ-coloring of G as desired. □

While the maximum degree of a graph is a nice, clean parameter, it can often be entirely unrelated to $\chi(G)$ with perhaps the most egregious example of this being the star $K_{1,\Delta}$ which has chromatic number 2. Given this, one can ask if its possible to strengthen the bound of Theorem 4.1 by using some sort of "refinement" of the maximum degree Δ which, in particular, gives more reasonable bounds for stars. To this end and with Mantra 10 as motivation, we might ask ourselves what the best possible bound we could prove using the same argument as in Theorem 4.1, giving rise to the following parameter.

Definition 6. We define the *degeneracy* of a graph G to be the smallest integer d such that there exists an ordering $v_1, \ldots, v_n$ of $V(G)$ such that $|\{v_j \in N(v_i) : j < i\}| \leq d$ and we denote the degeneracy of G by $d(G)$.

With this definition the exact same proof of Theorem 4.1 gives the following.

Theorem 4.3. *If G is a graph, then*

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

Note that we always have $d(G) \leq \Delta$ via considering an arbitrary ordering of $V(G)$, so Theorem 4.1 is always at least as strong as Theorem 4.1. Moreover, it is an exercise to show that $d(G) = 1$ whenever G is a forest with at least 1 edge, meaning Theorem 4.3 is tight for all such graphs.

As an aside, the reader might feel that our definition of degeneracy is rather ad-hoc and specific only to the very particular proof we were trying to generalize. However, it turns out that degeneracy plays an important role in other areas such as Turán problems and that it has other (perhaps more natural) equivalent formulations. We touch on some of these connections in the exercises.

The last extension of Theorem 4.1 that we touch on asks if we can not only find some proper $(\Delta + 1)$-coloring but one which has some additional "nice" properties. This is perhaps natural to consider given that a closer look at our proof of Theorem 4.3 reveals that there is not just one proper $(\Delta + 1)$-coloring but in fact exponentially many, so we can perhaps be a bit more greedy with the sort of coloring we get at the end. There are various "nice" properties one could consider for colorings; the one we focus on will be the following.

Definition 7. We say that a proper k-coloring $c : V(G) \to [k]$ is *equitable* if $|c^{-1}(i)| \in \{\lfloor v(G)/k \rfloor, \lceil v(G)/k \rceil\}$ for all i. That is, each color is used as equal a number of times as possible.

Equitable colorings are a lot harder to come by compared to usual colorings. Indeed, the star $K_{1,\Delta}$ has exponentially many proper 3-colorings but none of them are equitable if $\Delta \geq 5$. However, it turns out that equitable colorings always exist at the threshold of $\Delta + 1$.

Theorem 4.4 (Hajnal-Szemerédi Version II)**.** *If G is a graph with maximum degree Δ then there exists an equitable proper $(\Delta + 1)$-coloring.*

This result turns out to be equivalent to our previous statement Theorem 2.11 of the Hajnal-Szemerédi Theorem, which is perhaps surprising at first glance but which is not too hard to prove; we leave this as an exercise. As before, we refrain from proving this result.

4.2 Lower Bounds and Perfect Graphs

For our lower bounds, we recall that $\alpha(G)$ denotes the largest size of an independent set of G and that $\omega(G)$ denotes the largest size of a clique of G.

Theorem 4.5. *For every graph G, we have*

$$\chi(G) \geq \omega(G),$$

and

$$\chi(G) \geq \frac{v(G)}{\alpha(G)}.$$

Proof. The first bound follows simply because the vertices making up the clique of size $\omega(G)$ of G must all be given colors that are distinct from each other. For the second bound, we observe that in any proper coloring $c : V(G) \to [k]$ that $c^{-i}(i)$ is an independent set of G for all i (otherwise c would have two adjacent vertices mapped to the same color i). In particular, we have

$$v(G) = \sum_{i=1}^{t} |c^{-1}(i)| \leq t\alpha(G),$$

and taking $t = \chi(G)$ gives the result.

□

Both of these bounds can easily seen to be tight for $G = K_n$. However, characterizing all cases of equality analogous to Brooks's Theorem seems difficult to do here. Indeed, $\chi(G) = v(G)/\alpha(G)$ holds if and only if $V(G)$ has a partition into maximum independent sets and offhand there does not seem to be a simple way to characterize this property. The case of $\chi(G) = \omega(G)$ is even more complex, as for any graph G' we can form a graph $G = G' \sqcup K_n$ with $n = v(G')$ and this trivially satisfies $\chi(G) = \omega(G)$ despite the structure of G being entirely arbitrary on half of its vertices. To avoid having to take into account silly constructions like these, we will want to shift to studying a certain class of graph families which are ubiquitous in structural graph theory.

Definition 8. We say that a family of graphs $\mathcal{G}$ is *hereditary* if it is closed under deleting vertices, that is, if for every $G \in \mathcal{G}$ we have $G - v \in \mathcal{G}$ for every vertex $v \in V(G)$. Equivalently, $\mathcal{G}$ is hereditary if for every graph $G \in \mathcal{G}$ all of the induced subgraphs of G are also in $\mathcal{G}$.

Many natural families of graphs are hereditary, such as those avoiding some graph F as either an induced or non-induced subgraph. Returning to our previous problem, we will now aim to characterize not the full family of graphs $\mathcal{G}$ with $\chi(G) = \omega(G)$ for all $G \in \mathcal{G}$ but simply the largest hereditary family of graphs $\mathcal{G}$ with this property. Equivalently, we aim to study the following type of graphs.

Definition 9. We say that a graph G is *perfect* if $\chi(G') = \omega(G')$ for every induced subgraph G' of G.

Again to be clear, the family of all perfect graphs is a hereditary family and it is the largest one satisfying $\chi(G) = \omega(G)$ for every graph in the family. Perfect graphs have a long history of study with this ultimately culminating in a full characterization of their structure.

Theorem 4.6 (Strong Perfect Graph Theorem)**.** *A graph G is perfect if and only if both G and its complement $\overline{G}$ do not contain an induced odd cycle of length at least 5.*

The fact that graphs must satisfy this property to be perfect is an exercise. The converse is tremendously difficult and was originally proven by Chudnovsky, Robertson, Seymour, and Thomas in 2006.

Do a proof of weak perfect graph theorem.

4.3 Coloring Variants

Here we look at some variants of the notion of proper colorings.

4.3.1 List Colorings

It is very common in coloring arguments to construct some proper k-coloring $c : V(G) \to [k]$ by inductively defining $c(v)$ for some vertex v and then constructing a coloring of $G - v$. However, when we do this we are no longer exactly looking for a proper k-coloring of $G - v$ but rather a coloring where each $u \notin N(v)$ is allowed to be any color in $[k]$ while $u \in N(v)$ are required to be colored from the set $[k] \setminus \{c(v)\}$, and because of this we can't directly apply any inductive statement that holds for proper k-colorings. The solution to this problem is to consider a more general notion of coloring which is preserved by us iteratively coloring a vertex of our graph. Specifically, we do this by assigning each vertex a list of "allowed colors" $L(v)$ which we can think of as being the subset of $[k]$ obtained after removing any of the colors from vertices we've already deleted from G in some sort of inductive step. More precisely, we have the following.

Definition 10. Given a graph G, a *list assignment* is a function L which assigns to each $v \in V(G)$ a set $L(v)$. A *proper L-coloring* is a map c from $V(G)$ which satisfies $c(v) \in L(v)$ for all $v \in V(G)$ and which has $c(u) \neq c(v)$ for all u, v with $uv \in E(G)$. We say that G is *k-choosable* if there exists a proper L-coloring for G for all L with $|L(v)| \geq k$ and we define the *list chromatic number* $\chi_\ell(G)$ to be the smallest k such that G is k-choosable.

As an example, observe that G has a proper k-coloring if and only if it has a proper L-coloring with $L(v) = [k]$ for all v. As such, G being k-choosable implise that it is k-colorable and hence

$\chi(G) \leq \chi_\ell(G)$ for every graph G. As such, the following is a direct strengthening of the results of the previous subsection.

Theorem 4.7. *If G is a graph of maximum degree Δ then*

$$\chi_\ell(G) \leq d(G) + 1 \leq \Delta + 1.$$

The proof of this is essentially identical to our previous arguments and we leave the details as an exercise to the reader.

While Theorem 4.7 is certainly at least as strong as our results upper bounding $\chi(G)$, it is not clear if this is a strict strengthening. That is, it is not clear whether there exists any graph with $\chi(G) \neq \chi_\ell(G)$. And indeed, intuitively it doesn't feel like this should be the case. That is, finding a proper L-coloring seems hardest to do when the lists $L(v)$ overlap as much as possible since otherwise it seems easier for us to avoid creating monochromatic edges. As such, it naively seems like the worst-case scenario for L is if $L(v) = [k]$ for all v which exactly recovers the notion of a proper k-coloring.

Perhaps surprisingly (or unsurprising given we've dedicated a whole subsubsection to this topic), there do in fact exist L which are strictly harder to properly color compared to the identically $[k]$ assignment, implying that $\chi(G) < \chi_\ell(G)$ for such graphs. Genuinely surprisingly, this holds even for bipartite graphs where $\chi(G)$ and $\chi_\ell(G)$ can be made arbitrarily far apart from each other.

Theorem 4.8. *For every integer $t \geq 2$, there exists a graph G with $\chi(G) = 2$ and $\chi_\ell(G) \geq t$.*

Sketch of Proof. We only prove this for $t = 3$ with the generalization of this argument being left as an exercise to the reader. For this, take $G = K_{3,3}$ say with bipartition $U = \{u_1, u_2, u_3\}$ and $V = \{v_1, v_2, v_3\}$. Define L by having $L(u_i) = L(v_i) = \{1, 2, 3\} \setminus \{i\}$. Observe now that if there exists a proper L-coloring c then $\{c(u_1), c(u_2), c(u_3)\}$ contains at least 2 colors since if this only contained one color i then this would contradict $c(u_i) \in L(u_i) = \{1, 2, 3\} \setminus \{i\}$. But this set containing at least two colors implies that $L(v_i) \subseteq \{c(u_1), c(u_2), c(u_3)\}$ for some i, namely the one whose color set $L(v_i)$ equals these two colors. This means that for any choice of $c(v_i) \in L(v_i)$ that there will exist some $u_j \in N(v_i)$ with $c(u_j) = c(v_i)$, contradicting this color being proper. We conclude that no proper L-coloring can exist for this choice of L, implying that $\chi_\ell(K_{3,3}) > 2$. □

Corollary 4.9. *There does not exist a function $f : \mathbb{N} \to \mathbb{N}$ such that $\chi_\ell(G) \leq f(\chi(G))$ for all graphs G. That is, the list-chromatic number can not be bounded by some function of the chromatic number.*

As a final remark, we note that in very recent years an even greater generalization of list coloring has appeared in the literature known alternatively as *correspondence coloring* or *DP-coloring* which in particular originated in the context of inductive proofs similar to our motivation for studying list colorings. Maybe say more about this at some point.

4.3.2 Fractional Relaxations and Fractional Colorings

Need to adjust the exposition given its new placement, see comments below for potential inspirational

Many of the max-min theorems we have discussed up to this point fit within the framework of linear or fractional relaxations of integer programs.

Broadly speaking, an integer program is a problem which aims to either maximize or minimize a linear function of the form $\sum c_i x_i$ where each x_i is an integer-valued variable satisfying some set of linear inequalities. As a very basic example, consider the integer program (I) which has variables $x_1, x_2 \in \mathbb{Z}$ defined by

$$\begin{aligned} \text{maximize } & x_1 + x_2 \\ \text{subject to } & x_1 \leq 1.5, \\ & x_2 \leq 1.5. \end{aligned}$$

In this example, it is not difficult to see that the optimal value of this integer program is 2 obtained by taking $x_1, x_2 = 1$.

For a more interesting example, for a given graph G consider the integer program (M) which for each $e \in E(G)$ has a variable $x_e \in \mathbb{Z}$ and which is defined by

$$\begin{aligned} \text{maximize } & \sum x_e \\ \text{subject to } & \forall e \in E(G),\ 0 \leq x_e \leq 1 \\ & \forall v \in V(G),\ \sum_{e : v \in e} x_e \leq 1. \end{aligned}$$

In this situation, because we demand $x_e \in \mathbb{Z}$ the first constraint is equivalent to saying $x_e \in \{0, 1\}$, and one can check that the second constraint makes it so that the set of e with $x_e = 1$ form a matching of G. As such, the optimum value of this integer program is simply the matching number $\mu(G)$.

In many cases it is useful to consider its *linear relaxation* or *fractional relaxation* of an integer program, which is defined by relaxing the constraint that each variable lies in $\mathbb{Z}$ to just being in $\mathbb{R}$. The advantage of this is that linear relaxations can often be solved using efficient algorithms (while integer programs famously can not), which in turn often yields bounds on the original integer program we care about.

For example, the linear relaxataion of (I) above has variables $x_1, x_2 \in \mathbb{R}$ and is defined by

$$\begin{aligned} \text{maximize } & x_1 + x_2 \\ \text{subject to } & x_1 \leq 1.5, \\ & x_2 \leq 1.5. \end{aligned}$$

In this case the optimum value is 3 and is obtained by taking $x_1, x_2 = 1.5$; note how this is strictly larger than the previous optimum value of 2 when x_1, x_2 were required to be integers.

Similarly the optimum value of the linear relaxation of (M) is a quantity $\mu^*(G)$ called the fractional matching number of G which can be larger than the usual matching number. For example, $\mu^*(K_3) = 1.5$ with the lower bound coming from taking $x_e = 1/2$ for all e.

In the two examples above the linear relaxation has a different optimum value than the original integer program, but this is not always the case. Indeed, an equivalent statement of König's Theorem is that $\mu^*(G) = \mu(G)$ whenever G is bipartite. Similarly, given a network N one can construct an integer program whose optimum value equals the maximum flow obtained from an integral flow, and the Integral Max-Flow Min-Cut Theorem implies that this optimum value equals the optimum value of its linear relaxation provided the capacity function is integral valued.

Let us focus now on the particular case of the chromatic number of G. Naively we might think of defining an integer program whose variables correspond to vertices of G and where a variable is given value i if the vertex is colored i, but this does not fundamentally work because i is really just a symbol and does not hold any meaning as an integer. After some more thought one might realize that an equivalent way to define a coloring is as a partition of G into independent sets, and this is the approach we will take. To this end, given a graph G we define an integer program (C) with variables $x_I \in \mathbb{Z}$ for each independent set I under the conditions

$$\begin{aligned} \text{minimize } & \sum x_I \\ \text{subject to } & \forall I,\ 0 \le x_I \le 1, \\ & \forall v \in V(G),\ \sum_{I: v \in I} x_I = 1. \end{aligned}$$

That is, we want to find the smallest number of independent sets with the property that every vertex is in exactly one such independent set. We can then define the *fractional chromatic number* $\chi^*(G)$ (also sometimes denoted $\chi_f(G)$) to be the optimum value of the linear relaxation of this integer program. Insert picture of C_5 with a 2-fold coloring.

There are a number of different and equivalent ways one can think about the fractional chromatic number. One intuitive way is to think of coloring each vertex but now we can color a vertex e.g. half red and half blue (which is represented by having some independent sets I, I' containing v with $x_I, x_{I'} = \frac{1}{2}$). This can be made more precise by considering something known as the k-fold chromatic number $\chi_k(G)$, which is defined to be the smallest integer n such that one can assign to each vertex of G a set in $\binom{[n]}{k}$ such that adjacent vertices are given disjoint sets. For example, $\chi_1(G) = \chi(G)$. It is not difficult to prove that $\chi^*(G) \ge \chi_k(G)/k$ for all k, and in fact it turns out that

$$\chi^*(G) = \inf_k \frac{\chi_k(G)}{k}.$$

Yet another way to view these colorings is through homomorphisms. Indeed, let $K_{n;k}$ denote the Kneser graph which is defined to have vertex set $\binom{[n]}{k}$ where two sets are adjacent if and only if they are disjoint (noting that such graphs appeared in our linear algebra proof of the Erdős-Ko-Rado Theorem). It is not difficult to see that $\chi_k(G) \le n$ if and only if G has a homomorphism to $K_{n;k}$, and as such $\chi^*(G)$ can again be defined as $\inf n/k$ where the infimum ranges over all Kneser graphs $K_{n;k}$ which G has a homomorphism to.

By definition we have $\chi^*(G) \le \chi(G)$. It is thus natural to ask which lower bounds for $\chi(G)$ continue to hold for the potentially smaller quantity $\chi^*(G)$. One example is the following classic bound.

Proposition 4.10. *For every graph G we have*

$$\chi^*(G) \ge \frac{v(G)}{\alpha(G)}.$$

Proof. Let x_I be a set of variables satisfying the linear relaxation of the chromatic number integer program. Because of the constraints of the program, we have

$$v(G) = \sum_v \sum_{I:v\in I} x_I = \sum_I |I| x_I \le \sum_I \alpha(G) x_I.$$

We conclude that if x_I satisfies the conditions of the program then $\sum x_I \ge v(G)/\alpha(G)$, meaning this must also hold for $\chi^*(G)$ which is the infimum of all such summations. □

This bound has a number of nice consequences. For example, one can prove that $\chi^*(G) = v(G)/\alpha(G)$ whenever G is vertex-transitive, which is the case for many nice families of graphs. In particular, $\chi^*(K_n) = n$ which in turn implies that $\chi^*(G) \ge \omega(G)$ since at least $\omega(G)$ "fractional colors" are needed to color all the vertices of a largest clique. Yet another consequence is that $\chi^*(G)$ and $\chi(G)$ can be arbitrarily far apart from each other with Kneser graphs being one such example (though it takes quite a bit of work to determine precisely what $\chi(K_{n;k})$ is).

4.3.3 Edge Colorings

All of the colorings we have considered up to this point involve colorings of the vertices of G. What if we were to consider colorings of its edges instead? While a natural idea, it is not immediately clear what a "proper" edge coloring should be. Motivated by the idea that a vertex coloring is proper if no two vertices which share an edge in common are given the same color, we might consider edge colorings to be proper if no two edges which share a vertex in common are given the same color.

Definition 11. Given a graph G, we say that a function $c' : E(G) \to [k]$ is a *proper k-edge coloring* if $c'(e) \ne c'(f)$ for any distinct edges e, f with $e \cap f \ne \emptyset$. We define the *chromatic index* $\chi'(G)$ to be the smallest integer k such that G has a proper k-edge coloring.

Proper edge colorings of a graph G are in fact equivalent to proper vertex colorings of an appropriate auxiliary graph of G.

Lemma 4.11. *Given a graph G, define the line graph $L(G)$ to be the graph with vertex set $E(G)$ where two distinct edges e, f are adjacent to each other in $L(G)$ if $e \cap f \ne \emptyset$. A function $c' : E(G) \to [k]$ is a proper k-edge coloring of G if and only if it is a proper k-coloring of $L(G)$.*

We can use this connection to proper colorings to immediately conclude some very strong bounds on $\chi'(G)$.

Proposition 4.12. *If G is a graph with maximum degree Δ, then*

$$\Delta \leq \chi'(G) \leq 2\Delta - 1.$$

Proof. Indeed, observe that $\omega(L(G)) \geq \Delta$ as the Δ edges incident to a vertex of maximum degree in G form a clique in the line graph $L(G)$. On the other hand, the maximum degree of $L(G)$ is at most $2\Delta - 2$ as every edge uv in G is incident to at most $2\Delta - 2$ edges other than uv itself (since each of u, v are incident to at most $\Delta - 1$ other edges respectively). The bounds now follow immediately from Theorem 4.1. □

One can use Brooks's Theorem to improve the upper bound of this proposition by 1 for $\Delta \geq 3$, but we choose not to do so here since a substantially stronger bound holds.

Theorem 4.13 (Vizing's Theorem)**.** *If G is a graph with maximum degree Δ then $\chi'(G) \in \{\Delta, \Delta + 1\}$.*

We omit the proof due to Guantao literally teaching a full course on edge colorings right now.

Despite Vizing's Theorem determining χ' up to an additive error of 1 for every graph G there is still a lot that can be said about edge colorings especially in the context of multigraphs, though we will not go into this further here.

4.4 Clique Numbers and Chromatic Numbers

A major theme of structural graph theory is to determine when a given parameter of a graph G can be bounded by a function of another parameter. For example, we saw that $\chi_\ell(G)$ can not be upper bounded by a function of its natural lower bound $\chi(G)$ while $\chi'(G)$ can be very strongly upper bounded by its natural lower bound $\Delta(G)$. For $\chi(G)$, the natural question to ask in view of Theorem 4.5 is whether $\chi(G)$ is upper bounded by a function of its clique number $\omega(G)$. As a first step, we need to figure out if an analog of Theorem 4.8 holds in our setting.

Question 4.14. *Is it true that for every t there exists a graph G with $\omega(G) = 2$ but $\chi(G) \geq t$?*

That is, do there exist triangle-free graphs with arbitraily large chromatic numbers? The answer to this question is immediately yes for $t = 3$ by considering odd cycles. One can also verify it for $t = 4$, though it likely will take you either a lot of trial and error or a computer (as the smallest such example is on 11 vertices), and these are approaches which will not generalize to, say, $t = 1000$. The difficulty in finding these constructions should suggest that either this is false for large t or that we need a more systematic scheme for forming our constructions. And indeed, we will in fact show that this question has a positive answer by coming up with a systematic way for constructing examples.

The motivation for our approach is as follows. Say we have some triangle-free graph G with chromatic number at least t, we want to build from this a new graph $M(G)$ which is triangle-free and which has chromatic number at least $t+1$. The simplest way to force chromatic number at least $t + 1$ is to add a new vertex w to G which is adjacent to all of $V(G)$ since the new vertex is forced to be given a coloring distinct from the t which we know must be used for G, but this approach completely fails to maintain that our graph is triangle-free. To get around this, for each $u_i \in V(G)$ we will create a new "duplicate" vertex v_i in such a way that we essentially force the color of v_i to be the same as the color of u_i and such that these duplicate vertices v_i form an independent set. If we can achieve this, then by adding a new vertex w adjacent to all of the duplicate vertices will achieve our desired goal. After pondering on this idea for a bit one might be led to the following operation.

Definition 12. Given a graph G with vertices $u_1, \ldots, u_n$, its *Mycielskian* $M(G)$ is a graph with vertex set $u_1, \ldots, u_n, v_1, \ldots, v_n, w$ such that:

- $u_i u_j \in E(M(G))$ and $u_i v_j \in E(M(G))$ if and only if $u_i u_j \in E(G)$,
- $v_i w \in E(M(G))$ for all i, and
- $u_i w \notin E(M(G))$ and $v_i v_j \notin E(M(G))$ for all i, j.

Insert picture of $G = K_2$ and also maybe $G = C_5$..

That is, $M(G)$ is formed by taking G, duplicating each vertex so that v_i has the same set of neighbors as u_i in G, and then adding a new vertex w adjacent to all the duplicated vertices. Crucially, this operation does precisely what we want it to do.

Proposition 4.15. *For every graph G, $\chi(M(G)) = \chi(G) + 1$ and $M(G)$ is triangle-free whenever G is triangle-free.*

Proof. For triangle-freeness, we observe that no triangle in $M(G)$ can involve two v_i vertices since such vertices are never adjacent, and as such no triangle can involve w whose only neighbors are v_i vertices. As such, if there is a triangle it must either be of the form u_i, u_j, u_k or v_i, u_j, u_k, but such vertices form a triangle in $M(G)$ if and only if u_i, u_j, u_k form a triangle in G, proving this half of the result.

For ease of notation let $t = \chi(G)$. To prove $\chi(M(G)) \leq t + 1$ we construct an explicit proper $(t+1)$-coloring for $M(G)$ as follows. Start with an arbitrary proper t-coloring c' of G. Now define $c : V(M(G)) \to [t+1]$ by having $c(u_i) = c(v_i) = c'(u_i)$ and $c(w) = t + 1$. That is, we duplicate the coloring of c' on both the u and v vertices and then give w a completely new color. Any edge involving w will be monochromatic because w is the only vertex with color $c(w)$. One can also check that if some edge $u_i u_j$ or $v_i v_j$ were monochromatic under c then the edge $u_i u_j$ would be monochromatic under c' which we assumed not to be the case. This shows c is a proper coloring, proving the bound.

We now prove that $\chi(M(G)) \geq t+1$, and for this we assume for contradiction that there exists some proper t-coloring c of $M(G)$.

Claim 4.16. *For every color $s \in [t]$, there exists some u_i with $\{c(u_j) : u_j \in N_G(u_i)\} = [t] \setminus \{s\}$.*

Proof. Assume this was false for some s, we aim to use this to contradict that G has chromatic number t. To this end, define a coloring $c' : V(G) \to [t] \setminus \{s\}$ by having $c'(u_i) = c(u_i)$ whenever $c(u_i) \neq s$ and otherwise take $c'(u_i)$ to be an arbitrary color in $[t] \setminus (\{s\} \cup \{c(u_j) : u_j \in N_G(u_i)\})$, noting that such a color exists by hypothesis. We claim that this is a proper coloring. Indeed, the only way an edge u_iu_j can be monochromatic under c' is if, say, $c(u_i) = s$, but in this case we must have $c(u_j) \neq s$ since c is proper coloring and hence $c'(u_j) = c(u_j) \neq c'(u_i)$ by construction. We have thus shown that G can be properly colored using only $t-1$ colors, contradicting $\chi(G) = t$. □

With this claim we see that $\{c(v_1), \ldots, c(v_n)\} = [t]$ since for each u_i as in the claim we must have $c(v_i) = s$. But this means $c(w)$ will equal the color of one of its neighbors v_i, a contradiction to c being a proper coloring. □

Corollary 4.17. *For all $t \geq 2$ there exists a triangle-free graph with chromatic number t.*

Proof. Take $G_2 = K_2$ and iteratively define $G_{i+1} = M(G_i)$. The proposition immediately implies that G_t satisfies the conditions of the corollary. □

A natural followup now is to ask to what extent we can strengthen this result. For example, what if our graph is both C_3-free and C_5-free (the two smallest certificates for whether a graph has chromatic number 2 or not), can we find graphs of arbitrarily large chromatic number in this case? Note that the Mycielskian $M(G)$ will be ineffective for this problem since any edge in G creates a C_5 in $M(G)$. It is natural then to go back to our motivation for $M(G)$ and see if one can modify it to get rid of C_5's as well, but I do not know of any way to make this work.

Ultimately it turns out that there do exist explicit constructions of graphs with no C_3, C_4, or C_5 which have arbitrarily large chromatic numbers due to Tutte. However, these graphs are tremendously large and as far as we know these particular constructions do not generalize to the next natural followup question of asking if there exist graphs with large chromatic number which avoid all of C_3, C_5, and C_7. Ultimately, this problem does indeed have a positive answer in a very strong sense. To this end, we recall that the *girth* of a graph is the length of its shortest cycle.

Theorem 4.18 (Erdős-Hajnal)**.** *For all integers $\ell, t \geq 2$ there exists a graph G with girth at least ℓ and $\chi(G) \geq t$.*

This result says in a very strong sense that $\chi(G)$ is a "global" parameter of G, in the sense that it implies there exist graphs which G locally look like a tree (in the sense that G restricted to the vertices within distance $g/2$ of a given vertex is a tree) but nevertheless needs an arbitrarily large number of colors to actually color the whole graph.

There is no known family of "elementary" graphs[8] which satisfies Theorem 4.18. However, it is relatively easy to give a non-constructive proof of this result through the probabilistic method, and as such we will postpone the proof of this result until Section 6.

At this point we have more than proved that one can not upper bound $\chi(G)$ by a function of $\omega(G)$ for arbitrary graphs G, but what about if we turn our attention away from all graphs and restrict to some nice family of graphs instead?

[8]There do exist explicit constructions due to Lubotzky, Phillips, and Sarnak, but these are highly complicated and rely on quite a bit of algebra and number theory.

Definition 13. We say that a family of graphs $\mathcal{G}$ is χ*-bounded* if there exists a function $f : \mathbb{N} \to \mathbb{N}$ such that $\chi(G) \leq f(\omega(G))$ for all $G \in \mathcal{G}$.

For example, Theorem 4.18 says that the family of all graphs is not χ-bounded. On the other hand, the family of perfect graphs by definition are χ-bounded with $f(n) = n$, and Vizing's Theorem implies that the family of line graphs is χ-bounded with respect to $f(n) = n + 1$. There are many open questions regarding which families of graphs are χ-bounded as well as determining optimal values for the function f with the biggest open question being the following.

Conjecture 4.19 (Gyárfás-Sumner)**.** *For every tree T, the family of graphs $\mathcal{G}_T$ which do not contain an induced copy of T is χ-bounded.*

This conjecture was originally made by Gyárfás in 1975 (and independently by Sumner later) and despite receiving a lot of attention, the only trees we know of for which this problem is solved is when T is a star, path, or has radius 2.

4.5 Exercises

1. We begin with some warmups.

   (a) Recall that a map $\phi : V(G) \to V(H)$ is a homomomorphism if $\{\phi(u), \phi(v)\} \in E(H)$ whenever $\{u, v\} \in E(G)$. Prove that a graph G has a proper t-coloring if and only if there exists a homomorphism $\phi : V(G) \to K_t$ [1].

   (b) Prove that if $c : V(G) \to [t]$ is a proper coloring then $c^{-1}(i)$ is an independent set of G. Prove that if $c' : E(G) \to [t]$ is a proper edge coloring then $(c')^{-1}(i)$ is a matching of G [1].

$* * *$

2. Prove that the number of proper t-colorings of a graph G is at least $\prod_{v \in V(G)}(t - \deg(v))$. In particular, every n-vertex graph G with maximum degree Δ has at least 2^n proper $(\Delta + 2)$-colorings [1+].

3. Prove that there exists some $C > 1$ such that every n-vertex graph G with maximum degree Δ has at least C^n proper $(\Delta+1)$-colorings. That is, G does not just have 1 proper $(\Delta + 1)$-coloring but at least exponentially such colorings (Hint: the proof we have in mind yields that there are at least $2^{\frac{\Delta}{\Delta+1}n}$ such colorings, in particular giving the result with $C = 2^{2/3}$) [2+].

4. Prove that one can equivalently define the degeneracy $d(G)$ of a graph G to be the smallest integer d such that every subgraph $G' \subseteq G$ has $\delta(G') \leq d$, i.e. has a vertex of degree at most d [2].

5. Prove that $e(G) \leq d(G)v(G)$ for all graphs G. In particular, graphs with bounded degeneracy have at most a linear number of edges [1+].

6. Let us consider the degeneracy of various types of graphs.

   (a) Prove that a graph G has $d(G) = 1$ if and only if G is a forest with at least one edge [1+].

   (b) Prove that if G is a graph with maximum degree Δ then $d(G) = \Delta$ if and only if G is regular. In particular, note that this gives an easy proof of Brooks's Theorem for graphs G which are not regular [1+].

   (c) Prove that if G is a planar graph then $d(G) \le 5$. In particular, note that this implies $\chi(G) \le 6$ for planar graphs (Hint: you may assume without proof the fact that planar graphs have $e(G) \le 3v(G) - 6$ provided $v(G) \ge 3$) [1+].

7. Here we briefly showcase how degeneracy appears in other graph theoretic contexts.

   (a) Prove that for all d there exists a graph F with $d(F) = d$ such that $\mathrm{ex}(n, F) = \Theta(n^{2-1/d})$ (Hint: you may assume without proof anything I claimed in the chapter on Turán Problems) [2-].

   (b) Prove that if F is a graph with $d(F) = d$, then $\mathrm{ex}(n, F) = O(n^{2-\frac{1}{4d}})$; a major open problem of Erdős conjectures that in fact $\mathrm{ex}(n, F) = O(n^{2-\frac{1}{d}})$ should hold, which is best possible by the previous part [3+].

   (c) Recall that $R_2(F)$ denotes the smallest number N such that any red-blue edge coloring of K_N contains a monochromatic copy of F. Prove that if $d(F) = d$ then $R_2(F) = O_d(v(F))$ [4-].

8. Prove that our two stated versions of the Hajnal-Szemerédi Theorem (Theorem 2.11 and Theorem 4.4) are equivalent to each other [2-].

* * *

9. Prove that if a graph G is perfect, then G and $\overline{G}$ do not contain an induced odd cycle of length at least 5 [2].

10. Determine which of the following families of graphs are hereditary: all graphs, regular graphs, planar graphs, trees, forests [1].

11. Prove that for every hereditary family $\mathcal{G}$ that there exists a family of graphs $\mathcal{F}$ such that $G \in \mathcal{G}$ if and only if G does not contain any graph of $\mathcal{F}$ as an induced subgraph. Prove that there exists a hereditary family $\mathcal{G}$ such that the family $\mathcal{F}$ can not be taken to be finite [2-].

* * *

12. Formally prove that $\chi_\ell(G) \le d(G) + 1$ for every graph G [1].

13. We consider the list chromatic number of complete bipartite graphs.

    (a) Complete our proof of Theorem 4.8 by showing that for all t there exists some n_t such that $\chi_\ell(K_{n_t,n_t}) \geq t$ [2].

    (b) Give an alternative proof by showing that $\chi_\ell(K_{t-1,(t-1)^{t-1}}) \geq t$ [2].

14. Prove that if G is a regular graph, then its fractional matching number satisfies $\mu^*(G) = v(G)/2$ [2-].

15. Write an integer program whose optimum value is $\mathrm{ex}(n, K_3)$ and prove that the optimum value of its linear relaxation is asymptotically larger than $n^2/4$ [2].

16. Prove that $\chi^*(G) \geq \chi_k(G)/k$ for all k [1+].

17. Prove that if G is vertex-transitive (meaning for every $u, v \in V(G)$ there exists an isomorphism $\phi : V(G) \to V(G)$ with $\phi(u) = v$), then $\chi^*(G) = v(G)/\alpha(G)$ [2].

18. Use Brooks's Theorem to prove that $\chi'(G) \leq 2\Delta - 2$ for every graph G with maximum degree $\Delta \geq 3$ [1+].

* * *

19. Prove that every K_3-free graph on at most 10 vertices has chromatic number at most 3, meaning that $M(C_5)$ is the smallest triangle-free graph with chromatic number 4. Your proof should be human readable and checkable, i.e. it can not be of the form "I generated every graph on at most 10 vertices on my computer and verified that this is true" [3-].

20. Given a graph G and an integer $k \geq 1$, we define the Zykov graph $Z(G, k)$ by taking k disjoint copies $G_1, \ldots, G_k$ of G, and then for each of the $v(G)^k$ sequences $\vec{x} = (x_1, \ldots, x_k) \in V(G_1) \times \cdots \times V(G_k)$ we add a new vertex $v_{\vec{x}}$ whose neighborhood equals $\{x_1, \ldots, x_k\}$.

    Prove that $Z(G, k)$ is triangle-free whenever G is, and that $\chi(Z(G, k)) = \chi(G)+1$ provided $k \geq \chi(G)$. As such, these graphs give another explicit family of triangle-free graphs with arbitrarily large chromatic numbers [2].

21. Prove Markov's inequality whenever X is a non-negative discrete random variable [1+].

22. In this exercise we will partially motivate the exact statement of the Gyárfás-Sumner conjecture. To this end, for each graph F let $\mathcal{G}_F$ be the family of graphs which does not contain F as an induced subgraph.

    (a) Prove that if F contains a cycle then $\mathcal{G}_F$ is not χ-bounded [2-].

    (b) Prove that if F_1, F_2 are forests then $F_1 \sqcup F_2$ is χ-bounded if and only if F_1 and F_2 are both χ-bounded (Hint: inductively define $f(1)$, then $f(2)$, then $f(3)$, and so on) [2+].

(c) Conclude that to determine which graphs F are such that $\mathcal{G}_F$ is χ-bounded it suffices to do so in the case when F is a tree [1].

5 Coloring Edges and Ramsey Theory

Turán's original motivation for the Turán problem came from another area of extremal combinatorics known as Ramsey theory I thought this was true but I'm failing to find a reference. In a very abstract sense, Ramsey theory (which extends far beyond just that of graphs) aims to prove that every sufficiently large structure contains relatively simple and orderly substructures. The original problem, as well as the namesake of the theory, comes from the following foundational result of Ramsey[9] from **REF**.

Definition 14. A *red-blue edge coloring* of a graph G is a map $\chi : E(G) \to \{\text{red}, \text{ blue}\}$. We say that such a coloring has a *monochromatic* K_n if there exists a subgraph of G isomorphic to K_n such that either every edge of the subgraph is colored red or if every edge of the subgraph is colored blue.

Theorem 5.1 (Ramsey's Theorem). *For all integers $n \geq 1$, there exists a (finite) N such that every red-blue edge coloring of K_N contains a monochromatic K_n.*

Equivalently, this says that for all integers $n \geq 1$, there exists some (finite) N such that every N-vertex graph G either contains a clique of size n or an independent set of size n (as can be seen by coloring the edges of K_N red if they belong to G and blue otherwise). That is, large graphs can not simultaneously have arbitrarily large clique and independent numbers.

The original proof of Ramsey's Theorem does not give explicit bounds on the size of N, and the central problem ins Ramsey Theory is to get better bounds on this quantity.

Definition 15. We define the *(diagonal) Ramsey number* $R(n)$ to be the smallest integer N such that every red-blue edge coloring of K_N contains a monochromatic K_n.

There are many variants of this classical Ramsey number $R(n)$, several of which we will discuss below.

5.1 Classical Bounds

Let us start by working some small examples to give a little intuition for the problem in general. It is immediate that $R(1) = 1$ and $R(2) = 2$, so the first non-trivial case of the problem is to determine[10] $R(3)$.

Proposition 5.2. *We have $R(3) = 6$.*

Proof. The lower bound comes from giving a coloring of the edges of K_5 which does not contain a triangle. The unique way to do this is to take a $C_5 \subseteq K_5$ and color its edges red with the

[9]Funnily enough Ramsey was not a combinatorialist but rather a logician, and to this day there is still a lot of work on Ramsey theoretic problems from the perspectives of both logic and combinatorics.

[10]Colloquially this result is known as the "party problem" due to the following interpretation of its statement: if there are 6 people at a party, then there exist 3 people there who either all know each other or who all do not know each other.

remaining edges (which also form a C_5) being colored blue. It is easy to check that such a coloring has no monochromatic triangle.

For the upper bound, consider an arbitrary red-blue coloring of the edges of K_6 and assume for contradiction that this did not contain a monochromatic triangle. Let u be an arbitrary vertex, and observe that u has 5 total edges incident to it each of which is given one of 2 colors, so by the pigeonhole principle at least 3 of the edges of u all have the same color, say without loss of generality that the edges uv_1, uv_2, uv_3 are all colored red. Now if any edge v_iv_j is colored red then u, v_i, v_j would form a red triangle, so we can assume that all of the edges v_iv_j are colored blue. But in this case v_1, v_2, v_3 forms a blue triangle, again yielding a contradiction. $\square$

At its core, the reason that the upper bound proof worked is that if a red-blue coloring does not contain a monochromatic K_3, then the "red neighborhood" of any vertex u can not contain either a red K_2 nor a blue K_3. Building on this idea leads to the following definition.

Definition 16. Given integers m, n, we define $R(m, n)$ to be the smallest integer N such that if every edge of K_N is colored either red or blue, then there either exists a red K_m or a blue K_n.

For example, one can check that $R(2, 3) = 3$ which is implicitly what we used in our upper bound proof for $R(3)$. Generalizing this idea gives the following observation of Erdős and Szekeres.

Lemma 5.3 (Erdős-Szekeres)**.** *For all $m, n \geq 2$, we have*

$$R(m, n) \leq R(m-1, n) + R(m, n-1).$$

Proof. Let $N = R(m-1, n) + R(m, n-1)$ and assume for contradiction that there exists a red-blue edge coloring of K_N which does not contain a red K_m nor a blue K_n. Let u be an arbitrary vertex and let V_R denote the set of vertices v such that uv is colored red, and similarly define V_B. Note that $|V_R| + |V_B| = N - 1 = R(m-1, n) + R(m, n-1) - 1$, and that we must either have $|V_R| \geq R(m-1, n)$ or $|V_B| \geq R(m, n-1)$ (since otherwise $|V_R| + |V_B| \leq R(m-1, n) + R(m, n-1) - 2$).

First consider the case that $V_R \geq R(m-1, n)$. By definition of $R(m-1, n)$, the coloring on $K_N[V_R]$ must contain either a red K_{m-1} or a blue K_n. The latter case can not happen by assumption of our coloring, and if the former happens then this K_{m-1} together with u would form a red K_m, again giving a contradiction. A similar conclusion holds if $|V_B| \geq R(m, n-1)$, proving the result. $\square$

Using this recurrence relation together with the boundary condition $R(1, n) = R(n, 1) = 1$ gives the following.

Theorem 5.4. *For all $m, n \geq 1$, we have*

$$R(m, n) \leq \binom{m+n-2}{m-1}.$$

Indeed, by induction on $m+n$ we have that

$$R(m,n) \le R(m-1,n) + R(m,n-1) \le \binom{m+n-3}{m-2} + \binom{m+n-3}{m-1} = \binom{m+n-2}{m-1},$$

with the last step being Pascal's identity. Finally, taking $m = n$ in this bound gives bounds for diagonal Ramsey numbers.

Corollary 5.5. *For all $n \ge 1$, we have*

$$R(n) \le \binom{2n-2}{n-1} \le 4^n.$$

Let us turn now to lower bounds, starting with an elementary bound.

Lemma 5.6. *We have $R(n) \ge (n-1)^2 + 1$.*

Proof. Color the edges of $R_{(n-1)^2}$ via breaking up the vertex sets in to $n-1$ parts $V_1, \ldots, V_{n-1}$ each of size $n-1$ and coloring all the edges within each part red and all the edges between two parts blue. It is easy to see that this avoids monochromatic copies of K_n. □

Note that in this coloring that the blue edges form a copy of the Turán graph $T_{n-1}(n-1)$ and I think there's some connection here but I forget the details. It was believed for some time that $R(n)$ should grow polynomialy like in this lemma here, but Erdős disproved this in a very strong form.

Theorem 5.7. *We have*

$$R(n) \ge (1+o(1)) \frac{n}{e\sqrt{2}} 2^{n/2}.$$

This is a strange bound and not one should necessarily expect to understand how to prove even if you work out the right proof idea. Indeed, our proof will utilize the following.

Mantra 11. First figure out how your proof works using an abstract set of parameters, then go back and choose whatever parameters you need in order for the arithmetic to go through

Let us see this in action.

Proof. To partially motivate the idea of the argument, we observe that it is very easy to show $R(n) \ge n+1$ for $n \ge 3$. Indeed, there are only two colorings of K_n which contain a monochromatic K_n, and as long as $n \ge 3$ we can find a coloring which avoids one of these two bad ones. To get our stated lower bound, we will similarly use an elementary counting argument to bound the number of "bad" colorings of K_N and then argue that if N is not too large then there are more total colorings than bad colorings, proving that there exists some coloring which is not bad.

From now on we fix an integer N which we will determine later once we see how the numbers work out. For each subset $S \subseteq [N]$ of size n, let B_S denote the set of edge colorings of K_N which have a monochromatic K_n on S. Because the total number of edge-colorings of K_N is

$2^{\binom{N}{2}}$ and because a coloring avoids monochromatic K_n's if and only if it does not lie in any B_S set, we see that there exists an edge-coloring of K_N avoiding monochromatic K_n's if and only if

$$2^{\binom{N}{2}} > |\bigcup_{S\in\binom{[N]}{n}} B_S|.$$

It thus remains to show that this latter set is small. Using elementary arguments we have

$$|\bigcup_{S\in\binom{[N]}{n}} B_S| \le \sum_{S\in\binom{[N]}{n}} |B_S| = \binom{N}{n} 2^{1+\binom{N}{2}-\binom{n}{2}}$$

where this last step used that every coloring in B_S has 2 choices for how it can act on the edges of S (either all red or all blue) together with $2^{\binom{N}{2}-\binom{n}{2}}$ choices for the remaining edges. As such, we will succeed if

$$\binom{N}{n} 2^{1-\binom{n}{2}} < 1.$$

To get a handle on this, we use the well-known binominal inequality $\binom{m}{k} \le (em/k)^k$ to conclude that it suffices to have N such that

$$2\left(\frac{eN2^{(n-1)/2}}{n}\right) < 1,$$

and in particular the result holds provided $N < 2^{1/n} \cdot \frac{n}{e\sqrt{2}} 2^{-n/2}$, and picking such an N gives the desired bound. □

This counting argument is all well and good, but we can give a more modern perspective by rewriting our proof in the language of probability.

Alternative Proof. Let N be an integer to be determined later and consider a uniform random red-blue edge coloring of K_N. Let X be the random variable which is equal to the number of monochromatic K_n's that are in the random coloring of K_N. Crucially, we observe that if $\mathbb{E}[X] < 1$, then $R(n) > N$. Indeed, because X is integer valued, the only way $\mathbb{E}[X] < 1$ is possible is if there exists some coloring of K_N such that $X = 0$, i.e. a coloring without any monochromatic copies of K_N.

To get a handle on $\mathbb{E}[X]$, for each $S \in \binom{[N]}{n}$ we let $\mathbb{1}_S$ denote the indicator random variable for $K_N[S]$ being monochromatic. That is, $\mathbb{1}_S$ is the random variable defined by having $\mathbb{1}_S = 1$ if $K_N[S]$ is monochromatic and $\mathbb{1}_S = 0$ otherwise. With this $X = \sum \mathbb{1}_S$, so by linearity of expectation we have

$$\mathbb{E}[X] = \sum \mathbb{E}[\mathbb{1}_S] = \sum \Pr[\mathbb{1}_S = 1] = \binom{N}{n} 2^{1-\binom{n}{2}},$$

as can be checked by a simple counting argument. Thus in total, we conclude $R(n) > N$ provided $\binom{N}{n}2^{1-\binom{n}{2}} < 1$, which as we showed in the previous version of the proof happens for $N = (1 + o(1))\frac{n}{e\sqrt{2}}2^{-n/2}$. □

While both the counting argument and the probabilistic argument for [theorem] are effectively equivalent to each other, the prespective of "thinking probabilistically" has proven to be the more useful in general. Indeed, it is hard at this point not to find an important result in Ramsey theory where the lower bound (and sometimes even the upper bound) does not use some amount of ideas or techniques motivated by probability theory. Since we are not assuming the reader has any knowledge of probability we will not dwell on this point any further at this point, though the interested reader is invited to go to [probabilistic methods section] for much more on this perspective.

We note that in both cases of our argument, the lower bound for $R(n)$ we gave was non-constructive, i.e. we did not explicitly construct a coloring of K_N which avoids monochromatic K_n's, we only showed that such a coloring must exist. It is a major open problem to find a constructive argument which gives anywhere close to these bounds here.

Open Problem 5.8. *For some $c > 1$, find "explicit" red-blue edge colorings of K_{c^n} which avoid monochromatic K_n's.*

Observe that our proof not only shows that constructions should exist for $c = \sqrt{2}$, but in fact a more careful inspection shows that for any $c < \sqrt{2}$ that *almost every* coloring should work. Nevertheless, how to explicitly find such a coloring problem remains quite elusive.

The results we have mentioned in this sections are all classical, and the reader might wonder what is the current state of the art. For the lower bound, the only improvement over [result] is an argument due to Lovász using a slightly more involved probabilistic approach that gives a lower bound of [whatever], improving the bound of [result] by a multiplicative factor of [whatever].

For the upper bound, modest results showing bounds of the form $4^{n-o(n)}$ for an increasing series of $o(n)$ functions were obtained over the years until a recent major breakthrough by [authors in year] who proved that $R(n) \leq$???, and since then some further optimizations of their argument has yielded a bound of $R(n) \leq$???. At present this is all that is known for diagonal Ramsey numbers despite decades of hard work from an armada of talented mathematicians.

In addition to the diagonal Ramsey numbers $R(n)$, a lot of work has been put into studying the assymetric case $R(m, n)$. In particular, the study of these numbers when m is fixed and n tends towards infinity is referred to as "off-diagonal" Ramsey numbers. These problems are essentially equivalent to asking: how large can $\alpha(G)$ be if G is K_m-free and contains a given number of vertices? Indeed, more exposition, also comment on how $m = 3, 4$ are reasonably well understood due to complex probabilistic arguments.

5.2 More Colors and Arithmetic Ramsey Theory

There are a ton of variants for Ramsey numbers that one can consider. One of the immediate ones to consider is using more than just two colors. To this end, we define the *multi-color Ramsey number* $R_q(n)$ to be the smallest number N such that every q-coloring of the edges of K_N contains a monochromatic copy of K_n. Similar to [before] one can show that these numbers exist. In particular, we leave it as an exercise to prove the following bounds for the first non-trivial case of $n = 3$.

Theorem 5.9. *We have*

$$2^q < R_q(3) \leq 3 \cdot q!$$

Another direction is to consider coloring combinatorial objects other than graphs. One natural choice would be the integers $[N]$, from which we can ask if there exists a monochromatic subset satisfying some sort of arithmetic condition. One classical result due to Schur is as follows.

Theorem 5.10 (Schur)**.** *For all $q \geq 1$, there exists a finite number N_q such that any q-coloring of $[N]$ contains a monochromatic solution to the equation $x+y=z$, i.e. there exist three integers x, y, z with $x+y=z$ which are all assigned the same color.*

Proof. We will in fact prove that

$$N_q \leq R_q(3),$$

following a common theme in Ramsey theory of upper bounding one Ramsey problem by a function of another. To prove this, we will start with some coloring $\chi : [N] \to [q]$ and then use this to construct an auxiliary coloring $\chi' : E(K_N) \to [q]$ in such a way that monochromatic triangles under χ' correspond to monochromatic solutions to $x+y=z$ under χ. There are a couple of plausible ways one might try and define χ'. For example, given the edge $xy \in E(K_N)$ it is perhaps natural try coloring this edge to be the same color as either $\min(x, y)$ or $\max(x, y)$, but neither of these are really "compatible" with the goal of finding a solution to $x+y=z$.

With a bit more thought, one might come up with the (correct) idea of defining $\chi'(xy) = \chi(|x-y|)$. To see why this does what we want, assume that χ' has a monochromatic triangle on $u < v < w$. This implies that $\chi(v-u), \chi(w-v), \chi(w-u)$ all have the same color. Moreover, we have $(v-u)+(w-v)=(w-u)$, so taking $x=v-u,\ y=w-v$, and $z=w-u$ gives a monochromatic solution under χ. In total this implies that if $N \geq R_q(3)$ and χ is an arbitrary coloring then, because χ' must contain a monochromatic triangle since $N \geq R_q(3)$, χ contains a monochromatic solution to $x+y=z$. This proves $N_r \leq R_q(3)$, and in particular that this number is finite. □

A lot more can be said about this area known as arithmetic Ramsey theory. Perhaps the most famous result in this direction is Van der Waerden's Theorem.

Theorem 5.11 (Van der Waerden's Theorem)**.** *For all k, q, there exists a finite number $N_{k,q}$ such that any r-coloring of $[N_{k,q}]$ contains a monochromatic k-term arithmetic progression. That is, there exist integers $a, d \geq 1$ such that $a, a+d, \ldots, a+(k-1)d$ are all given the same color.*

Proving this is not so easy, and the bounds for $N_{k,q}$ are horrendous even in the case of $q=2$. In fact, an even stronger statement than Van der Waerden's Theorem is known to be true.

Theorem 5.12 (Szemerédi's Theorem)**.** *Every subset of $[N]$ which does not contain a k-term arithmetic progression has size $o(N)$.*

To see this implication, observe that every r-coloring of $[N]$ contains a subset of size at least $r^{-1}N$ which, by Szermedi's Theorem, must contain a k-term arithmetic progression whenever N is sufficiently large. This is an example of a general phenomenon where Turán results (which bound how dense a structure can be before it contains a given substructure) often upper bound Ramsey results (which bound how large a structure can be with the property that it can be partitioned into r substructures avoiding a given substructure) simply because one of the partition elements in a Ramsey result must have relatively large density.

5.3 Hypergraph Ramsey

Insert, mostly stepping up lemma and the general uppper bound/discussion on the gap in tower heights.

5.4 Ramsey Without Colors

We will omit this for time unless requested by popular by demand. Broadly speaking it will be around the theme that Ramsey isn't just about saying that colored objects contain things. Some examples include monotone sequences and convex sets.

5.5 Exercises

1. Let's look at some small Ramsey numbers:

   (a) Prove that $R(3,4) = 9$ [2]

   (b) Prove that $R(4) \leq 18$ [1].

   (c) Prove that $R(4) = 18$ [3].

   (d) Determine[11] $R(5)$ [5].

2. Prove that every n-vertex graph has a clique or independent set on at least $\frac{1}{2}\log_2(n)$ vertices [1+].

3. Recall that a tournament is a digraph obtained by giving an orientation to each edge of a complete graph, and that a tournament is transitive if one can order its vertices $v_1, \ldots, v_n$ in such a way that $v_i \to v_j$ if and only if $i < j$. Prove that every tournament on n vertices contains a transitive tournament of size at least $\lfloor \log_2(n) \rfloor + 1$ [2-].

* * *

4. Prove for all $n, q \geq 2$ that $R_q(n) \leq q^{qn}$ [2-].

[11] Currently the best known bounds are $43 \leq R(5) \leq 46$. The fact that this is still open should demonstrate how hard determining $R(n)$ exactly is. Indeed, Erdős once said something to the effect of: if aliens came to Earth and demanded we tell them what $R(5)$ was in the next 10 years or they would destroy us, then we should dedicate all our resources to this problem. If instead they ask for $R(6)$, then we should instead dedicate all our resources to fighting the aliens because we have no hope of doing what they ask.

5. Let us look at the multi-color Ramsey number $R_q(3)$.

   (a) Prove that $R_q(3) > 2^q$ [2-].

   (b) Prove that $R_q(3) \leq 3 \cdot q!$, noting that this is best possible for $q = 2, 3$ [2].

   (c) Improve this upper bound to $R_q(3) \leq \lfloor e \cdot q! \rfloor + 1$, which as far as we know is still the best known upper bound [3-].

6. For every graph F and integer q, define $R_q(F)$ to be the smallest integer N such that any q-edge coloring of K_N contains a monochromatic copy of F. Prove that if $\text{ex}(n, F) = O(n^{2-\alpha})$ for some $\alpha > 0$, then $R_q(F) = O(q^{1/\alpha})$ (Hint: concretely if you assume $\text{ex}(n, F) \leq Cn^{2-\alpha}$ then you should be able to prove something like $R_q(F) \leq (4Cq)^{1/\alpha}$, for example) [2-].

7. One of the most important results in general Ramsey theory is the Hales-Jewett Theorem which is a sort of "high-dimensional tic-tac-toe" theorem that goes as follows: Insert statement, exercise is to derive Van der Waerden from this.

* * *

8. We say that a graph G is K_n-Ramsey if any red-edge edge coloring of G contains a monochromatic copy of K_n, and we define the *size Ramsey number* $\hat{R}(n)$ to be the smallest number of edges in a graph which is K_n-Ramsey.

   (a) Observe that $R(n)$ can be defined to be the smallest number of *vertices* in a graph which is K_n-Ramsey, motivating this definition [1].

   (b) Prove that $\hat{R}(n) \leq \binom{R(n)}{2}$ [1+].

   (c) Prove that $\hat{R}(n) = \binom{R(n)}{2}$; noting crucially that this equality holds despite us largely not understanding what $R(n)$ is (Hint: prove that any K_n-Ramsey graph must have chromatic number at least $R(n)$) [2+].

Part II

Methods

6 Probabilistic Methods

We will only cover a small portion of this topic since GSU has a whole class dedicated to it. We encourage anyone interested in digging further to read either the standard text by Alon and Spencer, as well as my own notes here.

6.1 Deletion Arguments

One the most important development in extremal combinatorics has been the idea of using probabilistic tools to solve extremal problems. We've already seen one example of this in Theorem 5.7 where we used a uniform random edge-coloring to prove exponential lower bounds on the Ramsey number $R(n)$. Another classical result in this direction is the following.

Lemma 6.1. *Every graph G has a bipartite subgraph with at least $\frac{1}{2}e(G)$ edges.*

Proof. Let $U \subseteq V(G)$ be obtained by including each vertex independently and with probability $\frac{1}{2}$, and let $V = V(G) \setminus U$. Let G' be the graph which consists of every edge $e \in E(G)$ with one vertex in U and one vertex in V. It is easy to check that $\mathbb{E}[e(G')] = \frac{1}{2}e(G)$, so such a (bipartite) subgraph exists. □

In both this proof and our proof for lower bounding $R(n)$, we proceeded by constructing some random object and then showing with positive probability that it has the sort of properties that we want. A very useful variant of this approach is something known as a deletion argument (also referred to in the literature as alteration arguments).

Roughly speaking, such arguments involve constructing some random object which "almost" has the properties we want. The idea then is that we can simply delete a small number of elements from this object in order to turn it into an object which genuinely has the properties we want without changing the structure of the original object by too much. For example, a simple version of this approach gives a general lower bound for Turán numbers of arbitrary graphs F.

Theorem 6.2. *If F is a graph with v vertices and e edges with $e \geq 2$, then*

$$\mathrm{ex}(n, F) = \Omega(n^{2-\frac{v-2}{e-1}}).$$

In this proof and for the rest of the section we will base our random construction off the most important object in probabilistic combinatorics, namely the Erdős-Renyi random graph model. For an integer $n \geq 1$ and a real number $0 \leq p \leq 1$ we let $G_{n,p}$ denote the random n-vertex graph obtained by including each edge independently and with probability p. Thus $G_{n,1} = K_n$ with probability 1, $G_{n,0}$ is the empty graph with probability 1, and $G_{n,1/2}$ is equally likely to be any n-vertex graph.

Proof. As is often the case in the probabilistic method, our proof begins by considering the random graph $G_{n,p}$ with p a quantity to be determined later. Let X denote the number of copies of F in $G_{n,p}$, which is a quantity we will get a handle of through indicator variables.

To this end, for each copy F' of F in K_n, let $1_{F'}$ be the indicator variable which is 1 if F' is a subgraph of $G_{n,p}$ and 0 otherwise. Note then that $X = \sum 1_{F'}$. Moreover, we have $\Pr[1_S = 1] = p^e$, so by linearity of expectation we find

$$\mathbb{E}[X] = \sum_{F'} \mathbb{E}[1_{F'}] = \sum_{F'} p^e \le p^e n^v,$$

with this last step using that there are at most n^v copies of F in K_n since each copy can be specified by a map from $V(F)$ to $V(K_n)$.

Informally, this inequality suggests that if p is significantly smaller then $n^{-v/e}$, then we expect $G_{n,p}$ not to have any copies of F. On the other hand, the expected number of edges in $G_{n,p}$ at this point is $p\binom{n}{2} \approx n^{2-v/e}$, suggesting that $G_{n,p}$ gives an F-free graph with about $n^{2-v/e}$ edges for this range of p. It is not difficult to make this argument precise to give a bound of $\mathrm{ex}(n, F) = \Omega(n^{2-v/e})$, but an even simpler argument can be made to work by choosing p somewhat above $n^{v/e}$.

Observe that when $p \gg n^{-v/e}$, the calculation above suggests that $G_{n,p}$ will contain copies of F (at least in expectation), so $G_{n,p}$ will not be an F-free graph for this range of p. However, we can get around this by observing that if $G \subseteq G_{n,p}$ is obtained by deleting an edge from each copy of F in $G_{n,p}$, then G will be F-free by construction. Moreover, the number of edges we will have is $e(G) \ge e(G_{n,p}) - X$ since at most X of the original edges from $G_{n,p}$ are deleted. Using linearity of expectation gives

$$\mathbb{E}[e(G)] \ge \mathbb{E}[e(G_{n,p}) - X] \ge p\binom{n}{2} - p^e n^v \ge \frac{1}{4}pn^2 - p^e n^v. \tag{2}$$

At this point we want to choose p so that the above expression is maximized. Intuitively Make this a mantra, possibly with some partial justification if only for this one particular example this will happen when both terms on the rightside of (2) are roughly equal to each other, i.e. when $pn^2 \approx n^v p^e$. This suggests taking $p \approx n^{\frac{2-v}{e-1}}$. And indeed, after playing around for a bit, one sees that, for example, taking $p = 2^{-3} n^{\frac{2-v}{e-1}}$ and plugging it into (2) gives $\mathbb{E}[e(G)] \ge 2^{-6} n^{2-\frac{2-v}{e-1}}$ Maybe spell out more. Because G is a (random) F-free graph with at least this many edges in expectation, there must exist some deterministic F-free graph with at least this many edges, proving the result. □

Theorem 6.2 can fail to be effective if we consider F with, say, a bunch of isolated vertices. However, a simple observation allows one to improve upon Theorem 6.2 in cases like these.

Corollary 6.3. *For every graph F with $e(F) \ge 2$ we define the 2-density*

$$m_2(F) := \max_{F' \subseteq F,\ e(F') \ge 2} \frac{e(F') - 1}{v(F') - 2}.$$

For any F with $e(F) \ge 2$ we have

$$\mathrm{ex}(n, F) = \Omega(n^{2 - \frac{1}{m_2(F)}}).$$

Proof. If $m_2(F) = 1$ then we only need to prove $\mathrm{ex}(n, F) = \Omega(n)$ which is trivial via considering an n-vertex star if $F \neq K_{1,t}$ and considering an n-vertex matching otherwise. We can thus assume $m_2(F) > 1$ from now on.

Let $F' \subseteq F$ be any subgraph obtaining the maximum in $m_2(F)$, which under the assumption of $m_2(F) > 1$ implies that $e(F') \geq v(F')$. Thus by Theorem 6.2 we have

$$\mathrm{ex}(n, F) \geq \mathrm{ex}(n, F') = \Omega(n^{2-\frac{v(F')-2}{e(F')-1}}) = \Omega(n^{2-\frac{1}{m_2(F)}}),$$

where here this first inequality used the fact that any F'-free graph is also F-free. □

A somewhat more involved application of the deletion method can be used to prove Theorem 4.18, which we recall says that for all $\ell, t \geq 2$ there exists a graph with girth at least ℓ and chromatic number at least t. For this proof we will use Markov's inequality, which we recall says that if X is a non-negative random variable then $\Pr[X \geq t] \leq \mathbb{E}[X]/t$ for all real t.

Theorem 4.18. The naive idea we want to try is to pick some values for n and p such that with high-probability $G_{n,p}$ simultaneously has few (or even 0) cycles of length at most ℓ and has large chromatic number. Let us address each of these obstacles in turn.

First, let X_i denote the number of cycles of length i in G and let $X_{<\ell} = \sum_{i=3}^{\ell-1} X_i$. Observe that $\mathbb{E}[X_i] \leq p^i n^i$ as the total number of cycles of length i in K_n is at most n^i and the probability that any given cycle C survives into $G_{n,p}$ is exactly p^i (i.e. this is the probability that $G_{n,p}$ independently keeps all i edges of C). By linearity of expectation we find that

$$\mathbb{E}[X_{<\ell}] \leq \sum_{i=3}^{\ell-1} p^i n^i \leq (\ell - 1) \max\{pn, (pn)^{\ell-1}\}. \tag{3}$$

We now turn to studying $\chi(G_{n,p})$, and a priori it is not so clear how to approach this. The key insight is that we only care about proving lower bounds for this chromatic number, so it suffices to bound some general lower bound for χ which might be simpler to analyze. In particular, by Theorem 4.5 it suffices to show that $n/\alpha(G_{n,p})$ is large, i.e. that $\alpha(G_{n,p})$ is small, which is much simpler to do. Indeed, if we let Y_a denote the number of independent sets of $G_{n,p}$ of size a then by linearity of expectation and the basic inequality $1 - x \leq e^{-x}$, we find

$$\mathbb{E}[Y_a] = (1-p)^{\binom{a}{2}} \cdot \binom{n}{a} \leq e^{-p\binom{a}{2}} \cdot n^a = (ne^{-p(a-1)/2})^a \leq (ne^{-pa/2})^a. \tag{4}$$

Heuristically, what this tells us is that if $e^{-pa/2} \ll n^{-1}$, i.e. if $p \gg \frac{\log(n)}{a}$, then $G_{n,p}$ with high probability will not contain any independent sets of size at least a, which if true would imply that $\chi(G_{n,p}) \geq n/a$. This gives a good lower bound if $a \ll n$, and hence heuristically we need $p \gg \log(n)/n$ in order for us to conclude that $G_{n,p}$ has large chromatic number. Unfortunately for this range of p, (3) suggests that the number of short cycles in $G_{n,p}$ could be as large as $(\log n)^\ell$. In total this suggests (the true fact that) $G_{n,p}$ does not simultaneously have high girth and high chromatic number for any choice of p.

The saving grace to this approach is the observation that although $G_{n,p}$ does not have 0 short cycles for $p \gg \log n/n$, it does have *few* of them. In particular, if we take $G_{n,p}$ and delete a vertex from each of its short cycles then this graph will by definition have large girth and also be very close to $G_{n,p}$ if $G_{n,p}$ has few short cycles.

With all this motivation in mind, let $p = C \log n/n$ with n, C sufficiently large integers so that the following inequalities hold. Let A_1 be the event that $X_{<\ell} \le n/2$. By Markov's inequality and (3) we have

$$\Pr[A_1] = 1 - \Pr[X_{<\ell} > n/2] \ge 1 - \frac{(\ell-1)(pn)^{\ell-1}}{n/2} \ge 1 - \frac{(\ell-1)C^{\ell-1}(\log n)^{\ell-1}}{n/2} > \frac{1}{2},$$

with this last inequality holding for n sufficiently large in terms of C and ℓ. Similarly let A_2 denote the event that $\alpha(G_{n,p}) < n/2t$. By Markov's inequality and (4) we have

$$\Pr[A_2] = 1 - \Pr[Y_{n/2t} \ge 1] \ge 1 - (ne^{-pn/4t})^{n/2t} = 1 - (n \cdot n^{-C/4t})^{n/2t} > \frac{1}{2},$$

with the last inequality holding for $C > 4t$ and n sufficiently large. From this, we conclude that $\Pr[A_1 \cap A_2] > 0$, i.e. with positive probability both A_1 and A_2 occur, i.e. there exists an n-vertex graph G such that it has at most $n/2$ cycles of length less than ℓ and $\alpha(G) < n/2t$. Define G' by taking G and deleting 1 vertex from each cycle of length less than ℓ. By assumption we have $v(G') \ge n/2$ and $\alpha(G') \le \alpha(G) \le n/2t$, and as such

$$\chi(G') \ge \frac{v(G')}{\alpha(G')} \ge \frac{n/2}{n/2t} = t,$$

proving the result. □

So far we have applied the deletion argument to a very natural random object, namely that of $G_{n,p}$. However, the deletion argument can often be amplified by considering more complex random objects. We will look at some examples of this in the coming subsections.

6.2 Dependent Random Choice

When using the probabilistic method, it is often the case that the simplest possible way of generating a given random object is enough to get the job done. However, there are many instances where a more carefully chosen random variable can be chosen in order to give stronger bounds. One basic example of this is the following observation, which we leave as an exercise to the reader.

Lemma 6.4. *Given a non-empty graph G, let $v_1, v_2 \in V(G)$ be random vertices where*

- *v_1 is chosen uniformly at random, and*
- *v_2 is chosen by first uniformly at random choosing an edge e of G and then uniformly at random choosing v_2 to be one of the vertices of e. Equivalently, one can uniformly choose a vertex v_2' and then let v_2 be a uniform random neighbor of v_2'.*

Then,

$$\mathbb{E}[\deg(v_1)] \le \mathbb{E}[\deg(v_2)].$$

That is, if we want to randomly select a vertex with high expected degree in G, then it is always better (at least in theory) to work with the more complicated random variable v_2 over v_1. At a high-level, the reason v_2 gives vertices that tend to be incident to more edges is because we literally defined v_2 to be incident to an edge. Another intuitive reason this works is that our definition of v_2 is more "robust" in the sense that the degree of v_2 will be entirely unaffected by the addition of isolated vertices to G, while in contrast v_1 will perform strictly worse with such extra vertices.

The discussion above captures the central idea behind the dependent random choice method: if we want a random variable X to have some property P, then it can be helpful to construct X in such a way that the property P is "built into" the definition of X somehow (e.g. if we want a vertex to be incident to many edges, then we pick a vertex incident to an edge).

In what follows we look at some more examples of this philosophy due mostly to Fox and Sudakov [**?**] all aimed around the theme of X being a set of vertices and P being the property that X has many common neighbors. To this end, throughout this section we let $N(S)$ denote the common neighborhood of a set of vertices S, i.e. we let $N(S) := \{u : u \in N(v)\ \forall v \in S\}$. Our guiding example for these results will be the following.

Lemma 6.5. *Let G be an n-vertex graph with average degree at least d. For any choice of integers m, r, t, there exists a set $U \subseteq V(G)$ such that every r-subset of U has at least m common neighbors, and such that*

$$|U| \geq \frac{d^t}{n^{t-1}} - \binom{n}{r}\left(\frac{m}{n}\right)^t.$$

Proof. The format of the result suggests how we might try to prove it: we'll randomly construct a set W which will have expected size at least d^t/n^{t-1}, after which we'll use the method of alterations to delete from W a set of a bad vertices that will in expectation have size at most $\binom{n}{r}(m/n)^t$, in total giving a final set U with the desired properties and desired size.

The absolute simplest way we could try and make such a random set W is by including each vertex independently and with some probability p. However, this procedure is doomed to fail if G is, say, a clique on roughly $\sqrt{dn}$ vertices, since in this case almost all of the vertices of W will be bad with high probability. To get around this, we will follow the same philosophy of Theorem 6.4: we will choose some auxiliary set T in a (simple) random way and then define W in terms of this auxiliary set. Moreover, we will do this in such a way that our construction for W "biases" it towards having many common neighbors in the same way that defining v_2 in Theorem 6.4 to be incident to an edge "biased" it towards being incident to many edges.

With this motivation in mind, let T be the random set obtained by uniformly at random selecting t vertices with repetition (i.e. each vertex is equally likely to be the ith vertex added to T, and in total T has size at most t), and define $W = N(T)$ (note that defining W to be the common neighborhood of a set "biases" it towards having many common neighbors). All that remains now is a basic alternatoins argument bounding the size of W and the number of "bad" events.

Observe that the probability of a given vertex v being included in W is exactly $(\deg(v)/n)^t$, so by linearity of expectation and convexity we find that

$$\mathbb{E}[|W|] = \sum_v \left(\frac{\deg(v)}{n}\right)^t \geq \frac{d^t}{n^{t-1}}.$$

Now define a set of vertices $S \subseteq V(G)$ of size r to be *bad* if $|N(S)| < m$. Crucially, we note that the probability W contains a given bad set S is at most $(m/n)^t$ since $S \subseteq W$ if and only if $T \subseteq N(S)$. Thus the expected number of bad sets of W is at most $\binom{n}{r}(m/n)^t$. Taking U to be the set obtained by deleting a vertex from each bad set of W, we see that U has the desired properties by construction, and that it has the desired size in expectation, proving that such a U exists. ☐

We can use Lemma 6.5 to quickly prove some nice bounds on Turán numbers through the following basic embedding lemma.

Lemma 6.6. *Let F be a bipartite graph with bipartition $A \cup B$ with $|A| = a,\ |B| = b$ such that the vertices in B all have degree at most $r \leq a$. If G is a graph which contains a set U such that $|U| = a$ and such that every r-subset of U contains at least $a + b$ common neighbors, then G contains F as a subgraph.*

Proof. In order to show that G contains F as a subgraph, we show that there exists an injective homomorphism ϕ from $V(F)$ to $V(G)$, which we construct as follows. Choose $\phi|_A$ to be an arbitrary bijection onto U. Iteratively for each $v \in B$ that has yet to be assigned, choose $\phi(v)$ to be any common neighbor of $\phi(N_F(v))$ which has yet to be assigned by ϕ. Note that there exist at least $a + b$ common neighbors of $\phi(N_F(v))$ by hypothesis, so there certainly exists one such vertex which has yet to be assigned. This mapping gives the result. ☐

With this we can quickly prove the following, the statement and proof for which comes from Alon, Krivelevich, Sudakov [?], though this result can also be obtained by using an earlier result of Füredi's [?].

Theorem 6.7 ([?, ?])**.** *If F is a bipartite graph with bipartition $A \cup B$ with $|A| = a$ and $|B| = b$ such that the vertices of B all have degree at most r, then*

$$\mathrm{ex}(n, F) \leq (a + b)n^{2-1/r}.$$

Observe that this result generalizes the Kővári-Sós-Turán Theorem, at least in terms of order of magnitude.

Proof. Assume for contradiction that G is an n-vertex F-free graph with average degree at least $d := 2(a + b)n^{1-1/r}$. By Lemma 6.6, we would be done if we could find a set U of size at least a such that every r-subset has at least $m := a + b$ common neighbors. By Lemma 6.5, for any choice of t we can find a set U with these properties of size at least

$$\frac{d^t}{n^{t-1}} - \binom{n}{r}\left(\frac{m}{n}\right)^t \geq (2a + 2b)^t n^{1-t/r} - (a + b)^t n^{r-t}.$$

Taking $t = r$ (which is chosen so that the two terms in the difference above are as close as possible to each other), we find that there exists such a set of size at least

$$(2a + 2b)^r - (a + b)^r \geq a.$$

We have thus found our desired set U, which together with Lemma 6.6 gives a copy of F in G, a contradiction. □

Theorem 6.7 is one of the few general upper bounds that are known for bipartite Turán problems and is a special case of a stronger conjecture of Erdős. For this, we recall that a graph F is *r-degenerate* if every subgraph of F contains a vertex of degree of at most r.

Conjecture 6.8 (Erdős [?]). *If F is a bipartite graph which is r-degenerate, then* $\mathrm{ex}(n, F) = O(n^{2-1/r})$.

This conjecture remains wide open in general, but it is possible to prove a weak version of this result using dependent random choice. Indeed, mimicing the proof of Theorem 6.6 gives the following result which suggests that something resembling Theorem 6.5 might be of use here.

Lemma 6.9. *Let G be a graph which contains vertex sets U_1, U_2 such that for each $k \in \{1, 2\}$, every subset of at most r vertices in U_k contains at least m common neighbors in U_{3-k}. Then G contains every r-degenerate bipartite graph on m vertices.*

Proof. Let F_1 be an m-vertex r-degenerate bipartite graph on $V_1 \cup V_2$. By definition this means that there exists a vertex $v_1 \in F_1$ such that $\deg_{F_1}(v_1) \leq r$, and that there is some $v_2 \in F_2 := F_1 - v_1$ with $\deg_{F_2}(v_2) \leq r$ and so on. We now define a map $\phi : V_1 \cup V_2 \to U_1 \cup U_2$ with $\phi(V_i) \subseteq U_i$ as follows. Iteratively assume we have defined $\phi(v_m), \phi(v_{m-1}), \ldots, \phi(v_{q+1})$ and that $v_q \in V_i$. Since $S := N(v_q) \cap \{v_m, \ldots, v_{q+1}\}$ has at most r vertices by assumption, the set $\phi(S) \subseteq U_{3-i}$ has at least m common neighbors, so choose $\phi(v_q)$ to be any of these vertices that has yet to be assigned. It is not difficult to see that this gives the desired embedding. □

Motivated by this lemma, we prove the following variant of Lemma 6.5.

Lemma 6.10. *Let $r, m \geq 2$ and let G be an n-vertex graph with at least $mn^{1-1/6r}$ edges. Then G contains two subsets U_1, U_2 such that, for $k = 1, 2$, every subset of r vertices in U_k has at least m common neighbors in U_{3-k}.*

Proof. The rough strategy of the proof is as follows. We will first apply Lemma 6.5 directly to obtain a large set U_1 such that every q-subset of U_1 (with $q > r$) has at least m common neighbors. We then mimic the proof of Lemma 6.5 by choosing a random set $T \subseteq U_1$ of size t and letting $U_2 = N(U_1)$. By choosing an appropriate value of t, the set U_2 will satisfy the condition. Moreover, if $q - t \geq r$, then for any r-subset $S \subseteq U_1$, the set $S \cup T$ has at least m common neighbors, all of which in particular lie in $N(T) = U_2$, so U_1 will also have the desired property.

We now being the formal argument. Apply Lemma 6.5 using $q := 3r$ for the parameters r, t in that lemma to get a set U_1 such that every subset of size $3r$ has at least m common neighbors and such that

$$|U_1| \geq \frac{d^{3r}}{n^{3r-1}} - \binom{n}{3r}(m/n)^{3r} \geq m^{3r}n^{1/2} - m^r/(3r)! \geq mn^{1/2}.$$

Now let T be a set obtained by including $t = 2r$ vertices uniformly at random from U_1 with replacement, and let $U_2 = N(T)$. The probability that U_2 contains a set of r vertices which have fewer than m common neighbors in U_1 is at most

$$\binom{n}{r}(m/|U_1|)^{2r} \le \frac{1}{r!} < 1,$$

and in particular there exists a choice of T such that no r-subset of U_2 has fewer than m common neighbors. Note that for any r-subset $S \subseteq U_1$, the set $S \cup T$ has size at most $3r$ vertices, so by construction S has at least m common neighbors which lie in $N(T) = U_2$. Thus U_1, U_2 gives the desired result. □

Combining these two lemmas immediately gives the following.

Theorem 6.11. *If F is an m-vertex r-degenerate graph, then*

$$\mathrm{ex}(n, F) \le mn^{2-1/6r}.$$

We note that one can optimize the proof of Lemma 6.10 to improve the exponent of this theorem slightly (notably by using $(3-2\sqrt{2})r$ instead of $3r$ throughout). However, the end result is still weaker than the best known bound of $\mathrm{ex}(n, F) \le m^{1/2r}n^{2-1/4r}$ due to Alon, Krivelevich, and Sudakov [**?**], with their proof more or less being a slight refinement of the argument we gave.

6.3 Random Algebraic Constructions

A recent trend of probabilistic combinatorics is to consider certain random algebraic objects, with the key insight here generally being that the algebraic nature of these objects forces certain structural conditions that can be exploited through random sampling. We look at one such instance of this phenomenon applied to studying multicolor Ramsey numbers.

Recall that $R_q(n)$ denotes the smallest N such that any q-edge coloring of K_N contains a monochromatic K_n. We did things differently in class compared to what's written in the text, but in any case we have $R_q(n) \le q^{qn}$. By modifying the probabilistic argument showing $R_2(n)$ is at least roughly $2^{n/2}$ one can easily prove $R_q(n)$ is at least roughly $q^{n/2}$. However, by using a product construction due to Lefmann one can prove a stronger lower bound roughly of the form $2^{qn/4}$.

Here we give a significant improvement to this lower bound of Lefmann by utilizing the following key observation. The initial idea for this lemma can be seen in Conlon and Ferber [**?**], though it was first really used by Wigderson [**?**] and then generalized by Sawin [**?**].

Lemma 6.12. *Let G be graph with no clique of size n, and let p be the probability that vertices $v_1, \ldots, v_n \in V(G)$ chosen independently and uniformly at random form an independent set. Then for all $q \ge 2$, we have*

$$R_q(n) \ge p^{-(q-2)/n}2^{(n-1)/2}.$$

Note that when $q = 2$ this recovers the usual lower bound for Ramsey numbers from the random coloring.

Proof. Let N be an integer to be determined later, and let $f_1, \ldots, f_{q-2} : V(K_N) \to V(G)$ be chosen independently and uniformly at random. Define a coloring $\chi : E(K_N) \to [\ell]$ in the following way: for distinct $x, y \in V(K_n)$, if there exists i such that $f_i(x)f_i(y) \in E(G)$, then set $\chi(xy)$ to be the minimum i with this property. Otherwise, set $\chi(xy)$ to be $q-1$ or q with probability 1/2 each. That is (as Wigderson notes in his paper), this coloring comes from covering K_N with $q-2$ randomly permuted blowups of G and then randomly using two colors to deal with any uncovered vertices.

We first observe that there is no monochromatic K_t in any color $i \leq q-2$. Indeed, if $\{x_1, \ldots, x_n\}$ were such a clique then this would imply $\{f_i(x_1), \ldots, f_i(x_n)\}$ forms a clique in G (since $\chi(x_j x_k) = i$ implies $f_i(x_j)f_i(x_k) \in E(G)$).

It remains to show that, with positive probability, there is no monochromatic K_n in color $i \in \{q-1, q\}$. Observe that a clique K_n in K_N has all of its edges colored by $q-1$ or q if and only if each f_i maps K_t to an independent set of G, and the probability that this happens is exactly p^{q-2} by hypothesis, and given this the probability that this K_n will be monochromatic is $2^{1-\binom{n}{2}}$. In total then, the expected number of monochromatic cliques will equal $\binom{N}{n} p^{q-2} 2^{1-\binom{n}{2}}$, and this will be less than 1 provided $N \leq p^{-(q-2)/n} 2^{(n-1)/2}$. Thus there exists a coloring of this size with no monochromatic clique, giving the desired result. $\square$

Observe that the p in Theorem 6.12 roughly corresponds to the number of independent sets of size at most n in G, so we need to find a graph with small clique number and not too many small independent sets. To this end, let $V \subseteq \mathbb{F}_2^n$ be the set of vectors v with $v \cdot v = 0$ (i.e. vectors with even Hamming weight), and let G be the graph on V where two vectors u, v are adjacent if and only if $u \cdot v = 1$.

Lemma 6.13. *If n is even, then the graph G contains no clique of size n.*

Proof. Assume for contradictiont that there exist distinct vectors $v_1, \ldots, v_n \in V$ with $v_i \cdot v_j = 1$ for all $i \neq j$ (and $= 0$ for $i = j$ by definition of V). We claim that these vectors are linearly independent. Indeed, if there exists $\alpha_i \in \{0, 1\}$ with $\sum \alpha_i v_i = 0$, then by taking the dot product of v_j on both sides we find $\sum_{i \neq j} \alpha_i \equiv 0$ for all i, and it is not difficult to show that this implies $\alpha_i = 0$ for all i (here we need that n is even, else $\alpha_i = 1$ for all i would work). However, V is a $n-1$ dimensional subspace, so it contains no set of n linearly independent vectors, proving the result. $\square$

Lemma 6.14. *The probability p that a uniformly random tuple $(v_1, \ldots, v_n) \in V^n$ is such that $\{v_1, \ldots, v_n\}$ is independent in G is at most $2^{-3n^2/8 + o(n^2)}$.*

Proof. Let X be the set of tuples $(v_1, \ldots, v_n) \in V^n$ such that $v_i \cdots v_j = 0$ for all i, j, so our goal is to upper bound $|X|/|V|^n = |X| 2^{-n^2+n}$. Define the rank of a tuple in X to be the rank of the smallest subspaces containing every vertex of the tuple. We claim that the number of tuples in X of rank r is at most

$$n! \left(\prod_{i=0}^{r-1} 2^{n-i} \right) \cdot 2^{(n-r)r} = n! 2^{nr - \binom{r}{2} + nr - r^2}. \tag{5}$$

Indeed, possibly by reordering the tuple (giving us the factor of $n!$) we can assume the first r vectors are linearly independent, and given $v_1, \ldots, v_i$ with $0 \le i < r$, the number of choices for a v_{i+1} which is linearly independent of $v_1, \ldots, v_i$ is exactly 2^{n-i}. After this every vector must lie in the span of $v_1, \ldots, v_r$, giving exactly 2^r choices for the remaining $n-r$ vectors.

We next claim that there exists no tuple in X of rank larger than $n/2$. Indeed, if S is the span of the vectors in a tuple of X, then note that $S \subseteq S^\perp$ since $v_i \cdot v_j = 0$ for all i, j. From linear algebra we have $n = \dim S + \dim S^\perp \ge 2 \dim S$, proving the claim.

It is not hard to prove that (5) is increasing for $r \le n/2$, so plugging in $r = n/2$ and using the pigeonhole principle gives an upper bound for $|X|/(n/2)$ of the form $2^{5n^2/8+o(n^2)}$, giving the desired bound on $|X|/|V|^n$. □

Putting all these lemmas together gives the following.

Corollary 6.15. *For $q \ge 2$ we have*

$$R_q(n) \ge \left(2^{\frac{3q}{8}-\frac{1}{4}}\right)^{n-o(n)}.$$

This bound stood as the best for about a year until Sawin [?] realized one could do somewhat better by replacing the algebraic graph G described above with a purely random graph, namely $G_{n,p}$ with $p \approx .455$. Thus, although the initial breakthrough for multicolor Ramsey numbers came from a random algebraic approach, the method was later subsumed by a simpler random model. This sort of thing happens somewhat often with proofs using the random algebraic method. Because of this, some mathematicians are of the opinion that any time the random algebraic method is used, there exists a simpler random model which gives better results. I don't personally believe that this is true, and even if it were, the fact that random algebraic methods consistently give initial breakthroughs to longstanding open problems makes them worth considering in my eyes.

6.4 Exercises

1. Prove that $e(G') > \frac{1}{2}e(G)$ whenever $e(G) > 0$ [2].

2. We can use probability to give another proof of Turán's Theorem.

   (a) (Caro-Wei) Prove that if G is an n-vertex, then

   $$\alpha(G) \ge \sum_{x \in V(G)} \frac{1}{\deg(x)+1}.$$

   (Hint: construct a random independent set I in such a way that $\Pr[x \in I] = \frac{1}{\deg(x)+1}$) [2+].

   (b) Conclude $\mathrm{ex}(n, K_r) \le (1 - \frac{1}{r-1})\frac{n^2}{2}$, which asymptotically matches Turán's Theorem [1+].

3. Let G be a graph with m edges and N copies of a graph F.

   (a) Prove (without using probability) that G contains an F-free subgraph with at least $m - N$ edges. [1+]

   (b) Prove that if $N \geq m$ and $e(F) \geq 2$, then G contains an F-free subgraph with at least

   $$\Omega\left(\frac{m^{1+\frac{1}{e(F)-1}}}{N^{\frac{1}{e(F)-1}}}\right)$$

   edges [2].

   c What does this result imply when $G = K_n$? [1]

   We note that our intended proof of (b) uses probability to "boost" the weak deterministic bound from (a) to get a substantially stronger bound. This is a common application of the probabilistic method, as the next exercise aims to show as well.

4. Prove that there exists $\varepsilon > 0$ such that $R(3, n) = \Omega(n^{1+\varepsilon})$. What's the best value of ε you can find using this method? [2+].

5. Given a graph G, define the crossing number $cr(G)$ to be the minimum number of crossing pairs of edges in any embedding of G into $\mathbb{R}^2$. For example, $cr(G) = 0$ if and only if G is planar.

   (a) Prove (without using probability) that $cr(G) \geq e(G) - 3v(G)$ (Hint: use Euler's formula) [2].

   (b) (Crossing lemma) Prove that there exists some $C > 0$ such that if G is an n-vertex graph with $m \geq Cn$ edges then

   $$cr(G) = \Omega\left(\frac{m^3}{n^2}\right).$$

   [2+]

* * *

6. Prove Theorem 6.4 [1+].

7. Possibly add some more related to the new subsections I added.

7 Linear Algebra Methods

Roughly speaking, the *linear algebra method* in combinatorics works as follows:

1 Associate a "linear algebraic object" M to your problem (e.g. a matrix or a list of vectors).

2. Determine algebraic information about M (e.g. its rank, eigenvalues, eigenvectors),

3. Use this algebraic information to conclude something about your original problem.

The linear algebra method applies to a broad range of problems. We only scratch the surface here, and we refer the reader to books by Babai and Frankl and by Matoušek for a more thorough treatment of this versatile method.

7.1 Introduction to Spectral Graph Theory

Within the context of graph theory, perhaps the most natural linear algebraic object to consider is the adjacency matrix. To this end, given a graph G we define its adjacency matrix $A(G)$ to be the symmetric matrix whose rows and columns are indexed by $V(G)$ where $A(G)_{u,v} = 1$ if $u \sim v$ in G and $A(G)_{u,v} = 0$ otherwise. We write A instead of $A(G)$ whenever G is clear from context. Insert picture of an example.

A priori, A is simply a convenient way to encode the graph G, and as such there is no reason to really study A as a linear operator. Surprisingly, the algebraic properties of A contains a tremendous amount of combinatorial information about G. Because of this, there is a large area known as *spectral graph theory* which centers around studying algebraic properties of both A as well as many other types of matrices that can be associated to graphs. We will only get a glimpse of this area here and refer the reader to Section 13 for more on this.

We begin with a classical connection between A and combinatorial properties of G, namely that of closed walks. For this, we recall that a sequence of vertices $(w_1, \ldots, w_{k+1})$ of a graph G is called a *walk* of length k if $w_{i+1} \in N(w_i)$ for all $1 \le i \le k$, and we say that this walk is *closed* if $w_1 = w_{k+1}$. For the rest of this section, we make frequent use of the standard linear algebra fact that every real symmetric matrix (such as A) with n rows and columns has n real eigenvalues as well as a orthonormal basis of eigenvectors.

Lemma 7.1. *If G is an n-vertex graph and if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of its adjacency matrix A, then the number of closed walks of G of length k equals $\sum_i \lambda_i^k$.*

Proof. We first observe that if G is a graph and if $u, v \in V(G)$, then the number of walks of length k from u to v is $A^k_{u,v}$. Indeed, by definition of matrix multiplication, we have

$$A^k_{u,v} = \sum A_{uw_2} \cdots A_{w_k v},$$

where the sum ranges over all sequences of vertices $w_2, \ldots, w_k$. A given term of this sum will be 1 if th sequence $(u, w_2, \ldots, w_k, v)$ defines a walk and will be 0 otherwise, showing that $A^k_{u,v}$ is the desired amount.

From this observation, we see that the number of closed walks of length k is exactly

$$\sum_{u \in V(G)} A^k_{u,u} = \mathrm{Tr}(A^k) = \sum_i \lambda_i^k,$$

where here the first equality used the definition of the trace of a matrix, and the second equality used both that the trace of a square matrix equals the sum of its eigenvalues and that raising a square matrix to a power k raises all of its eigenvalues to the power k as well. □

The first non-trivial cases of $k = 2, 3$ of Theorem 7.1 gives the following.

Corollary 7.2. *If G is an n-vertex graph and if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of its adjacency matrix A, then*

$$2e(G) = \sum_i \lambda_i^2,$$

and

$$6t(G) = \sum_i \lambda_i^3,$$

where here $t(G)$ denotes the number of triangles of G.

Proof. Observe that (w_1, w_2, w_3) is a closed walk of G of legnth 2 if and only if it is of the form (u, v, u) with $uv \in E(G)$. It follows that the number of closed walks of length 2 is exactly $2e(G)$ (since there is one for each orientation of each edge), giving the first result by the previous lemma.

For the second result, we observe that a sequence of vertices (u, v, w, u) defines a closed walk of length 3 in G if and only if u, v, w are distinct and all adjacent to each other, i.e. if and only if the vertices u, v, w form a triangle. Moreover, each triangle contributes 6 such walks (we have 3 choices for which vertex of the triangle to start on and then 2 choices for the second vertex), giving the second result. □

In addition to using Theorem 7.1 to determine combinatorial information about G from A, we can also use it to gain algebraic information about A from G.

Corollary 7.3. *If G is a non-empty graph and if $\lambda_{\max}, \lambda_{\min}$ are the largest and smallest eigenvalues of A, then $\lambda_{\max} > 0 > \lambda_{\min}$.*

Proof. By Theorem 7.1 (or simply by definition of A and the trace), we have that

$$\sum \lambda_i = 0.$$

On the other hand, we have

$$\sum \lambda_i^2 = 2e(G) > 0.$$

These two statements are only possible if there exists some eigenvalue which is positive and some eigenvalue which is negative, proving the result. □

The statements above all hold for arbitrary graphs, but more can be said for particular kinds of graphs. In particular, spectral graph theory tends to be at its strongest for regular graphs due to the following key observation.

Lemma 7.4. *If G is a d-regular graph, then the all 1's vector $\mathbf{1}$ is an eigenvector of A with eigenvalue d.*

Proof. For all $u \in V(G)$ we have

$$(A\mathbf{1})_u = \sum_v A_{u,v}\mathbf{1}_v = \sum_{v \in N(u)} 1 = d = d\mathbf{1}_u,$$

proving that $A\mathbf{1} = d\mathbf{1}$. □

One famous application of spectral graph theory to regular graphs comes from Hoffman's bound for the independence number of G.

Theorem 7.5 (Hoffman's Ratio Bound)**.** *Let G be a non-empty n-vertex d-regular graph and A its adjacency matrix. Then*

$$\alpha(G) \le \frac{-\lambda_{\min}}{d - \lambda_{\min}} \cdot n,$$

where $\lambda_{\min}$ denotes the smallest eigenvalue of A.

Maybe spell out some of the linear algebra details a bit more here and also maybe give more intuition for the steps, eg emphasize the idea (maybe even as a mantra) of using characteristic vectors and about using eigenbasis to analyze size of things, etc..

For example, if $G = K_n$ then one can check that the eigenvalues of A are $d = n-1$ with multiplicity 1 and -1 with multiplicity $n-1$. As such, the Ratio Bound gives an upper bound of $\alpha(K_n) \le 1$ which is best possible.

Proof. Let I be an independent set of size $\alpha := \alpha(G)$ and let x be the vector indexed by $V(G)$ with $x_u = 1$ if $u \in I$ and $x_u = 0$ otherwise. Observe that because I is an independent set, we have

$$x^T A x = \sum_{u,v} x_u A_{u,v} x_v = \sum_{u,v \in I} A_{u,v} = 0.$$

Let $y_1, \ldots, y_n$ be an orthonormal eigenbasis for A with eigenvalues $\lambda_1, \ldots, \lambda_n$. Since G is regular, the all 1's vector $\mathbf{1}$ is an eigenvector with eigenvalue d, so we can assume $y_1 = \mathbf{1}/\sqrt{n}$ and $\lambda_1 = d$. Writing $x = \sum c_i y_i$ for some real numbers c_i, we see that

$$\alpha = x^T x = \sum c_i^2,$$

and

$$\alpha/\sqrt{n} = \langle x, y_1 \rangle = c_1.$$

Putting all of this together, we find

$$0 = x^T Ax = x^T \sum c_i\lambda_i y_i = \sum c_i^2\lambda_i = (\alpha^2/n)d + \sum_{i\neq 1} c_i^2\lambda_i$$
$$\geq (\alpha^2/n)d + \sum_{i\neq 1} c_i^2\lambda_{\min} = (\alpha^2/n)d + (\alpha - \alpha^2/n)\lambda_{\min}.$$

Dividing both sides by α and rearranging gives

$$\alpha(\lambda_{\min} - d)/n \geq \lambda_{\min}.$$

Dividing both sides by $(\lambda_{\min} - d)/n$ (which is negative because $\lambda_{\min} < 0$ since G is non-empty) gives the result. $\square$

Hoffman's Ratio Bound is effective for a number of graphs. In particular, one can use this to prove the Erdős-Ko-Rado Theorem, which is the fundamental theorem of extremal set theory.

Theorem 7.6 (Erdős-Ko-Rado)**.** *Let $\mathcal{F} \subseteq \binom{[n]}{r}$ be a collection of r-element subsets of $[n]$ which is intersecting, i.e. which is such that $F \cap F' \neq \emptyset$ for all $F, F' \in \mathcal{F}$. If $n \geq 2r$, then*

$$|\mathcal{F}| \leq \binom{n-1}{r-1}.$$

Note that this result is best possible by considering $\mathcal{F}$ to consist of all sets containing the element 1.

Sketch of Proof. Define an auxiliary graph G which has vertex set $\binom{[n]}{r}$ where we have $F \sim F'$ if and only if $F \cap F' = \emptyset$. From this, we see that a family $\mathcal{F}$ is intersecting if and only if it is an independent set of G. One can show that G has $\binom{n}{r}$ vertices, that it is regular with degree $\binom{n-r}{r}$, and (less trivially via using $n \geq 2r$) that the smallest eigenvalue of its adjacency matrix equals $-\frac{r}{n-r}\binom{n-r}{r}$. In total this implies that any intersecting family $\mathcal{F}$ satisfies

$$|\mathcal{F}| \leq \alpha(G) \leq \binom{n}{r} \frac{\frac{r}{n-r}\binom{n-r}{r}}{\binom{n-r}{r} - \frac{r}{n-r}\binom{n-r}{r}} = \frac{r}{n-r}\binom{n}{r} = \binom{n-1}{r-1},$$

proving the result. $\square$

7.2 Beyond the Adjacency Matrix

While the adjacency matrix is perhaps the most natural matrix to associate to a graph G, there are many different types of matrices that could be considered which each have their own sets of advantages and disadvantages. In particular, many results and proofs for the adjacency matrix continue to hold word for word for a slightly broader class of matrices which can sometimes be useful to consider. To illustrate this fact, we consider another classical result relating the eigenvalues of A to combinatorial properties of G.

Lemma 7.7. *For any graph G, the largest eigenvalue $\lambda_{\max}$ of the adjacency matrix A satisfies*

$$\lambda_{\max} \leq \Delta(G).$$

Proof. Let x be an eigenvector of A corresponding to $\lambda_{\max}$ and let $v \in V(G)$ be such that $|x_v|$ is maximized. Then by our definitions, we have

$$|\lambda_{\max} x_v| = |(Ax)_v| = |\sum_u A_{v,u} x_u| \leq \sum_{u \sim v} |x_u| \leq \deg(v)|x_v| \leq \Delta(G)|x_v|.$$

This shows $|\lambda_{\max}| \leq \Delta$, proving the result. □

Examining this proof, we see that we hardly used any of the properties of A in our argument. In particular, word for word the same argument gives the following.

Lemma 7.8. *Let G be a graph and M any symmetric matrix such that $M_{u,v} = \pm 1$ if $uv \in E(G)$ and $M_{u,v} = 0$ otherwise. Then the largest eigenvalue $\lambda_{\max}$ of M satisfies*

$$\lambda_{\max} \leq \Delta(G).$$

A priori it's not clear whether Theorem 7.8 is actually interesting or if it is just a generalization for generalization's sake. Surprisingly, this result plays a key role in a beautiful proof of Huang's solving a 30 year problem known as the sensitivity conjecture. Recall here or elsewhere that Q_n has vertex set $\{0,1\}^n$ with two bistrings being adjacent if they differ in exactly 1 position.

Theorem 7.9 (Huang [?])**.** *Let Q_n be the hypercube graph on 2^n vertices. If $V \subseteq V(Q_n)$ is a subset of size $2^{n-1}+1$, then the induced subgraph $Q_n[V]$ has maximum degree at least $\sqrt{n}$.*

This result is sharp in several ways. First, it is easy to find subsets of size 2^{n-1} such that $Q_n[V]$ is the empty graph, so in order to get any non-trivial lower bound on the maximum degree one needs V to have size at least $2^{n-1}+1$. Second, Chung et. al. [?] proved that there exist choices of V such that $Q_n[V]$ has maximum degree $\lceil \sqrt{n} \rceil$, so this bound is essentially best possible.

It was shown by Gotsman and Linial [?] that proving a result of this form is equivalent to showing that two notions of "sensitivity" for Boolean functions are equivalent, which led to a great deal of interest in resolving it. Nevertheless, it remained unanswered for 30 years until Huang came up with the following remarkable proof.

The key idea is to define the $2^n \times 2^n$ matrix B_n recursively by

$$B_0 = [0], \qquad B_n = \begin{bmatrix} B_{n-1} & I \\ I & -B_{n-1} \end{bmatrix},$$

where here I denotes the identity matrix of dimension 2^{n-1}. Observe that if the negative sign in the definition of B_n wasn't there, then this would just define the adjacency matrix of Q_n. Thus this is a sort of "twisted adjacency matrix" which has -1's in some of the positions where there are usually 1's. This choice of signings turns out to spread out the spectrum of B_n in a nice way.

Lemma 7.10. *The spectrum of B_n consists of $\pm\sqrt{n}$ each occurring with multiplicity 2^{n-1}.*

Proof. It is straightforward to prove by induction that $B_n^2 = nI$, which implies that every eigenvalue λ of B_n satisfies $\lambda^2 = n$. Thus $\sigma(B_n)$ consists of $\pm\sqrt{n}$, and each must appear with equal multiplicity because $\mathrm{Tr}(B_n) = 0$. □

We will also need a basic fact from linear algebra.

Theorem 7.11 (Cauchy interlacing theorem). *Let B be a real symmetric $n \times n$ matrix and C an $m \times m$ principal sumbmatrix of B with $m \leq n$. If B has eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ and C has eigenvalues $\mu_1 \geq \cdots \geq \mu_m$, then for all i*

$$\lambda_i \geq \mu_i \geq \lambda_{i+n-m}.$$

We note that one easy way to remember these inequalities is by considering the case when B is diagonal, in which case these upper bounds for μ_i are achieved if C drops the $n - m$ smallest diagonal entries of B, while dropping the $n - m$ largest diagonal entries makes it so that all of the lower bounds are achieved.

Shockingly, we have everything we need for our proof of Huang's Theorem.

Proof of Theorem 7.9. Let $B = B_n$ be as described above. Let $V \subseteq V(Q_n)$ be any subset of size $2^{n-1} + 1$ and let C be the submatrix of B indexed by the rows and columns corresponding to V. Let $G = Q_n[V]$. Observe that C satisfies the conditions for M of Lemma 7.8 since B is a (symmetrically) signed version of the adjacency matrix. By Lemma 7.8, the Cauchy interlacing theorem, and the previous lemma, we conclude that

$$\Delta(G) \geq \lambda_1(C) \geq \lambda_{2^{n-1}}(B) = \sqrt{n},$$

proving the result. □

7.3 Beyond Matrices

The linear algebra method extends far beyond just using eigenvalues of matrices to solve problems. To illustrate this, we briefly look at what is perhaps the most famous application of the linear algebra method, though this will require us to briefly leave the world of graph theory and enter the related world of extremal set theory.

Consider the following (somewhat whimsical) setup. The city of Oddtown has a number of clubs, each of which follows the following odd set of rules: each club must have an odd number of people, and every two distinct clubs must have an even number of people in common.

The main question now becomes: if Oddtown has n people, what's the maximum number of clubs it can have? Equivalently, if $\mathcal{F} = \{F_1, F_2, \ldots, F_m\} \subseteq 2^{[n]}$ is a set system such that $|F_i|$ is odd for all i and such that $|F_i \cap F_j|$ is even for all $i \neq j$, then what is the maximum size of $\mathcal{F}$?

A very simple construction is to take $F_i = \{i\}$ for all i, which trivially satisfies the stated conditions. However, it is far from the only construction. For example, if n is even, then one can also take each F_i to be either $\{i\}$ or $[n] \setminus \{i\}$, and there are many, many more constructions achieving a bound of n (in fact, there's close to 2^{n^2} non-isomorphic constructions due to Szegedy [**?**, Exercise 1.1.14]).

Given all of these constructions, it seems plausible that (1) the true answer is indeed n, and (2) proving this might be difficult (since we have to come up with an argument that somehow deals with all of these constructions in a unified way). Fortunately, the linear algebra method manages to give a unified approach for all of these constructions in an extremely elegant way. More generally, if a given problem has many distinct looking extremal constructions, then it is often the case that the linear algebra method can come in handy.

Theorem 7.12 (Oddtown)**.** *Let $\mathcal{F} \subseteq 2^{[n]}$ be a set system such that $|F|$ is odd for all $F \in \mathcal{F}$ and such that $|F \cap F'|$ is even for all $F \neq F' \in \mathcal{F}$. Then $|\mathcal{F}| \leq n$.*

Proof. Given a set $F \subseteq [n]$, define its characteristic vector $\chi_F \in \mathbb{F}_2^n$ by having $(\chi_F)_i = 1$ if $i \in F$ and $(\chi_F)_i = 0$ otherwise. Note crucially that for any F, F', the dot product satisfies

$$\langle \chi_F, \chi_{F'} \rangle = |F \cap F'| \mod 2.$$

We claim that $\{\chi_F : F \in \mathcal{F}\}$ is a set of linearly independent vectors. Indeed, say we had

$$\sum_{F \in \mathcal{F}} \lambda_F \chi_F = 0.$$

Take any $F' \in \mathcal{F}$ and apply the dot product on both sides to get

$$\sum_{F \in \mathcal{F}} \lambda_F \langle \chi_F, \chi_{F'} \rangle = 0.$$

By the observation above and the hypothesis of the theorem, we see $\langle \chi_F, \chi_{F'} \rangle = 0$ if $F \neq F'$ and that $\langle \chi_{F'}, \chi_{F'} \rangle = 1$. Thus the above says $\lambda_{F'} = 0$, and as $F' \in \mathcal{F}$ was arbitrary, we conclude that these vectors are indeed linearly independent.

Since we have $|\mathcal{F}|$ linearly independent vectors in $\mathbb{F}_2^n$, we must have $|\mathcal{F}| \leq n$, giving the result. □

While the above technically is a proof without the use of matrices, we note that one can write an essentially equivalent proof in the language of matrices. However, for many generalizations of oddtown, the most natural way to use this argument is through the language of vectors (with these vectors typically being some set of low degree polynomials). We will maybe explore this further in the exercises.

7.4 Exercises

Throughout this and maybe earlier we define the spectrum $\sigma(M)$ of a real symmetric matrix M to be the multiset of eigenvalues of A and we let $\lambda_{\max}, \lambda_{\min}$ denote the largest and smallest eigenvalues of A.

1. Prove that if G is connected and has diameter d, then A has at least $d + 1$ distinct eigenvalues (Hint: it suffices to show that the minimum polynomial of A has large degree) [2].

2. Prove that if G is a graph with average degree $\bar{d}$, then $\lambda_{\max} \geq \bar{d}$ [2-].

3. (Wilf's Theorem) Prove that if G is a graph, then $\chi(G) \leq \lambda_{\max} + 1$; note that by Theorem 7.7 this bound is always at least as strong as the classic bound $\chi(G) \leq \Delta(G) + 1$ (Hint: prove this by induction on $v(G)$ via using the previous problem) [2+]. I can't really ask this without recalling the Raleigh quotient.

4. (Hoffman's Bound for the Chromatic Number) Prove that if G is a graph, then

$$\chi(G) \geq 1 - \frac{\lambda_{\max}}{\lambda_{\min}}.$$

Note that this result is immediate from Hoffman's Ratio Bound if G is d-regular (assuming the easy to prove fact that $d = \lambda_{\max}$), so the difficulty is in proving this for non-regular graphs [3-].

5. Bipartite graphs turn out to have nice characterizations in terms of their spectrum.

   (a) We say that a matrix M has spectrum symmetric about 0 if the number of eigenvalues it has equal to λ is the same as the number of eigenvalues it has equal to $-\lambda$ for all λ.

   Prove that a graph is bipartite if and only if the spectrum of its adjacency matrix is symmetric about 0 [2].

   (b) Prove that a connected graph is bipartite if and only if $\lambda_{\min} = -\lambda_{\max}$ [2].

6. Prove that if A is the adjacency matrix of $K_{s,t}$, then $\sigma(A)$ has eigenvalues equal to $\sqrt{st}$ and $-\sqrt{st}$ with the rest equal to 0 [1+].

7. Prove that there exist two graphs G_1, G_2 with adjacency matrices A_1, A_2 such that $\sigma(A_1) = \sigma(A_2)$ and such that G_1 is connected while G_2 is not connected (Hint: there exist examples where G_1, G_2 have 5 vertices each) [2-].

   In general, two graphs with $\sigma(A_1) = \sigma(A_2)$ are called *cospectral*. Such graphs are important in spectral graph theory since they tell us the limitations of what can be determined by the spectrum of the adjacency matrix. For example, this result shows that one can not determine whether G is connected or not from $\sigma(A)$ alone.

* * *

8. There are various ways to generalize Hoffman's bound, here's one direction which changes how we measure the "size" of an independent set. Given a graph G, a vector x indexed by $V(G)$, and a set of vertices I, define $|I|_x = \sum_{i \in I} x_i^2$, and define $\alpha_x(G) = \max_I |I|_x$ where I ranges over all independent sets of G.

   Prove that if G is a graph and if M is a (not necessarily symmetric) matrix with rows and columns indexed by $V(G)$ such that $M_{u,v} = 0$ whenever $u \not\sim v$ and such that M has a basis of eigenvectors. If $\lambda_{\min}$ is the smallest eigenvalue of M, and if x is a unit eigenvector of M with eigenvalue $\lambda > \lambda_{\min}$, then

$$\alpha_x(G) \le \frac{-\lambda_{\min}}{\lambda_{\min} - \lambda}$$

[1+].

We note that this result can be used to prove a variant of the Erdős-Ko-Rado theorem, see reference.

9. Determine a larger class of matrices for which the results from the first couple of problems hold, eg diameter.

8 Supersaturation

Need to move stability discussion to its own chapter now

Recall that the Turán number $\mathrm{ex}(n, F)$ is the maximum number of edges an n-vertex graph can have without containing any copies of a given graph F. There are a number of "refinements" of the Turán number that one can consider, with two of the most important being the following:

- (Supersaturation Problem) What is the minimum number of copies of a graph F that a graph with n vertices and m edges must have?

- (Stability Problem) What does an n-vertex F-free graph with $m \approx \mathrm{ex}(n, F)$ edges "look like"?

The Supersaturation Problem contains the Turán problem since the answer to the supersaturation problem is 0 if and only if $m \le \mathrm{ex}(n, F)$. As such, the Supersaturation Problem is interesting only in the range $m > \mathrm{ex}(n, F)$. In contrast, the Stability Problem only makes sense for $m \le \mathrm{ex}(n, F)$.

Supersaturation and Stability results are nice not only because they are interesting in their own right, but also because one can often find applications of these results to other problems of interest. We explore some of these results and applications in the coming subsections.

8.1 Supersaturation for General Graphs and Erdős-Stone-Simonovits

We will begin by discussing supersaturation results for general graphs F when our host graph G is relatively dense, and as an application of this we will give another proof of the Erdős-Stone-Simonovits Theorem. To start we prove a very weak version of the Erdős-Stone-Simonovits Theorem.

Lemma 8.1. *For every graph F, the limit*

$$\pi(F) := \lim_{n\to\infty} \frac{\mathrm{ex}(n, F)}{\binom{n}{2}}$$

exists, and we refer to this limit as the Turán density of F.

Note that the Erdős-Stone-Simonovits Theorem says the stronger fact that the Turán density is equal to $\frac{\chi(F)-2}{\chi(F)-1}$.

Proof. Our proof uses a "local averaging" argument which sums over the total number of edges appearing in induced subgraphs of G of a given size. To this end, observe that for any n-vertex graph G we have

$$\sum_{v\in V(G)} e(G-v) = (n-2)e(G),$$

since each edge xy is counted by exactly $n-2$ terms in this sum. In particular, if G is an extremal F-free graph then this equality together with the observation $\mathrm{ex}(n-1,F) \geq e(G-v)$ gives

$$n\mathrm{ex}(n-1,F) \geq (n-2)\mathrm{ex}(n,F),$$

which for $n \geq 3$ is equivalent to

$$\frac{\mathrm{ex}(n-1,F)}{\binom{n-1}{2}} \geq \frac{\mathrm{ex}(n,F)}{\binom{n}{2}}.$$

In total this implies that the sequence of numbers $\mathrm{ex}(n,F)/\binom{n}{2}$ for $n \geq 3$ is a non-increasing sequence which is trivially bounded below by 0, which from basic real analysis implies that its limit $\pi(F)$ exists, proving the result. □

A similar local averaging argument allows us to prove the following supersaturation result.

Theorem 8.2. *For every graph F and $\varepsilon > 0$, there exists some $\delta, m > 0$ such that any n-vertex graph G with $n \geq m$ and $e(G) \geq (\pi(F)+\varepsilon)\binom{n}{2}$ contains at least $\delta n^{v(F)}$ copies of F.*

Note that any n-vertex graph trivially contains at most $n^{v(F)}$ copies of F, so this result says that any graph which has asymptotically more edges than the Turán number for F contains a constant proportion of all possible copies of F.

Proof. The intuition for our argument is as follows: we will take some large (but fixed) integer m and look at all of the induced subgraphs of G on m vertices. The theorem statement suggests that a constant proportion of these induced subgraphs should have a copy of F, and this certainly holds if the induced subgraph has more than $\mathrm{ex}(m,F) \approx \pi(F)\binom{m}{2}$ edges. If this fails for most induced subgraphs, then this will imply that $e(G)$ can't be much more than $\pi(F)\binom{n}{2}$ since most of its induced subgraphs have few edges, a contradiction to our hypothesis.

To make this precise, let m be an integer such that $\mathrm{ex}(m,F) \leq (\pi(F)+\frac{\varepsilon}{4})\binom{m}{2}$, which exists by definition of $\pi(F)$, and from now on we assume $n \geq m$.

Claim 8.3. *There exist at least $\frac{\varepsilon}{4}\binom{n}{m}$ m-element subsets V such that $e(G[V]) > (\pi(F)+\frac{\varepsilon}{4})\binom{m}{2}$.*

Proof. If this were not the case then

$$\binom{n-2}{m-2}e(G) = \sum_{V\in\binom{V(G)}{m}} e(G[V]) \leq \left(1-\frac{\varepsilon}{4}\right)\binom{n}{m}\cdot\left(\pi(F)+\frac{\varepsilon}{4}\right)\binom{m}{2} + \frac{\varepsilon}{4}\binom{n}{m}\cdot\binom{m}{2}$$
$$\leq \left(\pi(F)+\frac{\varepsilon}{2}\right)\binom{n}{m}\binom{m}{2}.$$

Using the combinatorial identity $\binom{n}{m}\binom{m}{2} = \binom{n}{2}\binom{n-2}{m-2}$ we see that this implies $e(G) \leq (\pi(F) + \frac{\varepsilon}{2})\binom{n}{2}$, a contradiction. □

By definition of m, each of the $\frac{\varepsilon}{4}\binom{n}{m}$ induced subgraphs $G[V]$ as in this claim contain a copy of F. Observe that each copy of F is contained in at most $\binom{n-v(F)}{m-v(F)}$ of these induced subgraphs, so in total the number of copies of F is at least

$$\varepsilon/4\binom{n}{m}\binom{n-v(F)}{m-v(F)}^{-1} = \Omega_\varepsilon(n^{v(F)}),$$

proving the result. □

Can maybe note how the exact supersaturation function is some weird saw tooth-fractally curve which differs in what's predicted for bipartite.

This result can be used to show that Turán densities are preserved after taking a certain blowup operation.

Theorem 8.4. *Given a graph F and an integer t, define the blowup $F[t]$ to be the graph obtained by replacing each vertex $v_i \in F$ with an independent set V_i and by replacing each edge $v_iv_j \in E(F)$ with a complete bipartite graph between V_i and V_j. For all F and $t \geq 1$, we have*

$$\pi(F[t]) = \pi(F).$$

This result quickly implies Erdős-Stone-Simonovits: if F is a t-vertex graph with chromatic number r, then $F \subseteq K_r[t]$, and hence

$$\pi(F) \leq \pi(K_r[t]) = \pi(K_r) = \frac{r-2}{r-1},$$

with the first equality using this theorem and the second using Turán's Theorem. On the other hand, the Turán graph shows $\pi(F) \geq \frac{r-2}{r-1}$, so this gives $\pi(F)$ for all graphs.

To prove this, we need a moderate generalization of the Kővári-Sós-Turán Theorem.

Lemma 8.5. *For all $r, t \geq 1$, there exists a constant $C_{r,t}$ such that for any r-vertex graph F, if G is an n-vertex graph with more than $4tn^{r-t^{-r+1}}$ copies of F, then G contains a copy of $F[t]$.*

Note that when $F = K_2$ this simply says that the Turán number of $K_{t,t}$ is at most $O(n^{2-1/t})$, and more generally in the language of the generalized Turán number this says that $\mathrm{ex}(n, F, F[t]) = O(n^{r-t^{-r+1}})$.

Proof. For ease of notation we prove the result only for $F = K_r$, though the exact same sort of argument will go through for general F. We prove the result by induction on r, the case $r = 1$ being trivial. The case $t = 1$ is vacuously true, so we will assume $t \geq 2$ throughout.

Let G be a graph as in the hypothesis and assume for contradiction that G is $K_r[t]$-free. Let $\mathcal{P}$ consist of the set of pairs (T, K) where K is a copy of K_{r-1} in G and T is a set of t vertices which are all adjacent to every vertex of K. Observe that no set T can be in more than $4tn^{r-1-t^{-r+2}}$ pairs, since it were then by induction the set of at least this many K_{r-1}'s forming a pair with T would contain a copy of $K_{r-1}[t]$ which together with T would form a copy $K_r[t]$ in G, a contradiction. Because there are at most n^t choices for T, this implies

$$|\mathcal{P}| \leq 4tn^{r-1-t^{-r+2}+t}.$$

For each K which is a copy of K_{r-1}, let $\deg(K)$ denote the number of K_r's that K is in. It follows from this and convexity that

$$|\mathcal{P}| = \sum_K \binom{\deg(K)}{t} \ge n^{r-1}\binom{n^{1-r}\sum_K \deg(K)}{t}.$$

Now observe that $\sum_K \deg(K)$ is equal to r times the total number of K_r's in G, so by hypothesis we have that

$$n^{r-1}\binom{n^{1-r}\sum_K \deg(K)}{t} \ge n^{r-1}\binom{4tn^{1-t^{-r+1}}}{t} \ge n^{r-1}(2tn^{1-t^{-r+1}})^t = (2t)^t n^{r-1+t-t^{-r+2}},$$

where here this last inequality used that $\binom{\alpha}{t} \ge (\alpha-t)^t \ge (\frac{1}{2}\alpha)^t$ for any $\alpha \ge 2t$. This lower bound for $|\mathcal{P}|$ is strictly greater than our upper bound under our assumption of $t \ge 2$, giving the desired contradiction. □

Proof of Theorem 8.4. Because $F \subseteq F[t]$, we trivially have $\pi(F) \le \pi(F[t])$ for all t. We aim to show that for all $\varepsilon > 0$, there exists some n_0 such that every n-vertex $F[t]$-free graph G with $n \ge n_0$ and $e(G) \ge (\pi(F)+\varepsilon)\binom{n}{2}$ contains a copy of F, from which it will follow that $\pi(F[t]) \le \pi(F)$, proving the result.

Fix $\varepsilon > 0$ and let n_0 be a large integer to be determined later. If G is an n-vertex $F[t]$-free graph with $n \ge n_0$ and $e(G) \ge (\pi(F)+\varepsilon)\binom{n}{2}$, then by Theorem 8.2, G contains at least $\delta n^{v(F)}$ copies of F. If $n \ge n_0$ is sufficiently large then $\delta n^{v(F)} > 4tn^{v(F)-t^{-v(F)+1}}$, implying by Theorem 8.5 that G contains a copy of $F[t]$, a contradiction. We conclude that $\pi(F[t]) \le \pi(F)$ as desired. □

As an aside, we note that everything we've done here holds with almost identical proofs for the setting of r-uniform hypergraphs.

8.2 Supersaturation for Bipartite Graphs and Sidorenko's Conjecture

Theorem 8.2 shows that any graph G with more than $\mathrm{ex}(n,F)+\varepsilon n^2$ edges contains many copies of F, but this often does not say much for bipartite graphs which all have $\mathrm{ex}(n,F)=o(n^2)$. In general not much is known about supersaturation for general bipartite graphs, which is in part due to the fact that we do not know what $\mathrm{ex}(n,F)$ is for general bipartite graphs F, making it unclear what range of values for $e(G)$ we should be looking at for this problem. Nevertheless, a striking conjecture of Erdős and Simonovits makes a guess for what the answer should be for arbitrary bipartite graphs.

Conjecture 8.6 (Erdős-Simonovits Supersaturation Conjecture I)**.** *For every graph F and $c > 0$, there exists some $c' > 0$ such that if G is an n-vertex graph with $e(G) > (1+c)\mathrm{ex}(n,F)$ then G contains at least*

$$c'\left(\frac{e(G)}{n^2}\right)^{e(F)} n^{v(F)}$$

copies of F.

The motivation for this is that the random graph $G_{n,p}$ with $p = e(G)/n^2$ has about $e(G)$ edges with high probability and in expectation has about $p^{e(F)}n^{v(F)} = (e(G)/n^2)^{e(F)}n^{v(F)}$ copies of F. As such, this conjecture essentially says that if G is any graph with slightly more than $\text{ex}(n, F)$ edges, then it should contain about as many copies of F as we would expect to see in a random graph of the same density.

Theorem 8.6 is very wide open and, given just how strong it is, could easily be false. More progress, however, has been made on the following weakening of Theorem 8.6 which was also made in this same paper of Erdős and Simonovits.

Conjecture 8.7 (Erdős-Simonovits Supersaturation Conjecture II)**.** *If F is bipartite, then there exists some $\alpha > 0$ and C, c' such that if G is an n-vertex graph with $e(G) > Cn^{2-\alpha}$, then it contains at least $c'\left(\frac{e(G)}{n^2}\right)^{e(F)} n^{v(F)}$ copies of F.*

That is, relatively dense graphs should have about as many copies of F as a random graph of the same density. A large body of work has been dedicated to this conjecture, though most of these results are phrased in terms of an equivalent formulation of the conjecture in the language of homomorphisms.

Recall that a homomorphism from a graph F to a graph G is a map $\phi : V(F) \to V(G)$ such that $\phi(x)\phi(y) \in E(G)$ whenever $xy \in E(F)$. We let $\text{Hom}(F, G)$ denote the set of homomorphisms from F to G and $\text{hom}(F, G) := |\text{Hom}(F, G)|$. We define the *homomorphism density* by

$$t_F(G) = \frac{|\text{Hom}(F, G)|}{v(G)^{v(F)}},$$

which equivalently can be interpreted as the probability that a random map $\phi : V(F) \to V(G)$ is a homomorphism. We say that a graph F is *Sidorenko* if for every graph G,

$$t_F(G) \geq t_{K_2}(G)^{e(F)}.$$

Equivalently, F being Sidorenko says for every n-vertex graph G that

$$\text{hom}(F, G) \geq \left(\frac{2e(G)}{n^2}\right)^{e(F)} n^{v(F)},$$

since $\text{hom}(K_2, G) = 2e(G)$. The definition of being Sidorenko is again motivated by thinking about the case when $G = G_{n,p}$, since in this situation we expect to have $t_{K_2}(G) \approx p$ and $t_F(G) \approx p^{e(F)}$. As such, saying a graph F is Sidorenko essentially means that every graph G has at least as many homomorphisms of F as one would expect in a random graph of the same density. Analogous to the supersaturation conjectures of Erdős and Simonovits above, Sidorenko conjectured that this phenomenon should hold for every bipartite graph.

Conjecture 8.8 (Sidorenko's Conjecture)**.** *A graph F is Sidorenko if and only if F is bipartite.*

The fact that non-bipartite graphs are not Sidorenko is an easy exercise, so the difficulty lies in showing bipartite graphs are all Sidorenko.

While the statements of Sidorenko's Conjectures and Theorem 8.7 are spiritually similar, it should perhaps come as a surprise that they are in fact equivalent statements. We will show one direction of this equivalence here and leave the other as an exercise.

Proposition 8.9. *If F is a non-empty bipartite graph which is Sidorenko, then it satisfies the conditions of Theorem 8.7.*

Proof. We will prove that Theorem 8.7 holds with $\alpha = 1/e(F)$ and $C = v(F)^2$. To this end, let G be an n-vertex graph with $e(G) > v(F)^2 n^{2-1/e(F)}$.

Let $\mathrm{Hom}^*(F, G)$ denote the set of homomorphisms from F to G which are injective. Note that each $\phi \in \mathrm{Hom}^*(F, G)$ corresponds to a copy of F in G and that each copy of F corresponds to at most $v(F)!$ injective homomorphisms, so it suffices to prove that $|\mathrm{Hom}^*(F, G)|$ is large.

Because F is Sidorenko, we have by definition that

$$|\mathrm{Hom}(F, G)| = t_F(G) n^{v(F)} \geq t_{K_2}(G)^{e(F)} n^{v(F)} = (2e(G))^{e(F)} n^{v(F)-2e(F)}.$$

On the other hand, the number of homomorphisms from F to G which are not injective is trivially at most, say, $v(F)^2 n^{v(F)-1}$ (since there are at most $v(F)^2$ ways to specify a pair of vertices of F that map to the same vertex, and given this there are at most $n^{v(F)-1}$ ways to specify the image of the map). As such, we have

$$\begin{aligned} |\mathrm{Hom}^*(F, G)| \geq |\mathrm{Hom}(F, G)| - v(F)^2 n^{v(F)-1} &\geq (2e(G))^{e(F)} n^{v(F)-2e(F)} - v(F)^2 n^{v(F)-1} \\ &\geq \frac{1}{2}(2e(G))^{e(F)} n^{v(F)-2e(F)}, \end{aligned}$$

with this last step using $e(G) \geq v(F)^2 n^{2-e(F)}$. This is exactly the bound we aimed to show, proving the result. ☐

The problem of showing that a given bipartite graph F is Sidorenko is quite difficult in general. For example, while it is an easy exercise to prove that stars are Sidorenko, proving that every tree is Sidorenko is fairly difficult to do. That being said, we can give a nice proof showing that even cycles are Sidorenko by using tools from linear algebra.

Theorem 8.10. *Every even cycle $C_{2\ell}$ is Sidorenko.*

Proof. Let G be an n-vertex graph. Observe that the number of homomorphisms from $C_{2\ell}$ to G is exactly equal to the number of closed walks of length 2ℓ in G, and as such by Theorem 7.1 we have

$$|\mathrm{Hom}(C_{2\ell}, G)| = \sum \lambda_i^{2\ell},$$

where λ_i is the ith eigenvalue of the adjacency matrix of G. We will prove the result by giving an effective lower bound on the largest eigenavlue λ_1. For this, we recall the standard linear algebra fact that the largest eigenvalue of a real symmetric matrix A can be characterized by

$$\lambda_1 = \max_x \frac{x^T A x}{x^T x},$$

where the maximum ranges over all vectors x. In particular, taking x to be the all 1's vector shows that

$$\lambda_1 \geq \frac{2e(G)}{n},$$

and hence

$$|\mathrm{Hom}(C_{2\ell}, G) \geq \lambda_1^{2\ell} \geq \left(\frac{2e(G)}{n}\right)^{2\ell} = \left(\frac{2e(G)}{n^2}\right)^{2\ell} n^{2\ell},$$

which is equivalent to saying that $C_{2\ell}$ is Sidorenko. ☐

8.3 Applications of Bipartite Supersaturation

We now give some applications of supersaturation results for bipartite graphs to Turán problems. Our proofs will all involve counting various objects in graphs, and our counting will be most effective whenever our host graph G is close to regular. A general result of Erdős and Simonovits [?] allows us to reduce to the case of being almost regular.

Definition 17. We say a graph G is K-almost regular for some $K > 0$ if $\Delta(G) \leq K \cdot \delta(G)$, i.e. if the degrees of any two vertices of G are within a multiplicative factor of K from each other.

The following lemma, essentially due to Erdős and Simonovits, allows us to more or less always assume the graphs we are working with for bipartite Turán problems are K-almost regular.

Lemma 8.11. *Let $0 < \varepsilon < 1$ and $C \geq 1$ be real numbers. If G is an n-vertex graph with $e(G) \geq Cn^{1+\varepsilon}$ and with n sufficiently large in terms of ε, then there exists a subgraph $G' \subseteq G$ with $e(G') \geq \frac{2}{5}Cv(G')^{1+\varepsilon}$, with $v(G') = \Omega_\varepsilon(n^{\varepsilon\frac{1-\varepsilon}{1+\varepsilon}})$ which is K-almost regular with $K = 40 \cdot 2^{\varepsilon^{-2}}$.*

The exact dependencies on ε here are not so important; the point is that G' has at least the same relative density as G, is K-almost regular for some K depending only on ε, and has a large number of vertices whenever G does.

Proof. We begin by discussing some of the high-level intuition for our forthcoming proof, where for convenience we let $d(G)$ denote the average degree of G. First of all, one can notice that if $\Delta(G) \approx d(G)$, then we can do our usual "iteratively delete vertices of degree at most $d(G)/2$ " to end up with a subgraph G' which has $\delta(G') \approx d(G)$, and which definitionally has $\Delta(G') \leq \Delta(G) \approx d(G)$, proving the result. It remains then to handle the case when $\Delta(G) \gg d(G)$, for which there's two subcases to consider. First, if very few edges of G are incident to vertices of degree much larger than $d(G)$ (e.g. if G has a small number of dominating vertices), then we can simply remove these high-degree vertices, get back a graph with $\Delta(G') \approx d(G)$ and then repeat the argument above of removing low degree vertices to get what we want. The difficult subcase then is if most of the edges of G are incident to high-degree vertices, which is what happens if e.g. $G = K_{s,n-s}$ for s small. For this particular example of G, the best thing we can do is throw out most of the low-degree vertices to end up with a $K_{s,s}$. More generally, we will throw away most of the low degree vertices of G while keeping the high-degree vertices to obtain a new graph $G_1 \subseteq G$. In general G_1 will not satisfy the conditions that we want, but we can then iterate the argument above to either find a subgraph of G_1 solving the problem or some new $G_2 \subseteq G_1$ obtained by deleting low-degree vertices. One can argue that at some point these sequence of G_i graphs must terminate, allowing us to find the desired subgraph. As a final technical aside: while in theory the amount of vertices we throw away at each step as well as the cutoff for what $\Delta(G) \approx d(G)$ means could be chosen in a clever way depending on the structure of G, we will choose these numbers to be the same for every graph of the same order to simplify our analysis.

We now proceed with the formal details. Fix some C, ε as in the lemma statement, and with some foresight we let $t = 2^{1+\varepsilon^{-2}}$ though for now the reader should just think of t as a large number depending on ε. Given an n-vertex graph G we let $B_t(G)$ denote the set of $\lceil n/2t \rceil$ vertices of G which have the largest degrees.

Claim 8.12. *If G is an n-vertex graph with $e(G) \geq Cn^{1+\varepsilon}$ and if the number of edges of G incident to $B_t(G)$ is at most $\frac{1}{2}Cn^{1+\varepsilon}$, then there exists a subgraph $G' \subseteq G$ which is $20t$-almost regular and has $e(G') \geq \frac{2}{5}Cn^{1+\varepsilon}$ and $v(G') \geq \frac{2}{5t}n$.*

Proof. Let $G'' = G - B_t(G)$ and then define G' by iteratively deleting vertices from G'' which have degree at most $\frac{1}{10}Cn^{\varepsilon}$. By construction, we have

$$e(G') \geq e(G'') - \frac{1}{10}Cn^{\varepsilon} \cdot n \geq \frac{1}{2}Cn^{1+\varepsilon} - \frac{1}{10}Cn^{1+\varepsilon} = \frac{2}{5}Cn^{1+\varepsilon},$$

and we also have $\delta(G') \geq \frac{1}{10}Cn^{\varepsilon}$. On the other hand, we claim that every vertex in G' has degree at most $2Ctn^{\varepsilon}$ in G (and hence also in G'). Indeed, if there existed some $v \in V(G') \subseteq V(G) \setminus B_t(G)$ with $\deg_G(v) > 2Ctn^{\varepsilon}$, then this would imply $\deg_G(w) > 2Ctn^{\varepsilon}$ for every $w \in B_t(G)$ by definition, and as such the number of edges in G incident to $B_t(G)$ would be at least

$$\frac{1}{2}\sum_{w \in B_t(G)} \deg_G(w) > \frac{1}{2} \cdot 2Ctn^{\varepsilon}\lceil n/2t \rceil \geq \frac{1}{2}Cn^{1+\varepsilon},$$

a contradiction to the case that we are in. We conclude that every vertex in G' has degree at most $2Ctn^{\varepsilon}$ and hence the graph is $20t$-almost regular. Moreover, we trivially have

$$v(G') \geq 2e(G')/\Delta(G') \geq \frac{2}{5t}n$$

□

Claim 8.13. *If G is an n-vertex graph with $e(G) \geq Cn^{1+\varepsilon}$ and if the number of edges of G incident to $B_t(G)$ is at least $\frac{1}{2}Cn^{1+\varepsilon}$, then there exists a subgraph $G' \subseteq G$ on $2\lceil n/2t \rceil$ vertices with $e(G') \geq \frac{1}{4t}Cn^{1+\varepsilon}$.*

Proof. We prove this result probabilistically (though it can also easy be proven in a deterministic way). Let $R \subseteq V(G) \setminus B_t(G)$ be a uniform random subset of $\lceil n/2t \rceil$ vertices, and let $G'' = G[B_t(G) \cup R]$. Observe that if e is an edge incident to $B_t(G)$, then it is contained in G'' with probability 1 if $e \subseteq B_t(G)$ and otherwise it lies in G'' with probability exactly $\frac{\lceil n/2t \rceil}{n}$ since this occurs if and only if the vertex of e in $V(G) \setminus B_t(G)$ is included in R. Thus by hypothesis of the case we are in, we have

$$\mathbb{E}[e(G'')] \geq \frac{1}{2}Cn^{1+\varepsilon} \cdot \frac{\lceil n/2t \rceil}{n} \geq \frac{1}{4t}Cn^{1+\varepsilon}.$$

In particular, there exists some specific instance G' of the random graph G'' which has at least this many edges and which has $2\lceil n/2t \rceil$ vertices by construction, proving the result. □

It remains to iteratively apply these claims to prove the result. To this end, let G_0 be an n_0-vertex graph with $e(G_0) \geq Cn_0^{1+\varepsilon}$ and n_0 sufficiently large. Iteratively if G_i does not satisfy the first condition of the claim we let $G_{i+1} \subseteq G_i$ be the subgraph guaranteed by the second claim.

Claim 8.14. *If n_0 is sufficiently large then there exists some $i \leq (1-\varepsilon)\frac{\log(n_0)}{\log(t/4)} + \frac{\log(4)}{\log(t/4)}$ such that the first claim applies to G_i.*

Proof. The key observation to make is that $v(G_i) \approx n_0/t^i$. More precisely, one can inductively prove that $v(G_i) \le n_0/t^i + 3$ for all i. Moreover, in the range of i we consider one can show (as we implicitly do after this claim) that for n_0 sufficiently large in terms of ε we have $n_0/t^i \ge 3$, meaning that we have $v(G_i) \le 2n_0/t^i$.

With this observation in mind one can inductively prove that $e(G_i) \ge (4t)^{-i} C n_0^{1+\varepsilon}$ for all i. Because we trivially have $e(G_i) \le v(G_i)^2$ this implies $4t^{-2i} n_0^2 \ge (4t)^{-i} C n_0^{1+\varepsilon}$ which means $4(t/4)^{-i} \ge C n_0^{-1+\varepsilon} \ge n_0^{-1+\varepsilon}$. Taking logarithms and rearranging gives $i \le (1-\varepsilon)\log(n_0)/\log(t/4) + \log(4)/\log(t/4)$, meaning this must hold for any G_i which exists, and the fact that no larger $G_{i'}$ exists must mean that some G_i of this form is such that the first claim applies, proving the result. □

Letting $i \le (1-\varepsilon)\log(n_0)/\log(t/4)$ be such that we can apply the first claim to G_i, we obtain $G_i' \subseteq G_i$ which is $20t$-almost regular, has

$$e(G_i') \ge \frac{2}{5} C v(G_i)^{1+\varepsilon} \ge \frac{2}{5} C v(G_i')^{1+\varepsilon},$$

and which has

$$v(G_i') \ge \frac{2}{5t} v(G_i) \ge \frac{2}{5t} t \cdot t^{-i} n_0 \ge \frac{2}{5t} 4^{-\log(t)/\log(t/4)} \cdot n_0^{1-(1-\varepsilon)\log(t)/\log(t/4)},$$

where this second inequality used that we can inductively prove $v(G_i) \ge t^{-i} n_0$ for all i and the last inequality used our bound on i. Roughly now we get our desired exponent in n_0 provided t is large. To get a particularly nice expression, we take $t = 2^{1+\varepsilon^{-2}}$ which makes this exponent in n_0 equal to

$$1 - (1-\varepsilon)\frac{1+\varepsilon^{-2}}{-1+\varepsilon^{-2}} = 1 - (1-\varepsilon)\frac{1+\varepsilon^2}{1-\varepsilon^2} = 1 - \frac{1+\varepsilon^2}{1+\varepsilon} = \varepsilon\frac{1-\varepsilon}{1+\varepsilon}.$$

Thus G_i' has $v(G_i') = \Omega_\varepsilon(n^{\varepsilon\frac{1-\varepsilon}{1+\varepsilon}})$ and is K-almost regular with $K = 20t = 40 \cdot 2^{\varepsilon^{-2}}$ as desired. □

We now use this together with our result for Sidorenko even cycles to prove the following.

Proposition 8.15. *We have* $\mathrm{ex}(n, C_6) = O(n^{4/3})$.

Proof. In view of Theorem 8.11 and Lemma saying we can find bipartite subgraphs on at least half the edges, it suffices to show that there exists some C such that if G is an n-vertex bipartite graph with $e(G) \ge Cn^{4/3}$ which is K-almost regular with K as in Theorem 8.11 with $\varepsilon = 1/3$, then G contains a C_6. Let G be such a graph and assume for contradiction that G contains no C_6.

Claim 8.16. *Let C_4' be the 5-vertex graph obtained by adding a leaf to C_4. Then* $|\mathrm{Hom}(C_4', G)| \ge \frac{1}{6!}|\mathrm{Hom}(C_6, G)|$.

Proof. Because G is C_6-free, every homomorphism in $\mathrm{Hom}(C_6, G)$ is non-injective. Writing the vertices of C_6 as $v_1 v_2 \cdots v_6$, any non-injective $\phi \in \mathrm{Hom}(C_6, G)$ must either have $\phi(v_i) = \phi(v_{i+2})$ or $\phi(v_i) = \phi(v_{i+3})$ for some $1 \le i \le 6$ (with these indices written modulo 6). But this latter case would imply that $\phi(v_i), \phi(v_{i+1}), \phi(v_{i+2})$ forms a triangle in G, a contradiction to it being bipartite. Thus every $\phi \in \mathrm{Hom}(C_6, G)$ must have $\phi(v_i) = \phi(v_{i+2})$ for some i.

Roughly speaking, each such homomorphism ϕ as above "corresponds" to a homomorphsim of C_4'. More precisely, if we write the vertices of C_4' as $u_1, \ldots, u_4, u_1'$ with $u_1 \cdots u_4$ the cycle and u_1' adjacent to u_1, then the map $\phi' : V(C_4') \to V(G)$ defined by $\phi'(u_1) = v_i,\ \phi'(u_1') = v_{i+1}$ and $\phi'(u_j) = v_{i+j+1}$ is a homomorphism of C_4'. Since each homomorphism of C_4' can trivially come from at most 6! homomorphisms of C_6 in this way we obtain our desired bound. ☐

Observe that $|\mathrm{Hom}(C_4', G)| \leq |\mathrm{Hom}(C_4, G)|\Delta(G)$. Using this, the fact that G is K-almost regular, and that C_6 is Sidorenko, we in total find that

$$|\mathrm{Hom}(C_4, G)| \geq \frac{|\mathrm{Hom}(C_4', G)|}{\Delta(G)} \geq \frac{(2Cn^{-2/3})^6 n^6/6!}{2KCn^{1/3}} \geq C^4 n^{5/3},$$

with this last step holding for C sufficiently large.

Claim 8.17. *G has at least $\frac{1}{2\cdot 4!}C^4 n^{5/3}$ copies of C_4.*

Proof. We aim to show that at most $\frac{1}{2}C^4 n^{5/3}$ elements of $\mathrm{Hom}(C_4, G)$ are non-injective, which together with our lower bound for $|\mathrm{Hom}(C_4, G)|$ above and the fact that each copy of C_4 is the image of at most 4! homomorphisms gives the result.

To see this, observe that if a homomorphism of C_4 is not injective then its image is either P_3 or K_2. The total number of P_3's in G is at most $n\Delta(G)^2 \leq 4K^2C^2n^{5/3}$, and for C sufficiently large this is smaller than $\frac{1}{4\cdot 4!}C^4 n^{5/3}$ and hence the number of homomorphisms from C_4 to P_4's in G is at most $\frac{1}{4}C^4 n^{5/3}$. The same bound can be obtained for homomorhpisms with K_2 as its image, in total giving the desired result. ☐

The key idea now is that we will try to use these many copies of C_4 to find two C_4's in G which share an edge and which are otherwise vertex disjoint, as this gives a C_6 (with an extra edge) as a subgraph of G. To this end, by the pigeonhole principle we observe that there is some edge $e \in E(G)$ which is contained in many C_4's, namely at least

$$\frac{\frac{1}{2\cdot 4!}C^4 n^{5/3}}{Cn^{4/3}} \geq C^2 n^{1/3}$$

for C sufficiently large. Let $\tilde{C}_4$ be any copy of C_4 containing e and let u, v be the vertices of $\tilde{C}_4$ which are not in e. Note that the number of C_4's containing both e and, say, u is at most $\Delta(G) \leq 2KCn^{1/3} < \frac{1}{2}C^2 n^{1/3}$ with the same bound holding for C_4's contianing e and v. In total we conclude that there exists some C_4 containing e which is otherwise disjoint from $\tilde{C}_4$, and this C_4 together with $\tilde{C}_4$ gives a C_6, a contradiction.

☐

Let us reiterate the high-level steps of this proof: we started with the fact that C_6 is Sidorenko, meaning there are many C_6 homomorphisms. From here we used assumptions of our graph to get many C_4' homomorphisms, then C_4 homomorphisms, with us ultimately being able to say we have many C_4 copies (i.e. that there are few P_3, K_2 homomorphisms). We then managed to stitch together these C_4 copies into a C_6, giving the result. We emphasize that this sort of framework of iteratively relating homomorphism counts between increasingly "degenerate" graphs F is a common way to proceed with this method.

Another application of supersaturation gives the best known bounds on the Turán number of the hypercube Q_3, which we recall is defined to be the graph on bitstrings of length 3 where two strings are adjacent if they differ in exactly 1 position.

Theorem 8.18. *We have* $\mathrm{ex}(n, Q_3) = O(n^{8/5})$.

Proof. The key insight for this problem is in figuring out a good way to "represent" Q_3. In particular, we crucially observe that Q_3 consists of a C_6 together with two vertices (e.g. 000 and 111) which are each adjacent to either the odd or even vertices of this C_6. As such, a graph G will contain a copy of Q_3 precisely if we can find two vertices u, v such that the bipartite graph between their neighborhoods $N(u) \setminus \{v\}$ and $N(v) \setminus \{u\}$ contains a C_6. Because we know $\mathrm{ex}(n, C_6) = O(n^{4/3})$, it will suffice for us to find many edges between $N(u)$ and $N(v)$, and we observe that such edges xy exactly correspond to P_4's $uxyv$ in G. Thus we wish to find two vertices u, v which are the ends of many P_4's. We will do this by showing the stronger fact that there exist adjacent vertices u, v such that uv lies in many C_4's (and hence u, v are the endpoints of many P_4's).

As before, it suffices to show that there is a large integer C such that if G is an n-vertex bipartite graph with $e(G) \geq Cn^{8/5}$ which is K-almost regular then G contains a copy of Q_3. For some slight ease in notation we will assume for convenience that $e(G) = Cn^{8/5}$. By something we did for homework, the number of C_4's in such a graph is at least $\Omega(e(G)^4 n^{-4})$. By the pigeonhole principle, there exists some edge $uv \in G$ such that the number of C_4's it is in is at least $\Omega(e(G)^3 n^{-4}) \geq C^2 n^{4/5}$ assuming C is sufficiently large.

Because G is bipartite, the sets $N(u) \setminus \{v\}$ and $N(v) \setminus \{u\}$ are disjoint and independent sets. Let $B = G[N(u) \cup N(v) \setminus \{u, v\}]$. Observe that $e(B)$ is exactly the number of C_4's containing uv as an edge and hence $e(B) \geq C^2 n^{4/5}$. On the other hand,

$$v(B) \leq 2\Delta(G) \leq 4KCn^{3/5}.$$

Thus for C sufficiently large we have, say, $e(B) \geq C^{1/2} v(B)^{4/3}$. Because $\mathrm{ex}(m, C_6) = O(m^{4/3})$ this implies that B contains a C_6 if C is sufficiently large. This C_6 together with u, v gives a copy of Q_3 in G, proving the result. □

Actually, a closer look at our proof here shows that we have in fact proven the following stronger result.

Theorem 8.19. *Let Q_3' be Q_3 after adding a "long diagonal" (e.g. an edge between 000 and 111). Then* $\mathrm{ex}(n, Q_3') = O(n^{8/5})$.

Indeed this follows because the Q_3 we found in G has the property that the two vertices u, v playing the roles of 000 and 111 are adjacent to each other.

As a final aside, we note that supersaturation results have becoming increasingly important in recent years due to them being a key ingredient in applying the method of hypergraph containers. We will explore such applications later in Section 17.

8.4 Exercises

1. Prove that stars $K_{1,t}$ are Sidorenko [1+].

2. Prove that paths of the form P_{2^r+1} for some integer $r \geq 0$ are Sidorenko [2].

3. In this exercise we explore something known as the "tensor product trick" which allows one to transform weaker bounds into stronger bounds almost by magic.

   (a) Given two graphs G, H, define the tensor product $G \otimes H$ to be the graph with vertex set $V(G) \times V(H)$ where $(u, v) \sim (u', v')$ if and only if $uu' \in E(G)$ and $vv' \in E(H)$. Prove that for any graphs F, G, H that $t_F(G \otimes H) = t_F(G)t_F(H)$ [1+].

   (b) Prove that if F is a graph and if there exists some $c > 0$ such that $t_F(G) \geq ct_{K_2}(G)^{e(F)}$ for all graphs G, then F is Sidorenko [2+].

4. Another operation which plays well with homomorphism densities is blowups.

   (a) Recall that given a graph G, the blowup $G[n]$ is the graph defined by replacing each vertex of G by an independent set of size n and each edge by a $K_{n,n}$ between the corresponding blownup sets of vertices. Prove that for all graphs F, G and integers $n \geq 1$ that $t_F(G[n]) = t_F(G)$ [1+].

   (b) Prove that a bipartite graph is Sidorenko if and only if it satisfies the Erdős-Simonovits Supersaturation Conjecture II [2].

5. This exercise considers supersaturation for trees.

   (a) Prove that if T is a tree and if G is an n-vertex graph with minimum degree $d \geq v(T)$, then the number of copies of T in G is at least $\frac{(d-v(T)+1)^{v(T)-1}n}{v(T)!}$ [1+].

   (b) Unfortunately the usual average degree to minimum degree lemma is not enough to conclude supersaturation for graphs of large average degree because the number of vertices of our subgraph to drop considerably. We thus require a modified version which gives a stronger minimum degree if the subgraph we obtain has few vertices.

   With this in mind: prove that if G is an n-vertex graph and if $b \geq 1$ is real, there exists a subgraph $G' \subseteq G$ with $v(G') > 0$ and minimum degree at least

$$2^{-b}\left(\frac{v(G')}{n}\right)^{1/b}\frac{e(G)}{v(G')}.$$

   (Hint: Prove that for all non-negative integers r there exist $G_r \subseteq G$ with at most $2^{-br}n$ vertices and at least $2^{-r}e(G)$ edges) [2].

   (c) Compare the result we get above with $b = 1$ with that of Theorem 1.12 [1].

(d) Prove that for every tree T, if G is an n-vertex graph with $e(G) \geq 8v(T)n$, then G contains at least $\Omega_T(e(G)^{v(T)-1}n^{2-v(T)})$ copies of T [1+].

6. Give an alternative proof showing $\text{ex}(n, Q_3) = O(n^{8/5})$ by using supersaturation for P_4 as established in the previous problem in place of C_4 supersaturation [1+].

7. Show that for all $K \geq 1$ and $\varepsilon > 0$ that there exist n-vertex graphs G with $e(G) = \Theta(n^{\varepsilon})$ such that every K-almost regular induced subgraph $G' \subseteq G$ has $v(G') = O(n^{\varepsilon})$. That is, the quantitative bound of $\Omega(n^{\varepsilon\frac{1-\varepsilon}{1+\varepsilon}})$ from Theorem 8.11 is not far from best possible [2-].

9 Regularity and Removal Lemmas

This chapter is centered around *Szemerédi's regularity lemma* (or simply *the regularity lemma* for short), which was originally a lemma proven by Szemerédi in his proof of Szemerédi's Theorem but which has since been recognized as a powerful and fundamental tool in graph theory with many applications to Turán problems, Dirac problems, and much more. We will only scratch the surface on what can be said here, and we refer the interested reader to the book "Graph Theory and Additive Combinatorics" by Zhao for a more thorough treatment.

9.1 The Regularity Lemma and its Applications

Informally, the regularity lemma says that for every graph G, there exists a partition of $V(G)$ into a bounded number of parts such that the graph between most pairs of parts "looks like" a random graph. We need some definitions to make this precise.

Definition 18. Let G be an n-vertex graph, $A, B \subseteq V(G)$ sets of (not necessarily disjoint) vertices, and $\varepsilon > 0$ a real number.

- We define $e(A, B) = |\{(x, y) \in A \times B : xy \in E(G)\}$. Note that $e(A, B)$ equals the number of edges between A and B if these sets are disjoint, and $e(A, B)$ equals twice the number of edges within A if $A = B$.

- We define the *density* of the pair (A, B) by
$$d(A, B) = \frac{e(A, B)}{|A||B|}.$$
Note that $0 \leq d(A, B) \leq 1$ for all A, B.

- We say that the pair (A, B) is *ε-regular* if for any $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$, we have
$$|d(X, Y) - d(A, B)| \leq \varepsilon.$$
That is, a pair is ε-regular if all large subsets of A, B have roughly the same density as the pair itself. Note that this sort of property is what we would expect to see if we constructed a random graph on $A \cup B$ by keeping each edge independently and with probability $d(A, B)$, and as such we can think of the graphs on ε-regular pairs as being "peudorandom."

- We say that a partition $V_1 \cup \cdots \cup V_m$ of $V(G)$ is an *ε-regular partition* if
$$\sum_{(V_i,V_j) \text{ which are not } \varepsilon\text{-regular}} e(V_i, V_j) \leq \varepsilon n^2.$$
.

Theorem 9.1 (Szemerédi's Regularity Lemma)**.** *For all $\varepsilon > 0$, there exists a number $M(\varepsilon)$ such that every graph G has an ε-regular partition $V_1 \cup \cdots \cup V_m$ with $m \leq M(\varepsilon)$*

To reiterate, this informally says that every graph G can be partitioned into a bounded number of parts such that the graph between most pairs of parts looks like a random graph, in the sense that most pairs are ε-regular. It is also worth noting that while it is very good that the number of parts is bounded by the function $M(\varepsilon)$, the function $M(\varepsilon)$ is horrendously large in terms of ε. In particular, the implicit constants in any proof which uses regularity lemma will be horrendously large as well. While we will not get into it here, a result of Gowers shows that $M(\varepsilon)$ must necessarily be quite large for the theorem to hold. In a similar spirit, one can cook up constructions which show that this result is not true if we replace the condition that most edges are in ε-regular pairs with the condition that *all* edges are in ε-regular pairs..

We now give a proof of the regularity lemma. We emphasize, however, that the reader may first find it helpful to actually read through some of the applications of this result first to get a feel for these strange definitions and only then come back to review its proof.

Proof. We'll just give a sketch due to time constraints.

Roughly speaking, the idea of the proof is to start with some arbitrary partition $\mathcal{P}$ of G. Iteratively if we have some partition $\mathcal{P}$ which is not ε-regular, then we can use this fact to produce a "refinement" $\mathcal{P}'$ of $\mathcal{P}$ which improves upon $\mathcal{P}$ in some measurable way. To capture this improvement, we associate to each partition $\mathcal{P}$ of G a certain parameter $q(\mathcal{P})$ lying between 0 and 1 called the "energy" of $\mathcal{P}$. Because the energy can not exceed 1, we can have at most a bounded number of improvement steps, at which point $\mathcal{P}$ must be ε-regular by construction.

Somewhat more precisely, for sets $U, V \subseteq V(G)$ we define

$$q(U,V) = \frac{|U||V|}{v(G)^2} d(U,V)^2,$$

and for a partition $\mathcal{P}$ for $V(G)$ we define

$$q(\mathcal{P}) = \sum_{U,V\in\mathcal{P}} q(U,V).$$

It is straightforward to check that $0 \le q(\mathcal{P}) \le 1$. Most importantly, we have the following.

Claim 9.2. *If $\mathcal{P}$ is a partition which is not ε-regular, then there exists a refinement $\mathcal{P}'$ (i.e. a partition such that for each $U' \in \mathcal{P}'$ there exists $U \in \mathcal{P}$ with $U' \subseteq U$) such that $|\mathcal{P}'| \le 4^{|\mathcal{P}|}$ and such that $q(\mathcal{P}') \ge q(\mathcal{P}) + \varepsilon^5$.*

Assuming this claim, we can prove the result by starting with $\mathcal{P}_1 = \{V(G)\}$ and then iteratively if $\mathcal{P}_i$ is not ε-regular then we let $\mathcal{P}_{i+1}$ be the refinement of $\mathcal{P}_i$ given by the claim. By the claim and properties of q we must have $1 \ge q(\mathcal{P}_i) \ge i\varepsilon^5$ and hence this process terminates at some $i \le \varepsilon^{-5}$, implying that $\mathcal{P}_{i'}$ is ε-regular at this point. Moreover, one can bound $|\mathcal{P}_{i'}|$ as, say, an iterated exponential of 4's of height $i' \le \varepsilon^{-5}$, so taking $M(\varepsilon)$ to be this quantity gives the result.

To prove this claim, the idea is that for each pair $V_i, V_j \in \mathcal{P}$ with $i \neq j$ which is not ε-regular we define $A_{i,j} \subseteq V_i$ and $A_{j,i} \subseteq V_j$ to be two sets witnessing that this pair is not ε-regular, i.e. which is such that $|d(V_i,V_j) - d(A_{i,j},A_{j,i})| > \varepsilon$. Similarly if (V_i,V_i) is not ε-regular then we let $A_{i,i}, A_{i,*} \subseteq V_i$ denote a pair of sets witnessing this fact. Now for each i and each binary vector

x indexed by the $A_{i,j}$ sets for all $j \in \mathbb{N} \cup \{*\}$ which exist, we let V_i^x be the set of vertices v with $v \in A_{i,j}$ if and only if $x_{A_{i,j}} = 1$. One can show that the refinement $\mathcal{P}'$ consisting of all the V_i^x sets for all i, x satisfies $q(\mathcal{P}') \geq q(\mathcal{P}) + \varepsilon^5$. Moreover, each $V_i \in \mathcal{P}$ gets partitioned into at most $2^{|\mathcal{P}|+1}$ pieces (since there are at most $|\mathcal{P}|+1$ total sets $A_{i,j}$ with the $+1$ coming from $A_{i,*}$). We thus have $|\mathcal{P}'| \leq |\mathcal{P}| 2^{|\mathcal{P}|+1} \leq 4^{|\mathcal{P}|}$, proving the result. □

We now turn to applications of the regularity lemma, which in general all go through the same three basic steps:

(1) Take an ε-regular partition $V_1 \cup \cdots \cup V_m$ for your graph G as guaranteed by the regularity lemma.

(2) "Clean" the graph G by deleting a small number of "poorly behaved" edges, e.g. by deleting all edges between any pairs (V_i, V_j) which are either not ε-regular, have low density, or which have $|V_i|$ relatively small.

(3) Solve the problem for the cleaned graph, often by invoking known results from extremal combinatorics or using "counting" lemmas.

One basic and very important example of this framework comes from the following result known as the triangle removal lemma (or simply the removal lemma depending on context).

Theorem 9.3 (Triangle Removal Lemma). *For all $\varepsilon > 0$, there exists a $\delta > 0$ such that if G is an n-vertex graph with at most δn^3 triangles, then G can be made triangle-free by deleting at most εn^2 edges.*

Essentially, this says that any graph which is close to being triangle-free (in the sense that it has $o(n^3)$ triangles) can be made triangle-free by deleting at most $o(n^2)$ edges.

Proof. Fix some $\varepsilon > 0$, and with foresight we define $\varepsilon' = \min\{\frac{1}{4}, \varepsilon/4\}$ and $\delta = \frac{1}{8}(\varepsilon')^6 M(\varepsilon')^{-3}$ where $M(\varepsilon')$ is as in the statement of the regularity lemma. Let G be an n-vertex graph with at most δn^3 triangles. With our framework above in mind, we begin by applying the regularity lemma to obtain an ε'-regular partition $V_1, \ldots, V_m$ and then we define our "cleaned" subgraph $G' \subseteq G$ by deleting all edges between pairs (V_i, V_j) such that either (V_i, V_j) is not ε'-regular, or $d(V_i, V_j) \leq 2\varepsilon'$, or $\min\{|V_i|, |V_j|\} \leq \varepsilon' m^{-1} n$.

Claim 9.4. *It suffices to show that the graph G' is triangle-free.*

Proof. Observe that the number of edges we deleted going from G to G' is certainly at most

$$\varepsilon' n^2 + \sum_{i,j} 2\varepsilon' |V_i||V_j| + \sum_{i : |V_i| \leq \varepsilon' m^{-1} n} |V_i| n \leq \varepsilon' n^2 + 2\varepsilon' + \varepsilon' n = 4\varepsilon' n^2 \leq \varepsilon n^2,$$

where this first inequality used that the number of terms in the sum is at most m. As such, if we assume the hypothesis of the claim then we see that we can indeed remove at most εn^2 edges from G so that the resulting graph is triangle-free. □

Assume for contradiction that G' contains a triangle, say with each of its vertices coming from parts V_i, V_j, V_j and we emphasize that we do not require that these integers i, j, k to be distinct from each other. This implies that there is at least one edge in G' between each of these three parts, which by definition of G' implies that all of these pairs of parts are ε'-regular, have density at least $2\varepsilon'$, and that each of these parts has size at least $\varepsilon' m^{-1} n$.

Roughly, the intuition now is that if these ε'-regular graphs were truly random graphs of density at least $2\varepsilon'$, then with high probability G' would have at least roughly $\varepsilon'^3|V_i||V_j||V_k|$ edges, contradicting our assumption that G had few triangles. We now translate this probabilistic thinking to the ε'-regular setting of G' through a series of claims.

Claim 9.5. *The set of vertices $V_i' \subseteq V_i$ which have at least $\varepsilon'|V_j|$ neighbors in V_j and at least $\varepsilon'|V_k|$ neighbors in V_k satisfies $|V_i'| \geq (1-2\varepsilon')|V_i|$.*

Proof. Indeed, if we let $X \subseteq V_i$ denote the set of vertices with less than $\varepsilon'|V_j|$ neighbors in V_j, then

$$d(X, V_j) < \frac{\varepsilon'|X||V_j|}{|X||V_j|} = \varepsilon'.$$

Because $d(V_i, V_j) \geq 2\varepsilon'$ and (V_i, V_j) is ε'-regular, this inequality is only possible if $|X| \leq \varepsilon'|V_i|$. An analogous argument shows that the number of vertices of V_i with less than $\varepsilon'|V_k|$ neighbors in V_k is at most $\varepsilon'|V_i|$, proving the claim. □

Claim 9.6. *Every $x_i \in V_i'$ as defined above is contained in at least $\frac{1}{2}(\varepsilon')^3|V_j||V_k|$ triangles.*

Proof. By definition of V_i', the sets $V_j' := N_{G'}(x_i) \cap V_j$ and $V_k' := N_{G'}(x_i) \cap V_k$ both have at least an ε' proportion of the vertices of V_j, V_k. Using this and the fact that the pair (V_j, V_k) is ε'-regular, we find that

$$d(V_j', V_k') \geq d(V_j, V_k) - \varepsilon' \geq \varepsilon',$$

which by definition means that the number of pairs of adjacent vertices $(x_j, x_k) \in V_j' \times V_k'$ is at least $\varepsilon'|V_j'||V_k'| \geq (\varepsilon')^3|V_j||V_k|$. Since each of these pairs (x_j, x_k) forms a triangle with x_i and since each such triangle arises from at most 2 pairs (x_j, x_k), we find that the number of triangles using x_i is as stated. □

By combining these two claims, we see that for any choice of $\varepsilon' \leq \frac{1}{4}$ that the number of triangles in G' is at least

$$\frac{1}{4}(\varepsilon')^3|V_i||V_j||V_k| \geq \frac{1}{4}(\varepsilon')^6 m^{-3} n^3 \geq \frac{1}{4}(\varepsilon')^6 M(\varepsilon')^{-3} n^3 = 2\delta n^3,$$

with the first inequality using that every part surviving in G' has size at least $\varepsilon' m^{-1} n$. This in turn implies that $G \supseteq G'$ contains at least $2\delta n^3$ triangles, a contradiction to our assumption that it contains at most δn^3 triangles, proving the result. □

As an aside, because our proof used the regularity lemma, the dependencies of ε and δ we obtain are very bad. There do exist regularity-free proofs of the triangle removal lemma due to Fox which gives nearly optimal bounds on these dependencies, but even these nearly optimal bounds are still quite large.

The triangle removal lemma is, in addition to simply being a nice statement, an incredibly important tool in its own right. We will see some application of this in the next subsection. For now , we continue looking at applications of the regularity lemma by proving the Erdős-Stone-Simonovits Theorem which we recall below.

Theorem 9.7. *For any graph F with at least one edge, we have*

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

Proof. The key observation due to Simonovits is that we can reduce our problem to studying complete r-partite graphs where every part has size t, and we let $K_{r;t}$ denote this graph.

Claim 9.8. *It suffices to prove for all $r, t \geq 2$ that*

$$\mathrm{ex}(n, K_{r;t}) \leq \left(\frac{r-2}{r-1} + o(1)\right)\binom{n}{2}.$$

Proof. Assume this is true and consider any graph F with $r := \chi(F)$. The lower bound comes from considering $G = K_{r-1;\lfloor n/(r-1)\rfloor}$. For the upper bound, we observe that $F \subseteq K_{r;v(F)}$ since any $\chi(F) = r$ implies that F is r-partite, and as such it is certainly contained in a complete r-partite graph with every part of size F. As such,

$$\mathrm{ex}(n, F) \leq \mathrm{ex}(n, K_{r;v(F)}) \leq \left(\frac{r-2}{r-1} + o(1)\right)\binom{n}{2},$$

proving the result. ☐

We now prove this upper bound on $\mathrm{ex}(n, K_{r;t})$. This was originally done by Erdős and Stone, though we emphasize that their proof was not based on regularity like ours is. Many of the details here are completely analogous to what we did in our proof of the triangle removal lemma, so we will a bit terse in our exposition whenever the parallels are clear.

Fix some r, t. Proving this asymptotic bound is equivalent to showing that for any $\delta > 0$, we have $\mathrm{ex}(n, K_{r;t}) \leq (\frac{r-2}{r-1} + \delta)\binom{n}{2}$ for all sufficiently large n. To this end, fix some $\delta > 0$, let $d > 0$ be some small constant in terms of δ, r, t, and let $\varepsilon > 0$ be a very small constant in terms of d, r, t (we won't specify it exactly, but we will want $\varepsilon \approx (d/2)^{rt}$).

Let G be an n-vertex graph with at least $(\frac{r-2}{r-1} + \delta)\binom{n}{2}$ edges and $V_1 \cup \cdots \cup V_m$ an ε-partition of G as guaranteed by the regularity lemma. Let $G' \subseteq G$ be the subgraph defined by deleting all edges between V_i and V_j for any pair (V_i, V_j) which is either not ε-regular, or has $d(V_i, V_j) \leq d$, or has $\min\{|V_i|, |V_j|\} \leq \varepsilon m^{-1} n$.

As in our proof of the triangle removal lemma, we observe that the number of edges we delete when going from G to G' is at most

$$\varepsilon n^2 + dn^2 + \varepsilon n^2 \leq \frac{\delta}{2} n^2,$$

with the last step using our assumption of ε, d being sufficiently small in terms of δ. In particular, G' is an n-vertex graph with at least $(\frac{r-2}{r-1} + \delta/2)\binom{n}{2}$ edges, so for n sufficiently large Turán's

Theorem guarantees that G' contains a K_r. Possibly by relabeling our parts we can assume that the r vertices of this K_r lie in parts $V_1, \ldots, V_r$ where again we allow the possibility that $V_i = V_j$ for some $i \neq j$. The existence of this K_r implies that there exist edges in G' between each of the parts V_i, V_j for all distinct $1 \le i, j \le r$, so by construction of G' this implies that each of these pairs (V_i, V_j) is ε-regular and has $d(V_i, V_j) \ge d$.

Intuitively at this point we will try and proceed as follows to build our copy of $K_{r;t}$: we begin by selecting some vertex $x_{1;1} \in V_1$ which has at least $(d-\varepsilon)|V_j|$ in each of the other parts, with most vertices in V_1 satisfying this property. We then pick $x_{1;2} \in V_1$ which has at least $(d-\varepsilon)|V_j \cap N(x_{1;1})|$ neighbors in each set $V_j \cap N(x_{1;1})$, which again is satisfied by most choices of vertices in V_1 provided ε is much smaller compared to d. We then pick $x_{1;3}$ to have many neighbors within the common neighborhoods $N(x_{1;1}) \cap N(x_{1;2})$ and continue in this way until we have selected vertices $x_{1;1}, \ldots, x_{1;t}$ which have many common neighbors in each of the other parts. From here we pick some $x_{2;1}$ in the intersection of this common neighborhood and V_2 and proceed in a similar way until we have eventually constructed a full copy of $K_{r;t}$. The following gives a formal framework for this approach to work.

Claim 9.9. *For n sufficiently large, we have that if $\widetilde{V}_1, \subseteq, \widetilde{V}_r$ are subsets of $V_1, \ldots, V_r$ such that $(d-\varepsilon)^t|\widetilde{V}_i| \ge 2r\varepsilon|V_i|$ for all i, then for all i and integers $1 \le s \le t$ there are at least $2^{-s}|\widetilde{V}_i|^s$ tuples $(x_1, \ldots, x_s) \in \widetilde{V}_i^s$ of distinct vertices such that $|\bigcap_{s'=1}^{s} N(x_{s'}) \cap \widetilde{V}_j| \ge (d-\varepsilon)|\widetilde{V}_j|$ for all $j \neq i$*

Proof. We prove the result by induction on s, the base case $s = 0$ being trivial. Let $(x_1, \ldots, x_{s-1})$ be an arbitrary tuple satisfying the conditions for $s-1$, let $\widetilde{V}_j^* = \bigcap_{s'=1}^{s-1} N(x_{s'}) \cap \widetilde{V}_j$ and let X_j denote the set of $x \in \widetilde{V}_i$ such that $|N(x) \cap \widetilde{V}_j^*| < (d-\varepsilon)|\widetilde{V}_j|$. Then

$$d(X_j, \widetilde{V}_j^*) < d - \varepsilon \le d(V_i, V_j) - \varepsilon.$$

Because (V_i, V_j) is an ε-regular pair and $|\widetilde{V}_j^*| \ge (d-\varepsilon)^{s-1}|\widetilde{V}_j|\varepsilon|V_j| \ge \varepsilon|V_j|$ by hypothesis, we must have $|X_j| < \varepsilon|V_i|$. Using this and the requirement that our tuples have distinct vertices, we find that the number of choices for x which we can append to $(x_1, \ldots, x_{s-1})$ while satisfying the conditions of the claim is at least

$$|\widetilde{V}_i \setminus (\bigcup X_j)| - s + 1 \ge |\widetilde{V}_i| - (r-1)\varepsilon|V_i| - t \ge |\widetilde{V}_i| - r\varepsilon|V_i| \ge \frac{1}{2}|\widetilde{V}_i|,$$

with the middle inequality holding for n sufficiently large in terms of ε since $|V_i| \ge \varepsilon m^{-1} n \ge \varepsilon M(\varepsilon)^{-1} n$. Since we inductively assumed the number of choices for $(x_1, \ldots, x_{s-1})$ was at least $2^{1-s}|\widetilde{V}_i|^{s-1}$ we conclude the result. □

To prove the result, we first apply this claim with $\widetilde{V}_j = V_j$ for all j to find a tuple $(x_{1;1}, \ldots, x_{1;t}) \in V_1^t$ satisfying these conditions. Iteratively given that we have constructed $(x_{i';1}, \ldots, x_{i';t})$ for all $1 \le i' < i \le r$, we apply the claim with $\widetilde{V}_j = V_j \bigcap_{1 \le i' < i,\ 1 \le s \le t} N(x_{i';s})$ for $j \ge i$ and $\widetilde{V}_j = V_j$ for $j < i$ to find a tuple $(x_{i;1}, \ldots, x_{i;t}) \in \widetilde{V}_i$. Note that iteratively we always apply the claim after assuming $|\widetilde{V}_j| \ge (d-\varepsilon)^{rt}|V_j|$, so the hypothesis of the claim will hold provided ε is sufficiently small in terms of d, r, t. We conclude that this process terminates with distinct vertices such that $x_{i;s} \sim x_{i';s'}$ whenever $i < i'$ by construction, giving our copy of $K_{r;t}$ as desired. □

As an aside, we note that our proof in fact shows the somewhat stronger fact that having $e(G) \geq (\frac{r-2}{r-1} + \delta)\binom{n}{2}$ implies that G contains $\Omega_\delta(n^{rt})$ copies of $K_{r;t}$).

The exact mechanics of our proofs for both the triangle removal lemma and the Erdős-Stone-Simonovits Theorem are very similar to each other, in that they both rely on showing that if there exists a set of r parts whose pairs are all ε-regular and have high density, then one can find many copies of K_r. Arguments of this form are very common with the regularity lemma, and as such it can be useful to record these facts into "counting lemmas", a general version of which is the following.

Lemma 9.10 (Graph Counting Lemma). *For every graph F with vertex set $\{v_1, \ldots, v_r\}$ and for every real number $\delta > 0$, there exists some $\varepsilon > 0$ such that the following holds: if G is a graph and $V_1, V_2, \ldots, V_{v(F)} \subseteq V(G)$ are such that (V_i, V_j) is ε-regular with $d(V_i, V_j) \geq \delta$ whenever $v_iv_j \in E(F)$, then the number of homomorphisms ϕ from F to G with $\phi(v_i) \in V_i$ is at least*

$$(1-\delta) \prod_{v_iv_j \in E(F)} (d(V_i, V_j) - \delta) \prod_i |V_i|.$$

To be clear, this lemma only guarantees many homomorphisms of F and not necessarily many copies of F. However, the number of homomorphisms of F which are not injective is at most $O(v(G)^{v(F)-1})$, so this result guarantees many injective homomorphisms (and hence copies of F) whenever $|V_i| = \Omega(v(G))$ for all i. The proof of the graph counting lemma is spiritually similar to the proofs we have done up to this point, so we leave its proof as an exercise. One application of this counting lemma is the following.

Lemma 9.11. *For every graph F with at least one edge, let* $\mathrm{Hom}(F)$ *denote the family of graphs H for which there exists a homomorphism $\phi : V(F) \to V(H)$. For all $\delta > 0$, there exists some n_0 such that if G is an n-vertex F-free graph with $n \geq n_0$, then G can be made* $\mathrm{Hom}(F)$*-free by removing at most δn^2 edges.*

We note that in our proof of the Erdős-Stone-Simonovits Theorem we implicitly showed this is true if instead of forbidding all of $\mathrm{Hom}(F)$ we forbid only $K_{\chi(F)} \in \mathrm{Hom}(F)$. Again we leave the details of this result as an exercise.

9.2 Applications of the Removal Lemma

We now discuss applications of the triangle removal lemma, which we recall says that any graph with $o(n^3)$ triangles can be made triangle-free by deleting at most $o(n^2)$ edges. We begin with a Turán type problem. To this end, given graphs H and F we define the *generalized Turán number* $\mathrm{ex}(n, H, F)$ to be the maximum number of copies of H in an n-vertex F-free graph. Note that $\mathrm{ex}(n, K_2, F) = \mathrm{ex}(n, F)$.

Theorem 9.12. *We have* $\mathrm{ex}(n, K_3, K_4 - e) = o(n^2)$ *where $K_4 - e$ denotes the graph obtained from K_4 after deleting an edge.*

That is, every n-vertex graph where every edge is contained in at most one triangle has $o(n^3)$ triangles. This is equivalent to saying that every n-vertex graph where every edge is contained in exactly one triangle has $o(n^3)$ triangles, which is the most common way this theorem appears in the literature.

Proof. Let G be any n-vertex $(K_4 - e)$-free graph. By definition every edge of G is contained in at most one triangle, implying that the number of triangles is at most $e(G)/3 = O(n^2) = o(n^3)$. By the triangle-removal lemma one can delete $o(n^2)$ edges of G to make it triangle-free. But by definition of G, each edge removed destroys at most one triangle in G, implying that G must have had $o(n^2)$ triangles to begin with, proving the result. □

This bound of $\mathrm{ex}(n, K_3, K_4 - e) = o(n^2)$ is best possible in that there exists a construct showing that $\mathrm{ex}(n, K_3, K_4 - e) \geq n^{2-o(1)}$ Maybe see exercises for more on this.

We now give some applications of the removal lemma to areas outside of combinatorics. We begin with the original motivation for the regularity lemma, namely in determining how large a subset $A \subseteq [n]$ can be if it contains no k-term arithmetic progression, i.e. no k distinct integers $a_1, \ldots, a_k \in A$ such that $a_{i+1} - a_i = a_{j+1} - a_j$ for all i, j. The simplest non-trivial case of $k = 3$ was originally solved by Roth using Fourier analysis. A substantially simpler proof can be given using the removal lemma.

Theorem 9.13 (Roth's Theorem). *If $A \subseteq [n]$ contains no 3-term arithmetic progression, then $|A| = o(n)$.*

Again the bound of $|A| = o(n)$ is best possible here as for all k there exist sets $A \subseteq [n]$ without k-AP's with size $|A| \geq n^{1-o(1)}$.

Proof. Let $A \subseteq [n]$ be such that it contains no 3-term arithmetic progression, which we crucially observe is equivalent to saying that no distinct $x, y, z \in A$ satisfy $x + y = 2z$ since in this case x, z, y would be a progression with common difference $z - x = y - z$. We now wish to construct a graph G_A whose edges are defined based on A such that G_A inherits some nice properties because A is 3-AP free. After a lot of thought, one might be led to the following idea for constructing G_A:

- The graph $G_A$is tripartite with parts $V_1 = [n],\ V_2 = [2n]$, and $V_3 = [3n]$,
- We have $v_1 \in V_1$ adjacent to $v_2 \in V_2$ if and only if there exists $a \in A$ with $v_1 + a = v_2$,
- We have $v_2 \in V_2$ adjacent to $v_3 \in V_3$ if and only if there exists $a \in A$ with $v_2 + a = v_3$, and
- We have $v_1 \in V_1$ adjacent to $v_3 \in V_3$ if and only if there exists $a \in A$ with $v_1 + 2a = v_3$.

We emphasize that the adjacency condition for V_1, V_3 is defined differently compared to the other cases. Crucially, this graph does inherit nice properties whenever A is 3-AP free.

Claim 9.14. *If $A \subseteq [n]$ is 3-AP free, then $(v_1, v_2, v_3) \in V_1 \times V_2 \times V_3$ is the vertex set of a triangle in G_A if and only if $v_2 = v_1 + a$ and $v_3 = v_1 + 2a$ for some $a \in A$.*

Proof. It is straightforward to check that every triple of vertices $(v_1, v_1 + a, v_1 + 2a)$ with $a \in A$ is a triangle. Assume now that (v_1, v_2, v_3) forms a triangle in G_A. By definition this means $v_2 - v_1 := a \in A$, $v_3 - v_2 := b \in A$, and $v_3 - v_1 := 2c$ for some $c \in A$. As such we have

$$a + b = (v_2 - v_1) + (v_3 - v_2) = v_3 - v_1 = 2c.$$

Because A is 3-AP free, this equality is only possible if at least two of a, b, c are equal to each other, but one can check that this is only possible if $a = b = c$, proving the claim. □

From this claim we conclude that the number of triangles in G_A is exactly $|A|n$ when A is 3-AP free since a triangle is uniquely identified by picking $v_1 \in V_1$ and the $a \in A$ such that $v_2 = v_1 + a$. This claim also implies that every edge of G_A is contained in at most one triangle since, for example, any edge v_1v_2 with $v_1 \in V_1, v_2 \in V_2$ can only be in a triangle with $v_3 = v_1 + 2(v_2 - v_1) \in V_3$. By Theorem 9.12 we conclude that

$$|A|n = o(n^2),$$

and hence that $|A| = o(n)$, proving the result. □

It is perhaps tempting giving the simplicity of this argument to try and use some sort of removal lemma to try and prove Szemerédi's Theorem that sets $A \subseteq [n]$ without k-AP's have $|A| = o(n)$, but this turns out to be substantially harder to do with perhaps the "simplest" such proof being those that rely on the difficult machinery of hypergraph removal lemmas.

The last application we consider is from property testing. In this setup, we want to quickly determine whether a given graph G either has some desired property or if it is far from having this property. To this end, we say that an n-vertex graph G is *ε-close* to having a property $\mathcal{P}$ if there exist sets of edges $E, E' \subseteq K_n$ with $|E|, |E'| \le \varepsilon n^2$ such that $G + E - E'$ has property $\mathcal{P}$ and we say that G is *ε-far* otherwise. That is, G is ε-far from $\mathcal{P}$ if we can not get G to satisfy $\mathcal{P}$ even after changing up to $2\varepsilon n^2$ of its edges. Determining precisely whether a graph satisfies a property or is far from it can take quite a bit of time if n is large. Remarkably, the property of triangle-freeness can be tested with arbitrarily high probability after checking only $O(1)$ vertices of G.

Theorem 9.15. *For all $\varepsilon, c > 0$ there exists a randomized algorithm which runs in time $O_{\varepsilon,c}(1)$ which correctly determines whether a given graph G is either triangle-free or ε-far from being triangle-free with probability at least $1 - c$.*

Proof. Let C be some large (but fixed) integer depending on ε, c to be determined later. Our algorithm goes as follows: we uniformly at random pick three vertices $v_1, v_2, v_3 \in V(G)$ and test if these vertices form a triangle. We repeat this process independently for a total of C times. If none of these vertices form a triangle then we output that G is triangle-free, and otherwise if some v_1, v_2, v_3 form a triangle then we output that G is ε-far from being triangle-free.

This algorithm always correctly identifies that G is triangle-free, so it remains to check that it correctly identifies G as being ε-far from triangle-free with high probability. And indeed, observe that G being ε-far from triangle-free means that G can not be made triangle-free by removing at most εn^2 edges. The contrapositive of the triangle removal lemma then implies that

G must contain at least δn^3 triangles for some δ depending only on ε. As such, the probability that three uniform random vertices of G form a triangle is at least δ, and hence the probability that the algorithm above fails to find any triangle in its C trials is at most

$$(1-\delta)^C,$$

and this quantity can be made less than c by taking C sufficiently large in terms of ε, c, proving the result. □

9.3 Variants

There are many variants of both the regularity lemma and the removal lemma. One of the most commonly used variants is a convenient version of the regularity lemma which guarantee that each part in the partition has nearly the same size.

Theorem 9.16 (Equitable Regularity Lemma)**.** *For all $\varepsilon > 0$, there exists a number $M(\varepsilon)$ such that every graph G has an ε-regular partition $V_1 \cup \cdots \cup V_m$ with $\varepsilon^{-1} \le m \le M(\varepsilon)$ and with the property that $||V_i| - |V_j|| \le 1$ for all i, j.*

Observe that we added the additional hypothesis here that $m \ge \varepsilon^{-1}$, which can sometimes be a convenient feature to have in proofs.

There also exist a number of variants for the removal lemma, such as by extending it to hold for arbitrary graphs (as we discuss in the exercises), as well as to settings where we care about induced copies. We will not discuss these further here and instead refer the reader again to the great book by Yufei Zhao on this topic.

9.4 Exercises

1. Prove that if G is an n-vertex graph with less than $\varepsilon^3 n^2$ edges, then $V_1 = V(G)$ is an ε-regular partition with only one part. Because of this, ε-regular partitions are only interesting and useful in the case when G has $\Theta(n^2)$ edges [1].

2. Prove the graph counting lemma [2-].

3. Using the graph counting lemma, prove the graph removal lemma: for every graph F and $\varepsilon > 0$ there exists some $\delta > 0$ such that if G is an n-vertex graph with at most $\delta n^{v(F)}$ copies of F, then G can be made F-free by removing at most εn^2 edges of G [2-].

4. Behrend famously showed that there exist 3AP-free sets $A \subseteq [n]$ of size $n^{1-o(1)}$. In this exercise we walk through the idea of this construction.

   (a) (Motivating idea) Prove that if $x, y, z \in \mathbb{R}^d$ are vectors with the same norm and with $x + y = 2z$, then $x = y = z$ [2-].
   With the observation, the idea will be for us to take a large set of points $A' \subseteq \mathbb{R}^d$ lying on a sphere (meaning that A' has no "3AP's") and then translating A' into a set $A \subseteq [n]$ such that A preserves the no 3AP property.

(b) Prove for all m, d that there exists a set of points $A' \subseteq \{0, 1, \ldots, m-1\}^d$ and an integer ℓ such that $\sum x_i^2 = \ell$ for all $x \in A'$ and such that $|A'| \geq m^d/(dm^2 + 1)$ [1+].

(c) For integers m, d, define a map $\phi : \{0, 1, \ldots, m-1\}^d \to \mathbb{Z}_{\geq 0}$ by having $\phi(x) = \sum_{i=1}^d x_i (2m)^{i-1}$. Prove that if $x, y, z \in [m]^d$ have $\phi(x) + \phi(y) = 2\phi(z)$, then $x + y = 2z$ (Hint: if desired you may use the fact that every number has a unique base b representation for all b) [2].

(d) Prove that there exist $A \subseteq [n]$ with no 3AP such that

$$|A| \geq n e^{-C\sqrt{\log n}},$$

for some $C > 0$ [2].

5. Prove that $\text{ex}(n, K_3, K_4 - e) \geq n^2 e^{-C\sqrt{\log n}}$ for some $C > 0$ [1+].

6. Prove that $\text{ex}_{lin}(n, C_3^{(3)}) = n^{2-o(1)}$, which we recall is the maximum number of edges in an n-vertex linear 3-graph which contains no loose triangle [2-].

7. The Turán Problem asks us to maximize the number of edges in an F-free graph, while the Ramsey Problem essentially asks us to minimize the independence number in an F-free graph. The *Ramsey-Turán Problem* combines these two trains of thought by bounding $e(G)$ when G avoids a graph and has a given independence number.

   Towards this end, prove that if G is an n-vertex K_3-free graph then $e(G) \leq \alpha(G)n$. In particular, if $\alpha(G) = o(n)$ then $e(G) = o(n^2)$, which differs sharply from the bound $\text{ex}(n, K_3) = \Theta(n^2)$ occurring when there is no restriction on $\alpha(G)$ [1+].

8. The first non-trivial case of the Ramsey-Turán Problem mentioned above is when we avoid K_4. Towards this end, throughout this problem let G be an n-vertex graph and $V_1, \ldots, V_m$ an ε-regular partition for some $0 < \varepsilon < \frac{1}{2}$. We note that for each part of the problem, you may choose to prove the result with each function of ε replaced by some other function of ε if you find this to be easier to prove.

   (a) Prove that if (V_i, V_j) is an ε-regular pair with $|V_i|, |V_j| \geq \frac{1}{2}M(\varepsilon)^{-1}n$ and $d(V_i, V_j) \geq \frac{1}{2} + 2\varepsilon$ and if $\alpha(G) < \frac{1}{4}\varepsilon M(\varepsilon)^{-1}n$, then G contains a K_4 (Hint: why can't G just be a complete bipartite graph on $V_i \cup V_j$?) [2].

   (b) Let V_i, V_j, V_k be three (possibly non-distinct) sets in the partition such that all three of their pairs are ε-regular. Prove that if all three of these sets have size at least $\frac{1}{2}M(\varepsilon)^{-1}n$, if all three of their pairs have density at least 2ε, and if $\alpha(G) < \frac{1}{2}\varepsilon^2 M(\varepsilon)^{-1}n$, then G contains a K_4 [2].

   (c) Conclude that if V_i is such that (V_i, V_i) is ε-regular and if $|V_i| \geq \frac{1}{2}M(\varepsilon)^{-1}n$, $d(V_i, V_i) \geq 2\varepsilon$, and if $\alpha(G) < \frac{1}{2}\varepsilon^2 M(\varepsilon)^{-1}n$ then G contains a K_4 [1].

(d) Put these pieces together to prove the following fact: for all $\varepsilon' > 0$ there exists some $\varepsilon > 0$ such that every n-vertex K_4-free graph G with $\alpha(G) < \frac{1}{2}\varepsilon^2 M(\varepsilon)^{-1} n$ satisfies

$$e(G) \leq \left(\frac{1}{8} + \varepsilon'\right) n^2.$$

Note that this is significantly smaller than the Turán bound $\text{ex}(n, K_4) \sim \frac{1}{3}n^2$ which has no condition on $\alpha(G)$ (Hint: take an equitable ε-regular partition and delete edges involved with irregular pairs and with $d(V_i, V_j) < 2\varepsilon$. Use (a) and (c) to conclude that there must be many pairs V_i, V_j with $i \neq j$ which are ε-regular and which have density at least 2ε; what does this imply in turn?) [2+].

(e) Prove that there exist n-vertex K_4-free graphs G which have $\alpha(G) = o(n)$ and $e(G) = (\frac{1}{8} - o(1))n^2$ [3+].

9. Put this earlier but Prove Theorem 9.11 (you may assume the Graph Counting Lemma as well as the Equitable Regularity Lemma if desired) [2-].

10 Stability

Recall that the stability problem asks what an n-vertex graph with $e(G) \approx \mathrm{ex}(n, F)$ must look like. Perhaps the most natural starting point to explore such results is the case $F = K_r$ since at the moment these are the only graphs for which we actually know $\mathrm{ex}(n, F)$ for all values of n. In this setting we have the following elegant result of Füredi.

Theorem 10.1 (Füredi)**.** *If $t \geq 0$ is an integer and if G is a K_r-free n-vertex graph with $e(G) \geq e(T_{r-1}(n)) - t$, then G can be made $(r-1)$-partite by deleting at most t edges.*

Let us emphasize a few features of this result. First, this applies for *arbitrary* choices of t and n, which is somewhat uncommon for stability results which typically require one or both of these quantities to be large. In particular, the $t = 0$ case of this result implies Turán's Theorem (since the only $(r-1)$-partite graph with at least $e(T_{r-1}(n))$ edges is $T_{r-1}(n)$ itself). Finally, the bound of needing t edges to remove to be $(r-1)$-partite is not too far from optimal. For example, if G is obtained by taking a C_5 and duplicating each vertex $n/5$ times, then $n^2/25$ edges must be removed from G to make this graph bipartite, and Füredi's Theorem gives the only slightly weaker bound of roughly $n^2/4 - n^2/5 = n^2/20$.

Proof. As noted above, the proof we give of this result must necessarily imply a proof of Turán's Theorem. To help motivate our approach, we will sketch the (new) proof of Mantel's Theorem that our general approach will imply, and the exact argument we give is often referred to as an Erdős degree majorization argument..

Let G be an n-vertex triangle-free graph and let $v_1 \in V(G)$ be a vertex of maximum degree. Let $V_1' := N(v_1)$ and $V_1 = V(G) \setminus V_1'$. Letting $e(S, T)$ denote the number of edges between two sets S and T, we see that

$$\sum_{u \in V_1} \deg(u) = e(V_1, V_1') + 2e(G[V_1]),$$

simply because every edge counted by $e(V_1, V_1')$ has exactly 1 vertex in V_1 while every edge in $G[V_1]$ has 2 vertices in V_1. Crucially, because G is triangle-free, there can not exist any edges within $V_1' = N(v_1)$, meaning that $e(V_1, V_1') + e(G[V_1]) = e(G)$. In total this implies

$$e(G) \leq \sum_{u \in V_1} \deg(u) - e(G[V_1]) \leq |V_1||V_1'| - e(G[V_1]) \leq \lfloor n^2/4 \rfloor - e(G[V_1]),$$

with this first inequality using that v_1 is a vertex of maximum degree and that $\deg(v_1) = |V_1'|$ by definition of $V_1' = N(v_1)$, and the second inequality used that $|V_1|, |V_1'|$ are non-negative integers with $|V_1| + |V_1'| = n$. This inequality together with the observation $e(G[V_1]) \geq 0$ gives $e(G) \leq \frac{1}{4}n^2$, proving Mantel's Theorem. In fact, a closer look shows we get something stronger: if we assume $e(G) \geq \lfloor n^2/4 \rfloor - t$ then this implies $e(G[V_1]) \leq t$, and if we delete these at most t edges from G then this gives a bipartite graph with bipartition $V_1 \cup V_1'$.

The argument above shows how to prove the result for $r = 3$, and we now sketch out the details of how to iterate this for larger r. Let $V_0' = V(G)$. Iteratively given that we have defined V_{i-1}', if $V'_{i-1} \neq \emptyset$ then we let $G_{i-1} = G[V'_{i-1}]$ and let v_i be a vertex of G_{i-1} of maximum degree in G_{i-1}. Let $V'_i = N_{G_{i-1}}(v_i)$ and $V_i = V'_{i-1} \setminus V'_i$. By the same reasoning as above we have

$$e(V_i, V'_i) + 2e(G[V_i]) = \sum_{u \in V_i} \deg_{G_{i-1}}(u) \leq |V_i||V'_i|.$$

This process produces some non-empty sets $V_1, \ldots, V_s$. Note that we must have $s \leq r-1$, as otherwise the vertices $v_1, \ldots, v_r$ would form a K_r in G. Observe that every edge of G either lies between V_i and V'_i for exactly one i or within $G[V_i]$ for exactly one i. As such, the above implies that

$$e(G) = \sum_{i=1}^{s} e(V_i, V'_i) + e(G[V_i]) \leq \sum_{i=1}^{s} |V_i||V'_i| - \sum_{i=1}^{s} e(G[V_i]).$$

Observe now that $\sum_{i=1}^{s} |V_i||V'_i|$ is exactly the number of edges in the complete s-partite graph with parts $V_1, \ldots, V_s$, and as such this quantity is at most $e(T_s(n)) \leq e(T_{r-1}(n))$. Because $e(G) \geq e(T_{r-1}(n)) - t$, we have $\sum e(G[V_i]) \leq t$. Deleting these at most t edges from G gives an s-partite graph, which in particular is $(r-1)$-partite since $s \leq r-1$, proving the result. □

By using Theorem 9.11 we can bootstrap this stability result for K_r to arbitrary F, giving the following result of Erdős and Simonovits.

Theorem 10.2. *For every non-empty graph F and $\varepsilon > 0$, there exists some n_0 such that if G is an n-vertex F-free graph with*

$$e(G) \geq \frac{\chi(F) - 2}{\chi(F) - 1}\binom{n}{2} - \varepsilon n^2,$$

then G can be made $(\chi(F) - 1)$-partite by removing at most $3\varepsilon n^2$ edges.

Proof. By Theorem 9.11, we can remove at most εn^2 edges from G to obtain a graph G' which is Hom(F)-free. In particular, F has a homomorhpism to $K_{\chi(F)}$ by definition of $\chi(F)$, so G' is $K_{\chi(F)}$-free. Since $e(G') \geq \frac{\chi(F)-2}{\chi(F)-1}\binom{n}{2} - 2\varepsilon n^2 \geq e(T_{\chi(F)-1}(n)) - 2\varepsilon n^2$, we have by Theorem 10.1 that G' can be made $(\chi(F) - 1)$-partite by removing at most $2\varepsilon n^2$ edges. In total then G can be made $(\chi(F) - 1)$-partite by removing at most $3\varepsilon n^2$ edges, proving the result. □

We now give an example of the stability method, which broadly speaking takes a known asymptoticly tight bound for a given extremal problem and turns it into an *exact* tight bound by using a stability theorem which allows one to gradually turn an approximation to the extremal object into the extremal object itself.

Specifically, we aim to use the partite subgraph guaranteed by Theorem 10.2 to show that $\mathrm{ex}(n, F) = e(T_{\chi(F)-1}(n))$ for large n whenever F has a critical edge, which we recall means that $\chi(F - e) < \chi(F)$ for some edge e. Broadly speaking, we will do this by taking an extremal graph G which is almost partite by Theorem 10.2, after which we will zoom in on some induced subgraph $G' \subseteq G$ which is genuinely partite and almost complete between its parts. Finally, we iteratively add vertices back to G' while maintaining these properties, showing that G must have had these properties as well and hence satisfy $e(G) \leq e(T_{\chi(F)-1}(n))$.

Formally, the condition for the induced subgraph G' as described above is the following.

Definition 19. Given a graph G', we say that a partition $V_1, \ldots, V_r$ of $V(G')$ is an ε-almost complete partition if every $x \in V_i$ has $|N(x) \cap V_j| \geq |V_j| - \varepsilon v(G')$ for every $j \neq i$.

Motivated by our sketch above, we show that ε-almost complete partitions are partite and can have new vertices effectively added to them whenever certain conditions are met.

Lemma 10.3. *Let F be a graph with a critical edge and with $\chi(F) = r + 1$ and G' an F-free graph which has an ε-almost complete partition $V_1, \ldots, V_r$.*

(i) If $|V_i| \geq \varepsilon(r-1)v(F)v(G') + v(F)$ for all i, then each V_i is an independent set of G'.

(ii) If a new vertex x is added to G' in such a way that the result graph is still F-free, then there exists some i such that x has at most $\varepsilon(r-1)v(F)v(G') + v(F)$ neighbors in V_i.

Proof. For (i), assume for contradiction that there exists an edge xy in, say, V_1. We use this to show that there exist sets $U_i \subseteq V_i$ of size $v(F)$ for each i such that every edge is present between every U_i and U_j and such that $x, y \in U_1$. Observe that such sets necessarily contain a copy of F, so we will establish our desired contradiction once this is done.

We build our sets U_i inductively, with $U_1 \subseteq V_1$ taken to be an an arbitrary set of $v(F)$ vertices containing x, y. Iteratively given that we have chosen $U_1, \ldots, U_{i-1}$, we have by definition of ε-almost complete partitions that each vertex in each of these sets is adjacent to all but at most $\varepsilon v(G')$ vertices of V_i. As such, we can take $U_i \subseteq V_i$ to be any set of $v(F)$ vertices which avoids these at most $\varepsilon v(G') \cdot (i-1)v(F) \leq \varepsilon(r-1)v(F)v(G')$ non-neighbors, which we can do by assumption on $|V_i|$. This proves (i).

For (ii), assume for contradiction that such an x is adjacent to at least $\varepsilon(r-1)v(F)v(G')+v(F)$ vertices V_i' in each V_i. By replicating the same argument as above, we can find $U_i \subseteq V_i'$ of size $v(F)$ which are all complete to each other which together with x will give a copy of F, a contradiction to G' plus x being F-free. □

There are a few more things we need in order to make effective use of Theorem 10.3. For example, to use (i) we need to argue that the almost partite subgraph of G guaranteed by Theorem 10.2 has part sizes on the order of $\Omega(n)$. This can be done by showing unbalanced partite graphs have too few edges to be anywhere near extremal.

Lemma 10.4. *If G is an r-partite graph on n vertices with $r \geq 2$ and if one of its parts has size at most $n/r - \varepsilon n$, then*

$$e(G) \leq (1 - 1/r)\frac{n^2}{2} - \frac{1}{2}\varepsilon^2 n^2.$$

Proof. It is not too difficult to argue that $e(G)$ is maximized under these conditions if one of it is complete with one of its parts having size exactly $n/r - \varepsilon n$ and the rest having size $n/r + \varepsilon n/(r-1)$. With this we have

$$\begin{aligned} e(G) &= (r-1)(n/r - \varepsilon n)(n/r + \varepsilon n/(r-1)) + \binom{r-1}{2}(n/r + \varepsilon n/(r-1))^2 \\ &= (1 - 1/r)n^2/2 - \frac{r-2}{2r}\varepsilon n^2 - \frac{r}{2r-2}\varepsilon^2 n^2, \end{aligned}$$

and this is at most the desired quantity.

□

Going back to Theorem 10.3; when we apply (ii) we will want $G' \cup \{x\}$ to continue having an ε-almost complete partition so that we can repeatedly apply the lemma. This strategy is doomed to fail if G contains a vertex of small degree, but we can make it work if we add an assumption of large minimum degree.

Proposition 10.5. *Let F be a graph with a critical edge and with $\chi(F) = r+1$. There exists some n_0 such that if G is an n-vertex F-free graph with $n \geq n_0$ and $\delta(G) \geq \delta(T_r(n))$, then $e(G) \leq e(T_r(n))$ with equality only if $G \cong T_r(n)$.*

Proof. Fix some $\varepsilon > 0$ sufficiently small and n_0 sufficiently large so that the following calculations hold. Let G be an n-vertex F-free graph with $\delta(G) \geq \delta(T_r(n))$ and $e(G) \geq e(T_r(n))$. In particular, $e(G) \geq \frac{\chi(F)-2}{\chi(F)-1}\binom{n}{2} - \varepsilon n^2$, so by Theorem 10.2 there exists some partition $V_1, \ldots, V_r$ of $V(G)$ such that $\sum e(G[V_i]) \leq 3\varepsilon n^2$. We begin by showing that this partition is closed to balance and close to complete.

We claim that $|V_i| \geq n/r - 3\sqrt{\varepsilon}n$ for all i. Indeed, if this were not the case then Theorem 10.4 implies

$$e(G) \leq 3\varepsilon n^2 + (1-1/r)\frac{n^2}{2} - \frac{1}{2}(3\sqrt{\varepsilon})^2 n^2 \leq (1-1/r)\frac{n^2}{2} - 1.5\varepsilon n^2,$$

a contradiction to $e(G) \geq e(T_r(n))$.

Let $X_i \subseteq V_i$ be the set of vertices x which have $|N(x) \cap V_j| < |V_j| - \sqrt{\varepsilon}n$ for some j. Then we have that

$$e(G) \leq 3\varepsilon n^2 + e(T_r(n)) - |X_i| \cdot \sqrt{\varepsilon}n,$$

which implies $|X_i| \leq 3\sqrt{\varepsilon}n$ for all i since otherwise $e(G) < e(T_r(n))$ and there would be nothing to prove.

Now let $V_i' = V_i \setminus X_i$ for all i. Note that by construction, each $x \in V_i'$ has

$$|N(x) \cap V_j'| \geq |N(x) \cap V_j| - |X_j| \geq |V_j| - 4\sqrt{\varepsilon}n.$$

Thus $V_1', \ldots, V_r'$ is a $4\sqrt{\varepsilon}$-almost partition of $G' = G[\bigcup V_i']$. In particular, this is a $6\sqrt{\varepsilon}n$-almost partition of G', and for technical reasons this will be slightly more convenient for us to work with.

We now aim to iteratively add vertices from $\bigcup_i X_i$ back to G' while maintaining the property of having an almost partition. To this end, let $x \in \bigcup X_i$ be arbitrary. Because G is F-free, adding any x to G' results in an F-free graph. By Theorem 10.3(ii), there exists some i such that x has at most $6\sqrt{\varepsilon}(r-1)v(F)v(G') + v(F) \leq 6rv(F)\sqrt{\varepsilon}n$ neighbors in V_i', so for all $j \neq i$,

$$\begin{aligned}
|N(x)\cap V_j'| &= |N(x)| - \sum_k |N(x)\cap X_k| - \sum_{k\neq j} |N(x)\cap V_k'|\\
&\geq \delta(G) - \sum_k |X_k| - \sum_{k\neq i,j} |V_k'| - 6r\sqrt{\varepsilon}v(F)n\\
&= \delta(G) - |V(G)| + |V_i'| + |V_j'| - 6r\sqrt{\varepsilon}v(F)n\\
&\geq \delta(G) - n + (n/r - 6\sqrt{\varepsilon}n) + |V_j'| - 6r\sqrt{\varepsilon}v(F)n,
\end{aligned}$$

with this last inequality using our lower bound on $|V_i|$ and upper bound on $|X_i|$.

Crucially, our minimum degree condition on G implies $\delta(G) \geq (1-1/r)n - O(1)$, so we can bound the above by, say,

$$|N(x)\cap V_j'| \geq |V_j'| - 20rv(F)\sqrt{\varepsilon}n.$$

We conclude then that G' together with x has a $20rv(F)\sqrt{\varepsilon}n$-almost partition $V_1'',\ldots,V_r''$, namely the one obtained by adding x to V_i'. Moreover, we observe that

$$|V_j''| \geq |V_j| - |X_j| \geq n/r - 6\sqrt{\varepsilon}n,$$

which for $n \geq n_0$ sufficiently large and ε small is at least $20rv(F)\sqrt{\varepsilon}(r-1)v(F)n + v(F)$. Thus by Theorem 10.3(i), each V_j'' set is an independent set, meaning x in fact has 0 neighbors inside V_i'. Using this, we can redo our calculation above to find for all $j \neq i$ that

$$\begin{aligned}
|N(x)\cap V_j''| &\geq \delta(G) - \sum_k |X_k| - \sum_{k\neq i,j} |V_k''|\\
&= \delta(G) - |V(G)| + |V_i''| + |V_j''|\\
&\geq \delta(G) - n + (n/r - 6\sqrt{\varepsilon}n) + |V_j''|,
\end{aligned}$$

where again we emphasize that the first inequality used our new-found fact that x has no neighbors in V_i'' . As such, we find that this new partition is in fact $6\sqrt{\varepsilon}$-almost complete.

At this point, the exact same calculations as above shows that we can iteratively add vertices of $\bigcup X_i$ to this new partition while maintaining that the graph has a $6\sqrt{\varepsilon}$-almost complete partition. In particular, after adding in all of the vertices of $\bigcup X_i$ we find that the original graph G has a $6\sqrt{\varepsilon}$-almost complete partition $U_1,\ldots,U_r$ where each set in the partition has size at least $n/r - 6\sqrt{\varepsilon}n$. By Theorem 10.3(i) each U_i is an independent set. This means G is in fact r-partite, so $e(G) \leq e(T_r(n))$ with equality only if $G \cong T_r(n)$. □

We can now finish our proof of Simonovits's critical edge theorem which we restate here.

Theorem 10.6. *If F is a graph with a critical edge, then there exists some n_0 such that $\mathrm{ex}(n,F) = e(T_{\chi(F)-1})$ for all $n \geq n_0$, and moreover every n-vertex F-free graph G with $e(G) = e(T_{\chi(F)-1}(n))$ has $G \cong T_{\chi(F)-1}(n))$.*

Proof. Let n_0 be sufficiently large for our following arguments to hold. Let $G_n = G$. Iteratively, if we have defined the m-vertex graph G_m and if $\delta(G_m) < \delta(T_{\chi(F)-1}(m))$, then we let G_{m-1} be the $(m-1)$-vertex graph obtained by deleting some vertex $v_m \in V(G_m)$ of degree less than $\delta(T_{\chi(F)-1}(m))$.

We claim that $e(G_m) \ge e(T_{\chi(F)-1}(m)) + n - m$ for all m for which G_m exists. Indeed, this trivially holds at $m = n$, and inductively it follows that

$$e(G_m) = e(G_{m+1}) - \deg_{G_{m+1}}(v_{m+1}) \ge e(T_{\chi(F)-1}(m+1)) + n - m - 1 - \delta(T_{\chi(F)-1}(m+1)) + 1$$
$$= e(T_{\chi(F)-1}(m)) + n - m,$$

with this last equality using that $T_{\chi(F)-1}(m)$ can be obtained from $T_{\chi(F)-1}(m+1)$ by deleting a vertex of minimum degree (i.e. by deleting a vertex from a largest part).

We claim now that there must exist some $m \ge \sqrt{n}$ such that $\delta(G_m) \ge \delta(T_{\chi(F)-1}(m))$. Indeed, for all m such that G_m exists we have

$$\binom{m}{2} \ge e(G_m) \ge e(T_{\chi(F)-1}(m)) + n - m \ge n - m.$$

This gives a contradiction if m is an integer such that $\binom{m}{2} + m < n$, and for n sufficiently large this holds if $m < \sqrt{n}$. We conclude that G_m can not exist with m so small, and therefore the process of defining the G_m must stop at some $m \ge \sqrt{n}$ meaning that we must have $\delta(G_m) \ge \delta(T_{\chi(F)-1}(m))$ at this point.

Fix m as in the claim above. By taking n sufficiently large we can assume $m \ge \sqrt{m}$ is sufficiently large. By Theorem 10.5 and our claims above, we have

$$e(T_{\chi(F)-1}(m)) \ge e(G_m) \ge e(T_{\chi(F)-1}(m)) + n - m,$$

with the first equality only if $G_m \cong T_{\chi(F)-1}(m)$. These two inequalities above can only hold if $m = n$, so in total we find that $e(G) = e(G_n) \le e(T_{\chi(F)-1}(n))$ with equality only if G is isomorphic to a Turán graph. □

This result determines the Turán number exactly for F with a critical edge, but what about for more general (non-bipartite) F? If F does not have a critical edge, then it is not difficult to show that $\mathrm{ex}(n, F) > e(T_{\chi(F)-1}(n))$ for large n. In fact one can prove the following general lower bound, where here we recall that for a family of graphs $\mathcal{F}$ we define $\mathrm{ex}(n, \mathcal{F})$ to be the maximum number of edges in an n-vertex $\mathcal{F}$-free graph, i.e. one which contains no element of $\mathcal{F}$ as a subgraph.

Proposition 10.7. *Given a graph F, define $\mathcal{M}_F$ to be the set of bipartite graphs F' which can be obtained by taking a $\chi(F)$-proper coloring of F and then deleting $\chi(F) - 2$ of its color classes. Then*

$$\mathrm{ex}(n, F) \ge e(T_{\chi(F)-1}(n)) + \Omega(\mathrm{ex}(n, \mathcal{M}_F)).$$

For example, if F is edge-critical then $\mathcal{M}_F$ contains a graph which is K_2 plus some isolated vertices, and hence Theorem 10.7 only gives the (correct) lower bound of $e(T_{\chi(F)-1}(n))$ for large n. In fact, a result of Simonovits[12] implies that $\mathrm{ex}(n, F) = e(T_{\chi(F)-1}(n)) + \Theta(\mathrm{ex}(n, \mathcal{M}_F))$ for

[12]This result is claimed in the literature, but as far as I know there is no available source which contains the proof of this result. Indeed, most works cite the paper "How to solve a Turán type extremal graph problem? (linear decomposition)" by Simonovits as a place where this result exists, but I personally have been unable to track this article down.

n large. As such, to solve the non-degenerate Turán problem in full one has to understand the Turán numbers for families of bipartite graphs.

10.1 Exercises

1. Prove that if F is not edge-critical then $\mathrm{ex}(n, F) \geq e(T_{\chi(F)-1}(n)) + 1$ whenever $n \geq \chi(F)$. Further, prove that if $\mathcal{M}_F$ is the set of bipartite graphs F' which can be obtained by taking a $\chi(F)$-proper coloring of F and then deleting $\chi(F) - 2$ of its color classes, then

$$\mathrm{ex}(n, F) \geq e(T_{\chi(F)-1}(n)) + \mathrm{ex}\left(\left\lceil \frac{n}{\chi(F) - 1} \right\rceil, \mathcal{M}_F\right)$$

[2-].

Part III

Bonus Topics

11 Random Graphs

TODO. Likely topics: thresholds, connectivity, spreadness theorems

12 Planar Graphs

TODO. Likely topics: Euler's formula, Wagner's Theorem characterizing planar graphs, 5-color theorem, minors, list colorings.

13 Spectral Graph Theory

In this section we give a brief crash course in spectral graph theory, which we recall is the study of matrices associated to graphs.

13.1 Laplacians

The most common matrix to associate to a graph is the adjacency matrix, but this is not always the best matrix to choose depending on your circumstances. Indeed, in one of our exercises we noted how the eigenvalues of the adjacency matrix can not be used to determine whether a graph is connected or not, so the adjacency matrix in general will not be useful for problems which measure how well-connected a graph is. For problems of this sort, a better choice is to use something known as a Laplacian matrix.

Given a graph G, we define $D = D_G$ to be its diagonal matrix of degrees, i.e. the diagonal matrix with $D_{v,v} = \deg(v)$. We define the *normalized Laplacian* $\mathcal{L}(G)$ of G by

$$\mathcal{L}(G) := I - D^{-1/2}AD^{-1/2} = D^{-1/2}(D-A)D^{-1/2}.$$

This definition only makes sense when G has no isolated vertices, though it is easy to tweak the definition to deal with this. For simplicity, we'll assume throughout this section that our graphs do not have isolated vertices. We will often write $\mathcal{L}$ for $\mathcal{L}(G)$ whenever G is understood, and we denote the eigenvalues of $\mathcal{L}$ by $\nu_1(\mathcal{L}) \le \nu_2(\mathcal{L}) \le \cdots \le \nu_n(\mathcal{L})$, or by $\nu_i(G)$ or ν_i. For ease of notation, we will sometimes write $d_u = \deg(u)$.

The definition for $\mathcal{L}$ is somewhat bizarre at first glance, and we'll attempt to motivate it soon. However, to streamline our conversation somewhat, we first begin by establishing some basic facts about $\mathcal{L}$ related to its Rayleigh quotient.

13.1.1 The Rayleigh Quotient of $\mathcal{L}$

An important fact from linear algebra is that if M is a real symmetric matrix, then its largest/smallest eigenvalue is equal to the maximum/minimum of $\frac{x^TMx}{x^Tx}$ where x is a non-zero vector. For the normalized Laplacian, this expression is a little complicated to write in terms of x, but it can be simplified significantly if we look at an appropriately normalized vector y.

Lemma 13.1. *Let x be a real vector and $y = D^{-1/2}x$. Then*

$$\frac{x^T\mathcal{L}x}{x^Tx} = \frac{\sum_{u\sim v}(y_u - y_v)^2}{\sum d_u y_u^2}.$$

Proof. Using $x = D^{1/2}y$ and the definition of $\mathcal{L}$, we see that

$$\frac{x^T\mathcal{L}x}{x^Tx} = \frac{y^T(D-A)y}{y^TDy}.$$

It's straightforward to see that the denominator is $\sum d_u y_u^2$. For the numerator, setting $z = (D-A)y$, we see that

$$z_u = d_u y_u - \sum_{u \sim v} y_v,$$

so

$$y^T z = \sum_u d_u y_u^2 - \sum_{u \sim v} y_u y_v,$$

and from here a little algebra gives the result. □

With this lemma as motivation, for any vector x we define its *harmonic vector* $y := D^{-1/2}x$. If x is an eigenvector for $\mathcal{L}$, then we say that y is a *harmonic eigenvector.*

A number of results about the spectrum of $\mathcal{L}$ can be obtained by looking at its Raleigh quotient.

Corollary 13.2. *Let $\nu_1 \leq \cdots \leq \nu_n$ be the eigenvalues of $\mathcal{L}$.*

(a) We have $\nu_1 = 0$. Moreover, the all 1's vector $\mathbf{1}$ is a corresponding harmonic eigenvector.

(b) We have $\nu_n \leq 2$ with equality iff G contains a bipartite component.

(c) If $n \geq 2$ then

$$\nu_2 = \min_{y: \sum d_u y_u = 0,\ y \neq 0} \frac{\sum_{uv \in E(G)} (y_u - y_v)^2}{\sum d_u y_u^2},$$

and this is positive iff G is connected.

Proof. For (a), observe that $\frac{\sum_{uv \in E(G)} (y_u - y_v)^2}{\sum d_u y_u^2} \geq 0$ for all y, so $\nu_1 \geq 0$ with equality being achieved by considering the all 1's vector.

For (b), we use the inequality $(a-b)^2 \leq 2a^2 + 2b^2$ (which is equivalent to saying $(a+b)^2 \geq 0$) to observe that

$$\frac{\sum_{u \sim v} (y_u - y_v)^2}{\sum d_u y_u^2} \leq \frac{\sum_{u \sim v} 2y_u^2 + 2y_v^2}{\sum d_u y_u^2} = 2,$$

where equality holds if and only if $y_u = -y_v$ for all $u \sim v$. This implies $\nu_n \leq 2$, and if some y achieves equality, then $\{u : y_u \neq 0\}$ defines a bipartite component whose bipartition is determined by the sign of y_u. On the other hand, if there is a bipartite component with parts U, V then taking $y_u = 1$ for $u \in U$, $y_v = 1$ for $v \in V$, and $y_w = 0$ otherwise shows that 2 is an eigenvalue.

For (c), we again recall from linear algebra that ν_2 will be the minimum of the Rayleigh quotient over vectors x that are orthogonal to the eigenvector for ν_1. By (a) we know that $D^{1/2}\mathbf{1}$ is an eigenvector for ν_1, and hence we want to consider x with

$$0 = \langle x, D^{1/2}\mathbf{1} \rangle = \langle D^{1/2}x, \mathbf{1} \rangle = \langle Dy, \mathbf{1} \rangle = \sum d_u y_u,$$

where $y = D^{1/2}x$. This gives the characterization of ν_2. Now assume $\nu_2 = 0$ and let y be a non-zero harmonic eigenvector for ν_2. It is straightforward to see that $\nu_2 = 0$ is only possible if $y_u = y_v$ whenever $u \not\sim v$, and from this it follows that $y_u = y_v$ whenever u, v are in the same connected component of G. If G had only one connected component then this would imply $\sum_u d_u = 0$ since y_u is the same non-zero value for all u, but this is only possible if G is empty, a contradiction to G being connected and $n \geq 2$. On the other hand, say G is disconnected with C, C' being non-empty components which partition $V(G)$. It is then straighforward to find real numbers a, b not both 0 such that $a \sum_{u \in C} d_u = b \sum_{u \in C'} d_u$, and thus defining the vector y which has $y_u = a$ for $u \in C$ and $y_u = b$ for $u \in C'$ shows $\nu_2 = 0$. □

With these preliminaries established, we can now move onto one of our main applications of the noramlized Laplacian.

13.1.2 Random Walks

Let G be an n-vertex graph and let $\pi \in \mathbb{R}^n_{\geq 0}$ be a non-negative vector with $\sum \pi_i = 1$. We define the *simple random walk* on G with initial distribution π as follows. For the 0th step of the walk, we start at vertex u with probability π_u. Given that you are each vertex u at step t, uniformly at random choose a neighbor of u and walk to that vertex. That is, move to vertex v with probability $1/\deg(u)$ if $u \sim v$. We let $\pi_u^{(t)}$ denote the probability that you are at vertex u after the tth step of the walk.

We wish to understand what these $\pi^{(t)}$ vectors converge to as t increases and how quickly this convergence happens. Towards this end, we first observe that $\pi^{(t+1)} = AD^{-1}\pi^{(t)}$ where A is the adjacency matrix of G and D is the diagonal matrix of degrees. In the language of Markov chains, we would say that the matrix

$$P := AD^{-1}$$

is the probability transition matrix for this stochastic process.

If P happened to have n orthogonal eigenvectors, then we could study how π converges by expanding it in a basis of eigenvectors. It is not obvious that this always works because P is not a symmetric matrix, but this does turn out to be true. Moreover, the eigenvectors/eigenvalues of P correspond to the eigenvectors/eigenvalues of $\mathcal{L}$.

Lemma 13.3. *If $x_1, \ldots, x_n$ are a set of orthogonal eigenvectors for $\mathcal{L}$ corresponding to $\nu_1, \ldots, \nu_n$, then $D^{1/2}x_1, \ldots, D^{1/2}x_n$ are eigenvectors for P corresponding to eigenvalues $1 - \nu_1, \ldots, 1 - \nu_n$.*

Proof. This follows from the observation that P is similar to the matrix

$$M = D^{-1/2}PD^{1/2} = D^{-1/2}AD^{-1/2} = I - \mathcal{L}.$$

Thus, if x_i is an eigenvector of $\mathcal{L}$ corresponding to ν_i, we have

$$PD^{1/2}x_i = D^{1/2}(I - \mathcal{L})x = (1 - \nu_i)D^{1/2}x_i = (1 - \nu_i)Dx_i.$$

□

As an aside, the matrix $M = D^{-1/2}AD^{-1/2}$ defined above has essentially the same spectrum as $\mathcal{L}$. As such, working with M is in principle the same as working with $\mathcal{L}$, but in practice both can be useful to consider depending on the context. For example, it is sometimes nicer to work with $\mathcal{L}$ since it is positive semidefinite, which implies e.g. that it has a nice Raleigh quotient; and sometimes it is nicer to work with M since it is non-negative, which implies e.g. that it has a Perron eigenvector.

From Theorem 13.3 and our observations about the Raleigh quotient of $\mathcal{L}$, we can quickly deduce the following convergence result for $\pi^{(t)}$. We note that in the statement of this corollary, the y_i vectors are the harmonic eigenvectors of $\mathcal{L}$.

Corollary 13.4. *Let G be a graph with degrees $d_1, \dots, d_n$, let $\hat{\pi} := (\frac{d_1}{\sum d_i}, \frac{d_2}{\sum d_i}, \cdots)$ and let $D^{1/2}y_1, \dots, D^{1/2}y_n$ be an orthogonal set of eigenvectors of $\mathcal{L}$ corresponding to the eigenvalues $\nu_1, \dots, \nu_n$ with $y_1 = \mathbf{1}$. Given a distribution π, let $c_i = \frac{\langle \pi, y_i \rangle}{\langle y_i, Dy_i \rangle}$. Then*

$$\pi^{(t)} = P^t \pi = \hat{\pi} + \sum_{i \neq 1} c_i (1 - \nu_i)^t D y_i.$$

Further, $\pi^{(t)} \to \hat{\pi}$ for all π if and only if G is connected and not bipartite.

The key takeaway here is that if G is connected and not bipartite, then every simple random walk converges to an (explicit) stable distribution $\hat{\pi}$. The conditions G connected and not bipartite are necessary, as otherwise one can cook up examples where some initial distribution π does not converge to $\hat{\pi}$.

Proof. Because the $D^{1/2}y_i$ vectors form an orthogonal basis, we can write $D^{-1/2}\pi = \sum c_i D^{1/2} y_i$ with

$$c_i = \frac{\langle D^{-1/2}\pi, D^{1/2}y_i \rangle}{\langle D^{1/2}y_i, D^{1/2}y_i \rangle} = \frac{\langle \pi, y_i \rangle}{\langle y_i, Dy_i \rangle}.$$

Thus $\pi = \sum c_i D y_i$. As each Dy_i vector is an eigenvector of P corresponding to the eigenvalue $(1 - \nu_i)$ by Theorem 13.3, we have

$$\pi^{(t)} = P^t \pi = \sum c_i (1 - \nu_i)^t D y_i.$$

Because $y_1 = \mathbf{1}$, we have

$$c_1 = \frac{\pi^T \mathbf{1}}{\mathbf{1}^T D \mathbf{1}} = \frac{\sum \pi_u}{\sum d_i} = \frac{1}{\sum d_i},$$

so $c_1 (1 - \nu_1)^t D y_1 = \hat{\pi}$, and combining this with the equality above gives the first result.

For the second result, note that we have $|1 - \nu_i| < 1$ for all $i > 1$ if and only if G is connected and not bipartite by Corollary 13.2, so in this case $\pi^{(t)}$ converges to $\hat{\pi}$. If G is not connected, then taking π to have $\pi_u = 1$ for some vertex u will not converge to $\tilde{\pi}$ since the support of $\pi^{(t)}$ will lie entirely in the connected component of u. Similarly if G is bipartite then taking $\pi_u = 1$ for some vertex u will cause $\pi^{(t)}$ to have all of its support in either one part or the other depending on the parity of t. $\square$

By using $(1 - \nu_i)^t \approx e^{-t\nu_i}$, it is possible to use this previous result to say something about the speed of convergence to π, the proof of which we omit.

Corollary 13.5. *Let G be a graph with maximum and minimum degrees $\Delta, \delta \geq 1$. Let*

$$\nu' := \min\{\nu_2, 2 - \nu_n\}.$$

Then for any distribution π, we have

$$\left\|\pi^{(t)} - \hat{\pi}\right\|_2 \leq e^{-t\nu'}\sqrt{\frac{\Delta}{\delta}},$$

i.e. the walk converges to $\hat{\pi}$ in the L_2 norm in roughly $t = \log(\sqrt{\Delta/\delta})/\nu'$ steps .

Note that if G is either bipartite or not connected then $\nu' = 0$, in which case the statement of Theorem 13.5 is trivial. Maybe mention this is called the spectral gap if that's correct.

The definition for ν' is somewhat inelegant, and one can get around this by using a *lazy random walk*. This is defined by performing a modified simple random walk where at each step, you choose to stay at your current vertex with probability $\frac{1}{2}$. Equivalently, this can be thought of as a simple random walk after adding $\deg(u)$ loops to each vertex of your graph. If $\tilde{\mathcal{L}}$ is the analog of the normalized Laplacian matrix for this non-simple graph, then it turns out $\tilde{\mathcal{L}} = \frac{1}{2}\mathcal{L}$ so all its eigenvalues are halved (i.e. this random walk is "twice as slow" as the original one). In particular, $\tilde{\nu}_n \leq 1$, which means the analog of ν' will always be equal to $1 - \tilde{\nu}_2$. Morally this happens because in a lazy random walk, being bipartite no longer prevents you from converging to a unique stationary distribution.

13.1.3 The Cheeger Inequality

Recall for $S, T \subseteq V(G)$ we let $e(S, T)$ denote the number of edges uv with $u \in S, v \in T$. We let $\overline{S} = V(G) \setminus S$ and define the volume of a set $\text{vol}(S) = \sum_{u \in S} \deg(u)$, which one can think of as the size of a set weighted by the degrees of its vertices.

We just showed that $\mathcal{L}$ can tell us how quickly a random walk on a graph G converges. Intuitively, a random graph will take a long time to converge if there exists a set $S \subseteq V(G)$ such that $\text{vol}(S), \text{vol}(\overline{S})$ are large but $e(S, \overline{S})$ is very small (for example, this happens if G consists of two disjoint cliques connected by a single edge). Thus intuitively $\mathcal{L}$ should be able to determine whether there is a small set of edges such that deleting these breaks the graph into two dense pieces, which is something that's useful to computer scientists for various reasons. And indeed, this intuition is correct. To be more precise, let

$$h_G(S) = \frac{e(S, \overline{S})}{\min\{\text{vol}(S), \text{vol}(\overline{S})\}}.$$

That is, this measures how many edges you have to delete in order to separate S from its complement, where you normalize by the total number of edges involving either S or $\overline{S}$. We define the *Cheeger ratio* or *conductance* of G by

$$h_G = \min_{S \neq \emptyset, V(G)} h_G(S).$$

Thus roughly h_G measures the fewest number of edges you need to delete to separate the graph into two (large) components. The following result shows that h_G is comparable to ν_2.

Theorem 13.6 (Cheeger's inequality)**.** *For any graph G, we have*

$$\frac{h_G^2}{2} \le \nu_2 \le 2h_G.$$

Proof Sketch. In view of Theorem 13.2, upper bounding ν_2 is equivalent to constructing a vector y with $\sum_u d_u y_u = 0$ such that the Raleigh quotient of $\mathcal{L}$ is small. Given that we want to bound ν_2 in terms of h_G, it makes sense to define y with respect to some partition $S, \overline{S}$ of $V(G)$. In particular, given a set S, one natural way to define y is to have y_u be some value on S and some other value on $\overline{S}$. Playing around with numbers, we are led to define the vector y given a set S by $y_u = \mathrm{vol}(S)^{-1}$ if $u \in S$ and $y_v = -\mathrm{vol}(\overline{S})^{-1}$ otherwise. Note that $\sum d_u y_u = 1 - 1 = 0$, so by Corollary 13.2 we have

$$\nu_2 \le \frac{\sum_{u\sim v}(y_u - y_v)^2}{\sum d_u y_u^2}.$$

Each non-zero term in the numerator is given by an edge uv which is counted by $e(S, \overline{S})$. Further, $\sum_{u\in S} d_u y_u^2 = \mathrm{vol}(S)^{-1}$ and $\sum_{v\in \overline{S}} d_v y_v^2 = \mathrm{vol}(\overline{S})^{-1}$. Define $\mathrm{vol}(G) = \mathrm{vol}(V(G)) = \mathrm{vol}(S) + \mathrm{vol}(\overline{S})$. Thus the Rayleigh quotient above equals

$$\frac{e(S,\overline{S})(\mathrm{vol}(S)^{-1} - \mathrm{vol}(\overline{S})^{-1})^2}{\mathrm{vol}(S)^{-1} + \mathrm{vol}(\overline{S})^{-1}} = \frac{e(S,\overline{S})(\mathrm{vol}(S) - \mathrm{vol}(\overline{S}))^2}{\mathrm{vol}(S)\mathrm{vol}(\overline{S})(\mathrm{vol}(\overline{S}) + \mathrm{vol}(S))} = \frac{e(S,\overline{S})(\mathrm{vol}(G) - 2\cdot\mathrm{vol}(S))^2}{\mathrm{vol}(S)\mathrm{vol}(\overline{S})\mathrm{vol}(G)}.$$

Without loss of generality we can assume $\mathrm{vol}(S) = \min\{\mathrm{vol}(S), \mathrm{vol}(\overline{S})\} \le \frac{1}{2}\mathrm{vol}(G)$, and hence $\overline{S} \ge \frac{1}{2}\mathrm{vol}(G)$. Thus the above expression is at most $\frac{e(S,\overline{S})(\mathrm{vol}(G))^2}{\frac{1}{2}\mathrm{vol}(S)(\mathrm{vol}(G))^2} = 2h_G(S)$. Taking S so that $h_G = h_G(S)$ gives the result.

The lower bound is more involved and we omit its full proof. The rough idea is to consider a harmonic eigenvector y corresponding to ν_2. It will then turn out that there exists a number γ such that $S = \{u : y_u < \gamma\}$ has $h_G(S)$ small. Intuitively this is because y corresponding to ν_2 means its Rayleigh quotient $\frac{\sum_{u\sim v}(y_u - y_v)^2}{\sum d_u y_u^2}$ must be small, which means there should be few edges u, v which don't have $y_u \approx y_v$. Thus the set S for some appropriately defined γ should have few edges from it to $\overline{S}$. For a complete proof we refer the reader to the book of Chung, which also goes into much more detail about the normalized Laplacian. □

13.1.4 The Combinatorial Laplacian

In addition to the adjacency matrix and normalized Laplacian matrix, one of the most studied matrices associated to a graph is the *combinatorial Laplacian*

$$L = L(G) = D - A.$$

For time constraints we will not discuss this matrix at great length and instead we will close only on the following nice fact.

Theorem 13.7 (Kirchoff's Matrix-Tree Theorem)**.** *Let G be an n-vertex graph and let $\tau(G)$ denote the number of spanning trees of G. Then*

$$\tau(G) = \frac{1}{n}\prod_{i=2}^{n} \lambda_i,$$

where $\lambda_1 \leq \cdots \leq \lambda_n$ are the eigenvalues of L.

For example, if $G = K_n$ then it is straightforward to check that the eigenvalues of L are 0 with multiplicity 1 and n with multiplicity $n-1$, so this implies that the number of spanning trees on $[n]$ is exactly n^{n-2}.

13.2 Regular Graphs

While the matrices $A, \mathcal{L}, L$ give different incomparable spectral information in general, these matrices are "equivalent" to each other whenever the graph G is regular. Indeed, if G is a d-regular graph then $\mathcal{L} = I - d^{-1}A$ and $L = dI - A$. As such, the spectrum of $\mathcal{L}, L$ are just linear transformations of the spectrum of A, so knowing the eigenvalues of one matrix is the same as knowing the eigenvalues of the other. That is, for regular graphs the adjacency matrix A can be used to prove anything that can be proven with either $\mathcal{L}$ or L (as well as many things beyond just this). As such, spectral graph theory is generally at its strongest when working with regular graphs. Here we consider a few special classes of regular graphs where even more can be said.

13.2.1 Strongly Regular Graphs

We say that G is a *strongly regular graph* with parameters[13] (n, d, ℓ, m) if G is an n-vertex d-regular graph such that any pair of adjacent vertices have ℓ common neighbors and any pair of non-adjacent vertices have m common neighbors. For brevity we will sometimes say that such a graph is an (n, d, ℓ, m)-srg. For example, K_n is an $(n, n-1, n-2, m)$-srg for any m, $K_{n,n}$ is a $(2n, n, 0, n)$-srg, and one can check that the Petesron graph is a $(10, 5, 0, 1)$-srg. Strongly regular graphs have many nice properties, and they have a particularly nice characterization of their eigenvalues.

Lemma 13.8. *If G is an (n, d, ℓ, m)-srg which is not the disjoint union of cliques, then the set of eigenvalues of A_G is exactly*

$$\{d, \frac{\ell - m + \sqrt{(\ell-m)^2 + 4(d-m)}}{2}, \frac{\ell - m - \sqrt{(\ell-m)^2 + 4(d-m)}}{2}\},$$

and the eigenvalue d appears with multiplicity 1.

Proof. The key observation is that

$$A^2 = \ell A + m(J - A - I) + dI$$

where J is the all 1's matrix. Indeed, this follows from $A^2_{u,v}$ being the number of walks form u to v of length 2, which is the number of common neighbors of u and v.

The fact that d is an eigenvalue follows from the fact that the all 1's vector is an eigenvector with this value. Consider now any eigenvector x with eigenvalue λ which is orthogonal to the all 1's

[13] Typically one writes λ, μ instead of ℓ, m, but this conflicts a little with our eigenvalue notation.

vector. Note that $Jx = 0$, so applying x to the expression above gives $\lambda^2 = (\ell - m)\lambda + (d - m)$, and solving for λ shows that λ must equal $\frac{\ell - m \pm \sqrt{(\ell-m)^2 + 4(d-m)}}{2}$. Thus the set of eigenvalues for A must lie within the proposed set, and moreover d appears with multiplicity 1 since any eigenvector orthogonal to it has a different eigenvalue[14].

It remains only to show that all three of these value occur as eigenvalues for G. To see this, we note that since G is not the disjoint union of cliques that it is connected with diameter at least 2, so A must have at least 3 distinct eigenvalues by an exercise you all did. □

Lot's of nice graphs turn out to be strongly regular graphs. One place they turn up is in the study of Moore graphs, which are motivated by the following observation.

Lemma 13.9 (Moore bound)**.** *If G is a d-regular graph with diameter k, then*

$$|V(G)| \le 1 + d \sum_{i=0}^{k-1} (d-1)^i.$$

Moreover, equality holds if and only if G has girth $2k+1$.

Proof. Pick any vertex x. It is not hard to show inductively that the number of vertices at distance $i \ge 1$ from x is at most $d(d-1)^i$, with this bound being tight if and only if x is not in a cycle of length at most $2i$. Summing up these bounds and using that every vertex is within distance k of some x gives the bound.

□

Any graph meeting the Moore bound is called a *Moore graph*. When $k = 1$ the unique Moore graph is K_n, so the first non-trivial case to consider is when $k = 2$.

Theorem 13.10 (Hoffman-Singleton REF)**.** *If G is a d-regular Moore graph of diameter 2, then $d \in \{2, 3, 7, 57\}$.*

Proof. We claim that such a G must be a $(d^2+1, d, 0, 1)$-srg. Indeed, every pair of adjacent vertices have 0 common neighbors (otherwise G would contain a triangle), and every pair of non-adjacent vertices have at most 1 common neighbor (otherwise G would contain a C_4) and at least 1 (otherwise G would have diameter larger than 2).

With this, we know that the eigenvalues of G are d and $\frac{-1 \pm \sqrt{4d-3}}{2}$. If r is the multiplicity of the eigenvalue $\frac{-1+\sqrt{4d-3}}{2}$ (and hence $d^2 - r$ is the multiplicity of the last one), then the fact that A has trace 0 implies

$$0 = 1 \cdot d + r\left(\frac{-1+\sqrt{4d-3}}{2}\right) + (d^2 - r)\left(\frac{-1-\sqrt{4d-3}}{2}\right) = d - \frac{1}{2}d^2 + \frac{1}{2}(2r - d^2)\sqrt{4d-3}. \quad (6)$$

We wish to show that this equation is impossible for all d outside of the values that we claimed. First of all, if $\sqrt{4d-3}$ is irrational, then we must have $2r - d^2 = 0$ for (6) to be rational. In this case (6) implies $d^2 = 2d$, and hence $d = 2$.

[14]Here we implicitly use that if G is not a disjoint union of cliques then the three possible values for eigenvalues are all distinct from each other, as can be checked by using the diameter argument we use at the end, for example.

If instead $\sqrt{4d-3}$ is rational, then $4d - 3 = s^2$ for some integer s. Solving for d gives $d = (s^2+3)/4$, and plugging this back into (6) and scaling both sides by 32 gives

$$0 = 8s^2 + 24 - s^4 - 6s^2 - 9 + 32(2r - d^2)s = [8s - s^3 - 6s + 32(2r - d^2)]s - 15.$$

Crucially, this implies that s is a divisor of 15. This means $s \in \{1, 3, 5, 15\}$, and plugging these values into $d = (s^2+3)/4$ recovers the stated possible values for d, proving the result. □

As an aside, Moore graphs of diameter 2 are known to exist for $d \in \{2, 3, 7\}$: these are C_5, the Peterson graph, and a graph known as the Hoffman-Singleton graph. It is a major problem to determine whether such a graph exists with $d = 57$ with incorrect proofs claiming that such a graph doesn't exist popping up from time to time.

13.2.2 Cayley Graphs

The indexing here isn't great/consistent throughout

Another important class of graphs for spectral graph theory are Cayley graphs. Given a group Γ, we say that a subset $S \subseteq \Gamma$ is closed under inverses if every $g \in S$ has $g^{-1} \in S$. For such Γ, S, we define the Cayley graph $C(\Gamma, S)$ to be the graph with vertex set Γ and where g, h are adjacent to each other if and only if $gh^{-1} \in S$ (or equivalently $hg^{-1} \in S$ since S is closed under inverses). The eigenvalues and eigenvectors of $C(\Gamma, S)$ are relatively simple to describe.

Lemma 13.11. *Recall that a function $\chi : \Gamma \to \mathbb{C} \setminus \{0\}$ is a character if χ is a group homomorphism where $\mathbb{C} \setminus \{0\}$ is given the group operation of multiplication. For any group Γ, character χ, and set $S \subseteq \Gamma$ closed under inverses, we have that the vector x defined by $x_g = \chi(g)$ is an eigenvector of $C(\Gamma, S)$ with eigenvalue $\sum_{s \in S} \chi(s)$.*

Proof. We observe that

$$(Ax)_g = \sum_{h: h \sim g} x_h = \sum_{h: hg^{-1} \in S} \chi(h) = \chi(g) \sum_{h: hg^{-1} \in S} \chi(hg^{-1}) = x_g \sum_{s \in S} \chi(s),$$

with this second to last step using that χ is a homomorphism. □

This lemma is particularly effective when Γ is a finite abelian group. Exposiate more on how in this case we get n orthogonal characters so in fact all eigenvalues are of this form.

Corollary 13.12. *Let $\Gamma = \prod_{i=1}^k \mathbb{Z}_{n_i}$ and let $\omega_k = e^{2i\pi/k}$. For each $g = (g_1, \ldots, g_k) \in \Gamma$, define the vector x_g by $(x_g)_h = \prod \omega_{n_i}^{g_i h_i}$ for each $h = (h_1, \ldots, h_k) \in \Gamma$. For any $S \subseteq \Gamma$, the vectors x_g form an orthogonal basis of eigenvectors for $C(\Gamma, S)$ with corresponding eigenvalue $\sum_{s \in S} \prod \omega_{n_i}^{g_i s_i}$.*

For example, consider the cycle graph C_n which can be written as the Cayley graph $C(\mathbb{Z}_n, \{1, -1\})$. This corollary then says that for each $j \in \{0, 1, \ldots, n-1\}$, C_n has eigenvalue

$$\omega_n^j + \omega_n^{-j} = e^{2ij\pi/n} + e^{-2ij\pi/n} = 2\cos(2j\pi/n).$$

As another example, the n-dimensional hypercube graph Q_n can be written as the Cayley graph $C(\mathbb{Z}_2^n, \{s_1, \ldots, s_n\})$ where s_j is the vector with a 1 in position j and 0's elsewhere. Since $\omega_2 = e^{i\pi} = -1$, we find for each binary string x that Q_n has eigenvalue

$$\sum_k (-1)^{x_k} = (n - |x|) - |x| = n - 2|x|,$$

where here $|x|$ denotes the number of 1's in x. Moreover, the multiplicity of $n - 2k$ is exactly $\binom{n}{k}$, i.e. the number of binary strings with exactly k 1's.

While the above are some rather simple examples of Cayley graphs, many more sophisticated examples are known. We in particular record the following construction of Kopparty for future use, which itself is based on an earlier construction of Alon.

Theorem 13.13. *For a prime p, let $S \subseteq \mathbb{Z}_p^3$ be defined by*

$$S = \{(xy, xy^2, xy^3) : x \in \mathbb{Z},\ p/3 < x < 2p/3,\ y \in \mathbb{Z}_p \setminus \{0\}\},$$

noting that this set is closed under inverses. Then the Cayley graph $C(\mathbb{Z}_p^3, S)$ is triangle-free and every eigenvalue $\lambda \neq |S|$ has $|\lambda| = O(p \log p)$ Then the Cayley graph

13.2.3 Expanders

Roughly speaking, *expanders* are (sparse) graphs G with the property that every "small" set of vertices S has many neighbors with these graphs playing a large role in both combinatorics and computer science. One of the strongest forms of expansion comes from graphs which satisfy certain spectral conditions.

To this end, we let λ_i denote the ith largest eigenvalue of G's adjacency matrix. We say that a graph G is an (n, d, λ)-graph if it is an n-vertex d-regular graph such that $\max_{i \neq 1} |\lambda_i(G)| \leq \lambda$. We will somewhat colloquially say that G is a "spectral expander" or a "pseudorandom graph" if G is an (n, d, λ)-graph with λ small relative to d, i.e. if G has a large spectral gap. The reason we care about such graphs is due to the famed expander mixing lemma.

Lemma 13.14 (Expander mixing lemma)**.** *If G is an (n, d, λ)-graph, then for any subsets $S, S' \subseteq V(G)$, we have*

$$\left| e(S, S') - \frac{d}{n}|S||S'| \right| \leq \lambda \sqrt{|S||S'|}.$$

Note that $\frac{d}{n}|S||S'|$ is exactly the number of edges between S and S' we would expect in a random graph with $p = d/n$, i.e. the expander mixing lemma says that (n, d, λ)-graphs behave like random graphs, hence the reason for calling such graphs "pseudorandom."

Proof. Let $\mathbf{1}_S$ be the indicator vector for S, i.e. $(\mathbf{1}_S)_v = 1$ if $v \in S$ and is 0 otherwise, and similarly define $\mathbf{1}_{S'}$. Define $M = A - \frac{d}{n}J$ and observe that $e(S, S') - \frac{d}{n}|S||S'| = \mathbf{1}_S^T M \mathbf{1}_{S'}$. Thus it suffices to determine how large $|\mathbf{1}_S^T M \mathbf{1}_{S'}|$ can be.

Let $x_1, \ldots, x_n$ denote an orthonormal set of eigenvectors for A corresponding to $\lambda_1, \ldots, \lambda_n$. Observe that x_1 is proportional to the all 1's vector, so

$$Mx_1 = Ax_1 - \frac{d}{n}Jx_1 = dx_1 - dx_1 = 0.$$

On the other hand, because each x_i with $i \neq 1$ is orthogonal to the all 1's vector, we have $Mx_i = \lambda_i x_i$ for all $i \neq 0$. We conclude that the eigenvalues of M are the same as A except with λ_1 replaced by 0. Thus by Cauchy-Schwarz and the fact that every eigenvalue of M has absolute value at most λ,

$$|\mathbf{1}_S^T M \mathbf{1}_{S'}| \le \|\mathbf{1}_S\|_2 \|M\mathbf{1}_{S'}\| \le \|\mathbf{1}_S\|_2 \cdot \lambda \|\mathbf{1}_{S'}\| = \lambda\sqrt{|S||S'|},$$

proving the result. □

Note that the smaller λ is, the more effective the expander mixing lemma is. As such it's natural to ask how small λ can be in terms of d, and the answer turns out to be roughly $\sqrt{d}$. One way to see this is by applying the expander mixing lemma with $S = \{v\}$ and $T = N(v)$. In this case the expander mixing lemma implies $d - d^2/n < \lambda\sqrt{d}$, which implies that λ must be asymptotically at least $\sqrt{d}$ if $d = o(n)$. One can improve this somewhat.

Theorem 13.15 (Alon-Boppana)**.** *If G is an (n, d, λ)-graph with d fixed, then $\lambda \ge 2\sqrt{d-1} - o(1)$.*

Graphs with $\lambda \le 2\sqrt{d-1}$ are called *Ramanujan graphs*. Such graphs are known to exist in certain cases and are of great interest. For our purposes, the main point is that (n, d, λ) graphs with $\lambda = O(\sqrt{d})$ are very good expanders.

Let us now discuss a few examples of (n, d, λ)-graphs with λ small in terms of d. First of all, it should be intuitive that a random d-regular graph ought to have λ small (since this means the graph is "psuedo-random"). And indeed, a difficult result of Friedman says that such graphs satisfy $\lambda = (2 + o(1))\sqrt{d-1}$ with high probability. Amongst the graph classes we have looked at, Lemma 13.8, implies the non-trivial eigenvalues of an (n, d, ℓ, m)-srg will have magnitude roughly $|\ell - m| + \sqrt{d}$, so these will be good expanders as long as ℓ, m are close to each other. Also the Cayley graph construction of Kopparty from Theorem 13.13 is a triangle-free (n, d, λ)-graph with $n \approx p^3$, $d = |S| \approx p^2$, and $\lambda \approx p \log p \approx d^{1/2} \log(d)$.

Let us explore some further properties of (n, d, λ)-graphs with an eye towards proving lower bounds for Ramsey numbers.

Lemma 13.16 (Alon-Rödl)**.** *If G is an (n, d, λ)-graph with $d \ge 1$, $\lambda > \frac{1}{2}$, and if $t \ge 2n \log^2 n/d$ is an integer, then the number of independent sets of size t in G is at most $(2e^2\lambda/\log^2 n)^t$.*

Proof. Let $B_1 = V$, and iteratively given an independent set $v_1, \ldots, v_{i-1}$, define B_i to be the set of vertices which are not adjacent nor equal to any of the v_j vertices. Let $C_i = \{v : |N(v) \cap B_i| \le \frac{d|B_i|}{2n}\}$, i.e. these are the sets of vertices which roughly have at most half the average number of neighbors in B_i.

Note that every (ordered) independent set $v_1, \ldots, v_t$ is obtained by iteratively picking $v_i \in B_i$. Moreover, such a sequence can have at most $s := \frac{2n}{d} \log n$ many $v_i \in B_i \setminus C_i$, since each time this happens we have $|B_{i+1}| \le (1-d/2n)|B_i|$. Thus typically $v_i \in B_i \cap C_i$, and it remains to show this set is small. To do this, observe by definition of C_i that $e(B_i, C_i) \le \frac{d|B_i||C_i|}{2n}$, so by the expander mixing lemma we have

$$\frac{d|B_i||C_i|}{2n} \le \frac{d|B_i||C_i|}{n} - e(B_i, C_i) < \lambda\sqrt{|B_i||C_i|} \implies \frac{4n^2\lambda^2}{d^2} > |B_i||C_i| \ge |B_i \cap C_i|^2,$$

so $|B_i \cap C_i| \le 2n\lambda/d$ for all i. In total then, the number of (unordered) sets of independent sets of size t is at most

$$\frac{1}{t!}\binom{t}{s} n^s (2n\lambda/d)^{t-s},$$

where $\binom{t}{s}$ chooses the positions where we (might) choose an element outside of C_i, n^s represents these choices for vertices outside of C_i, and where $1/t!$ goes from ordered tuples to unordered sets. Using $\binom{t}{s} \le 2^t$ and $t! \ge (t/e)^t$, we find that this quantity is at most

$$(2e/t)^t n^s (2\lambda n/d)^t (d/2\lambda n)^s = \left(\frac{4e\lambda n}{dt}\right)^t \left(\frac{d}{2\lambda}\right)^s.$$

Give more detail on this final derivation. □

For a graph F, let $R(F,t)$ denote the maximum number of vertices that a graph can have if it's F-free and contains no independent set of size t. Note that for all graphs on F vertices, $r(F,t) \le r(K_s,t) = r(s,t)$.

Theorem 13.17 (Mubayi-Verstraëte). *If there exists an F-free (n,d,λ)-graph, then with $t = \lceil 2n\log^2 n/d \rceil$ we have*

$$R(F,t) = \Omega\left(\frac{n}{\lambda}\log^2 n\right).$$

Proof. Let G_p be the random induced subgraph of G obtained by including each vertex independently and with probability p. Let $\mathcal{I}$ be the collection of independent sets of size t in G_p, and let G'_p be the graph obtained by deleting one vertex from each element of $\mathcal{I}$. By the previous lemma we have

$$\mathbb{E}[|V(G'_p)|] = \mathbb{E}[|V(G_p)| - |\mathcal{I}|] \le pn - p^t(2e^2\lambda/\log^2 n)^t.$$

Now if we take $p = (\log^2 n/2e^2\lambda)$ we find that this expectation is at least $pn - 1 = \Omega(\frac{n}{\lambda}\log^2 n)$, and hence there exists a graph with at least this many vertices which is F-free and has no independent set of size t by construction. □

For example, Kopparty's triangle-free Cayley graph which has $n \approx p^3, d \approx p^2$ and $\lambda \approx p\log p$ roughly implies that

$$R(3, p\log^2 p) = \Omega(p^2\log^2 p),$$

and by changing variables to some $N = p/\log^2 p$ we in total get

$$R(3,N) = \Omega\left(\frac{N^2}{\log^2 N}\right),$$

which very closely matches the general upper bound we proved of $R(3,N) \le \binom{N+1}{2} \approx N^2$.

Possibly insert comments on the true asymptotics of $N^2/\log N$ and/or the limits of how strong of a bound we can construct from this method/current bounds on K_s-free pseudorandom graphs.

14 Flows

Kőnig's Theorem from Section 2.4 is what one might informally refer to as a "max-min theorem", in that it says that a quantity defined in terms of a maximum (i.e. the maximum size of a matching in a bipartite graph) is equal to another quantity defined in terms of a minimum which naturally upper bounds the first quantity (i.e. $\tau(G)$). There are a number of other max-min theorems in graph theory of this form, with most of these results being essentially equivalent to each other. In the next two subsections we discuss two of these: the Max-Flow Min-Cut Theorem related to flows, and Menger's Theorem related to connectivity.

Maybe insert motivation for the problem, eg transporting water or traffic; in fact original application was how much supplies Russian railroads could transport.

14.1 Basics of Flows

In this subsection we shift our attention from graphs towards digraphs. Formally a digraph D is a pair of sets (V, E) where $E \subseteq \{(x, y) : x, y \in V,\ x \neq y\}$. We refer to the elements of V as vertices and the elements of E as directed edges or arcs, and we denote these sets by $V(D)$ and $E(D)$. We will often denote the directed edge (x, y) simply by xy, where here we emphasize that (unlike for undirected graphs) xy does not mean the same thing as yx. For a vertex v we define its in-neighborhood $N^-(v) = \{u : (u, v) \in E(D)\}$ and its out-neighborhood $N^+(v) = \{w : (v, w) \in E(D)\}$.

We will be interested in directed graphs together with some additional information. For this, when considering functions of the form $g : E(D) \to \mathbb{R}_{\geq 0}$, we will usually write $g(u, v)$ as shorthand for $g((u, v))$.

Definition 20. A *network* is a quadruple $N = (D, s, t, c)$ where D is a finite digraph, s, t are distinct vertices of D (sometimes referred to as the *source* and *terminal* vertices), and $c : E(D) \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ is a function called the *capacity function*. A *flow* for a network is any function $f : E(D) \to \mathbb{R}_{\geq 0}$ satisfying the following:

- For every $v \in V(D) \setminus \{s, t\}$ we have
$$\sum_{u \in N^-(v)} f(u, v) = \sum_{w \in N^+(v)} f(v, w),$$
i.e. the total flow into a vertex equals the total flow leaving the vertex; and

- We have $f(u, v) \leq c(u, v)$ for all $uv \in E(D)$.

For a flow f we define
$$\mathrm{val}(f) := \sum_{w \in N^+(s)} f(s, w) - \sum_{u \in N^-(s)} f(u, s),$$
and we let
$$\mathrm{val}(N) := \sup_f \mathrm{val}(f),$$

where the supremum ranges over all flows f.

That is, we define the value of the flow to be the total amount of flow leaving the source. This turns out to be equivalent to defining $\mathrm{val}(f)$ to be the total amount of flow entering the terminal vertex, and in fact the following more general fact is true.

Lemma 14.1. *Given a network N, we say that a pair of sets (S, T) is an st-cut if $T = V(D)\setminus S$ and if $s \in S$ and $t \notin S$. For any such cut and flow f, we have*

$$\mathrm{val}(f) = \sum_{(v,w)\in E(D)\cap S\times T} f(v,w) - \sum_{(w,v)\in E(D)\cap T\times S} f(w,v).$$

Proof. Let f be any flow of N. Then

$$\begin{aligned}\mathrm{val}(f) = \sum_{w\in N^+(s)} f(s,w) - \sum_{u\in N^-(s)} f(u,s) &= \sum_{v\in S}\left(\sum_{w\in N^+(v)} f(v,w) - \sum_{u\in N^-(v)} f(u,v)\right)\\ &= \sum_{v\in S}\left(\sum_{w\in N^+(v)\cap T} f(v,w) - \sum_{u\in N^-(v)\cap T} f(u,v)\right),\end{aligned}$$

where the first equality used that $\sum_{w\in N^+(v)} f(v,w) - \sum_{u\in N^-(v)} f(u,v) = 0$ for each $v \neq s$ by definition of f being a flow, and the second equality used that any $f(v,w)$ with $w \notin T = V(D)\setminus S$ appearing as a positive term in the sum will also appear as a negative term $-f(v,w)$, so only those arcs with exactly one vertex in S and T do not cancel each other out. Unwinding the definitions shows that this is exactly the quantity we wished to show $\mathrm{val}(f)$ equal to. □

This observation in turn gives a natural upper bound on $\mathrm{val}(N)$.

Lemma 14.2. *If N is a network and if (S, T) is an st-cut, then*

$$\mathrm{val}(N) \le c(S,T) := \sum_{(v,w)\in E(D)\cap S\times T} c(v,w).$$

Proof. By Theorem 14.1 we have for any flow f that

$$\mathrm{val}(f) = \sum_{(v,w)\in E(D)\cap S\times T} f(v,w) - \sum_{(w,v)\in E(D)\cap T\times S} f(w,v) \le \sum_{(v,w)\in E(D)\cap S\times T} c(v,w),$$

where the inequality used that flows satisfy $f(u,v) \ge 0$ (allowing us to drop all of the negative terms from the sum) and $f(v,w) \le c(v,w)$. This shows $\mathrm{val}(f) \le c(S,T)$ for all f, and hence that $\mathrm{val}(N) = \sup \mathrm{val}(f) \le c(S,T)$ as well. □

The Max-Cut Min-Flow Theorem, originally proved by Ford and Fulkerson, says that this simple upper bound on $\mathrm{val}(N)$ in terms of $c(S,T)$ is in fact best possible.

Theorem 14.3 (Max-Cut Min-Flow Theorem)**.** *For every network N we have*

$$\mathrm{val}(N) = \min c(S,T),$$

where the minimum is over all st-cuts of N.

We will not give a formal proof of this version of the Max-Cut Min-Flow Theorem. Instead, we will prove a slight variant for networks with integral-valued capacities which will be more relevant for our applications.

Theorem 14.4 (Integral Max-Flow Min-Cut)**.** *If N is a network with $c : E(D) \to \mathbb{Z}_{\geq 0} \cup \{\infty\}$ and $\mathrm{val}(N) < \infty$, then there exists an integral-valued flow $f : E(D) \to \mathbb{Z}_{\geq 0}$ such that*

$$\mathrm{val}(f) = \min c(S,T),$$

where the minimum is over all st-cuts of N.

Note that such a flow must have $\mathrm{val}(f) = \mathrm{val}(N)$ since $\mathrm{val}(f) \leq \mathrm{val}(N)$ by definition of N and $\mathrm{val}(N) \leq \min c(S,T)$ by Theorem 14.2.

Proof. Our construction of f will be algorithmic: we will start with a flow f_0 which is identically 0, and iteratively if $\mathrm{val}(f_i) < \min c(S,T)$ then we will construct an integral-valued flow f_{i+1} with $\mathrm{val}(f_{i+1}) \geq \mathrm{val}(f_i) + 1$. Because $\mathrm{val}(N) < \infty$, this process must end at some integral-valued flow f_n with $\mathrm{val}(f_n) = \min c(S,T)$, giving the result.

It remains then to show that we can find f_{i+1} with $\mathrm{val}(f_{i+1}) \geq \mathrm{val}(f_i) + 1$ whenever $\mathrm{val}(f_i)$ is not too large. The most natural way to do this is by adjusting some of the flows of f_i, noting that we will be only be able to increase the flow of an edge with $f(u,v) < c(u,v)$ and decrease the flow of an edge with $f(u,v) > 0$. Putting this intuition together gives the following idea.

Claim 14.5. *Given a flow f, we say a sequence of distinct vertices $(w_0, \ldots, w_a)$ is an augmented path from w_0 to w_a if for all $0 \leq i < a$ we either have $(w_i, w_{i+1}) \in E(D)$ and $f(w_i, w_{i+1}) < c(w_i, w_{i+1})$, or $(w_{i+1}, w_i) \in E(D)$ and $f(w_{i+1}, w_i) > 0$.*

If f is an integral-valued flow and if there exists an augmented path from s to t then there exists an integral-valued flow f' with $\mathrm{val}(f') \geq \mathrm{val}(f) + 1$.

Sketch of Proof. Define f' by having $f'(w_i, w_{i+1}) = f(w_i, w_{i+1}) + 1$ whenever $(w_i, w_{i+1}) \in E(D)$, by having $f'(w_{i+1}, w_i) = f(w_{i+1}, w_i) - 1$ whenever $(w_{i+1}, w_i) \in E(D)$, and $f'(e) = f(e)$ for all other edges e. Note that f' is non-negative and upper bounded by c by assumption of our augmented path and the fact that c, f are integral valued. Similarly it is not difficult to see that the flow into each $v \neq s,t$ equals the flow out of the vertex since this condition held for f, so f' is a flow. Moreover, f' either increases the flow on an edge out of s if $(w_0, w_1) = (s, w_1) \in E(D)$ or decreases the flow of an edge into s if $(w_1, w_0) = (w_1, s) \in E(D)$, so $\mathrm{val}(f') = \mathrm{val}(f) + 1$, proving the claim. □

If an augmented path as in this claim exists then we can define f_{i+1} as desired, so we can assume that no such path exists. We now aim to construct an st-cut (S,T) with $\mathrm{val}(f_i) = c(S,T)$, which will show that we have in fact already constructed the desired flow f_i. Motivated by our definition of augmented paths, we define $S \subseteq V(D)$ to be the set of vertices v such that there exists an augmented path from s to v, and we let $T = V(D) \setminus S$. Note that by assumption we have $t \notin S$, so (S,T) defines an st-cut. Thus by Theorem 14.1

$$\mathrm{val}(f_i) = \sum_{(v,w) \in E(D) \cap S \times T} f_i(v,w) - \sum_{(w,v) \in E(D) \cap T \times S} f_i(w,v).$$

We claim that $f_i(v,w) = c(v,w)$ for all $(v,w) \in E(D) \cap S \times T$. Indeed, if $f_i(v,w) < c(v,w)$ then there exists an augmented path from s to w, namely by taking the augmented path from s to v (which exists since $v \in S$) and then either appending w to the end if w is not in the path already or shortening the path to end at w if it does appear. This implies $w \in S$, a contradiction to $(v,w) \in S \times T$. The same argument implies that $f_i(w,v) = 0$ for $(w,v) \in E(D) \cap T \times S$. We conclude then that

$$\mathrm{val}(f_i) = \sum_{(v,w) \in E(D) \cap S \times T} c(v,w) = c(S,T).$$

This implies $\mathrm{val}(f_i) \geq \min c(S',T')$ where the minimum runs over all st-cuts, and we trivially have $\mathrm{val}(f_i) \leq \min c(S',T')$ by Theorem 14.2, so f_i gives the desired integral-valued flow.

□

Our proof above gives a fairly efficient algorithm for constructing integral-valued flows with $\mathrm{val}(f) = \mathrm{val}(N)$ whenever the capacity function is integral-valued. The same sort of argument works if c is rational-valued, but there are known examples where this naive algorithm fails if c can take on irrational values. Nevertheless, one can algorithmically prove the Max-Flow Min-Cut Theorem for arbitrary c by using a different algorithm due to Edmonds and Karp. One can also prove this non-algorithmically, as we briefly sketch out below.

Sketch of Proof of Max-Flow Min-Cut Theorem. If $\mathrm{val}(N) = \infty$ then there is nothing to show. Otherwise, one can show using real analysis that there exists a flow f with $\mathrm{val}(f) = \mathrm{val}(N)$. It is not difficult to show that if there existed an augmented path from s to t then one could construct a flow with higher value than f, a contradiction. As such, we can define S,T as we did in our previous proof and conclude that $\mathrm{val}(f) = c(S,T)$, proving the result. □

There are many variants of the Max-Flow Min-Cut Theorem. One such example which is important to us places constraints on the amount of flow each vertex can receive as opposed to constraining how much flow an edge can take on.

Definition 21. A *vertex-network* is a quadruple $N = (D,s,t,c)$ where D is a digraph with finite vertex set, s,t are vertices of D, and $c : V(D) \to \mathbb{R}_{\geq 0} \cup \{\infty\}$. A *flow* for a network is any function $f : E(D) \to \mathbb{R}_{\geq 0}$ satisfying the following:

- For every $v \in V(D) \setminus \{s,t\}$ we have
$$\sum_{u \in N^-(v)} f(u,v) = \sum_{w \in N^+(v)} f(v,w).$$
- We have $\sum_{u \in N^-(V)} f(u,v) \leq c(v)$ for all $v \neq s,t$.

We define $\mathrm{val}(f)$ and $\mathrm{val}(N)$ exactly as before.

Before we gave an upper bound on $\mathrm{val}(N)$ in terms of the sum of capacities of a set of edges, namely a set of edges E such that $D - E$ has no directed path from s to t. Similarly we have the following.

Lemma 14.6. *Given a vertex-network N, we say that a set of vertices $V \subseteq V(D)$ is a vertex-cut if $D - V$ has no directed path from s to t. For such a set, we have*

$$\mathrm{val}(N) \leq c(V) := \sum_{v \in V} c(v).$$

The proof of this is similar to the proof of Theorem 14.2. With this we can state vertex-capacity versions of both versions of the Max-Flow Min-Cut Theorem. In particular, we will need the following.

Theorem 14.7 (Integral-Vertex Max-Flow Min-Cut)**.** *If N is a vertex-network with $c : V(D) \to \mathbb{Z}_{\geq 0} \cup \{\infty\}$, then there exists an integral-valued flow f with*

$$\mathrm{val}(f) = \min c(V),$$

where the minimum ranges over all vertex-cuts V.

Sketch of Proof. Define a digraph D' which has vertex set $\{v_-, v_+ : v \in V(D)\}$ and all arcs of the form $\{(u_-, v_+) : (u, v) \in E(D)\} \cup \{(v_-, v_+) : v \in V(D)\}$. That is, we effectively split each vertex of D in two. We now define a capacity function c' on $E(D)$ by having $c'(v_-, v_+) = c(v)$ and having $c'(u_-, v_+) = \infty$ for all other arcs. One can now apply[15] the integral Max-Flow Min-Cut Theorem to this new network defined by D', c' to get an integral-valued flow f' which one can then lift to a flow f on N which has the desired value (since capacities for edges of D' are the same as capacities for vertices of D). □

14.2 Applications of Flows

To see the power of these result, we will use this to give a quick proof of Hall's Theorem.

Sketch of Proof of Hall's Theorem. Let G be a bipartite graph with bipartition $U \cup V$ and $|U| = |V| = n$ which satisfies Hall's condition. Define a directed graph D by directing every edge of G to go from U to V and then two new vertices s, t to $V(G)$ with all directed edges of the form (s, u) for $u \in U$ and (v, t) for $v \in V$. Let $c : V(D) \to \mathbb{Z}_{\geq 0} \cup \{\infty\}$ be defined by having $v(s) = c(t) = \infty$ and $(w) = 1$ for all other vertices.

We claim that $\min c(W) = n$ where the minimum ranges over all vertex cuts W of D. Indeed, taking $W = U$ shows that this minimum is at most n. For the other direction, let W be an arbitrary cut. Observe that this being a cut means $N(U \setminus W) \subseteq V \cap W$. By Hall's condition we have

$$|W| = |U \cap W| + |V \cap W| \geq |U \cap W| + |N(U \setminus W)| \geq |U \cap W| + |U \setminus W| = n,$$

with this last inequality using Hall's condition.

[15]Technically this doesn't work because as written D' has two "sources" s_-, s_+ and similarly with t_-, t_+. One can easily show that the Max-Flow Min-Cut Theorem continues to hold with multiple sources and sinks, so this is fine

Now by Integral-Vertex Max-Flow Min-Cut there exists some integral flow f with respect to this network which has value n. It is not difficult to see that this implies that there exists n vertex disjoint edges of G which receive a flow of 1 from f, showing that G has a perfect matching. □

Here's a flow proof of a version of Menger's Theorem.

Theorem 14.8 (Menger's Theorem)**.** *Given a graph G and distinct vertices s, t, let* $\mathrm{path}(s,t)$ *denote the largest size of a set of paths $\mathcal{P}$ from s to t such that no two paths have any vertices other than s, t in common, and let* $\mathrm{cut}(s,t)$ *denote the smallest size of a set of vertices V such that $G - V$ has no path from s to t. Then* $\mathrm{path}(s,t) = \mathrm{cut}(s,t)$*.*

Sketch of Proof. That $\mathrm{path}(s,t) \leq \mathrm{cut}(s,t)$ follows from the fact that if V is a cut then each path of $\mathcal{P}$ must use a vertex from V and no two paths can use the same vertex in this set. To show the other direction, define a directed graph D which has vertex set $V(G)$ and all directed edges of the form $\{(u,v),(v,u) : uv \in E(G)\}$; that is, we replace each edge with two directed edges going in either direction. Define $c : V(G) \to \mathbb{Z}_{\geq 0} \cup \{\infty\}$ by having $c(s) = c(t) = \infty$ and $c(v) = 1$ otherwise. It is not difficult to see that a smallest vertex-cut of this network has size $\mathrm{cut}(s,t)$, so by Integral-Vertex Max-Flow Min-Cut there exists an integral-valued flow f with value equal to $\mathrm{cut}(s,t)$. It is not difficult to see that the edges with positive flow (i.e. with flow equal to 1) define paths which are pairwise disjoint on $V(G) \setminus \{s,t\}$ (since each such vertex can receive at most one edge with flow equal to 1), giving the desired result. □

14.3 Exercises

1. Our version of Menger's Theorem involves paths which are internally-vertex disjoint, but it is equally natural to consider an edge disjoint version.

   (a) For a network N, we say that a set of directed edges $E \subseteq E(D)$ is an *st*-edge cut if $D - E$ has directed path from s to t, i.e. if there exists no sequence of vertices $(w_1, w_2, \ldots, w_a)$ with $w_1 = s,\ w_a = t$, and $w_i w_{i+1} \in E(D) \setminus E$ for all $1 \leq i < a$. For such a set of directed edges we define $c(E) = \sum_{e \in E} c(e)$. Prove that

   $$\min c(E) = \min c(S,T),$$

   where the first minimum ranges over all *st*-edge cuts and the second minimum ranges over all *st*-cuts [2-].

   (b) Given a graph G and distinct vertices s, t, let $\mathrm{path}'(s,t)$ denote the largest size of a set of paths $\mathcal{P}$ from s to t such that no two paths have any edges in common, and let $\mathrm{cut}'(s,t)$ denote the smallest size of a set of edges E such that $G - E$ has no path from s to t. Prove that $\mathrm{path}'(s,t) = \mathrm{cut}'(s,t)$ [2].

2. Insert some small finite digraph and ask to find max flow.

15 Hypergraphs

TODO . Likely topics: generalized KST, Fisher's inequality, Kruskal-Katona, sunflowers, traces and expansions

16 Scattered Gems

Here I collect some small sporadic results which are quite nice but which don't necessarily have a clean connection to any of the other chapters.

16.1 Maximal Independent Sets

Given a graph G, a set of vertices $I \subseteq V(G)$ is said to be a *maximal independent set* (or MIS for short) if it is an independent set which is maximal with respect to set inclusion, meaning that $I \cup \{v\}$ is not an independent set for every $v \notin I$. Moon and Moser independently considered the problem of maximizing the number of MIS's in a graph with a given number of vertices. To this end, let $\mathrm{mis}(G)$ denote the number of MIS's in a graph G.

Theorem 16.1 (Moon-Moser)**.** *If G is an n-vertex graph, then*

$$\mathrm{mis}(G) \le m(n) := \begin{cases} 3^{n/3} & n \equiv 0 \mod 3, \\ 4 \cdot 3^{(n-4)/3} & n \equiv 1 \mod 3, \\ 2 \cdot 3^{(n-2)/3} & n \equiv 2 \mod 3. \end{cases}$$

Moreover, this bound is best possible for all $n \ge 3$.

To give some intuition for these bounds, we observe that one simple construction for a graph with many MIS's is to take the disjoint union of n/r copies of K_r whenever $r|n$. Indeed, such a graph can easily be seen to have $\mathrm{mis}(G) = r^{n/r}$. Now amongst *real numbers* this quantity is maiximized when $r = e$, but since our r is an integer one can check that the best one can do is $r = 3$. In particular, when $n \equiv 0 \mod 3$ then a disjoint union of K_3's gives an n-vertex graph with $3^{n/3}$ MIS's, and the Moon-Moser Theorem is best possible. For other values of $n \mod 3$ one can take a disjoint union of K_3's together with either 1 or 2 disjoint copies of K_2 to show that the bound of $m(n)$ is best possible.

Our proof of the Moon-Moser Theorem will be based on a very nice observation due to Wood, where here given a graph G and a vertex v we define the closed-neighborhood $N[v] := N(v) \cup \{v\}$.

Lemma 16.2 (Wood's Lemma)**.** *For any graph G and $v \in V(G)$ we have*

$$\mathrm{mis}(G) \le \sum_{u \in N[v]} \mathrm{mis}(G - N[u]).$$

Proof. Let I be an MIS of G. Note that we must have $I \cap N[v] \neq \emptyset$, otherwise $I \cup \{v\}$ would be a larger independent set. Moreover, if $u \in I$ then $I - u$ must be an MIS of $G - N[u]$ as any $w \in V(G) \setminus N[u]$ which would make $(I - u) \cup \{w\}$ independent in $G - N[u]$ would also imply that $I \cup \{w\}$ is independent in I. In total then every MIS of G can be written as the disjoint union of some $u \in N[v]$ together with an MIS from $G - N[u]$, giving the desired bound. □

Proof of Moon-Moser. We prove the result by induction on n, the base case $n = 3$ being straightforward. Let G be an n-vertex graph. Let $\delta = \delta(G)$ and let v be a vertex of degree δ.

By Wood's Lemma and induction we have

$$\mathrm{mis}(G) \le \sum_{u\in N[v]} \mathrm{mis}(G - N[u]) \le (\delta+1)m(n-\delta-1),$$

with this last step using that $|N[v]| = \delta+1$ by our choice of v and that $|N[u]| = \deg(u)+1 \ge \delta+1$ for every u by definition of $\delta = \delta(G)$.

From here, a little bit of case analysis based on δ and $n \mod 3$ can be used to show $(\delta + 1)m(n-\delta-1) \le m(n)$. For example, by observing that $4\cdot 3^{(n-4)/3} \le m(n) \le 3^{n/3}$ for all n we find for $\delta \ge 3$ that

$$\mathrm{mis}(G) \le (\delta+1)m(n-\delta-1) \le (\delta+1)3^{(n-\delta-1)/3} \le 4\cdot 3^{(n-4)/3} \le m(n).$$

The remaining cases of $\delta \in \{1,2\}$ for each value of $n \mod 3$ can be handled with similar calculations, proving the result. □

There are a number of results bounding $\mathrm{mis}(G)$ under additional assumptions, and in particular under assumptions which force G to be "far" from the extremal construction of disjoint union of triangles. One natural direction for this is to consider the problem for triangle-free graphs..

Theorem 16.3 (Hujter-Tuza)**.** *If G is an n-vertex graph, then*

$$\mathrm{mis}(G) \le m_t(n) := \begin{cases} 2^{n/2} & n \equiv 0 \mod 2, \\ 5\cdot 2^{(n-5)/2} & n \equiv 1 \mod 2. \end{cases}$$

Moreover, this bound is best possible for all $n \ge 4$.

The construction showing this is best possible is to take a disjoint union of K_2's together with one C_5 if n is odd. We do not know a direct proof of Hujter-Tuza using Wood's Lemma (since the inequality $(\delta+1)m_t(n-\delta-1) \le m_t(n)$ does not hold for $\delta = 2$ and n odd), though we would not be surprised if a slightly more careful analysis gave this result. There are a number of nice applications of Hujter-Tuza and its variants, such as counting the number of maximal triangle-free graphs through the method of hypergraph containers.

Mention your own open problems in this area.

Part IV

Advanced Methods

Note: a more in-depth look at all of these topics can be found in my notes here.

17 Hypergraph Containers

Given a hypergraph H, we say a set of vertices I is an independent set if no edge of H is contained in I. We let $\alpha(H)$ denote the maximum size of an independent set of H, and we let $\mathcal{I}(H)$ denote the set of independent sets of H.

Many problems in extremal combinatorics can be stated in terms of independent sets of hypergraphs. For example, one can define $\mathcal{H}_n^{AP}$ to be the 3-uniform hypergraph on $[n]$ where every triple $S \subseteq [n]$ is a hyperedge if and only if S is a 3-term arithmetic progression. Thus Roth's theorem is equivalent to saying $\alpha(\mathcal{H}_n^{AP}) = o(n)$. Similarly one can define $\mathcal{H}_n^{\Delta}$ to be the 3-uniform hypergraph whose vertex set is $E(K_n)$ and whose hyperedges are triples of edges in K_n which form a triangle. Independent sets of $\mathcal{H}_n^{\Delta}$ are triangle-free subgraphs of K_n, so Mantel's theorem says $\alpha(H_n^{\Delta}) = \lfloor n^2/4 \rfloor$.

In this chapter we explore a powerful technique for upper bounding the total number of independent sets of a given hypergraph. The proof technique is based on a relatively simple idea: say for a given hypergraph H that we have a collection of sets of vertices $\mathcal{C}$ (called a set of *containers*) with the property that every independent set I of H is a subset of some $C \in \mathcal{C}$. Then the number of independent sets of H is trivially at most

$$\sum_{C \in \mathcal{C}} 2^{|C|} \le |\mathcal{C}| 2^{\max_{C \in \mathcal{C}} |C|}.$$

Observe that this gives effective upper bounds whenever we can find a collection of containers such that both the whole collection is small, and such that each $C \in \mathcal{C}$ is small (e.g. not much larger than the size of a largest independent set in H). Although this may seem like a complicated thing to do, the general machinery of hypergraph containers gives a routine way to find such collections provided one can prove an appropriate supersaturation result.

17.1 Graph Containers

While the general method of containers involves bounding independent sets of *hypergraphs*, one can get reasonably far by only considering independent sets of *graphs*. To this end we prove the following graph container lemma, which will be the main workhorse for the rest of this section. Recall that a collection $\mathcal{C}$ of subsets $C \subseteq V(G)$ is a set of *containers* for G if every independent set $I \in \mathcal{I}(G)$ is a subset of some $C \in \mathcal{C}$, and here we use the notation $\binom{n}{\le k} = \sum_{i \le k} \binom{n}{i}$.

Lemma 17.1. *Let G be an n-vertex graph and $t > 0$ a positive number. There exists a collection $\mathcal{C}$ of containers such that*

(a) $|\mathcal{C}| \le \binom{n}{\le n/t}$.

(b) $\Delta(G[C]) < t - 1$ *for all* $C \in \mathcal{C}$.

In other words, there exists a small set of containers $\mathcal{C}$ such that each $C \in \mathcal{C}$ is "small" in the sense that it induces a graph with small maximum degree.

Proof. Our proof will be algorithmic: we construct a (deterministic) algorithm which takes as input a set $I \subseteq V(G)$ and which outputs a pair $(S(I), A(I))$ such that $S(I) \subseteq I \subseteq S(I) \cup A(I)$, and we will ultimately use $\{S(I) \cup A(I) : I \in \mathcal{I}(G)\}$ as our set of containers. We now describe the algorithm.

Fix an arbitrary ordering of $V(G)$. As input we take in an independent set $I \subseteq V(G)$. We initially set $S = \emptyset$ and $A = V(G)$ (the former corresponds to a set of "selected" vertices which are in I, and the latter to the set of "available" vertices which could possibly be in I given the current stage of the algorithm). The algorithm proceeds as follows:

Step 1 If $\Delta(G[A]) < t - 1$, set $(S(I), A(I)) = (S, A)$ and end the algorithm. Otherwise proceed to Step 2.

Step 2 Let v be the vertex of maximum degree in $G[A]$, with ties being broken according to the fixed ordering of $V(G)$. If $v \notin I$, then set $A = A - v$ and repeat Step 1. Otherwise proceed to Step 3.

Step 3 Set $A = A - v - N_{G[A]}(v)$, $S = S \cup \{v\}$. Proceed to Step 1.

Let's reiterate what's going on here. It's not difficult to show inductively that we always have $I \subseteq S \cup A$, so $S \cup A$ serves as a container set for I, and we would like to trim this set as much as possible. We do this by selecting a vertex $v \in A \cap I$ and adding it to S. If v has large degree in $G[A]$, then v being in the independent set I means that its many neighbors are not, so we get to remove all of these vertices from A while maintaining $I \subseteq S \cup A$. In particular, since we keep going so long as $G[A]$ has large maximum degree, we know at each step of this process that we're removing many vertices from A.

With the algorithm above in mind, we define

$$\mathcal{C} = \{S(I) \cup A(I) : I \in \mathcal{I}(G)\},$$

which is a set of containers since $I \subseteq S(I) \cup A(I)$ at every step of the algorithm. Since we terminate the algorithm precisely when $\Delta(G[A(I)]) = \Delta(G[S(I) \cup A(I)]) < t - 1$ (the equality holds since $S(I)$ has no neighbors in $S(I) \cup A(I)$), (b) holds. It thus remains to verify (a). To do this, we note the following.

Claim 17.2. *Let I_1, I_2 be two independent sets and let $(S_1, A_1), (S_2, A_2)$ be their outputs from the algorithm. If $S_1 = S_2$, then $A_1 = A_2$.*

Proof. Consider the last point in the algorithm where the (S, A) sets are the same for both I_1, I_2 (noting that such a point in the algorithm must exist since we start with $(\emptyset, V(G))$). If Step 1 holds for this (S, A) pair then we would end with $(S_1, A_1) = (S_2, A_2) = (S, A)$ so we can assume this is not the case. Letting v be the vertex as in Step 2, if $v \notin I_1, I_2$ or $v \in I_1, I_2$ then the next (S, A) pair in the algorithm will be the same for both sets, so we can assume that, say, $v \notin I_1$ and $v \in I_2$. This implies that $v \notin S_1 \subseteq I_1$ and $v \in S_2$, a contradiction to $S_1 = S_2$. □

This claim implies that given $S(I)$, the container $S(I) \cup A(I)$ is uniquely determined[16]. In

[16] Because of this, S is often called a "certificate" or "fingerprint" of I.

particular, if we always have $|S(I)| \leq n/t$, then the number of containers will be at most $\binom{n}{\leq n/t}$. And indeed, each round of the algorithm has $\Delta(G[A]) \geq t-1$, so every time a vertex is added to S at least $1 + (t-1) = t$ vertices are removed from A. In particular, at most n/t vertices can be added to S, giving the result.

It might be clearer to write all these pieces as claims, eg that $S \cup A$ contains I at each step, that S determines A, that S is small, and that we have small maximum degree. □

Actually, a closer inspection of the proof gives the following.

Lemma 17.3. *Let G be a graph on n vertices and $t \in \mathbb{R}$. There is a collection $\mathcal{C}$ of containers and functions*

$$f : \mathcal{I}(G) \to \binom{V(G)}{\leq n/t}, \qquad g : \binom{V(G)}{\leq n/t} \to \mathcal{C}$$

such that the following hold.

(a) The function g is a surjection. In particular, $|\mathcal{C}| \leq \binom{n}{\leq n/t}$.

(b) We have $\Delta(G[C]) < t-1$ for all $C \in \mathcal{C}$.

(c) For every $I \in \mathcal{I}(G)$ we have

$$f(I) \subseteq I \subseteq g(f(I)).$$

Proof. Consider the exact same algorithm as before. Define $f(I) = S(I)$ and $g(S) = C(S)$ (if $S \neq S(I)$ for any I, then assign g arbitrarily). It's not hard to check that this works. □

The extra source of power of this lemma is that for each $I \in \mathcal{I}$ we are given some set $S = f(I)$ contained in I. In many examples this extra information is needed to get tight upper bounds when counting independent sets, though for pedagogical purposes we will often work with the simpler Lemma 17.1 to get close to tight results.

Our first application of Lemma 17.1 will be to count the number of independent sets in d-regular graphs. As a point of reference, it is not difficult to show that if G consists of $n/2d$ disjoint copies of $K_{d,d}$, then

$$|\mathcal{I}(G)| = (2^{d+1} - 1)^{n/2d} = 2^{n/2 + n/2d + o(n)}.$$

Thus for d-regular graphs, we can't possibly hope to prove an upper bound on $|\mathcal{I}(G)|$ stronger than roughly $2^{n/2}$ when d is large. We can prove that this is close to best possible using containers.

Theorem 17.4. *Let G be a d-regular n-vertex graph with $\log n \ll d \ll n/2$. Then*

$$|\mathcal{I}(G)| \leq 2^{n/2 + o(n)}.$$

In fact, it turns out that $|\mathcal{I}(G)| \leq (2^{d+1} - 1)^{n/2d}$ for all d-regular n-vertex graphs. This was proven for bipartite graphs by Kahn [?] using entropy, and the problem was solved in full by Zhao [?]. As far as I'm aware, the proof of Theorem 17.4 presented here is due to Balogh.

As a first step to proving Theorem 17.4, we will apply Lemma 17.1 to our graph G to get a collection of containers $\mathcal{C}$. We would like to conclude the result by the observation from (**??**):

$$|\mathcal{I}(G)| \leq \sum_{C \in \mathcal{C}} 2^{|C|} \leq |\mathcal{C}| 2^{\max_{C \in \mathcal{C}} |C|},$$

but there's an issue with this. Namely, Lemma 17.1 only tells us that each $C \in \mathcal{C}$ induces a graph in G with small maximum degree. For a general graph this tells us nothing about $|C|$, but fortunately in d-regular graphs, $G[C]$ having small maximum degree is only possible if C is small. The following states a precise version of the contrapositive of the previous sentence.

Lemma 17.5. *For any $\varepsilon > 0$, if G is a d-regular graph and $C \subseteq V(G)$ with $|C| = n/2 + \varepsilon n$, then $\Delta(G[C]) \geq 2\varepsilon d$.*

This lemma is a form of supersaturation: a d-regular graph can have a subset of size $n/2$ with $G[C]$ empty (e.g. if G is bipartite), but if C is just a bit larger than this, then it must have relatively high maximum degree. As we will see, supersaturation results are almost always a necessary ingredient for applying the method of containers.

Proof. Because the maximum degree is always at least the average degree, we have

$$\Delta(G[C]) \geq 2e(G[C])/|C| \geq 2e(G[C])/n$$

, so it will suffice to show that $e(G[C])$ is large. To do this, we let $\overline{C} = V(G) \setminus C$ and note that

$$d|C| = \sum_{v \in C} d(v) = 2e(G[C]) + e(C, \overline{C}) \leq 2e(G[C]) + d|\overline{C}|.$$

Because $|C| = n/2 + \varepsilon n$ and $|\overline{C}| = n/2 - \varepsilon n$, in total this implies

$$2e(G[C]) \geq 2\varepsilon dn.$$

Combining this with the observation at the start gives the result. □

Corollary 17.6. *For all t, if G is an n-vertex d-regular graph, then there exists a set of containers $\mathcal{C}$ with $|\mathcal{C}| \leq \binom{n}{\leq n/t}$ and $|C| \leq \frac{1}{2}n + \frac{t}{d}n$ for all $C \in \mathcal{C}$.*

Proof. Let $\mathcal{C}$ be a set of containers as guaranteed by Lemma 17.1. Because $\Delta(G[C]) < t-1 \leq t$, Lemma 17.5 implies that $|C| \leq \frac{1}{2}n + \frac{t}{d}n$. □

With this we can prove Theorem 17.4.

Proof of Theorem 17.4. At this point all we need to do is use (**??**) after applying Corollary 17.6 with a carefully chosen value of t. Note that

$$|\mathcal{C}| \approx \binom{n}{n/t} \approx 2^{n \log(t)/t},$$

and we already know $2^{\max |C|} \approx 2^{\frac{1}{2}n + \frac{t}{d}n}$. Thus to minimize $|\mathcal{C}| \cdot 2^{\max |C|}$ we should choose t so that $\frac{t}{d} \approx \log(t)/t$, and in particular $t = \sqrt{d \log n}$ is a reasonable choice. One can verify with a more formal argument that this does indeed give the desired result after applying (**??**). □

We note that the statement of Corollary 17.6 and the optimization of t in the proof of Theorem 17.4 is in some sense independent[17] of the problem of determining $|\mathcal{I}(G)|$ for G a d-regular graph. That is, these results are effective for other problems which involve counting independent sets of d-regular graphs.

For example, recall that a q-coloring of a graph G is a map $\chi : G \to [q]$ such that $\chi(u) \neq \chi(v)$ whenever $uv \in E(G)$. Equivalently, a q-coloring is a partition of $V(G)$ into independent sets $I_1, \ldots, I_q$, With this latter formulation, we can use containers to get an effective bound on the number of q-colorings of G, which we'll denote by $X_q(G)$.

Again, let's consider a test case to figure out how strong of a bound we could possibly hope to prove. Let G be $n/2d$ disjoint copies of $K_{d,d}$. We know that G has close to as many independent sets as it could possible have, so it seems plausible that it would have many q-colorings as well. In particular, one can prove that $X_q(G) \approx (q/2)^n$, and once again we can prove that this is essentially best possible.

Theorem 17.7 ([**?**]). *Let G be an n-vertex d-regular graph and q an integer such that $q^2 \log n \ll d$. Then*

$$X_q(G) \leq (q/2 + o(1))^n.$$

We note that a stronger result was proven by Galvin [**?**] with a somewhat more involved proof.

Proof. By the same reasoning as in Theorem 17.4, there exists a set of containers $\mathcal{C}$ for G such that $|\mathcal{C}| \approx 2^{\sqrt{\frac{\log n}{d}}n}$ and $|C| \approx \frac{1}{2}n$ for each $C \in \mathcal{C}$. Consider all vectors of the form $(C_1, \ldots, C_q)$ with $C_i \in \mathcal{C}$, noting that the number of such vectors is at most $|\mathcal{C}|^q = 2^{o(n)}$.

Observe that every q-coloring can be identified by a vector $(I_1, \ldots, I_q)$ where each I_j is an independent set and $\bigcup I_j = V(G)$. Each of these vectors is "contained" in some "container vector" $(C_1, \ldots, C_q)$ with $C_j \in \mathcal{C}$ in the sense that $I_j \subseteq C_j$ for all j. Thus it's enough to count how many q-colorings each container vector contains.

A naive upper bound for the number of q-colorings contained in $(C_1, \ldots, C_q)$ is roughly $2^{qn/2}$ since this is the number of ways to choose an independent set from each C_i. This bound is too weak, so we have to utilize the extra information that the I_j partition $V(G)$.

To this end, assume $V(G) = \{v_1, \ldots, v_n\}$. Given $(C_1, \ldots, C_q)$, let a_i be the number of containers C_j with $v_i \in C_j$. It's not difficult to see that the number of q-colorings contained in this vector is then at most $\prod a_i$, and by the AMGM inequality this is at most $(\sum a_i/n)^n = (\sum |C_j|/n)^n$. Each of the q containers has size at most roughly $n/2 + \sqrt{\frac{\log n}{d}}n$, so this gives the desired result. $\square$

17.2 Triangle-Free Graphs

While much can be done using graph containers, the real power in the method of hypergraph containers comes from the setting of hypergraphs. For simplicity we will only state without proof one such container lemma in the setting of 3-uniform hypergraphs.

[17] Ha.

Theorem 17.8 ([?]). *For every $c \geq 1$, there exists $\delta > 0$ such that the following holds. Let $q \in (0,1)$ and suppose H is a 3-uniform hypergraph such that*

$$\Delta_1(H) \leq c\frac{e(H)}{v(H)},$$

$$\Delta_2(H) \leq cq\frac{e(H)}{v(H)},$$

$$\Delta_3(H) \leq cq^2\frac{e(H)}{v(H)}.$$

Then there exists a collection of sets $\mathcal{C}$ such that every independent set of H is a subset of some $C \in \mathcal{C}$, and moreover, $|C| \leq (1-\delta)v(H)$ for all $C \in \mathcal{C}$ and $|\mathcal{C}| \leq \binom{v(H)}{\leq 2q \cdot v(H)}$.

Note that for simple hypergraphs (i.e. those without repeated edges) we have $\Delta_3(H) = 1$, so this last bound is equivalent to lower bounding the average degree of H by $c^{-1}q^{-2}$. One important consequence of Theorem 17.8 is the following.

Theorem 17.9. *For all $n, \varepsilon > 0$, there exists a collection of n-vertex graphs $\mathcal{C}$ such that*

(a) Every triangle-free graph $G \subseteq K_n$ is a subgraph of some $C \in \mathcal{C}$,

(b) Every $C \in \mathcal{C}$ has less than εn^3 triangles, and

(c) We have $|\mathcal{C}| = n^{O_\varepsilon(n^{3/2})}$.

That is, there exists a small set of nearly triangle-free graphs which contains every triangle-free graph.

Proof. Start with $\mathcal{C} = \{K_n\}$, and note that $\mathcal{C}$ trivially satisfies (a). Iteratively proceed as follows. If every $C \in \mathcal{C}$ has less than εn^3 triangles then output the current collection $\mathcal{C}$. Otherwise, let $C \in \mathcal{C}$ be such that it contains at least εn^3 triangles. Form a 3-graph H with vertex set $E(C)$ where three edges of C form a hyperedge in H if they form a triangle. Note that $e(H) \geq \varepsilon n^3$ and $v(H) = e(C) \leq n^2$. Every edge is contained in at most n triangles, so $\Delta_1(H) \leq n \leq \varepsilon^{-1}\frac{e(H)}{v(H)}$. We also have $\Delta_2(H) = \Delta_3(H) = 1 \leq \varepsilon^{-1}(n^{-1/2})^2\frac{e(H)}{v(H)}$. With this we see that we can apply Theorem 17.8 with $q = n^{-1/2}$ and $c = \varepsilon^{-1}$. This gives a collection of containers $\mathcal{C}'$ for C, i.e. subgraphs $C' \subseteq C$ such that every triangle-free subgraph of C is contained in some $C' \in \mathcal{C}$. Remove C from $\mathcal{C}$ and add every $C' \in \mathcal{C}'$ to $\mathcal{C}$. Repeat this process.

Let $\mathcal{C}$ be the final collection that this algorithm produces. It is straightforward to show that (a) holds inductively, and (b) holds by construction. To show that the final collection is small, first note that each time we apply the container lemma, the number of new graphs we create is at most $\binom{v(H)}{\leq 2n^{-1/2}v(H)} = n^{O(n^{3/2})}$. Second, observe that each time we apply the container lemma to C, the graphs in $\mathcal{C}'$ have at most $(1-\delta)e(C)$ edges, where δ depends only on ε. Because we only iterate on C which have at least εn^2 edges (since they need at least εn^3 triangles), we iteratively apply the lemma at most some bounded number of times $b = b(\varepsilon)$ to reach any element in the final collection $\mathcal{C}$. Thus the total number of containers we create is $\left(n^{O(n^{3/2})}\right)^b = n^{O_\varepsilon(n^{3/2})}$ as desired. □

For Theorem 17.9 to be useful, we need to get a handle on graphs with at most εn^3 triangles. As is typical with containers, this will come from a supersaturation lemma. In particular, we will rely on a special case of result we proved in supersaturation section.

Lemma 17.10. *For every $\delta > 0$ there exists an $\varepsilon > 0$ such that if G is an n-vertex graph with $e(G) \ge (\frac{1}{2} + \delta)\binom{n}{2}$, then G contains at least εn^3 triangles.*

With this we can prove the following counting result.

Theorem 17.11. *The number of n-vertex triangle-free graphs is equal to*

$$2^{(1+o(1))n^2/4}.$$

Proof. The lower bound comes from considering all of the subgraphs of $K_{n/2,n/2}$. For the upper bound, fix some $\delta > 0$ and let ε be as in Theorem 17.10. Let $\mathcal{C}$ be the containers guaranteed by Theorem 17.9 with parameter ε. Because every triangle-free graph is a subgraph of some $C \in \mathcal{C}$, the number of triangle-free graphs is at most

$$\sum_{C\in\mathcal{C}} 2^{|C|} \le n^{O(n^{3/2})} \cdot 2^{\max_{C\in\mathcal{C}} e(C)},$$

Since each $C \in \mathcal{C}$ has less than εn^3 and at triangles, Theorem 17.10 implies $e(C) \le (\frac{1}{2} + \delta)\binom{n}{2}$ for all $C \in \mathcal{C}$. In total we get an upper bound of

$$2^{(\frac{1}{2}+\delta)\binom{n}{2}+O(n^{3/2}\log n)},$$

and letting δ tend towards 0 gives the result. □

While containers most directly allow one to solve problems that are equivalent to counting the number of independent sets of a hypergraph, there are other related problems which they're effective for. For example, we can use containers to solve probabilistic analogs of classical extremal results. To this end, given two graphs G, F, we let $\mathrm{ex}(G, F)$ denote the largest F-free subgraph of G. For example, $\mathrm{ex}(K_n, F) = \mathrm{ex}(n, F)$. We also say that a sequence of events A_n in a probability space holds with high probability (or whp for short) if $\Pr[A_n] \to 1$. With this in mind, the following result then can be viewed as a random version of Mantel's theorem.

Theorem 17.12. *Define $\mathrm{ex}(G_{n,p}, K_3)$ to be the largest triangle-free subgraph of $G_{n,p}$. We have $\mathrm{ex}(G_{n,p}, K_3) = (1 + o(1))pn^2/4$ whp provided $p \gg n^{-1/2}\log n$.*

Proof. The lower bound follows by considering $G_{n,p} \cap K_{n/2,n/2}$, which is always triangle-free and which has $(1 + o(1))pn^2/4$ edges whp. For the upper bound, fix $\delta > 0$, and let $\varepsilon > 0$ be as in Theorem 17.10. Let $\mathcal{C}$ be the set of containers given by Theorem 17.9 with parameter ε, and as before we have $e(C) \le (1/2+\delta)\binom{n}{2}$ for all $C \in \mathcal{C}$. Because every triangle-free graph is contained in some $C \in \mathcal{C}$, in order to have $\mathrm{ex}(G_{n,p}, K_3) \ge (1 + 4\delta)pn^2/4$, there must exist some $C \in \mathcal{C}$

such that $|G_{n,p} \cap C| \geq (1+4\delta)pn^2/4$. Let E_C be the event that this bound holds. Observe that $|G_{n,p} \cap C|$ is a binomial random variable with probability p and at most $(1+2\delta)n^2/4$ trials. By the Chernoff bound, we find $\Pr[E_C] \leq e^{-O_\delta(pn^2)}$. In total then, we have

$$\Pr[\mathrm{ex}(G_{n,p}, K_3) \geq (1+4\delta)pn^2/4] \leq \Pr\left[\bigcup_{C \in \mathcal{C}} E_C\right] \leq n^{O_\delta(n^{3/2})} \cdot e^{-O_\delta(pn^2)} \to 0,$$

with this last step holding by hypothesis on p. We conclude the reuslt by taking δ arbitrarily close to 0. $\square$

We note that for $p \ll n^{-1/2}$, a simple deletion argument shows that for $p \ll n^{-1/2}$ there exist triangle-free subgraphs with $(1+o(1))p\binom{n}{2}$ edges, and this is certainly best possible since $G_{n,p}$ has at most this many edges asymptotically. Thus the bound for p in Theorem 17.12 is almost optimal. In fact, we can obtain the optimal bound in this theorem by using a stronger container theorem analogous to lemma in container section.

Theorem 17.13. *For every $c \geq 1$, there exists $\delta > 0$ such that the following holds. Let $q \in (0,1)$ and suppose H is a 3-uniform hypergraph such that*

$$\Delta_1(H) \leq c\frac{e(H)}{v(H)},$$

$$\Delta_2(H) \leq cq\frac{e(H)}{v(H)},$$

$$\Delta_3(H) \leq cq^2\frac{e(H)}{v(H)}.$$

Then there exists $\mathcal{S} \subseteq \binom{V(H)}{\leq 2q \cdot v(H)}$ and functions $f : \mathcal{S} \to \binom{V(H)}{\leq (1-\delta)v(H)}$ and $g : \mathcal{I}(H) \to \mathcal{S}$ such that for every $I \in \mathcal{I}(H)$ we have

$$g(I) \subseteq I \subseteq f(g(I)) \cup g(I).$$

Moreover, $S \cap f(S) = \emptyset$ for all $S \in \mathcal{S}$, and if $I, I' \in \mathcal{I}(H)$ satisfy $g(I) \subseteq I'$, $g(I') \subseteq I$, then $g(I) = g(I')$.

This allows us to construct the following "strong" set of containers for triangle-free graphs.

Theorem 17.14. *Let $\mathcal{G}_n, \mathcal{T}_n$ denote the set of all n-vertex graphs and all n-vertex triangle-free graphs, respectively. For all $n, \varepsilon > 0$, there exists a set of graphs $\mathcal{S}$ with at most $O_\varepsilon(n^{3/2})$ edges, as well as functions $f : \mathcal{S} \to \mathcal{G}_n$ and $g : \mathcal{T}_n \to \mathcal{S}$ such that for every $G \in \mathcal{T}_n$, we have*

$$g(G) \subseteq G \subseteq f(g(G)) \cup g(G),$$

and such that $f(S)$ has less than εn^3 triangles for all $S \in \mathcal{S}$.

Proof. We start with $\mathcal{S}$ consisting only of the empty graph and define $g(G) = \emptyset$ and $f(\emptyset) = K_n$. Iteratively assume we have constructed some $\mathcal{S}, f, g$ satisfying all of the conditions except possibly that each $S \in \mathcal{S}$ has at most $O_\varepsilon(n^{3/2})$ edges and that $f(S)$ has less than εn^3 triangles (which holds for our initial step). If $f(S)$ has less than εn^3 triangles for all $S \in \mathcal{S}$ then we end the procedure. Otherwise, let S be such that $C = f(S)$ has at least εn^3 triangles. By repeating our computations from the proof of Theorem 17.9, we see that we can apply Theorem 17.13 to the 3-graph H encoding triangles of C, and we let $\mathcal{S}_C, f_C, g_C$ be the output of this theorem.

Claim 17.15. *Let $\mathcal{S}' := (\mathcal{S} \setminus \{S\}) \cup \{S_C \cup S : S_C \in \mathcal{S}_C\}$, define $g'(G) = g(G)$ if $g(G) \neq S$ and $g'(G) = g_C(G - S)$ otherwise, and define $f'(S') = f(S')$ if $S' \in \mathcal{S} \setminus \{S\}$ and $f'(S') = f_C(S' - S)$ otherwise. These maps are well defined and satisfy the conditions of the theorem except possibly that each $S \in \mathcal{S}'$ has at most $O_\varepsilon(n^{3/2})$ edges and that $f'(S)$ has less than εn^3 triangles.*

Proof. First observe that because $C \cap S = \emptyset$, each element of $\mathcal{S}_C$ (which is a subgraph of C) is disjoint from S. This implies that all of the elements $S_C \cup S$ for $S_C \in \mathcal{S}_C$ are distinct. Moreover, none of these elements are equal to any element of $\mathcal{S} \setminus \{S\}$. Indeed, if $S_C \cup S = S' \in \mathcal{S}$, then $\mathcal{S}$ would contain two elements with $S \subsetneq S'$. The last condition of Theorem 17.13 then implies that we must have $S = S'$. This all implies that g', f' are well defined maps, and it is not difficult to check that they inherit all of the other desired properties. □

With this we can keep applying Theorem 17.13 until we get $\mathcal{S}, f, g$ which satisfies all of the conditions except possibly that $e(S)$ is small. As in the proof of Theorem 17.9, one can check that each $S \in \mathcal{S}$ is obtained by applying Theorem 17.13 at most $O_\varepsilon(1)$ times, and each time its applied at most $O(n^{3/2})$ edges get added to S. With this we can conclude the result. □

We note that there exists a somewhat stronger version of Theorem 17.13 (and more generally **??**) which allows one to prove the previous result with less work. However, the theorem statement is somewhat more complicated conceptually (involving things called $(\mathcal{F}, \varepsilon)$-dense families), so for this exposition we have opted to use the simpler version. In any case, with this enhanced version of Theorem 17.9, we can improve upon our threshold for the random Mantel theorem by dropping a logarithmic term.

Theorem 17.16. *Define $\mathrm{ex}(G_{n,p}, K_3)$ to be the largest triangle-free subgraph of $G_{n,p}$. We have $\mathrm{ex}(G_{n,p}, K_3) = (1 + o(1))pn^2/4$ whp provided $p \gg n^{-1/2}$.*

Proof. The lower bound follows by considering $G_{n,p} \cap K_{n/2,n/2}$, which is always triangle-free and which has $(1 + o(1))pn^2/4$ edges whp. For the upper bound, fix $\delta > 0$, and let $\varepsilon > 0$ be as in Theorem 17.10. Let $\mathcal{S}, f, g$ be as in Theorem 17.14. Note that each $f(S)$ has at most $(1/4+2\delta)n^2$ edges by Theorem 17.10. For each $S \in \mathcal{S}$, let E_S be the event that $S \subseteq G_{n,p}$ and that $|f(S) \cap G_{n,p}| \geq (1+4\delta)pn^2/4$. Note that in order to have $\mathrm{ex}(G_{n,p}, K_3) \geq (1+4\delta)pn^2 + O_\varepsilon(n^{3/2})$, some E_S event must occur, and moreover that $\Pr[E_S] = p^{|S|} \cdot e^{-O_\delta(pn^2)}$. With this we have

$$\Pr[\mathrm{ex}(G_{n,p}, K_3) \geq (1 + 4\delta)pn^2/4 + O_\varepsilon(n^{3/2})] \leq \Pr\left[\bigcup_{S \in \mathcal{S}} E_S\right] \leq \sum_{s=0}^{O_\varepsilon(n^{3/2})} \sum_{S \in \mathcal{S}: |S| = s} p^s e^{-O_\delta(pn^2)}.$$

As the number of $S \in \mathcal{S}$ with $|S| = s$ is trivially at most $\binom{n^2}{s} \le (en^2/s)^s$, we find that the above is at most

$$\sum_{s=0}^{O_\varepsilon(n^{3/2})} (epn^2/s)^s e^{-O_\delta(pn^2)}.$$

One can check that the function $(epn^2/s)^s$ is increasing for $s \le pn^2$. Since we know $s \le C_\varepsilon n^{3/2}$ for some suitable C_ε, we get that the sum above is at most

$$C_\varepsilon n^{3/2} \cdot (eC_\varepsilon^{-1} pn^{1/2})^{C_\varepsilon n^{3/2}} e^{-O_\delta(pn^2)},$$

and this tends to 0 provided $pn^{1/2} \to \infty$ (since $pn^2 \gg n^{3/2} \log(pn^{1/2})$), proving the result. □

Note that in this proof, the main extra power we gained by utilizing Theorem 17.14 is that S must be contained in our subgraph. This makes it so that the $S \in \mathcal{S}$ with many edges "cost more", allowing us to gain.

We note that in general, it is very common that by using the weak container lemma, one ends up getting tight bounds up to a logarithmic factor, and this extra factor can usually be remedied by utilizing the strong container lemma in some straightforward (if slightly more tedious) way.

18 Entropy

Throughout this part we let log denote logarithms base 2, and we define $x \log x = 0$ whenever $x = 0$.

Let X be a discrete random variable, and for ease of notation we write $p_x = \Pr[X = x]$. The *binary entropy* of X is defined as

$$H[X] = -\sum_{x \in \mathrm{supp}(X)} p_x \log(p_x),$$

where the sum ranges over all x in the support of X, i.e. those x with $p_x > 0$. For example, if $p_1 = \frac{1}{2},\ p_2 = p_3 = \frac{1}{4}$ then

$$H[X] = -(\frac{1}{2}\log(\frac{1}{2}) - \frac{1}{4}\log(\frac{1}{4}) - \frac{1}{4}\log(\frac{1}{4})) = \frac{1}{2}\log(2) + \frac{1}{4}\log(4) + \frac{1}{4}\log(4) = \frac{3}{2}.$$

The definition for $H[X]$ should come as strange if one has never seen it before. Roughly speaking, $H[X]$ can be thought of as measuring how much "information" the random variable X carries in expectation. For example, one can easily check that $H[X] = 0$ if and only if X is deterministic, corresponding to the fact that knowing the outcome of a deterministic process gives no information. In the next remark we give a more robust motivation for why one might come up with the definition of $H[X]$, though the reader may want to simply accept this is a good definition and continue on with their lives.

Remark 18.1. *Intuitively, we want $H[X]$ to measure the expected amount of "information" or "surprise" we see when given the outcome X. To this end, we want to consider a function $f(p)$ which measures how surprised we should be upon seeing an event which happens with probability p. For example, we'll want $f(1) = 0$ since probability 1 events always happen and $f(0) = \infty$ since probability 0 events never occur. The final property that we claim is reasonable is that we should perhaps have $f(pq) = f(p) + f(q)$. Indeed, if we observe some event which happens with probability p and then another independent event which happens with probability q, then heuristically our total surprise from seeing these two events should be the sum of the surprise from each of the events, i.e. should be $f(p) + f(q)$. Since these two events happen together with probability pq by independence, we conclude that we should have $f(pq) = f(p) + f(q)$. Upon seeing this, a reasonable choice of our function f measuring our amount of surprise should be $f(p) = -\log(p)$ to maintain these various properties. With this, $H[X]$ is simply $\mathbb{E}[f(\Pr[X = x])]$, i.e. the expected surprise we get from X.*

We will discuss a number of important properties that the entropy function $H[X]$ has in due time. Perhaps the most important of these is the following observation, the name for which is entirely non-standard.

Lemma 18.2 (Fundamental Property of Entropy)**.** *If $\mathcal{X}$ is a set and $X \in \mathcal{X}$ is chosen uniformly at random, then*

$$|\mathcal{X}| = 2^{H[X]}.$$

Indeed, this follows immediately from unwinding the definitions of $H[X]$ and using $p_x = |\mathcal{X}|^{-1}$ for all $x \in \mathcal{X}$. With this observation, we see that bounding the size of any set $\mathcal{X}$ is equivalent to bounding the entropy $H[X]$ of a nice random variable X, and sometimes this entropic point of view gives us access to information-theoretic tools which can not normally be used to directly bound the size of $|\mathcal{X}|$.

???Comment on motivation/lack there of, then note fundamental property which is easy to prove and all that's needed in addition to subadditivity to prove the first thing.

18.1 A First Example

Before getting into the full power of the entropy method, let us begin with a straightforward example. For this we will need one basic fact about the entropy function.

Lemma 18.3 (Subadditivity)**.** *For any random variables $X_1, \ldots, X_n$, we have*

$$H[(X_1, \ldots, X_n)] \le \sum_i H[X_i].$$

Intuitively, intuitively says that the total information given by a vector of random variables $(X_1, \ldots, X_n)$ is no more than the sum of the information given by the individual variables. For example, if X_1 denotes the price of a certain stock a year from today and if X_2 denotes the price a year and one day from today, then both X_1 and X_2 individually carry quite a bit of information, but the vector (X_1, X_2) does not carry significantly more information than just X_1 itself since X_2 is likely to be very close to X_1 in value.

While not too difficult, we will omit proving **??** and will instead focus on showing how to use this for a typical application of the entropy method, namely to the problem of upper bounding (the sums of) binomial coefficients. To this end, for $p \in [0, 1]$ we define the binary entropy function

$$H(p) := -p \log_2(p) - (1-p) \log_2(1-p),$$

which is simply the entropy of a Bernoulli variable with probability of success p.

Proposition 18.4. *For all $k \le n/2$, we have*

$$\sum_{i=0}^{k} \binom{n}{i} \le 2^{H(k/n) \cdot n}.$$

We note that this bound is essentially tight, in the sense that if k is linear in n, then $\binom{n}{k} = 2^{(1+o(1))H(k/n) \cdot n}$.

Proof. Let $\mathcal{X}$ denote the set of binary strings of length n with at most k 1's. Note that

$$|\mathcal{X}| = \sum_{i=0}^{k} \binom{n}{i},$$

so by the Fundamental Property of Entropy, proving our desired bound $\log |\mathcal{X}| \le H(k/n) \cdot n$ is equivalent to showing $H[(X_1, \ldots, X_n)] \le H(k/n) \cdot n$ where $(X_1, \ldots, X_n) \in \mathcal{X}$ is chosen uniformly at random. For this we use the following.

Claim 18.5. *Each of the random variables X_i is a Bernoulli random variable with probability of success $p \le k/n \le 1/2$.*

Proof. By the symmetry of the problem, we see that each X_i is Bernoulli with the same probability of success p. By construction we deterministically have $\sum_i X_i \le k$ for all $(X_1, \ldots, X_n) \in \mathcal{X}$, and hence $pn = \sum_i \mathbb{E}[X_i] \le k$. We conclude that $p \le k/n$, and this is at most $1/2$ by hypothesis on k. ◻

The claim above implies that for all i,

$$H[X_i] = H(p) \le H(k/n),$$

with this last step using the easy to prove fact that $H(p)$ is increasing for $p \le 1/2$. This together with Subadditivity gives

$$H[(X_1, \ldots, X_n)] \le \sum H[X_i] = H(k/n) \cdot n,$$

proving the result. ◻

The framework in our proof above is typical for the entropy method: we (1) started with a uniform random object X and then (2) used entropy to upper bound the size of $\mathrm{supp}(X)$. Although this is the most common framework for using entropy, it is also possible to (1') start with a non-uniform random object X and then (2') use entropy to lower bound the size of $\mathrm{supp}(X)$. For this we will need to better understand the properties of the entropy function.

18.2 The Main Properties of Entropy

To derive the full power of the entropy method, we will need to work with more than just the Maximality Principle and Subadditivity. In particular, our forthcoming Theorem 18.6 lists out the main properties of entropy which we will need, and for this it will be convenient to establish some basic notation.

Given a random variable X and an event E, we define $X|E$ to be the random variable X conditioned on the event E. Given a pair of random variables X, Y we define

$$H[X|Y] := \mathbb{E}_{y \sim Y} H[X|Y = y] = -\sum_y \Pr[Y = y] \sum_x \Pr[X = x|Y = y] \log(\Pr[X = x|Y = y]).$$

The expression $H[X|Y]$ is referred to as *conditional entropy*. For convenience, we will often denote vectors of random variables $(X_1, \ldots, X_n)$ simply by $X_1, \ldots, X_n$, e.g. by writing $H[X_1, \ldots, X_n]$ instead of $H[(X_1, \ldots, X_n)]$. For an integer i we let $X_{<i} := (X_1, \ldots, X_{i-1})$. Slightly more generally, if X is a random vector indexed by a set S with a total ordering $<$, then we let $X_{<s}$ denote the elements of X indexed by $t < s$.

We now state our list of properties about the entropy function. The reader is not expected to memorize this right away, though it might be a good idea to what extent these properties agree with the intuition[18] that $H[X]$ measures the information of X.

Proposition 18.6. *The following properties hold for any random variable X.*

- *(Non-negativity) We have $H[X] \geq 0$ with equality if and only if X is deterministic.*
- *(Maximality Principle) We have*

$$H[X] \leq \log |\mathrm{supp}(X)|,$$

with equality if and only if X is uniformly distributed on $\mathrm{supp}(X)$*.*

- *(Chain Rule) For two random variables X, Y we have*

$$H[X, Y] = H[X] + H[Y|X].$$

More generally, given random variables $X_1, \dots, X_n$ we have

$$H[X_1, \dots, X_n] = \sum_i H[X_i|X_1, \dots, X_{i-1}].$$

- *(Subadditivity) For random variables $X_1, \dots, X_n$ we have*

$$H[X_1, \dots, X_n] \leq \sum_i H[X_i].$$

- *(Dropping Conditioning) For random variables X, Y, Z we have*

$$H[X|Y] \leq H[X],$$

with equality if and only if X is independent of Y. Similarly

$$H[X|Y, Z] \leq H[X|Y],$$

with equality if and only if X conditioned on Y has the same distribution as X conditioned on both Y and Z.

- *(Data Processing Inequality) If X, Y, Z are random variables such that Z is a function of Y, then*

$$H[X|Y] \leq H[X|Z].$$

[18] For example, Subadditivity says that the total information in the vector (X, Y) is at most the sum of the information of X and Y, Dropping Conditioning says that knowing less at the start can only lead to more information gained, and Data Processing says that one can't do anything to a random variable Y in order to reveal more information about X.

We will not prove Theorem 18.6 since it's a slight detour from our main goal of applying entropy to graph theory problems. Most proofs can be found in standard texts on entropy, e.g. the survey by Galvin [?] (which contains even more properties, especially around conditional entropy). The only exceptions might be proofs for the "only if" portion of Dropping Conditioning (which we will never use); as well as the Data Processing Inequality (whose name is not entirely standard), but this follows from observing that $H[X|Y,Z] = H[X|Y]$ by definition and that $H[X|Y,Z] \leq H[X|Z]$ by Dropping Conditioning.

We emphasize that there is plenty of redundancy in the list of properties above. For example Subadditivity follows from the Chain Rule and Dropping Conditioning, but this property is used so frequently that it will be useful to list it as a separate property. It is also worth noting that the function $H[X] = -\sum_x p_x \log(p_x)$ turns out to be essentially the unique function which satisfies the list of properties from Theorem 18.6. As such, $H[X]$ is essentially the unique function which captures how much "information" a random variable X has. For more on this see e.g. [?, Chapter 7].

Let us now apply these more sophisticated properties of entropy to solve a problem about counting walks in graphs. Recall that a walk of length k in G is a sequence of vertices $(x_1, \ldots, x_{k+1})$ such that $x_i \sim x_{i+1}$ for all $1 \leq i \leq k$. Equivalently, a walk of length k is a homomorphism from P_{k+1} to G. As such, the following result (commonly known as the Blakey-Roy Theorem though they weren't quite the first to prove it) is equivalent to Sidorenko's Conjecture for paths.

Theorem 18.7 (Blakey-Roy [?]). *If G is an n-vertex graph with $m \geq 1$ edges, then the number of walks of length k in G is at least*

$$2m(2m/n)^{k-2}.$$

Note that this bound is tight whenever G is regular.

Proof. Let $X = (X_1, \ldots, X_{k+1})$ be a random walk of length k in G chosen in the following *non-uniform* way: choose the pair (X_1, X_2) uniformly at random amongst all pairs such that $X_1 \sim X_2$, and given X_{i-1} for $i \geq 3$, we choose X_i uniformly at random amongst the neighbors of X_{i-1}. Observe that X is indeed always a walk of length k and that we have implicitly used $m \geq 1$ to guarantee (X_1, X_2) exists. We can then express its entropy as

$$\begin{aligned} H[X] &= \sum_{i=1}^{k+1} H[X_i | X_{<i}] \\ &= H[X_1] + H[X_1|X_2] + \sum_{i=3}^{k+1} H[X_i|X_{i-1}] \\ &= H[X_1, X_2] + \sum_{i=3}^{k+1} H[X_i, X_{i-1}] - H[X_{i-1}], \end{aligned} \tag{7}$$

where here the first and last equality used the Chain Rule, and the second used the equality case of Dropping Conditioning since X_i depends only on X_{i-1}.

We claim that (X_{i-1}, X_i) is uniformly random amongst the set of all pairs $(y, z) \in V(G)^2$ such that $y \sim z$. This is true for $i = 2$ by construction, so assume we have proven it true up to some value $i \geq 3$. In this case, for any $(y, z) \in V(G)^2$ such that $y \sim z$, we have by conditioning on every possible value that $X_{i-2} \sim X_{i-1}$ can take on,

$$\begin{aligned}\Pr[(X_{i-1}, X_i) = (y, z)] &= \sum_{x \in N(y)} \Pr[(X_{i-2}, X_{i-1}) = (x, y)] \cdot \Pr[X_i = z | (X_{i-2}, X_{i-1}) = (x, y)] \\ &= \sum_{x \in N(y)} \frac{1}{2m} \cdot \frac{1}{\deg(y)} = \frac{1}{2m},\end{aligned}$$

with this last equality using the hypothesis that (X_{i-2}, X_{i-1}) is distributed uniformly at random and that X_i is a uniform random neighbor of X_{i-1}. This establishes the claim.

With this claim, we have by the Maximization Principle that $H[X_i, X_{i-1}] = \log(2m)$ for all i, and also that $H[X_{i-1}] \leq \log(n)$ for all i. Using this with (7) gives

$$H[X] \geq (k-1)\log(2m) - (k-2)\log(n).$$

Now let W_k denote the set of walks of length k in G. Since X is a random element from W_k, we have $H[X] \leq \log|W_k|$, which combined with the lower bound for $H[X]$ above gives the desired result. $\square$

To emphasize, the proof above would theoretically have worked if we considered a uniform random element $\tilde{X} \in W_k$ instead of the non-uniform X described above, in the sense that the Maximization Principle gives

$$H[\tilde{X}] \geq H[X] \geq (k-1)\log(2m) - (k-2)\log(n).$$

However, it is not clear how one would prove the lower bound of $(k-1)\log(2m) - (k-2)\log(n)$ for $H[\tilde{X}]$ directly without going through X first. More generally, it can be useful when working with lower bounds from entropy to work with a distribution for X which is "natural" to the problem rather than one which is uniform. We also emphasize that to obtain tight examples, it is crucial that our choice of X is actually distributed uniformly whenever we are working with an extremal example (as otherwise the Maximization Principle shows that we can not hope to obtain a tight bound).

The exact same proof as above can easily be extended to prove Sidorenko's Conjecture for all trees at the cost of more complicated notation. A different and somewhat more involved entropy argument can be used to show Sidorenko's conjecture holds whenever F has a vertex which is adjacent to every vertex in the other partition set.

18.3 Random Chain Rules

This section concerns a strengthening of the chain rule. Historically, this strengthening was first used to give an entropy proof of the following result.

Theorem 18.8 (Brégman's Theorem [?]). *If G is a bipartite graph with bipartition $U \cup V$ such that $|U| = |V|$, then the number of perfect matchings of G is at most*

$$\prod_{u \in U} (\deg(u)!)^{1/\deg(u)}.$$

Observe that this bound is tight by considering disjoint unions of complete balanced bipartite graphs (possibly of differing sizes). This was originally proven by Brégman [?], and later Radhakrishnan [?] gave an elegant entropy-based proof that we present below.

Before stating the key lemma needed to prove Theorem 18.8, let's first try and prove this result naively from first principles and see where things go wrong. As usual, we start with a uniform random perfect matching M of G. We then translate M into a vector X indexed by U by having $X_u \in V$ be the unique neighbor of u in M. We fix some arbitrary ordering $<$ of the vertices of U, and then apply the chain rule to obtain

$$H[X] = \sum_u H[X_u | X_{<u}].$$

At this point, the naive entropy argument calls for upper bounding $H[X_u|X_{<u}]$ by log of the number of possible values X_u can take given the values of $X_{<u}$. To this end, we define $A_{<u} \subseteq N(u)$ to be the set of neighbors of u which do not appear in $X_{<u}$ (i.e. this is the set of available neighbors of u for M given the information in $X_{<u}$). The Maximality Principle then gives

$$\sum_u H[X_u|X_{<u}] \leq \sum_u \log(|A_{<u}|).$$

Unfortunately, for any given u and ordering $<$, we can't say anything about $|A_{<u}|$ other than $|A_{<u}| \leq \deg(u)$. Applying this worst-case bound for all u gives a trivial upper bound of $\prod_u \deg(u)$ in the end.

While it is true that worst case we can have $|A_{<u}| = \deg(u)$ for any given u, intuitively we should "typically" have $|A_{<u}| \approx \frac{1}{2}\deg(u)$, since for a "random" M we would expect around half of u's neighbors to be matched in M to vertices appearing before u in $<$ and half to be matched to vertices after u. Again, this intuition may not hold for a given M and $<$, but this intuition can be made precise if we consider a random ordering $<$ instead of a fixed one. To this end, one can consider the following random variant of the chain rule.

Lemma 18.9 (Random Chain Rule)**.** *Let X be a vector indexed by a set S and let $<$ be a random ordering of S. Then*

$$H[X] = \sum_{s \in S} \mathbb{E}_<[H[X_s|X_{<s}]].$$

Indeed, the proof of this follows from the fact that equality holds for any fixed $<$ (by the usual chain rule), and hence equality also holds when one takes expectations. With this we can quickly adapt our previous failed attempt to give Theorem 18.8.

Proof of Theorem 18.8. Let $<$ denote a uniform random ordering of U. Keeping all of the notation from the argument above, we have by the random chain rule that

$$H[X] = \sum_u \mathbb{E}_<[H[X_u|X_{<u}]] \leq \sum_u \mathbb{E}_<[\log(|A_{<u}|)]. \tag{8}$$

We claim that $|A_{<u}|$ is distributed uniformly at random amongst $[\deg(u)]$. Indeed, fix any perfect matching M (so now all the randomness lies in the much simpler random object $<$)

and let v denote the neighbor of u in M. We then observe that $|A_{<u|}| = i$ if and only if $i-1$ vertices of $N(u) \setminus \{v\}$ have their neighbors in M appear after u under $<$. Since $<$ gives a uniform random ordering on $N(u)$ regardless of our choice of M, we conclude that $|A_{<u}|$ is indeed equally likely to be any value in $[\deg(u)]$. Again, we emphasize that this result holds regardless of the fixed value of M, and hence $|A_{<u}|$ continues to be uniformly distributed even if we do not condition on M.

With this claim, we can write (8) above as

$$H[X] \le \sum_u \frac{1}{\deg(u)} \sum_{i=1}^{\deg(u)} \log(i) = \sum_u \frac{\log(\deg(u)!)}{\deg(u)}.$$

Exponentiating both sides gives the result. □

We emphasize that the power of the Random Chain Rule is that it gives us an extra source of randomness over the usual Chain Rule which we can exploit in the analysis of bounding $H[X]$. We show another example of this for the problem of counting 1-factorizations in K_n, which we recall are ordered partitions of $E(G)$ into perfect matchings. For example, K_4 has 6 different 1-factorizations, namely $(\{12, 34\}, \{13, 24\}, \{14, 23\})$ and all of its permutations.

I don't know if there's a relevant citation here.

Theorem 18.10. *The number of 1-factorizations of K_n when n is even is at most $((1 + o(1))n/e^2)^{\binom{n}{2}}$.*

Observe that this improves upon the trivial upper bound $(n-1)^{\binom{n}{2}}$, which is just the number of ways to partition the edge set of K_n into $n-1$ edge-disjoint graphs.

Proof. For this proof, it will be slightly more convenient to work with a base e notion of entropy rather than the usual base 2 notion. To this end, if X is a random variable with $p_x = \Pr[X = x]$, then we define

$$H_e[X] = -\sum_x p_x \log_e(p_x).$$

Note that $H_e[X] = \log_e(2)H[X]$, and in particular, essentially all of the properties for $H[X]$ continue to hold for $H_e[X]$.

Let M denote a uniformly random 1-factorization of K_n. We will think of M as assigning to each edge uv of K_n a color in $[n-1]$ such that the edges in color i form a perfect matching (equivalenlty, M is a proper edge coloring of K_n with $n-1$ colors). Let X be the vector indexed by $E(K_n)$ where X_{uv} equals the color assigned to uv by M.

As before, we will consider a uniformly random ordering $<$ on $E(K_n)$, but for technical reasons we will want to form this ordering in a slightly more complex way. To this end, assign to each edge uv a random weight w_{uv} chosen independently and uniformly from $[0, 1]$, then let $<$ be the ordering of $E(K_n)$ which has $uv < xy$ iff $w_{uv} < w_{xy}$. Again we emphasize that $<$ has the same distribution as if we just chose it to be uniformly at random, but it will be convenient for us to have these extra w_{uv} parameters to work with as yet another additional source of randomness for us to exploit in our analysis.

An application of the random chain rule then gives

$$H_e[X] = \sum_{uv} \mathbb{E}_<[H_e[X_{uv}|X_{<uv}]] \le \sum_{uv} \mathbb{E}_<[\log_e(|A_{<uv}|)], \tag{9}$$

where here A_{uv} denotes the set of colors that are "available" for uv given $X_{<uv}$; i.e. A_{uv} consists of the colors i that do not lie on any edge xy which intersects uv and which has $xy < uv$ under the coloring M.

It remains to estimate $\mathbb{E}_<[\log_e(|A_{<uv}|)]$, and for this, it will suffice to condition on the 1-factorization M and prove an upper bound that is independent of M. From now on we fix M, noting that $A_{<uv}$ will always contain the color c which M assigns to uv. Observe that a color $i \in [n-1] \setminus \{c\}$ will be in $A_{<uv}$ if and only if uv appears in the ordering before the two edges incident to uv which are colored i by M (note that exactly two such edges exist since M is a 1-factorization). Conditional on the value w_{uv}, the probability that this happens for any given i is $(1-w_{uv})^2$. Using this and Jensen's inequality gives

$$\mathbb{E}_<[\log_e(|A_{<uv}|)|w_{uv}, M] \le \log_e(\mathbb{E}_<[|A_{<uv}||w_{u,v}, M]) = \log_e(1+(n-2)(1-w_{uv})^2).$$

As w_{uv} was distributed uniformly at random in $[0,1]$, we find

$$\mathbb{E}_<[\log_e(|A_{<uv}|)|M] \le \int_0^1 \log_e(1+(n-2)(1-x)^2)dx.$$

Summing this over all edges uv together with (9) gives

$$H_e[X] \le \binom{n}{2} \cdot \int_0^1 \log_e(1+(n-2)(1-x)^2)dx = \binom{n}{2} \cdot (\log_e(n) - 2 + o(1)),$$

where this last equality follows from some fiddly integral analysis[19] (with the intuition being that the integrand is close to $\log_e(n(1-x)^2) = \log_e(n) + 2\log_e(1-x)$, and this ends up integrating to the desired value). Exponentiating both sides by e gives that the total number of 1-factorizations is at most $((1+o(1))n/e^2)^{\binom{n}{2}}$ as desired.

$\square$

As an aside, if in (9) we used the Chain Rule instead of the Random Chain Rule, then we would pessimistically use $|A_{<uv}| \le n-1$ for all uv, and this would give the trivial upper bound $(n-1)^{\binom{n}{2}}$.

[19] Slightly more precisely, one can argue that the integral evaluated from $1-n^{-.1}$ to 1 is at most $O(n^{-.1}\log n) = o(1)$, and outside of this range the difference between the integrand $\log_e(1+(n-2)(1-x)^2)$ and $\log_e(n(1-x)^2)$ is $o(1)$.

19 Absorption

Roughly speaking, extremal combinatorics studies how "large" a combinatorial object can be before it is guaranteed to satisfy a given property. Some properties are "local," in that they can be verified by looking at a small subset of our object (e.g. verifying that a graph has a triangle only requires looking at the edges between three particular vertices). Other properties are "global" in that they can only be verified by looking at the entire combinatorial object (e.g. verifying that a graph contains a Hamiltonian cycle requires looking at all of the vertices).

In this chapter we will look at what has been perhaps the biggest development in the study of global extremal problems: the absorption method. This method (along with its numerous variants) has served as a key component in many of the most recent breakthroughs in (global) extremal combinatorics, such as the existence of designs [**?**, **?**], the resolution of the Erdős-Farber-Lovász Conjecture regarding the edge-chromatic number of hypergraphs [**?**], the near resolution of the Ryser-Brualdi-Stein Conjecture regarding trasnversals in Latin squares [**?**], and many more.

Our guiding example will be studying a hypergraph generalization of the Dirac problem for perfect matchings. In the setting of graphs, every n-vertex graph G with n even and minimum degree $\delta(G) \geq n/2$ contains a perfect matching, with this bound on $\delta(G)$ being best possible by considering G to either be $K_{n/2-1,n/2+1}$ or the disjoint union of two cliques of size $n/2$. The analogous problem for hypergraphs, on the other hand, turns out to be substantially harder.

The first obstacle in addressing the hypergraph problem is the need for us to choose what we mean by the "minimum degree" of a hypergraph. Indeed, given an r-graph H and an integer $1 \leq i < r$, we can define the *minimum i-degree* $\delta_i(H)$ of H by $\delta_i(H) = \min_{S \subseteq \binom{V(H)}{i}} \deg_H(S)$, and for each possible value of i, r, one can consider the question of how large $\delta_i(H)$ needs to be to guarantee a perfect matching in H. Here we will focus only on the case of $i = r-1$, which turns out to be the easiest (though by no means easy!) case to solve. For convenience we will refer to $\delta_{r-1}(H)$ as the *minimum codegree* of H.

Given that a bound of $n/2$ is the threshold for the minimum codegree of graphs (i.e. for the case $r = 2$), it is perhaps natural to guess that the right answer in general should be something like n/r. Further evidence of this guess comes from the following result, which gives an almost full solution to this problem using a very simple proof, and which in particular recovers Dirac's result for $r = 2$.

Proposition 19.1. *For all $r \geq 2$, if H is an n-vertex r-graph with n divisible by r and $\delta_{r-1}(H) \geq n/r$, then H contains a matching of size at least $n/r - r + 2$.*

Proof. Let M be a largest matching in H and assume for contradiction that $|M| \leq n/r - r + 1$. This means that there are at least $r(r-1)$ vertices not in any edge of M, and we let $U_1, \ldots, U_r$ denote a collection of disjoint sets of uncovered vertices of size $r-1$.

Let $\mathcal{U}$ denote the set of edges of H which contain a U_i set. Observe that the minimum degree condition implies $|\mathcal{U}| \geq r \cdot n/r = n$, and also that the edges of $\mathcal{U}$ must all intersect an edge of M, as otherwise this edge of $\mathcal{U}$ together with M would form a larger perfect matching. In fact, each edge of $\mathcal{U}$ must intersect exactly one edge of M since each edge of $\mathcal{U}$ contains some set U_i of size $r-1$ which is disjoint from the edges of M and since the edges of M are pairwise disjoint. The pigeonhole principle then implies that there exists some $e \in M$ such that the number of edges $\mathcal{U}_e \subseteq \mathcal{U}$ intersecting e is at least

$$\left\lceil \frac{|\mathcal{U}|}{|M|} \right\rceil \geq \left\lceil \frac{n}{n/r-1} \right\rceil \geq r+1,$$

where here the first inequality used that $|M| < n/r$ by assumption.

We claim that $\mathcal{U}_e$ contains two disjoint edges. Indeed, define an auxiliary bipartite graph G between the r disjoint U_i sets and the r vertices of e by having $U_i \sim v$ if $U_i \cup \{v\} \in \mathcal{U}_e$. Since this bipartite graph has $|\mathcal{U}_e| \geq r+1$ edges and since each part of G has size r, it must have a matching of size 2, which exactly corresponds to two disjoint edges of $\mathcal{U}_e$. Adding these two edges to $M \setminus \{e\}$ gives a strictly larger matching than M, contradicting our hypothesis. □

At this point it feels like we're close to done. Indeed, given how easy the result above was to prove, it is perhaps natural to expect that with a little more work one could show that, say, a minimum codegree of the form $\delta_{r-1}(H) \geq n/r + O_r(1)$ is enough to guarantee the existence of a perfect matching covering all of the vertices. Surprisingly, it turns out that this intuition is very far from true. In particular, the following exact result of Rödl, Ruciński, and Szemerédi [**?**] shows that despite Theorem 19.1 suggesting a bound of n/r, the true codegree bound is much closer to $n/2$.

Theorem 19.2 ([**?**]). *For all $r \geq 2$, if H is an n-vertex r-graph with n sufficiently large and divisible by r and with*

$$\delta_{r-1}(H) \geq \begin{cases} n/2+3-r & r \equiv 0 \mod 4,\ n/r \equiv 1 \mod 2, \\ n/2+5/2-r & r \equiv 1 \mod 2,\ n \equiv 3 \mod 4, \\ n/2+3/2-r & r \equiv 1 \mod 2,\ n \equiv 1 \mod 4, \\ n/2+2-r & \textit{otherwise}, \end{cases}$$

then H contains a perfect matching. Moreover, all of these bounds are sharp.

Roughly speaking, the constructions showing that these bounds are best possible are obtained by splitting the vertex set of H into two parts U, V with $|U| \approx n/2$ an odd integer and then considering all edges that intersect U in an even number of vertices. It is not difficult to see that such an H can not have a matching covering all of U, and moreover that the minimum codegree is around $n/2 - r$ (since any given $(r-1)$-set S is either in an edge with all of $U \setminus S$ or all of $V \setminus S$).

The two results Theorem 19.2 and Theorem 19.1 illustrate a common and annoying phenomenon in the study of finding spanning subgraphs of (hyper)graphs: finding an "almost" spanning structure (e.g. a matching covering almost all of the vertices) is often much easier to do (both in terms of its proofs, as well as in the necessary degree hypothesis needed in its statement) than finding an "exact" spanning structure[20].

[20]Another example of this phenomenon that we have seen already is with Rödl's Theorem **??** showing the existence of asymptotically large partial Steiner systems, with it only being much later that genuine Steiner systems were constructed by [**?**, **?**].

While this disparity in difficulty can be frustrating at first, it is in fact a key element needed for the absorption method to work. In particular, because the codegree bound of $n/2$ is so much larger than what we need in order to cover almost all the vertices, one might consider a vague strategy for trying to solve the "exact" spanning problem of perfect matchings as follows:

(i) Start with a hypergraph H of minimum codegree $n/2$.

(ii) Choose a "small" set of vertices A which is "nice" and then delete these vertices from H.

(iii) Because A is "small", the minimum degree of H will still be at least n/r, and hence Theorem 19.1 guarantees that we can find a matching M in $H - A$ covering all but a "very small" set of leftover vertices L.

(iv) Because A is "nice" and because L is "very small", we can find a prefect matching on $H[A \cup L]$, which together with M forms a perfect matching of H.

Indeed, we will use exactly this approach to solve a somewhat weaker version of Theorem 19.2. For this, we need to make this last step (iv) and the definition of "nice" more precise. To this end, we make the following (non-standard) definition.

Definition 22. Given an r-graph H and reals a, ℓ, we say that a set of vertices $A \subseteq V(H)$ is an *(a, ℓ)-absorber* if $|A| \le a$ is a multiple of r and if for every $L \subseteq V(H) \setminus A$ with $|L| \le \ell$ a multiple of r, the induced subgraph $H[A \cup L]$ contains a perfect matching.

Our proof strategy above indicates that we will solve our problem if we can show that every r-graph H with minimum codegree at least $n/2$ has an (a, ℓ)-absorber with, say, $a \le (1/2 - 1/r)n$ and $\ell \le r(r-2)$. Of course, it is not at all clear at this point how we might construct such an A or what it might look like, so we will have to think more about this.

As a motivating example, consider the simplest non-trivial case that our leftover set L is just a single r-set which is not an edge. In this case we can ask ourselves: what is the simplest possible choice of A that will absorb L, i.e. such that $H[A]$ and $H[A \cup L]$ both have perfect matchings? In this very simplified problem, one easy solution is to use the following construction, where here the sets R, B play the roles of L, A in the discussion above.

Definition 23. Given an r-set $R = \{x_1, \ldots, x_r\}$, we call a set $B = \{x_1', \ldots, x_r'\} \subseteq V(H) \setminus R$ an *(absorbing) building block* for R if B is an edge of H and if $\{x_1, x_2', \ldots, x_r'\}$ and $\{x_1', x_2, \ldots, x_r\}$ are both edges of H.

A picture would be great here.

Indeed, one can easily check that these sets have the following properties that we need.

Lemma 19.3. *If B is a building block for R in an r-graph H, then $H[B]$ and $H[B \cup R]$ both have perfect matchings.*

Of course, we can only hope to use such building blocks if they actually exist in H. And indeed, provided our minimum codegree is sufficiently large, such sets will not only exist, but in fact make up a constant proportion of all possible r-sets of H.

Lemma 19.4. *Let H be an n-vertex r-graph with $n \geq 4r$ and $\delta_{r-1}(H) \geq (1/2+\varepsilon)n$ for some $\varepsilon > rn^{-1}$. If $R \subseteq V(H)$ is an r-set, then there exist at least $\varepsilon(2r)^{-r}n^r$ distinct building blocks for R.*

Proof. We will construct building blocks as follows: let $x_1' \in V(H) \setminus \{x_1\}$ be any vertex which is contained in an edge with $\{x_2, \ldots, x_r\}$, the number of which is at least $(1/2+\varepsilon)n - 1 \geq \frac{1}{2}n$. Next, choose $x_2', \ldots, x_{r-1}'$ to be arbitrary distinct vertices outside of the set $\{x_1'\} \cup R$, which can be done in at least $(n-2r)^{r-2} \geq 2^{2-r}n^{r-2}$ ways. Finally, choose $x_r' \notin R$ to be a vertex such that both $\{x_1', x_2', \ldots, x_r'\}$ and $\{x_1, x_2', \ldots, x_r'\}$ are edges of H, noting by our minimum codegree condition for both $\{x_1', x_2', \ldots, x_{r-1}'\}$ and $\{x_1, x_2', \ldots, x_{r-1}'\}$, that the number of such choices for x_r' is at least

$$2(1/2+\delta)n - n - r \geq \varepsilon n.$$

In total, the number of (not necessarily distinct) building blocks produced from this procedure is at least $\frac{1}{2}n \cdot 2^{2-r}n^{r-2} \cdot \varepsilon n \geq \varepsilon 2^{-r}n^r$, and since each block can be produced in at most $r! \leq r^r$ ways from this procedure we conclude the desired result. □

With this lemma in mind, our vague plan now will be as follows: we will try to construct our absorber A by taking a (carefully chosen) union of disjoint building blocks B. We then wish to say that for any (small) leftover set L, we can partition L into r-sets R_i such that for each i there exists an absorbing block $B_i \subseteq A$ for R_i, and moreover that we have $B_i \neq B_j$ for any $i \neq j$, at which point we would be done since we can find matchings in the $B_i \cup R_i$ sets as well as in any B_j which is not matched to any R_i. Again, it is not at all obvious that we can find an A such that this pairing of the R_i and B_i sets always exists regardless of what L is given to us, but the previous lemma gives some hope that this might work out if we choose our blocks B of A in a "random way." And indeed, this idea together with a little grit gives the following.

Proposition 19.5. *Let H be an n-vertex r-graph with $r \geq 2$ and $\beta > 0$ a real such that for every r-set $R \subseteq V(H)$, there exist at least βn^r absorption building blocks B for R. If $\alpha \leq \beta$ is a real number and if n is sufficiently large in terms of α, β, r, then there exists an $(\alpha n, 2^{-4}\alpha\beta n)$-absorber A of H.*

We emphasize that this proof basically boils down to taking A to be a random subset of building blocks together with a deletion argument, and as such the reader may wish to postpone the somewhat detailed proof and continue on to see how this result is used to complete our absorption argument.

Proof. A little thought shows that the pairing between R_i sets and B_i mentioned above is essentially asking for a suitable matching in an appropriately defined graph. To this end, define G to be the bipartite graph whose parts $\mathcal{B}, \mathcal{R}$ are disjoint copies of $\binom{V(H)}{r}$ and where we have $B \sim R$ in G whenever B is a building block of R.

Claim 19.6. *To prove the result, it suffices to find a subset $\mathcal{B}' \subseteq \mathcal{B}$ such that (1) $|\mathcal{B}'| \leq r^{-1}\alpha n$, (2) any two distinct vertices $B, B' \in \mathcal{B}'$ have $B \cap B' = \emptyset$, and (3) for any $\mathcal{R}' \subseteq \mathcal{R}$ with $|\mathcal{R}'| \leq 2^{-4}r^{-1}\alpha\beta n$, there exists a matching in $G[\mathcal{B}' \cup \mathcal{R}']$ such that every vertex of $\mathcal{R}'$ is covered.*

Proof. Assume that such a $\mathcal{B}'$ exists. Note that we may assume without loss of generality that $\mathcal{B}'$ contains no isolated vertices in G, as otherwise we could remove such vertices from $\mathcal{B}'$ while maintaining all of the other properties. In particular, this implies that each $B \in \mathcal{B}'$ is a building block for some R, and in particular that $H[B]$ has a perfect matching (i.e. that B is an edge). We aim to show then that $A := \bigcup_{B\in\mathcal{B}'} B$ is an $(\alpha n, 2^{-4}\alpha\beta n)$-absorber.

Note that $|A| \le \alpha n$ by (1) and this set has size a multiple of r by (2). Let $L \subseteq V(H) \setminus A$ be an arbitrary set of size $|L| \le 2^{-4}\alpha\beta n$ a multiple of r, and let $\mathcal{R}' = \{R_1, R_2, \ldots, R_t\}$ be an arbitrary partition of L into sets of size R, which has size at most $2^{-4}r^{-1}\alpha\beta n$ by definition. By (3), we can order the elements of $\mathcal{B}' = \{B_1, \ldots, \}$ such that $B_i \sim R_i$ in G for all $i \le t$, i.e. such that B_i is a building block for R_i, which by Theorem 19.3 implies that $H[B_i \cup R_i]$ has a perfect matching. This implies that $H[A \cup L]$ has a perfect matching, namely by taking the matchings from each $H[B_i \cup R_i]$ for $i \le t$ together with the matchings from each $H[B_i]$ for $i > t$ guaranteed by our assumption at the start of the proof (noting that this is indeed a matching since the B_i sets are disjont from both the $R_j \subseteq L \subseteq V(H) \setminus A$ sets as well as the other B_j sets by (2)). Since L was an arbitrary set of size $|L| \le 2^{-4}\alpha\beta n$ a multiple of r, we conclude that A is indeed an $(\alpha n, 2^{-4}\alpha\beta n)$-absorber. □

It thus remains to find a subset $\mathcal{B}'$ as in the claim. To this end, we note that our hypothesis on H implies that every $R \in \mathcal{R}$ has degree at least βn^r in G, which means that the graph is quite dense. Outside of this we have no real knowledge of what G looks like, so it is perhaps reasonable to try looking at a random subset of $\mathcal{B}$.

To this end, let $\mathcal{B}_p \subseteq \mathcal{B}$ be defined by including each vertex B in $\mathcal{B}_p$ independently with probability

$$p := \frac{1}{4} r^{-1}\alpha n^{1-r},$$

and let $G_p = G[\mathcal{B}_p \cup \mathcal{R}]$. Note that we will not be able to take $\mathcal{B}_p = \mathcal{B}'$ directly, as there is a small chance that $\mathcal{B}_p$ contains $B \ne B'$ which are not disjoint. To this end, we let X denote the set of pairs (B, B') with $B, B' \in \mathcal{B}_p$ distinct intersecting sets, and we let $\mathcal{B}' \subseteq \mathcal{B}_p$ be defined by taking $\mathcal{B}_p$ and deleting one element from each pair of X.

Claim 19.7. *Let E_1 denote the event that $|\mathcal{B}_p| - |\mathcal{B}'| \ge 2^{-4}r^{-1}\alpha\beta n$, let E_2 denote the event that $|\mathcal{B}_p| \ge \alpha n$, and let E_3 denote the event that there exists an $R \in \mathcal{R}$ with $\deg_{G_p}(R) \le 2^{-3}r^{-1}\alpha\beta n$. If $\Pr[E_1 \cup E_2 \cup E_3] < 1$, then the result holds.*

Proof. If $\Pr[E_1 \cup E_2 \cup E_3] < 1$, then there exists some instance of $\mathcal{B}_p$ and $\mathcal{B}'$ such that none of these three events occur. In this case, $|\mathcal{B}'| \le |\mathcal{B}_p| \le \alpha n$ by E_2 not holding, so (1) of Theorem 19.6 holds, and we automatically have (2) holding by definition of $\mathcal{B}'$. For (3), let $G' = G[\mathcal{B}' \cup \mathcal{R}]$. Observe that because E_3 and E_1 do not hold, for all $R \in \mathcal{R}$ we have

$$\deg_{G'}(R) \ge \deg_{G_p}(R) - (|\mathcal{B}_p| - |\mathcal{B}'|) \ge 2^{-4}r^{-1}\alpha\beta n.$$

Thus for any $\mathcal{R}' \subseteq \mathcal{R}$ of size at most $2^{-4}r^{-1}\alpha\beta n$, we can greedily find a matching in G' covering all of $\mathcal{R}'$ by iteratively choosing for each $R \in \mathcal{R}'$ an arbitrary neighbor in G' which has not already been selected, with the bound above implying that this process always terminates successfully. This verifies (3), proving the claim. □

It thus remains to show that these three events are unlikely. For E_1, we observe that

$$\mathbb{E}[|X|] = \sum_{B \in \mathcal{B}} \sum_{\substack{B' \neq B, \\ B' \cap B \neq \emptyset}} p^2 \leq \binom{n}{r} \cdot r \binom{n}{r-1} \cdot p^2 \leq p^2 n^{2r-1} = 2^{-4} r^{-2} \alpha^2 n,$$

and hence by Markov's inequality and the observation $2^{-4}r^{-1}\alpha\beta n \geq 2^{-3}r^{-2}\alpha^2 n$ by our hypothesis $r \geq 2$ and $\beta \geq \alpha$, we find that

$$\Pr[|\mathcal{B}_p| - |\mathcal{B}'| \geq 2^{-4}r^{-1}\alpha\beta n] \leq \Pr[|X| \geq 2^{-3}r^{-2}\alpha^2 n] \leq \frac{1}{2}.$$

For E_2, we note that $|\mathcal{B}_p|$ is a binomial random variable with mean $p|\mathcal{B}| = p\binom{n}{r}$, so the Chernoff bound and n being sufficiently large implies that

$$\Pr[|\mathcal{B}_p| \geq \alpha n] \leq \Pr[|\mathcal{B}_p| \geq 2p|\mathcal{B}|] < \frac{1}{4}.$$

Similarly, the degree of each $R \in \mathcal{R}$ in G_p is a binomial random variable, so the Chernoff bound implies

$$\Pr[\deg_{G_p}(R) \leq 2^{-3}r^{-1}\alpha\beta n] \leq \Pr[\deg_{G_p}(R) \leq \frac{1}{2}p \deg_G(R)] < \frac{1}{4n},$$

and taking a union bound shows that $\Pr[E_3] < 1/4$. Combining this with the bounds above gives the desired result. □

With all this we can prove an asymptotically tight version of Theorem 19.2 which was originally proven by Rödl, Ruciński, and Szemerédi [**?**] prior to their exact solution from [**?**].

Theorem 19.8 ([**?**]). *For all $r \geq 2$ and $\varepsilon > 0$, if H is an n-vertex r-graph with n sufficiently large and divisible by r and with $\delta_{r-1}(H) \geq (1/2 + \varepsilon)n$, then H contains a perfect matching.*

Proof. Taking $\alpha := (2r)^{-r}\varepsilon$, we have by Theorem 19.4 and Theorem 19.5 that H has an $(\alpha n, 2^{-4}\alpha^2 n)$-absorber A. Letting $H' = H - A$, we see that

$$\delta_{r-1}(H') \geq \delta(H) - |A| \geq n/2 \geq v(H')/2,$$

with this second inequality using $\alpha \leq \varepsilon$. By Theorem 19.1 there exists a matching M' of H' which covers all but a set L of at most r^2 vertices. Because A is an $(\alpha n, 2^{-4}\alpha^2 n)$-absorber and $r^2 \leq 2^{-4}\alpha^2 n$ for n sufficiently large, we have by definition of an absorber that $H[A \cup L]$ contains a perfect matching M'', which combined with M' gives a perfect matching of H. □