Chapter 1

Introduction

1.1 Prologue

Extremal Graph Theory, on the most general level, investigates the extremal (maximal or minimal) value of various graph parameters over the family of graphs having a particular property. It is a lively subject with a rich history, where numerous natural questions have beautiful answers. It is a field very much driven by problems; many of the interesting ones are still wide open and stimulate an abundance of research.

Each such problem has two sides: one is the construction of an extremal structure, the other is the proof of its optimality. In this course we are putting extra emphasis on explicit constructions of extremal graphs, which do not customarily feature in standard treatments of the field. These constructions often require useful tools from algebra, geometry, or discrete Fourier analysis; the other main objective of these notes is to highlight them.

We will consider two families of problems: *Turán-type problems* and *Ramsey-type problems*. In the first lecture we discuss the underlying ideas of both areas by looking more closely at a classical question from each genre.

The study of explicit constructions is intimately connected to the notion of quasi-random graphs. On the one hand, quasi-random graphs should imitate randomness as closely as they can in some sense: this is essential when the extremal example is random-like. On the other hand, one can also exploit the imperfect randomness of quasi-random graphs to find structures that are not typical at all. The Ramsey- and Turán-type questions we study eloquently demonstrate this duality.

In Turán-type problems the optimal solution is frequently obtained not by random methods, but rather by bumping into a beautiful and unique structure. More often than not, these structures are very much quasi-random but seriously deviate from random in one regard: the one that is the focus of the particular Turán-type problem. One can say that solving a Turán-type problem is like finding a needle in a haystack. The hay represents the random objects taking up almost all the space, while the needle represents the optimal solution we search for: rare, unique, and hard to come by.

Ramsey-type problems are the opposite in some sense: an optimal or nearly optimal solution is obtained by random methods; often *most* of the solutions are provably nearly optimal. In this case however finding an *explicit construction* poses a difficult problem. The deficient randomness of quasi-randomness makes the parameters of available explicit constructions lag significantly behind the random ones. Returning to the folklore metaphor about hay and needles, when investigating a Ramsey-type problem, we can imagine we are a horse standing in front of a huge haystack. We are hungry, we are trying to eat. We reach into the stack, pull out something: it's a needle. We reach again, pull out something: again a needle. Pull again, needle again... Clearly almost anything is edible, yet we are still unable to eat for some mysterious reason. Constructing good Ramsey-graphs explicitly: it is really like finding hay in a haystack.

Notation. For a graph G, $V(G)$ denotes the vertex set and $E(G)$ denotes the edge set. Even though formally $E(G) \subseteq \binom{V(G)}{2}$, we often write xy instead of $\{x, y\}$ for the edge connecting vertices x and y. Sometimes we write $G = (V, E)$ for a graph with vertex set V and edge set E. The number of vertices of G is denoted by $v(G)$, the number of edges by $e(G)$. For subsets $X, Y \subseteq V(G)$, we denote by $E_G(X)$ the set of edges of G with both endpoints in X and write $e_G(X) = |E_G(X)|$ for its cardinality. $e_G(X, Y)$ denotes the number of ordered pairs (x, y) such that $x \in X$, $y \in Y$ and $xy \in E(G)$. If X and Y are disjoint, then $e_G(X, Y)$ is just the number of edges between X and Y. The neighborhood of vertex v is denoted by $N_G(v)$; formally $N_G(v) := \{u \in V(G); uv \in E(G)\}$. The degree of v is denoted by $d_G(v) := |N_G(v)|$. The set of neighbors of v in the subset $X \subseteq V(G)$ is denoted by $N_G(v, X)$ and $d_G(v, X) := |N_G(v, X)|$ denotes the degree of v into X. Often, when the underlying graph is clear from the context, the subscript G is omitted. For a subset $S \subseteq V(G)$ of the vertices, $G[S]$ denotes the subgraph of G induced by the vertex set S.

The asymptotic notation. Most of the time we will be interested in asymptotic behavior of the encountered quantities. Hence, we start by recalling some definitions for abbreviating asymptotics.

Definition: Let $f, g : \mathbb{N} \to \mathbb{R}$ be functions. Then we write

- $f(n) = o(g(n))$ if
$$\frac{f(n)}{g(n)} \longrightarrow 0 \qquad \text{as } n \to \infty \ .$$
Sometimes we also write $f \ll g$.

- $f(n) = O(g(n))$ if
$$\exists\ N \in \mathbb{N}\ \ \forall\ n \geq N\ :\ \left|\frac{f(n)}{g(n)}\right| \leq C$$
for some constant $C \geq 0$.

- $f(n) = \Omega(g(n))$ if $g(n) = O(f(n))$.

- $f(n) = \Theta(g(n))$ if
$$f(n) = O(g(n)) \qquad \text{and} \qquad f(n) = \Omega(g(n)) \; .$$
Sometimes we also write $f \sim g$.

- $f(n) \approx g(n)$ if
$$\frac{f(n)}{g(n)} \longrightarrow 1 \qquad \text{as } n \to \infty \; .$$

- $f(n) \lessapprox g(n)$ if
$$\limsup_{n\to\infty} \frac{f(n)}{g(n)} \leq 1 \qquad \text{as } n \to \infty \; .$$

This being settled, in the following sections we concentrate on describing the type of questions which are important in this course.

1.2 Turán-type problems

Let us consider five isolated vertices. At most how many edges are we able to cram onto them if we are not allowed to create a triangle? In a first try one might end up with a cycle of length five. The graph C_5 has five edges and one cannot insert another edge in it without creating a triangle. In other words, C_5 is a *maximal* triangle-free graph with respect to addition of edges. More trial reveals that C_5 does not represent the *maximum* though: the complete bipartite graph $K_{2,3}$ is triangle-free and has six edges. Shortly we will see that $K_{2,3}$ is optimal: every graph on five vertices and seven edges contains a triangle.

The general question, about the maximum number of edges in an n-vertex triangle-free graph, was posed as a problem in a Dutch journal by W. Mantel in 1907. Correct solutions were submitted by five readers, including of course the poser, hence today the statement is referred to as *Mantel's Theorem.* Mantel showed that the complete bipartite graph $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}$ on n vertices has the largest number of edges among all graphs on n vertices not containing a triangle. Note that

$$e(K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}) = \left\lfloor \frac{n}{2} \right\rfloor \cdot \left\lceil \frac{n}{2} \right\rceil = \left\lfloor \frac{n^2}{4} \right\rfloor \; .$$

Theorem 1.1 (Mantel, 1907) *Any triangle-free graph on n vertices has at most $\lfloor \frac{n^2}{4} \rfloor$ edges.*

Proof. Let $G = (V, E)$ be a triangle-free graph on n vertices. Observe that the neighborhood $N(v) = \{u \in V \; : \; vu \in E\}$ of *any* vertex v must be an independent set. In other words, every edge of G has at least one of its endpoints in $W(v) := V \setminus N(v)$. Hence summing up the degrees of the vertices in $W(v)$ accounts for each edge at least once, twice only those which are entirely contained in $W(v)$. That is, for any vertex $v \in V$ we have that

$$e(G) = e(W(v)) + e(N(v), W(v)) \leq 2e(W(v)) + e(N(v), W(v)) = \sum_{x \in W(v)} d(x)$$

Estimating each degree further with the maximum degree Δ of G, we have

$$e(G) \leq \sum_{x \in W(v)} d(x) \leq \sum_{x \in W(v)} \Delta = (n - d(v)) \cdot \Delta,$$

To obtain the strongest estimate, it obviously makes sense to use a vertex $v \in V$ of degree as large as possible. For a vertex v of degree Δ the above estimate reduces to $e(G) \leq (n - \Delta)\Delta$. This is a quadratic polynomial in Δ that attains its maximum over the integers for $\Delta = \lceil \frac{n}{2} \rceil$. Substituting we find that

$$e(G) \leq \left(n - \left\lceil \frac{n}{2} \right\rceil\right) \cdot \left\lceil \frac{n}{2} \right\rceil = e(K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}) \ .$$

□

Remark: Observe that the proof also establishes the uniqueness of $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}$ as the unique extremal construction. Indeed, the first inequality is an equality if and only if $W(v)$ is an independent set, so the graph is bipartite with $N(v)$ being the other class. The second inequality is tight if and only if all $x \in W(v)$ have maximum degree, so the graph is complete bipartite, while the third inequality is tight if and only if the parts are as equal as possible.

Mantel's Theorem says that in order to "kill" all triangles of the complete graph one must delete roughly half of its edges (recall that the complete graph on n vertices has $\binom{n}{2} \approx n^2/2$ edges). What if, instead of forbidding triangles, one forbids the presence of K_4, or more generally the presence of k-cliques? Do we have to delete significantly fewer edges to achieve that? How many fewer? How does this fraction depend on k? In the early 40's Turán, unaware of the work of Mantel, arrived at this problem and generalized Mantel's Theorem. His motivation was Ramsey theory; this relationship is discussed in the next section.

Consider the following graph on n vertices. Partition the vertex set into $k-1$ parts $V_1, \ldots, V_{k-1}$ of equal size (if $k-1$ does not divide n take lower and upper integer parts, hence making a partition such that the cardinalities of any two parts differ by at most 1). Leave these $k-1$ parts independent in the graph, but introduce an edge between any two vertices from different parts. This is the *complete* $(k-1)$*-partite graph* with parts $V_1, \ldots, V_{k-1}$, which is also called the *Turán graph* and denoted by $T_{n,k-1}$.

The Turán graph (or, for that matter, any graph whose vertex set can be covered by $k-1$ independent sets) is K_k-free: it does not contain a subgraph isomorphic to K_k. Indeed, by the pigeonhole principle, among any k vertices of a $(k-1)$-partite graph there will be two which are in the same part and hence are *not* adjacent. Turán proved that the Turán graph is *the* unique graph containing the most edges among all K_k-free graphs. So to get rid of all k-cliques, one must delete roughly a $\frac{1}{k-1}$-fraction of the edges of K_n.

Theorem 1.2 (Turán, 1941) *Let G be a graph on n vertices not containing a k-clique K_k. Then*

$$e(G) \leq e(T_{n,k-1}) \approx \left(\frac{n}{k-1}\right)^2 \binom{k-1}{2} = \left(1 - \frac{1}{k-1}\right)\frac{n^2}{2} \ .$$

The exact formula for the number of edges of the Turán graph $T_{n,k-1}$ can of course be easily calculated for every n, but the general formula is not so interesting and the asymptotic expression above (which is tight when $k-1$ divides n) says much more.

There are many ways to prove Turán's Theorem; later we will look at two different arguments.

Up to now we have considered graphs that did not contain cliques. One can of course ask questions of this type for any forbidden subgraph H. This is exactly what Turán did in his letters to Erdős and via these questions he practically initiated the field of Extremal Graph Theory. To introduce the corresponding key definition, we say that G is *H-free* and write $H \not\subseteq G$ if G does not have a subgraph isomorphic to H.

Definition: The *Turán number* (or *extremal number*) $ex(n, H)$ of a graph H is defined as

$$ex(n, H) := \max \{e \ : \exists \text{ an } H\text{-free graph } G \text{ with } v(G) = n, e(G) = e\} \ .$$

As we have seen

$$ex(n, K_3) = \left\lfloor \frac{n^2}{4} \right\rfloor \qquad \text{and} \qquad ex(n, K_k) = e(T_{n,k-1}) \ .$$

Turán asked what if, instead of forbidding K_4 (which is the graph of the tetrahedron), we forbid some other platonic polyhedra? How many edges can a graph without an octahedron, or cube, or dodecahedron or icosahedron, have? These problems could seem somewhat arbitrary, but as it turns out they do contain some of the most interesting features of the area and at first sight the results certainly come as a surprise. When one is told that, after more than sixty years, the asymptotic answer is not known for only one of the five platonic polyhedra, one tends to guess that the outstanding question might be about the dodecahedron or the icosahedron, since their graphs are more "complicated". It comes then as a minor shock that "complicatedness", in an everyday-sense, has nothing to do with this problem being hard.

$$\begin{aligned} ex(n, \texttt{tetrahedron}) &= \frac{n^2}{3} + o(n^2) \\ ex(n, \texttt{octahedron}) &= \frac{n^2}{4} + o(n^2) \\ ex(n, \texttt{dodecahedron}) &= \frac{n^2}{4} + o(n^2) \\ ex(n, \texttt{icosahedron}) &= \frac{n^2}{3} + o(n^2) \\ ex(n, \texttt{cube}) &= o(n^2) \ . \end{aligned}$$

The only problem still open is the one about cube-free graphs! This is quite astounding considering the fact that we do know quite precisely at most how many edges a dodecahedron-free graph can contain. The above are all corollaries of the following general theorem of Erdős, Stone, and Simonovits.

Theorem 1.3 (Erdős, Stone, Simonovits) *For any graph H, we have*

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

The special feature of this theorem is that the Turán number of any non-bipartite graph H essentially depends only on the chromatic number of H. For bipartite graphs the leading term disappears and we only know that the Turán number is of subquadratic order.

The Turán number of even the simplest bipartite graphs is often not known and is the subject of vigorous research. For example the order of magnitude of the Turán numbers of the cube Q_3, or the eight-cycle C_8 or the complete bipartite graph $K_{4,4}$ are still a mystery. Here is a sample of what we do know:

$$\begin{aligned} ex(n, C_4) &= \Theta(n^{3/2}) \\ ex(n, C_6) &= \Theta(n^{4/3}) \\ ex(n, C_{10}) &= \Theta(n^{6/5}) \\ ex(n, K_{3,3}) &= \Theta(n^{5/3}) \\ ex(n, K_{4,7}) &= \Theta(n^{7/4}) \end{aligned}$$

$$\begin{aligned} \Omega(n^{3/2}) \leq \ & ex(n, Q_3) \leq O(n^{8/5}) \\ \Omega(n^{8/7}) \leq \ & ex(n, C_8) \leq O(n^{5/4}) \\ \Omega(n^{5/3}) \leq \ & ex(n, K_{4,4}) \leq O(n^{7/4}) \end{aligned}$$

Concluding this section we present some of the classic conjectures about the Turán number of simple bipartite graphs. In a significant portion of the course we will discuss these in depth.

Conjecture 1

$$ex(n, C_{2k}) = \Theta(n^{1+\frac{1}{k}}).$$

This conjecture is verified for $k = 2, 3, 5$. Recall that we do not know the order of magnitude of $ex(n, C_8)$.

Conjecture 2 *Let $s \geq t \geq 2$ be integers.*

$$ex(n, K_{t,s}) = \Theta\left(n^{2-\frac{1}{t}}\right).$$

This conjecture is known to be true for arbitrary s if $t = 2$ or 3, and for $s > (t-1)!$ when $t \geq 4$.

Conjecture 3

$$ex(n, Q_3) = \Theta(n^{8/5})$$

One general observation about Turán-type problems is that extremal graphs in all known cases have a very strong, pronounced structure. Nothing is accidental about them, but at the same time their existence often feels like coincidence. This seems valid not only in the simplest case, where the simple-minded Turán-graph provides the unique optimal structure for K_k-free graphs, but also for the optimal C_{10}-free or $K_{4,7}$-free structures we will encounter later, which are much more complex. As a prototype of this phenomenon, let us mention that for $k \geq 3$ a "typical" n-vertex graph is K_k-free only if its number of its edges is at most $o\left(n^{2-\frac{2}{k-1}}\right)$— much much smaller than the quadratic number of edges of the Turán graph.
Anybody who tried it would agree: constructing an extremal object for a Turán-type problem, as it is said in the Prologue, is like looking for a needle in a haystack. In the first part of these notes we will be after those needles.

Exercise 1.1 *Determine $ex(n, K_{1,4})$ and $ex(n, P_4)$ for every n.*

Exercise 1.2 *Let G be the "diamond" graph on 4 vertices: $V(G) = [4]$ and $E(G) = \{12, 13, 14, 23, 24\}$. Determine $ex(n, G)$ for every n.*

Exercise 1.3 *Show that for any tree T with t edges, $\frac{(t-1)n}{2} - o(n) \leq \mathrm{ex}(n, T) \leq (t-1)n$. In the special case of the star graph, $T = K_{1,t}$, show that the lower bound is correct.*

Exercise 1.4 *Let G be a graph on n vertices with $\left\lfloor \frac{n^2}{4} \right\rfloor + 1$ edges. Show that G contains at least $\left\lfloor \frac{n}{2} \right\rfloor$ triangles.*

Exercise 1.5 *For the octahedron,* $K_{2,2,2}$*, show that for all* $n \geq 3$ *we have the* strict *inequality* $\mathrm{ex}(n, K_{2,2,2}) > e(T_{n,2})$*. How large an* n*-vertex octahedron-free graph can you find?*

Exercise 1.6 *The TV remote of George requires two working batteries to function. Opening the drawer in which he keeps the batteries, he finds eight. He remembers that four of them work and four of them do not, but there is no way of telling them apart without testing them in the remote. How quickly can George guarantee to get his remote working?*

1.3 Ramsey-type problems

A standard combinatorial exercise is the following.

Proposition 1.4 *In a party of six people there always exists three who pairwise know each other or three who pairwise do not know each other.*

Proof. We establish that at least one of the two conclusions necessarily holds. Let Frank be one of the six in the party. By the pigeonhole principle, either Frank knows at least three other people in the party, or there are at least three others whom he does not know. Consider the first case — the second is handled similarly — and let Esther, George, and Paul be three people who know Frank. If two of them, say Esther and George, also know each other, then they together with Frank would be three who pairwise know each other. Otherwise, Esther, George and Paul would pairwise not know each other and we arrive at the second of the possible conclusions. □

Remark: 1. For the argument to make sense and the statement of the claim to be true we must assume that there are no movie-stars in the party, that is, the relation of "knowing" each other must be symmetric.
2. Note that the same claim is not true for five people: just consider the party of five where each person knows two others in a cyclic fashion.

One can translate the statement of the claim into graph theoretic language. To each person we assign a vertex and make two vertices adjacent (by an undirected edge) if the corresponding people know each other. In this context Proposition 1.4 says that in any graph on 6 vertices there are three vertices that are pairwise adjacent or three vertices that are pairwise non-adjacent. In other words, for any graph G on six vertices the clique number or the independence number has to be at least 3. The five-cycle C_5, whose clique number and independence number are both 2, shows that six vertices are necessary.

The generalization of this problem was introduced by the great British logician/philosopher/economist Frank Plumpton Ramsey[1] in the 1920's. The main definition of the

[1] Besides being a mathematician, Ramsey also published fundamental papers in philosophy and economics. He died at the age of 26.

present section is the one of the *Ramsey number*: for $k, l \in \mathbb{N}$ let

$$R(k,l) := \min\{n \ : \ \forall \text{ graph } G \text{ with } v(G) = n,\ \omega(G) \geq k \text{ or } \alpha(G) \geq l\} \ .$$

That is, the Ramsey number $R(k,l)$ is the smallest integer n such that any graph on n vertices contains a clique of size at least k or an independent set of size at least l.

By the above $R(3,3) = 6$; nevertheless, even the fact that $R(k,l)$ is finite is not obvious at first glance. Ramsey showed that the Ramsey numbers $R(k,l)$ indeed exist. It is known that $R(4,4) = 18$; however, the calculation of $R(5,5)$ already exceeds the capacities of not only mathematicians, but even today's computers!

Paul Erdős and George Szekeres arrived at the Ramsey problem independently, motivated by a problem in discrete geometry.[2] They also proved the existence of Ramsey numbers and gave a quantitative upper bound. Their proof is a generalization of the pigeonhole principle argument for Proposition Proposition 1.4.

Theorem 1.5 *For any $k, l \in \mathbb{N}$ we have*

$$R(k,l) \leq \binom{k+l-2}{k-1} \ .$$

Proof. We give a proof by induction on $k + l$: first consider the base cases. Trivially

$$\begin{aligned} R(k,2) = k &\leq \binom{k+2-2}{k-1} = k \\ R(2,l) = l &\leq \binom{2+l-2}{2-1} = l \ . \end{aligned}$$

Now consider a graph G on $v(G) = R(k-1,l) + R(k,l-1)$ vertices. (Note that both $R(k-1,l)$ and $R(k,l-1)$ exist by induction.) We will find in G a clique of size k or an independent set of size l.

Take any vertex $v \in V(G)$, and look at its neighborhood $N(v)$. By the pigeonhole principle, v either has at least $R(k-1,l)$ neighbors or v has at least $R(k,l-1)$ non-neighbors. (The sum of the number of neighbors and the number of non-neighbors is $v(G) - 1$.)

Let us assume first that $|N(v)| \geq R(k-1,l)$. Then either $\alpha(G[N(v)]) \geq l$ and we found an independent set of size l in G, or $\omega(G[N(v)]) \geq k-1$ and this $(k-1)$-clique of $G[N(v)]$ together with v forms a k-clique of G.

An analogous argument can be made in the second case, when the set $U := V(G) \setminus (N(v) \cup \{v\})$ of non-neighbors of v is large, i.e., if $|U| \geq R(k,l-1)$. Then either $\omega(G[U]) \geq k$, in which case we found a k-clique in G, or $\alpha(G[U]) \geq l-1$ and then this independent set of size $l-1$ together with v forms an independent set of size l in G.

Hence we showed the following recursion

$$R(k,l) \leq R(k-1,l) + R(k,l-1).$$

[2]The famous Happy Ending Problem

This of course very much reminds us of the recursion of binomial coefficients and fortunately the initial conditions are adjusted for the formula to be correct.

We have

$$R(k,l) \le R(k-1,l) + R(k,l-1) \le \binom{k+l-3}{k-2} + \binom{k+l-3}{k-1} = \binom{k+l-2}{k-1}.$$

□

Corollary 1.6 *For any $k \in \mathbb{N}$ we have*

$$R(k,k) \le \binom{2k-2}{k-1} < \frac{4^k}{\sqrt{k}} .$$

For some decades a reasonable lower bound eluded even such giants of Hungarian mathematics, as Erdős, Szekeres and Turán. Turán approached the question of the lower bound from a different angle and arrived at a family of problems which are now known as Turán-type problems. In a Ramsey graph we *forbid* the existence of a k-clique, while at the same time we would like to achieve a "small" independence number. Small independence number is naturally associated with having a lot of edges. So a plausible strategy could be to cram as many edges as possible onto the vertices, while still keeping the graph k-clique free and just hope for the best in terms of the independence number.

This is exactly the question we talked about in the previous section. Turán's optimal construction for the problem, the Turán-graph, is relevant to our discussion here as well, because it provides a nonlinear lower bound for $R(k,k)$. Obviously both the clique number and the independence number of the Turán graph $T_{(k-1)^2,k-1}$ on $(k-1)^2$ vertices are $(k-1)$; thus

$$R(k,k) > (k-1)^2 .$$

Turán himself believed for some time that the Turán-graph does provide the extremal construction for the Ramsey problem as well and tried to show that any graph on N vertices either contains a clique or an independent set of size $\sqrt{N}$. He told his conjecture to Erdős who disproved it in a very strong sense.

Let us fix an integer $N \in \mathbb{N}$. The strategy of Erdős is remarkably simple, yet was quite unusual at the time: count first the total number of graphs on N vertices, then compare it to the number of those N-vertex graphs whose largest clique or largest independent set is of order at least k, and hope that the former is at least one larger than the latter.

Clearly, there are $2^{\binom{N}{2}}$ (labeled) graphs on N vertices, since for every one of the $\binom{N}{2}$ pairs of vertices there are two choices: the pair is either an edge or not.

Now we want to upper bound the number of graphs G with $v(G) = N$ such that $\omega(G) \ge k$ or $\alpha(G) \ge k$. In order to obtain a graph with a clique of size at least k, one first has to pick k vertices, this can be done in $\binom{N}{k}$ ways, and introduce all the $\binom{k}{2}$ edges between any pair of them. Then one does not have to care about the status of the edges between the remaining $\binom{N}{2} - \binom{k}{2}$ pairs, so there are two choices for each of them: edge or non-edge. Hence we have

$$2^{\binom{N}{2}-\binom{k}{2}} \cdot \binom{N}{k},$$

as a crude overcount for the number of N-vertex graphs containing a k-clique.

Similarly, one can obtain the same upper bound for the number of graphs containing an independent set of size at least k.

Hence, the number of graphs G with $v(G) = N$ such that $\omega(G) \geq k$ or $\alpha(G) \geq k$ is at most

$$2 \cdot 2^{\binom{N}{2}-\binom{k}{2}} \cdot \binom{N}{k} .$$

If this number is less than the total number of graphs, then we will have convinced ourselves about the existence of a graph G on N vertices with $\omega(G) < k$ and $\alpha(G) < k$, implying a lower bound on $R(k,k)$. Hence we need an integer N, such that

$$2 \cdot 2^{\binom{N}{2}-\binom{k}{2}} \cdot \binom{N}{k}, \quad < \quad 2^{\binom{N}{2}}$$

$$\text{that is, } 2 \cdot \binom{N}{k} \quad < \quad 2^{\binom{k}{2}} .$$

Using the very rough estimate $2 \cdot \binom{N}{k} \leq N(N-1)\dots(N-k+1) < N^k$, we see that an integer N satisfying

$$N^k \leq 2^{\binom{k}{2}} = 2^{\frac{k(k-1)}{2}} \qquad \Leftrightarrow \qquad N \leq 2^{\frac{k-1}{2}} = \sqrt{2}^{k-1}$$

would work. Therefore, we showed that there is at least one graph on $N = \lfloor \sqrt{2}^{k-1} \rfloor$ vertices which contains neither a clique nor an independent set of order k, implying an exponential lower bound on the Ramsey number. Both of the ancient bounds,

$$\sqrt{2}^{k-1} < R(k,k) < 4^k ,$$

we derived here are still essentially the best known today! Most recently, Conlon [?] showed that $R(k,k)$ is smaller than 4^k by an *arbitrary* polynomial factor. Still, nobody can prove that $R(k,k) \leq 3.99^k$, say. Using the estimate $\binom{N}{k} \leq \left(\frac{Ne}{k}\right)^k$ in the lower bound proof, one obtains $\left(\frac{1}{e\sqrt{2}} + o(1)\right) k\sqrt{2}^k < R(k,k)$. There is no substantial improvement in the last sixty years here either. All what happened is that the constant factor $\sqrt{2}$ has moved up to the numerator some forty years ago [?], but it is unclear whether a lower bound of, say, $k^{1.1}\sqrt{2}^k$ would hold.

It is one of the most notorious open problem of combinatorics to determine the value of $\lim_{k\to\infty} \sqrt[k]{R(k,k)}$, provided this limit exists. Erdős offered prize money even for a proof showing the existence of the limit.

1.4 Basics of the probabilistic method

The above counting argument of Erdős is considered today as the introduction of a general proof technique in combinatorics, called the *probabilistic method.* The method is used to establish the *existence* of a certain type of object without actually constructing it. This approach could be particularly helpful when other, deterministic attempts to construct an object with desired properties prove to be unfruitful.

The simple, yet revolutionary idea is that one constructs an appropriate *probability space of objects*, rather than a particular instance of the desired object, such that in the probability space the desired objects occur with nonzero probability. The point in the choice of the probability space is that the probability of the desired objects should be *provably* nonzero. Miraculously, it is often the case that while concentrating on satisfying the desired property of one particular object proves to be hopelessly hard, proving that an overwhelming majority of objects have the desired property is relatively easy.

While other instances of this existence proof technique appeared earlier in several branches of mathematics, no one before used it as systematically as Erdős, who almost single-handedly developed it into a *method.*

For the moment we will be content with the simplest of conclusions.

- If the probability of objects having some property P is not 0, then *there exists* an object with property P.

Often it is more convenient to talk in terms of expected value.

- If the expected value of a random variable is C then *there exists* an object for which the variable's value is *at least* C and *there exists* an object for which the variable's value is *at most* C.

1.4.1 The union bound

The union bound: Let $A_1, \ldots, A_i, \ldots$ be any (finite or infinite) set of events. Then

$$\Pr\left[\bigcup_{i=1}^{\infty} A_i\right] \leq \sum_{i=1}^{\infty} \Pr[A_i].$$

In order to see the union bound at work let us reformulate the above Ramsey argument in the language of probability. This makes much sense conceptually, because thinking in terms of probability theory allows one to later apply the full machinery of the field. Even though every single fact of discrete probability can in principle be expressed just as counting, with some of the more tricky ones this could be extremely cumbersome, if not close to impossible to carry out. The more general point of view of probability has proven to be more and more fruitful ever since the groundbreaking counting result of Erdős.

First we define the "appropriate" probability space: the random graph $G(n, 1/2)$ with edge probability $1/2$. For each of the $\binom{n}{2}$ pairs of vertices we flip a fair coin, and we put the corresponding edge in the graph if the result of the coin flip is "head". Note that in this probability space all the $2^{\binom{n}{2}}$ labeled graphs on vertex set $[n] = \{1, \ldots, n\}$ occur with the same probability.

The calculations in the previous subsection can be used to estimate the probability that $G(n, 1/2)$ has clique number at least k.

Fix a subset $S \subseteq [n]$ of size k. Let A_S be the event that the vertices of S form a clique in $G(n, 1/2)$. For A_S to happen, all the $\binom{k}{2}$ coin flips corresponding to the pairs of S must turn out to be heads, while all other coin flips can be arbitrary. Hence

$$\Pr[A_S] = \frac{1}{2^{\binom{k}{2}}}.$$

The probability that the clique number is at least k is the union of A_S over all $S \subseteq [n], |S| = k$. We can then use the union bound,

$$\Pr[\omega(G(n,1/2)) \geq k] = \Pr\left[\bigcup_{S \in \binom{[n]}{k}} A_S\right] \leq \sum_{S \in \binom{[n]}{k}} \Pr[A_S] = \frac{\binom{n}{k}}{2^{\binom{k}{2}}} \leq \left(\frac{en\sqrt{2}}{k\sqrt{2}^k}\right)^k.$$

One can bound $\Pr[\alpha(G(n, 1/2)) \geq k]$ similarly. Hence, say for $n = \sqrt{2}^k$, the union bound gives

$$\Pr[\omega(G(n,1/2)) \geq k \text{ or } \alpha(G(n,1/2)) \geq k] \;\; \leq \;\; 2 \cdot \left(\frac{e\sqrt{2}}{k}\right)^k.$$

For large enough k this probability is strictly less than 1 and thus the *existence* of a k-Ramsey graph on $\sqrt{2}^k$ vertices is proved.

But we get more. Let us play with the numbers a bit. By the above calculation, for every $k \geq 12$, at least 99.999 percent of all graphs on $\sqrt{2}^k$ vertices have both their clique number and their independence number strictly less than k. So we **know** that almost every graph is "good" for us in this sense. But should we just aim to hold one of these good Ramsey-graph in our hands ... we are going to see how elusive they are! Returning to the haystack parabole of the introduction: in the second part of these notes we will play the role of the hungry horse and try to finally pull out a hay instead of a needle. We will fail, but experience a lot of nice mathematics in the process. We will see that the situation is the exact opposite of the one we encountered with Turán-type problems. Optimal structures of Turán-type problem are very non-random; to find them one tries to distance oneself from random. Optimal structures of Ramsey-type problems are random-like in some sense and our explicit constructions will try to imitate some of that randomness.

1.4.2 Linearity of expectation

Linearity of Expectation Let $X_1, \ldots, X_n$ be random variables. Then

$$\mathbf{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbf{E}[X_i] .$$

To demonstrate this method we prove the asymptotic version of Theorem 1.2 (Turán's Theorem). That is, for any graph G with clique number $\omega(G) < k$, we must estimate $e(G)$ from above. We turn this around and rather for any given edge number e we estimate the clique number from below. First we use the linearity of expectation to derive a *lower* bound on the clique number of *any* graph in terms of the degrees of its vertices. Since the number of edges is linked to the degrees via $e(G) = \frac{1}{2}\sum d(v)$, this will also imply a lower bound in terms of the number of edges using a standard convexity argument.

Theorem 1.7 *For any graph G, we have*

$$\omega(G) \geq \sum_{v \in V(G)} \frac{1}{n - d(v)} \geq \frac{n^2}{n^2 - 2e(G)}.$$

Proof. Let $V = \{v_1, \ldots, v_n\}$ be the vertex set of G. For each permutation π of $[n]$, let us design a clique $C(\pi)$ of G in the following fashion: a vertex $v_{\pi(i)} \in C(\pi)$ if and only if $v_{\pi(i)}$ is adjacent to $v_{\pi(j)}$ for every $j < i$. For example, $v_{\pi(1)}$ is always in $C(\pi)$, but $v_{\pi(2)}$ is in there only if it is adjacent to $v_{\pi(1)}$. By definition, $C(\pi)$ is a clique for every permutation π.

The idea is to calculate the expected value of the cardinality of $C(\pi)$ if we select a permutation π uniformly at random among all permutations of $[n]$. While calculating the expectation of $|C(\pi)|$ looks impossibly complicated at first sight, as is often the case, if we divide the problem into subproblems, then the linearity of expectation can come to the rescue.

Let X_i be the characteristic random variable of the event that $v_i \in C(\pi)$. (That is, $X_i = 1$ if $v_i \in C(\pi)$ and $X_i = 0$ if $v_i \notin C(\pi)$.) Then $|C(\pi)| = \sum_{i=1} X_i$.

We are interested in the probability that a vertex v_i is contained in $C(\pi)$. This event occurs if and only if all its non-neighbors come only *after* v_i in the permutation π. In other words, if we restrict π to v_i and its $n - 1 - d(v_i)$ non-neighbors, then v_i is the first element. This has probability $\frac{1}{n-d(v_i)}$, hence,

$$\mathbf{E}[X_i] = \Pr[v_i \in C(\pi)] = \frac{1}{n - d(v_i)}.$$

By the linearity of expectation

$$\mathbf{E}[|C(\pi)|] = \sum_{i=1}^{n} \mathbf{E}[X_i] = \sum_{i=1}^{n} \frac{1}{n - d(v_i)}.$$

Since the expected value of this random variable is $\sum_{i=1}^{n} \frac{1}{n-d(v_i)}$, there must exist a *particular* permutation σ (maybe more than one) for which $|C(\sigma)|$ is at least $\sum_{i=1}^{n} \frac{1}{n-d(v_i)}$. As $C(\sigma)$ is a clique, the first inequality in the statement is proved.

For the second inequality note that the function $x \to \frac{1}{x}$ is convex on the positive reals, so the average value of the function over a set of positive real numbers is larger than the value of the function at the average of that set (by Jensen's inequality):

$$\sum_{i=1}^{n} \frac{1}{n-d(v_i)} \geq n \frac{1}{n-\bar{d}(G)} = \frac{n^2}{n^2-2e(G)},$$

where $\bar{d}(G) = \frac{2e(G)}{n}$ is the average degree of G. □

Now Turán's Theorem follows immediately.

Proof. (*of Theorem 1.2*) Let G be K_k-free graph, that is, $\omega(G) \leq k-1$. Hence by the previous theorem $k-1 \geq \frac{n^2}{n^2-2e(G)}$, which rearranges to give

$$e(G) \leq \left(1 - \frac{1}{k-1}\right) \frac{n^2}{2},$$

which proves Theorem 1.2. □

Remark: 1. It is worthwhile to ponder for a moment why this strange proof of Turán's Theorem can work. We bound the clique number as the expected number of the random variable $|C(\pi)|$. Since this estimate is eventually tight, an overwhelming majority of $C(\pi)$ have to be maximum cliques. This in fact is true in the extremal extremal case, i.e., when G is the Turán graph: every single permutation σ gives a maximum clique, as $C(\sigma)$ contains the first vertex (according to π) from each of the $k-1$ vertex classes of the Turán graph.

2. The way $C(\pi)$ is created is in some kind of greedy way, but maybe not the greediest possible. Let us again imagine a process where the vertices come one by one according to the random permutation π and we decide online whether we put them in the set $C(\pi)$ (and not change our decision later). The greediest of approaches would put a vertex into $C(\pi)$ if it is adjacent to every vertex that is already put in $C(\pi)$. The $C(\pi)$ produced this way will be a clique alright, but the problem is with the analysis of its size. This process creates a complex system of dependencies between the decisions and the probability of a particular vertex v_i being in $C(\pi)$ might potentially depend on the structure of the whole graph (and not only the degree of v_i). For this reason we chose to strengthen the requirement for membership in $C(\pi)$: not only the current elements of $C(\pi)$, but *all* preceding vertices had to be adjacent to v_i. Observe that for the Turán graph the two algorithms produce the same $C(\pi)$.