Extremal graphs and the Erdős-Stone-Simonovits theorems

Cas Widdershoven

June 23, 2016

Abstract

The Erdős-Stone-Simonovits theorems say that any graph F can be embedded in any graph G when G has a sufficient amount of vertices and edges. We will discuss some extremal problems and look at the Turán numbers of triangles, quadrilaterals and some other graphs leading up to the proof of the Erdős-Stone-Simonovits theorems.

1 Introduction

In this paper we will look at the Erdős-Stone-Simonovits theorems. We will first look at some extremal problems, then we will move to extremal graph theory and in the end we will prove the Erdős-Stone-Simonovits theorems.

In section 2 we will state some preliminaries. These consist primarily of graph theoretic definitions and we will end with some notations for asymptotic behaviour. This section can be skipped by readers with a graph theoretic background. In section 3 we will look at some extremal problems. We will look at planarity of graphs, colourability of planar graphs and we will end with Van der Waerden's theorem. In section 4 we will be looking specifically at extremal graph theory. We will discuss some well-known results from the extremal graph theory starting at Mantel's theorem and Erdős-Rényi-Sós and leading to Kővári-Sós-Túran. These sections lead up to the Erdős-Stone and Erdős-Stone-Simonovits theorems in section 5. Here we will look at those theorems and prove them using the tools developed in earlier sections.

In this paper, [1] is heavily used. The main structure and definitions, unless stated otherwise, come from [1]. For Van der Waerden's theorem in section 3, [3] is used.

2 Preliminaries

We will be looking at simple graphs. First, we will state some definitions. Any proofs in this section are inspired by [1].

Definition 1. *A (simple, unordered) graph (V, E) is a set V of vertices and a set E of unordered pairs (v, w), $v, w \in V$.*

Definition 2. *The* neighbourhood *of a vertex v in a graph (V, E), denoted as $\mathcal{N}(v)$ is the set of all vertices u such that $(v, u) \in E$.*

Definition 3. *The* degree *of a vertex v in a graph (V, E), denoted as $d(v)$, is the amount of edges containing this vertex, id est, the cardinality of the neighbourhood of v.*

Theorem 1. *The sum of the degrees of all vertices, $\sum_{v \in V} d(v)$, is equal to $2 * |E|$.*

Proof. Any edge increases the degree of two vertices by one, or, equivalently, increases the sum of all degrees by two. $\square$

Definition 4. *A* walk *is a sequence of vertices $v_0..v_k$ such that for any $i < k$, (v_i, v_{i+1}) is an edge of the graph.*

Definition 5. *A* path *is a walk $v_0..v_k$ such that $v_i \neq v_j$ for any $i, j \in \{1..k\}, i \neq j$.*

Definition 6. *A* cycle *is a walk $v_0...v_k$ such that $v_0 = v_k$ and for any $i, j, i \neq j$ and i and j are not equal to 0 or k, $v_i \neq v_j$ and $v_0 \neq v_i$.*

Definition 7. *A* subgraph *of a graph $G = (V, E)$ is a graph (V', E') where $V' \subset V$ and $E' \subset E \cap \binom{V'}{2}$. The* induced subgraph *on $V' \subset V$ is the graph $(V', E \cap \binom{V'}{2})$.*

Definition 8. *A graph is* connected *if for any two vertices v, w there exists a path in the graph that starts with v and ends in w.*

Definition 9. *A* connected component *of a graph G is a maximal connected subgraph of G.*

Definition 10. *A* tree *is a connected, acyclic graph, id est, a connected graph without any cycles.*

Theorem 2. *A tree on n vertices has $n - 1$ edges.*

Proof. We will use induction on n. Let $T = (V, E)$ be a tree on $n = |V|$ vertices. For $n = 1$, this is obviously true.

Now pick any $r \in V$. Let $e_1, e_2, ..., e_k$ be the set of neighbours of r and let V_i be the sets of vertices such that the path from r to any $u_i \in V_i$ passes through e_i. Observe that these sets are disjoint, since if $u \in V_i \cap V_j$ (where $i \neq j$), then there is a path through e_i from r to u and a path from u to r through e_j and therefore there must exist a cycle in T, which is a contradiction with the fact that T is a tree.

Observe that $T|_{V_i}$ is also a tree and thus, according to our induction hypothesis, if we let n_i to be $|V_i|$, $T|_{V_i}$ has $n_i - 1$ edges. Now we see that the amount of vertices in this tree is $1 + \sum_i n_i$ and the amount of edges in this tree is $k + \sum_i (n_i - 1) = \sum_i n_i$. Thus, the amount of edges in T is $n - 1$. $\square$

Definition 11. *A* complete graph *on a vertex set V is the graph $(V, \binom{V}{2})$, the graph where any two vertices in V are joined by an edge. It is denoted as K_V (or K_n for a complete graph on n vertices, where the vertices don't matter).*

Definition 12. *A* discrete graph *is a graph without edges,* $G = (V, \emptyset)$.

Definition 13. *A* bipartite graph *is a graph* $(V_1 \cup V_2, E)$ *such that no two edges in* V_1 *or* V_2 *are joined by an edge, the only edges are in between* V_1 *and* V_2*. It is denoted as* K_{V_1,V_2} *(or, similarly,* $K_{m,n}$ *if the vertices don't matter).*

Definition 14. *A* complete bipartite graph *is a graph* $(V_1 \cup V_2, E)$ *such that it is bipartite and for any* $v \in V_1, w \in V_2$, $(v, w) \in E$.

Finally we will need to have some way to describe asymptotic behaviour for functions. These definitions come from [2]

Definition 15. *A function* f *is said to be* little-o of g *if there exists a constant* $n_0 \in \mathbb{R}$ *such that for all constants* $c \in \mathbb{R}$ *and all* $n \in \mathbb{R}, n > n_0$

$$|f(n)| \leq |c \cdot g(n)|.$$

This is denoted as $f(n) = o(g(n))$.

Definition 16. *A function* f *is said to be* big-o of g *if there exist constants* $c, n_0 \in \mathbb{R}$ *such that for all* $n \in \mathbb{R}, n > n_0$

$$|f(n)| \leq |c \cdot g(n)|.$$

This is denoted as $f(n) = O(g(n))$.

Definition 17. *A function is said to be* theta of g *if there exist constants* c_-, c_+ *and* $n_0 \in \mathbb{R}$ *such that for all* $n \in \mathbb{R}, n > n_0$

$$|c_- \cdot g(n)| \leq |f(n)| \leq |c_+ \cdot g(n)|.$$

This is denoted as $f(n) = \Theta(g(n))$.

3 Some extremal problems

In this section we will look at some extremal problems in order to get some feeling of what they are and how you can solve them. We will first look at planar graphs, then we will look at the chromatic number of planar graphs and we will end with Van der Waerden's theorem.

In order to look at planar graphs, we will first need to deduce a simple theorem called *Euler's formula.* This formula fixes the relative amounts of vertices, edges and *faces* in a planar graph.

Definition 18. *The* faces *of a planar graph are the connected components of the plane after you remove the curves and points that represent the graph from this plane.*

.

Theorem 3 (Euler's formula)**.** *Let $G = (V, E)$ be a connected planar graph and let F be the faces of this graph, then*

$$|V| - |E| + |F| = 2 \tag{1}$$

Proof. We will prove this using induction. If G is acyclic, then obviously, $|F| = 1$ and $|V| - |E| = 1$, so the equation is true.

Now, suppose G has a cycle. By removing one edge from this cycle, we merge the two opposite faces, so the new amount of faces $|F'|$ is $|F| - 1$, and thus $|V| - (|E| - 1) + |F'| = |V| - (|E| - 1) + (|F| - 1) = |V| - |E| + |F| - 1 + 1$ and induction thus tells us that $|V| - |E| + |F| = 2$. □

Theorem 4. *For a connected planar graph $G = (V, E)$, where $|V| > 2$, the amount of edges is bounded by $|E| \leq 3|V| - 6$.*

Proof. This can easily be checked if $|E| \leq 3$. Now observe that if $|E| \geq 3$, every face has at least three edges bounding it. Moreover, every edge bounds at most two faces, so this gives us the inequality $3|F| \geq 2|E|$. Now we can use Eulers formula and substitute to see that $|E| \leq 3|V| - 6$:

$$\begin{aligned} |V| - |E| + |F| &= 2 \\ 3|V| - 3|E| + 3|F| &= 6 \\ -3|E| + 2|E| &\geq 6 - 3|V| \\ |E| &\leq 3|V| - 6. \end{aligned}$$

□

This of course implies that the average degree of the vertices in any planar graph is strictly less than 6. We can actually improve this bound by requiring the graph to be free of larger cycles, by generalising the proof.

Theorem 5. *For a connected planar graph $G = (V, E)$, where $|E| > g$ and G has no cycles of length $< g$ for some $g \in \mathbb{N}$, the amount of edges is bound by*

$$|E| \leq \frac{g}{g-2}(|V| - 2) \tag{2}$$

Proof. Following the same way of thinking from general case, we can come to the conclusion that each face is bounded by at least g edges, which gives us the inequality $g|F| \geq 2|E|$. We can substitute this in Euler's formula as follows:

$$\begin{aligned} |V| - |E| + |F| &= 2 \\ g|V| - g|E| + g|F| &= g * 2 \\ -g|E| + g|F| &= g * 2 - g|V| \\ (-g + 2) * |E| &\leq g * (2 - |V|) \\ |E| &\leq \frac{g}{g-2}(|V| - 2) \end{aligned}$$

□

According to these criteria for planar graphs, we can find a couple of non-planar graphs:

- The complete graph K_5, since it has 5 vertices and 10 edges and thus $|E| \not\leq 3|V|-6$.
- The complete bipartite graph $K_{3,3}$, since it has no cycles of length 3 and thus we can apply theorem 5 with $g = 4$ and observe that $|E| \not\leq 2(|V| - 2)$.
- Of course, any graph containing a nonplanar graph as its subgraph.

Now we will look at the colouring of planar graphs.

Definition 19. *A* colouring *on r colours of a graph $G = (V, E)$ is a function $c : V \to \{1..r\}$. This colouring is called* proper *if no two adjacent vertices are of the same colour, or, more formally, if $(u, v) \in E$ implies that $c(u) \neq c(v)$ for any two $u, v \in V$.*

Definition 20. *The* chromatic number *of a graph G is the smallest r such that there exists a proper colouring of G on r colours.*

An important observation to make, is that any planar graph can be properly coloured with six colours.

Theorem 6. *The chromatic number of any planar graph is at most 6.*

Proof. Let $G = (V, E)$ be any planar graph. By theorem 4 we see that there exists a vertex v with degree at most 5. By removing this vertex we get another planar graph $G|_{V\setminus\{v\}}$ and by induction this graph can also be coloured with 6 colours. Since the degree of v was at most 5, at least one of the colours in $\{1..6\}$ has not been used by any of the neighbours of v, and thus we can colour v with this colour. □

In fact, we can prove something even stronger, namely that the chromatic number of any planar graph is at most 5.

Theorem 7. *The chromatic number of any planar graph is at most 5.*

Proof. Let G be any planar graph. If G has a vertex of degree at most 4, we can use the reasoning of theorem 6, so let's assume we have a vertex v with a degree of 5.

By induction, we can colour $G|_{V\setminus\{v\}}$ with at most 5 colours. If there is one of the five colours unused in the neighbours of v, we can simply colour v with that colour, so let's assume each neighbour of v has a different colour.

Now suppose we have an embedding of G into $\mathbb{R}^2$. Name the neighbours of v as follows: n_1 is any vertex, n_2 is the first vertex you encounter when you look counter-clockwise around v from n_1, n_3 is the second, etcetera. Look at the induced subgraph of G on the vertices coloured by $c(n_1)$ or $c(n_3)$ (excluding v). Now we can distinguish two cases:

- **Case 1:** There does not exist a path from n_1 to n_3. In this case we can simply swap $c(n_1)$ and $c(n_3)$ in the connected component containing n_3 in this induced subgraph and colour v as $c(n_3)$ to obtain a proper colouring on 5 colours.

- **Case 2:** There does exist a path from n_1 to n_3. In this case this path together with v separates n_2 from n_4 and therefore we can swap $c(n_2)$ and $c(n_4)$ on the connected component of n_2 in the subgraph induced on the vertices with the colours $c(n_2)$ and $c(n_4)$ and colour v as $c(n_2)$ in order to obtain a proper colouring on 5 colours.

□

We will end this section by proving Van der Waarden's theorem. This theorem isn't about graph theory, but it shows some of the techniques we shall see in extremal graph theory. This proof is based on [3]

Before stating the theorem, we need to know the definitions for a *colouring* of numbers and for an *arithmetic progression.*

Definition 21. *A* colouring *of the numbers of* 1 *up to* N *on* r *colours is a function* $f : \{1..N\} \to \{1..r\}$.

Definition 22. *An* arithmetic progression $\mathrm{AP}(\alpha, \delta, k)$ *is a set of numbers* $\{\alpha, \alpha+\delta, \alpha+2*\delta, ..., \alpha+(k-1)\delta\}$.

Now we are ready to state and prove the theorem.

Theorem 8 (Van der Waerden's theorem)**.** *For any* $r, k \in \mathbb{N}$ *there exists an* $N \in \mathbb{N}$ *such that any colouring of* N *numbers on* r *colours contains a monochromatic arithmetic progression, id est, there is an* $\mathrm{AP}(\alpha, \delta, k)$ *such that all of the elements in this progression are coloured the same.*

Proof. We will use induction on k. Let us call the number for which there must exist a monochromatic arithmetic progression of length k in any colouring on r colours $\omega(r, k)$.

It is easy to see that $\omega(r, 1)$ is finite; this is simply asking for one number with a colour, so $\omega(r, 1) = 1$. We can also easily see that $\omega(r, 2) = r + 1$, using the pigeonhole principle. We will now show that $\omega(r, k)$ is finite given that $\omega(r, k-1)$ is finite.

We will first need some more definitions.

Definition 23. *The* focus *of an arithmetic progression* $\mathrm{AP}(\alpha, \delta, k)$ *is the point* $\alpha + k\delta$.

Definition 24. *A collection* $\mathcal{AP}$ *of arithmetic progressions is called* focused *if every* $\mathrm{AP} \in \mathcal{AP}$ *has the same focus*

Definition 25. *A collection* $\mathcal{AP}$ *of arithmetic progressions is called* colour-focused *if:*

- *the collection is focused,*
- *every* $\mathrm{AP} \in \mathcal{AP}$ *is monochromatic and*
- *no two progressions in* $\mathcal{AP}$ *have the same colour.*

We will proof the theorem using the following lemma:

Lemma 1. *For all $s \leq r$ there exists an n such that for any colouring on n numbers with r colours there exists either a monochromatic arithmetic progression of length k or a colour-focused collection of arithmetic progressions of size s.*

It is easy to see that this lemma implies the theorem; it says that there exists either the k-term monochromatic progression that we want or a colour-focused collection of size r of $k-1$-term arithmetic progressions, which we can easily extend to a k-term arithmetic progression by looking at its focus (which has to be one of the r colours). All that remains, is proving the lemma.

Proof of the lemma. We will use induction on s to prove this lemma. Suppose $s = 1$, then all the lemma says is that there exists an n such that for any colouring on n numbers with r colours there exists either a monochromatic arithmetic progression of length k or one of length $k-1$. Since we had assumed that $\omega(r, k-1)$ was finite, this proves the base case.

Suppose, then, that $s \geq 2$, and suppose that n_{s-1} suffices for $s-1$. Then we can find a new arithmetic progression by looking at colouring blocks of size $2n_{s-1}$, intervals from d up to $d + 2_n s - 1$.

Naturally, we can colour a block of size $2n_{s-1}$ with r colours in $r^{2n_{s-1}}$ ways, namely by picking one out of r colours for each number in our block. We also know that $\omega(r', k-1)$ is finite for any r', so if we look at the numbers up to $2n_{s-1}\omega(r^{2n_{s-1}}, k-1)$ we know that there must exist an arithmetic progression of blocks of size $2n_{s-1}$ of length $k-1$ that are coloured exactly the same, id est, we know that there must exist $k-1$ intervals of 1 up to $2n_{s-1}$ that are coloured exactly the same and that are evenly spaced. We call the elements of this arithmetic progression B_i for $1 \leq i \leq k-1$.

If we now look at a block B_i we see that it consists of $2n_{s-1}$ elements. Therefore, we know that this block must contain a colour-focused collection of $s-1$ arithmetic progressions of length $k-1$. Moreover, we know that it also contains its focus, since all elements of that collection are in the interval of 1 up to n_{s-1}. If this focus has the same colour as any of the $s-1$ elements in the collection, then we have found a k-term arithmetic progression and we are done. If this focus has another colour, however, we can see that this focus is at the same location in all of the blocks B_i. This means that we have found a new arithmetic progression consisting of the focuses of the collections in B_i. Combining that with the i'th element of AP_i for each arithmetic progression $\mathrm{AP}_i \subset B_i$ in the colour-focused collection we had found, we have found a new colour-focused collection of arithmetic progressions of size s and this proves the lemma. □

□

4 Extremal graph theory

We will now be looking at extremal graph theory. The main question in extremal graph theory is "how many edges must a graph have in order to have some local property", i.e. some specific subgraph. Extremal graph theory stems from the early 1940's, when

Turán proved his theorem, which we will be looking at later. First, we will be looking at some simpler proofs, such as Mantel's theorem.

Theorem 9 (Mantel's theorem)**.** *Let $G = (V, E)$ be a graph on n vertices. Then, if $|E| > \lfloor n/2 \rfloor \lceil n/2 \rceil$, G contains a triangle*

Proof. Suppose we have a triangle-free graph $G = (V, E)$. We will modify this graph such that it contains at least as many edges (and it still is triangle-free). This way, we will show that the maximum amount of edges in a triangle-free graph is $\lfloor n/2 \rfloor \lceil n/2 \rceil$.

Let $v \in V$ be a vertex in G with maximum degree, and let S be the set of all neighbours of v. Then we define (V, E') as the complete bipartite graph $K_{V \setminus S, S}$. Naturally, this graph is triangle-free. We will now check that $|E'| \geq |E|$.

Suppose we have the vertex w in $K_{V \setminus S, S}$. Then this vertex is an element of $V \setminus S$ or it is an element of S. Suppose it is a vertex in $V \setminus S$, then it has $|S|$ neighbours, and since S is the set of neighbours of a vertex with the highest degree in G, the degree of w in the new graph must be higher than or equal to the degree of w in G. Suppose, then, that w is an element of S. Then, it was a neighbour of v in G, and since G was triangle-free, it could not have been a neighbour of any other vertex in S. The maximum degree this vertex could have in G is therefore $|V \setminus S|$, which is exactly the degree of the vertex in $K_{V \setminus S, S}$. Thus, $K_{V \setminus S, S}$ is a triangle-free graph with at least as many edges as G and any maximal triangle-free graph must be a $K_{V \setminus S, S}$ for some $S \subset V$. The amount of edges in such a graph is maximal if $|S| = \lfloor |V|/2 \rfloor$, which gives us the bound of $\lfloor n/2 \rfloor \lceil n/2 \rceil$. □

We can generalise this theorem by looking at arbitrary complete graphs. This gives us Turán's theorem.

Theorem 10 (Turán's theorem)**.** *Let $G = (V, E)$ be a graph on n vertices, then G contains a K_{r+1} whenever*

$$|E| > \frac{1}{2}\left(n^2 - (n \bmod r)\lceil n/r \rceil^2 - (r - (n \bmod r))\lfloor n/r \rfloor^2\right).$$

If $n = cr$ for some $r \in \mathbb{N}$, then obviously this equals $\frac{r-1}{r}\binom{n^2}{2}$

Proof. We will use a generalised version of the proof used for Mantel's theorem. This means we will first take $G = (V, E)$ to be a K_{r+1}-free graph which we will modify so that it has at least as many edges and it still is K_{r+1}-free. We will use this to prove the bound of $\frac{1}{2}\left(n^2 - (n \bmod r)\lceil n/r \rceil^2 - (r - (n \bmod r))\lfloor n/r \rfloor^2\right) \leq \frac{r-1}{r}\binom{n^2}{2}$. We will use induction on r to construct a K_{r+1}-free graph of maximum size. Observe that, if $r = 2$, then only the discrete graphs have the right amount of edges, and indeed these are the only ones without K_2 (an edge).

First, we see that the complete r-partite graph on n vertices with parts of as equal size as possible has exactly $\frac{1}{2}\left(n^2 - (n \bmod r)\lceil n/r \rceil^2 - (r - (n \bmod r))\lfloor n/r \rfloor^2\right)$ vertices, so this bound is tight. This bound is seen by seeing that any of the n vertices can be connected to any of the vertices that are not in the same part, and there are $n \bmod r$

parts of size $\lceil n/r \rceil$ and $r - (n \bmod r)$ parts of size $\lfloor n/r \rfloor$. This gives us $\frac{n^2}{2}$ minus the edges that are missing in each discrete part, or $\frac{n^2}{2} - \frac{(n \bmod r)\lceil n/r \rceil^2 + (r-(n \bmod r))\lfloor n/r \rfloor^2}{2}$ edges in total.

Let $v \in V$ again be a vertex in G with maximum degree. Now, the induced subgraph on the set S of neighbours of v must be K_r-free or else G would have contained a K_{r+1}.

Now look at the graph $G' = (V, E')$ defined as follows. The induction hypothesis tells us that there is a K_r-free graph on S with at least as many edges as the induced subgraph on S in G. Take $G'|_S$ as this graph on S with at least as many edges as $G|_S$. Next, add all edges between $V \setminus S$ and S and make $G'|_{V \setminus S}$ discrete.

We claim that G' has at least as many edges as G does, and is still K_{r+1}-free. First, we observe that G' is K_{r+1}-free, since $G'|_S$ is K_r-free and $G'|_{V \setminus S}$ is discrete.

Now, we show that G' has at least as many edges as G. Induction tells us that $G'|_S$ has at least as many edges as $G|_S$, so it suffices to show that the sum of the degrees of the vertices in $V \setminus S$ in G' is at least as large as the sum of the degrees of the vertices in $V \setminus S$ in G. This is easily shown by observing that we took S to be the neighbours of the vertex with the highest degree, and thus $|S|$ (and therefore $d(v)$ for any $v \in V \setminus S$ in G') is higher than the degree of any vertex of $V \setminus S$ in G.

Now we have shown that the largest K_{r+1}-free graphs are the r-partite complete graphs, since it is always preferable to split any K_{r+1}-free graph up into a discrete part and a K_r-free part, and thus to split it up in r discrete parts. Therefore, the maximum amount of edges in any K_{r+1}-free graph is $\frac{1}{2}\left(n^2 - (n \bmod r)\lceil n/r \rceil^2 - (r - (n \bmod r))\lfloor n/r \rfloor^2\right)$ (the amount of edges is maximalised if the parts are of equal size, since otherwise you could increase the amount of edges by moving a vertex from the biggest part to the smallest part). □

The proof of Turán's theorem marked the beginning of the extremal graph theory. One of the principal questions in extremal graph theory is "what is the maximum amount of edges a graph on n vertices can have without having some fixed graph F as a subgraph?" This number, $\lfloor n/2 \rfloor \lceil n/2 \rceil$ in the case of triangles, or $\frac{1}{2}\left(n^2 - (n \bmod r)\lceil n/r \rceil^2 - (r - (n \bmod r))\lfloor n/r \rfloor^2\right)$ in the case of complete graphs on $r+1$ vertices (with n as the amount of vertices in the graph), is called the Turán number of a graph F, and it is denoted as $\mathrm{ex}\,(n, F)$. The r-partite complete graph on n vertices with parts of size as equal as possible is called the *Turán graph* and it is denoted as $T_{n,r}$.

Similarly, one can ask the question "For a set of graphs $\mathcal{F}$, what is the maximum number of edges a graph on n vertices can contain such that it does not contain a subgraph F for every $F \in \mathcal{F}$?" This number is similarly called the Turán number of the collection $\mathcal{F}$ and denoted as $\mathrm{ex}\,(n, \mathcal{F})$.

We will now look at some other specific Turán numbers and in the next chapter we will address the general case, which is given by the Erdős-Stone-Simonovits theorem.

Theorem 11. *Let F be the $(k+1)$-vertex star, i.e., a graph with one vertex of degree k and k vertices of degree 1. Then* $\mathrm{ex}\,(n, F)$ *is* $\frac{k-1}{2} * n$.

Proof. Obviously, if any vertex in a graph G has degree k, then G contains F, so every vertex in G has at most degree $k-1$. The graph consisting of disjoint K_{k-1}s, let's call it H, contains only vertices of degree $k-1$ and therefore H has the most edges. The amount of edges in H is $\frac{k-1}{2} * n$. □

Theorem 12. *The Turán number of a path of length k is $\frac{k-1}{2} * n$,* $\mathrm{ex}\,(n, P_k) = \frac{k-1}{2} * n$

Proof. Suppose $G = (V, E)$ is a graph on n vertices and $|E| > \frac{k-1}{2} * n$. We can find a path of length k in this graph by looking at the subgraph obtained by successively removing vertices with degree less than $\frac{k-1}{2}$ until every vertex has a degree at least $\frac{k-1}{2}$. This gives us an induced subgraph $G' = (V', E')$ with n' vertices, all of which have degree $d(v) \geq \frac{k-1}{2}$ (and $|E| > \frac{k-1}{2} * n'$). Of this graph we will take a connected component with the maximum ratio of edges to vertices, $G'' = (V'', E'')$ (with n'' vertices).

We will now find a path of length k in G''. Consider a path $v_0v_1...v_t$ of maximum length in G''. If $t \geq k$, then we are done, so we will assume that $t < k$.

By maximality of the path, all neighbours of v_0 and v_t are in the path. Suppose that $(v_0, v_t) \in E''$, then we have a cycle $v_0v_1...v_tv_0$ in G'' and thus $v_{i+1}v_{i+2}...v_tv_0...v_i$ is also a path in E'' for all $0 \leq i \leq t$. This means that all neighbours of all vertices in the path must be part of the graph, and since G'' was a connected component of G', the path must consist of all vertices in G''. This leads to a maximum amount of edges in G'' of $\binom{t+1}{2} \leq (k-1) * n''$. This means that G' has a maximum amount of edges of $(k-1) * n'$, which is a contradiction.

Thus, (v_0, v_t) can't be an edge in G''. Now observe that, since all of the neighbours of v_0 and v_t are in $\{v_0, ..., v_t\}$, $d(v_0) \geq \frac{k-1}{2}$, $d(v_t) \geq \frac{k-1}{2}$ and $t < k$, there must be v_i and v_{i+1} such that $(v_0, v_{i+1}) \in E''$ and $(v_i, v_t) \in E''$. Now we have the cycle $v_0v_1...v_iv_tv_{t-1}...v_{i+1}v_0$ and we can, analogously to before, conclude that the path must cover all of G'', and thus that the maximum of edges in G' is $(k-1) * n'$, contradictory to our assumption.

This means that t is at least k, i.e., the graph G' (and therefore G), contains a path of length k. □

Theorem 13. *Let T be a tree on $k+1$ vertices. Then $\frac{k-1}{2} * n \leq ex(n, T) \leq (k-1) * n$.*

Proof. No K_k can contain T, of course, so that gives us the lower bound of $\frac{k-1}{2} * n$. The upper bound can be seen by observing that any graph on n vertices with $(k-1) * n$ or more edges contains at least one vertex of degree k, and obviously any tree on $k+1$ vertices can be embedded in such a graph (if n is big enough). □

The proof for Erdős-Rényi-Sós is based on their paper at [4]

Theorem 14 (Erdős-Rényi-Sós)**.**

$$\mathrm{ex}(n, K_{2,2}) \leq \frac{n + n\sqrt{4n-3}}{4} \tag{3}$$

Proof. We start by observing that in any $K_{2,2}$-free graph the maximum amount of neighbours two vertices can have in common, is one.

Now, we can take $G = (V, E)$ to be a $K_{2,2}$-free graph on n vertices. We take S to be the following:

$$S = \{(\{u, v\}, z) | z, u, v \in V, (z, u), (z, v) \in E\}. \tag{4}$$

We now know that $|S|$ is at most $\binom{n}{2}$. We also know that $|S| = \sum_z \binom{d(z)}{2}$, since S is all combinations of two neighbours for each neighbour. This gives rise to the following inequality:

$$\sum_z \binom{d(z)}{2} \leq \binom{n}{2} \tag{5}$$

First observe that Cauchy-Schwarz tells us

$$\left(\sum_z d(z)\right)^2 \leq n \sum_z d(z)^2 \tag{6}$$

Now we can get a lower bound for the number of edges $(\frac{1}{2}\sum_z d(z))$, let $m = \frac{1}{2}\sum_z d(z)$, then we can find a lower bound for $\frac{1}{2}\sum_z d(z)$ in terms of m and n:

$$\begin{aligned}
\sum_z \binom{d(z)}{2} &\leq \binom{n}{2} \\
\frac{1}{2}\sum_z d(z)^2 - \frac{1}{2}\sum_z d(z) &\leq \binom{n}{2} \\
n\sum_z d(z)^2 - n\sum_z d(z) &\leq 2n\binom{n}{2} \\
\left(\sum_z d(z)\right)^2 - n\sum_z d(z) &\leq n^2(n-1) \\
4m^2 - 2mn &\leq n^2(n-1) \\
\left(2m - \frac{1}{2}n\right)^2 - \frac{n^2}{4} &\leq n^2(n-1) \\
\left(2m - \frac{1}{2}n\right)^2 &\leq n^2(n-1) + \frac{n^2}{4} \\
2m - \frac{1}{2}n &\leq \sqrt{n^2\left(n - \frac{3}{4}\right)} \\
m &\leq \frac{n\sqrt{n-\frac{3}{4}}}{2} + \frac{n}{4} \\
m &\leq \frac{n + n\sqrt{4n-3}}{4}
\end{aligned}$$

which is an upper bound to our Turán number. Moreover, this bound is tight, we can construct a graph that has $\Theta\left(n^{3/2}\right)$ edges and no $K_{2,2}$ as follows:

Consider a finite, discrete, three dimensional field $\mathbb{F}_p^3$. We can now construct a graph with as its vertices the directions in the field, i.e., the elements $(a, b, c) \in \mathbb{F}_p^3 \setminus \{(0, 0, 0)\}$ where two elements (x, y, z) and (a, b, c) are considered the same if there exists a λ such that $(a, b, c) = \lambda(x, y, z)$. Two vertices in this graph will be connected by an edge if they are ortagonal to each other, i.e. (a, b, c) is connected to (a, b, c) if $ax+by+cz = 0$. Now it is obvious to see that any two vertices in this graph have exactly one vertex in common in their neighbourhood. The amount of vertices in this graph is $(p^3-1)/(p-1) = p^2+p+1$ (all elements in the field without $(0, 0, 0)$ and with each line reduced to a point) and the degree of each vertex in this graph is $(p^2-1)/(p-1) = p+1$. This gives us an amount of edges m of

$$\begin{aligned} m &= (p+1)(p^2+p+1) \\ m &= (p+1)*n \\ m &= \Theta(\sqrt{n})*n \\ m &= \Theta(n^{3/2}) \end{aligned}$$

□

For the proof of Kővári–Sós–Turán, [5] was used.

Theorem 15 (Kővári–Sós–Turán). *For every r, s we have*

$$\mathrm{ex}(n, K_{r,s}) \leq \frac{(s-1)^{1/r}}{2} n^{2-1/r} + \frac{nr}{2} \tag{7}$$

Proof. We will use a generalized version of the proof of theorem 14 for this one. We observe that any graph $G = (VE)$ is $K_{r,s}$-free if and only if any r vertices have at most $s-1$ neighbours in common. This gives rise to the following inequality, for a graph with n vertices and m edges:

$$\sum_z \binom{d(z)}{r} \leq (s-1)\binom{n}{r}. \tag{8}$$

We now label the vertices with 1 up to n as follows:

$$\begin{aligned} z_i &\geq r \quad \text{for } i \leq b \\ z_i &< r \quad \text{for } i > b \end{aligned}$$

for some $b \leq n$. Now we can rewrite equation (8) as follows:

$$\sum_{i=1}^{n} \binom{d(z_i)}{r} = \sum_{i=1}^{b} \binom{d(z_i)}{r} > \frac{1}{r!} \sum_{i=1}^{b} (d(z_i) - r)^r$$

We can now use Hölder's inequality, which says that $\sum_{k=1}^{L} x_k y_k \leq \left(\sum_{k=1}^{L} x_k^p\right)^{1/p} \left(\sum_{k=1}^{L} y_k^q\right)^{1/q}$, or, equivalently, for $q = 1 + \frac{1}{p-1}$,

$$\sum_{k=1}^{L} x_k^p \geq L^{1-p} \left(\sum_{k=1}^{L} x_k\right)^p.$$

This gives us

$$\begin{aligned}
\sum_{i=1}^{n}\binom{d(z_i)}{r} &> \frac{1}{r!}b^{1-r}\left(\sum_{i=1}^{b}(d(z_i)-r)\right)^r\\
\sum_{i=1}^{n}\binom{d(z_i)}{r} &> \frac{1}{r!}b^{1-r}\left(\sum_{i=1}^{n}d(z_i)-br-\sum_{i=b+1}^{n}d(z)\right)^r \geq \frac{1}{r!}b^{1-r}\left(\sum_{i=1}^{n}d(z_i)-nr\right)^r\\
\sum_{i=1}^{n}\binom{d(z_i)}{r} &> \frac{1}{r!}n^{1-r}\left(\sum_{i=1}^{n}d(z_i)-nr\right)^r.
\end{aligned}$$

This finally gives rise to the following inequality:

$$\begin{aligned}
\frac{1}{r!}n^{1-r}\left(\sum_{i=1}^{n}d(z_i)-nr\right)^r &\leq (s-1)\binom{n}{r} < \frac{s-1}{r!}n^r\\
n^{1-r}\left(\sum_{i=1}^{n}d(z_i)-nr\right)^r &< (s-1)n^r\\
\left(\sum_{i=1}^{n}d(z_i)-nr\right)^r &< (s-1)n^{2r-1}\\
\sum_{i=1}^{n}d(z_i)-nr &< (s-1)^{1/r}n^{2-1/r}\\
m &< \frac{(s-1)^{1/r}}{2}n^{2-1/r}+\frac{nr}{2}
\end{aligned}$$

□

5 Erdős-Stone-Simonovits

By now we have developed enough tools to look at the principal theorems of this paper, the theorems of Erdős-Stone and Erdős-Stone-Simonovits. We will start by stating the theorems and proving the Erdős-Stone-Simonovits theorems using Erdős-Stone and then we will conclude by proving Erdős-Stone.

Recall from section 3 the definitions of a *(proper) colouring* and the *chromatic number* of a graph.

Theorem 16 (Erdős-Stone-Simonovits)**.** *Let F be a graph, and let r be the chromatic number of F. Then the Turán number of F is given by:*

$$\mathrm{ex}(n,F)=\left(\frac{r-2}{r-1}+o(1)\right)\binom{n}{2}. \tag{9}$$

Or, even stronger,

Theorem 17 (Erdős-Stone-Simonovits). *Let $\mathcal{F}$ be a finite collection of graphs, and let r be the minimum chromatic number of $F \in \mathcal{F}$, then*

$$\mathrm{ex}(n, \mathcal{F}) = \left(\frac{r-2}{r-1} + o\,(1)\right)\binom{n}{2}. \tag{10}$$

In the proofs for these theorems, we will use the Erdős-Stone theorem.

Theorem 18 (Erdős-Stone). *For any integer $r \geq 2$, $t \geq 1$ and any $\epsilon > 0$, there exists an integer n_0 such that any graph on $n \geq n_0$ vertices with $\left(\frac{r-2}{r-1} + \epsilon\right)\binom{n}{2}$ edges must contain a $T_{rt,r}$.*

Now it is easy to prove Erdős-Stone-Simonovits.

Proof of theorem 16 (Erdős-Stone-Simonovits). Let F be a graph and suppose r is the chromatic number of this graph and m is the amount of vertices in this graph. The Turán number of this graph is of course greater than the amount of edges in $T_{n,r-1}$, the $(r-1)$-partite Turán graph, since F can't be embedded in $T_{n,r-1}$ (since this graph can be coloured in $r-1$ colours by assigning a colour to each part). Thus,

$$\mathrm{ex}(n, F) \geq \frac{1}{2}\left(n^2 - (n \bmod r)\lceil n/r\rceil^2 - (r - (n \bmod r))\lfloor n/r\rfloor^2\right). \tag{11}$$

Erdős-Stone now tells us that there exists an n_0 for our r and $t = m$ such that for any graph G on $n \geq n_0$ vertices with more than $\left(\frac{r-2}{r-1} + \epsilon\right)\binom{n}{2}$ edges, G contains a copy of $T_{rt,r} = T_{rm,r}$. Of course, F can be embedded in $T_{rm,r}$, since one can simply put all vertices of the same colour in the same part of $T_{rm,r}$, each part is big enough to hold all vertices of F and every two nodes that are in different parts of $T_{rm,r}$ are connected. Therefore, we know that

$$\mathrm{ex}(n, F) \leq \left(\frac{r-2}{r-1} + \epsilon\right)\binom{n}{2}. \tag{12}$$

This proves the theorem that $\mathrm{ex}(n, F) = \left(\frac{r-2}{r-1} + o(1)\right)\binom{n}{2}$. □

Proof of theorem 17 (Erdős-Stone-Simonovits). Given a finite collection of graphs $\mathcal{F}$, suppose r is the least chromatic number of any $F \in \mathcal{F}$. Then for sufficiently large n Erdős-Stone tells us that there is a copy of $T_{rt,r}$ in any graph on n vertices with $\left(\frac{r-2}{r-1} + \epsilon\right)\binom{n}{2}$ edges with $t > |V_F|$ where V_F is the set of all vertices of the graph $F \in \mathcal{F}$ with the least chromatic number. As we have shown before, this graph can be embedded in $T_{rt,r}$ and thus for any graph G on n vertices with $\left(\frac{r-2}{r-1} + \epsilon\right)\binom{n}{2}$ edges for a sufficiently large n, we can find a graph $F \in \mathcal{F}$ that can be embedded in G. □

Now all that is needed to conclude the proof, is the proof of Erdős-Stone's theorem.

Proof of theorem 18 (Erdős-Stone). We will use the following lemma to prove Erdős-Stone:

Lemma 2. *For any integer $r \geq 2$, $t \geq 1$ and any $\epsilon' > 0$ there exists an integer m_0 such that any graph G on $m \geq m_0$ vertices and with all degrees greater than or equal to $\left(\frac{r-2}{r-1} + \epsilon'\right) m$ contains a copy of $T_{rt,r}$.*

We will first prove Erdős-Stone using this lemma, and then we will prove the lemma to conclude the proof.

Let $G = (V, E)$ be a graph on n vertices with more than $\left(\frac{r-2}{r-1} + \epsilon\right)\binom{n}{2}$ edges. We can construct a graph that meets all constraints of the lemma by recursively removing a vertex with degree less than $\left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right)$ times the amount of vertices in the graph, i.e., we can create a decreasing sequence

$$G_n \subset G_{n-1} \subset ... \subset G_{m+1} \subset G_m \tag{13}$$

where G_{i-1} is obtained from G_i by removing a vertex of degree less than $\left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right) * i$.

The lemma tells us that there is an m_0 such that if $m' > m_0$, any graph on m' vertices with degrees higher than $\left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right) * m'$ contains a copy of $T_{rt,r}$ for any $r \geq 2, t \geq 1$ and any $\epsilon' = \frac{\epsilon}{2} > 0$. We claim that our G_m suffices these requirements. By definition, the degrees are high enough, so all we need to prove, is that $m > m_0$ if n is large enough.

This can be easily seen by observing that the amount of edges removed from G_n (with $|E|$ edges) can be at most $\sum_{i=m+1}^{n}\left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right) * i$, and thus:

$$\begin{aligned}
\left(\frac{r-2}{r-1} + \epsilon\right) * \binom{n}{2} &\leq |E| \leq \binom{m}{2} + \left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right) * \sum_{i=m+1}^{n} i \\
\left(\frac{r-2}{r-1} + \epsilon\right) * \binom{n}{2} &\leq |E| \leq \binom{m}{2} + \left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right) * (\frac{1}{2}(n-m)(m+n+1)).
\end{aligned}$$

We can now move n to the left hand side to see that there is an n_0 such that $n > n_0$ implies that $m > m_0$.

$$\begin{aligned}
\left(\frac{r-2}{r-1} + \epsilon\right) * \binom{n}{2} &\leq \binom{m}{2} + \left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right) * (\frac{1}{2}(n-m)(m+n+1)) \\
\left(\frac{r-2}{r-1} + \epsilon\right) * \binom{n}{2} &\leq \binom{m}{2} + \left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right) * (\frac{1}{2}(n^2+n-m^2-m)) \\
\left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right) * \left(\binom{n}{2} - \frac{1}{2}(n^2+n)\right) + \frac{\epsilon}{2}\binom{n}{2} &\leq \binom{m}{2} - \left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right)\left(\frac{m^2+m}{2}\right) \\
\frac{\epsilon}{2}\binom{n}{2} - \left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right) * n &\leq \binom{m}{2} - \left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right)\left(\frac{m^2+m}{2}\right) \\
\frac{\epsilon}{4}n^2 + O(n) &\leq \frac{m^2-m}{2} - \left(\frac{r-2}{r-1} + \frac{\epsilon}{2}\right)\left(\frac{m^2+m}{2}\right) \\
\frac{\epsilon}{4}n^2 + O(n) &\leq \left(1 - \frac{r-2}{r-1} - \frac{\epsilon}{2}\right) m^2/2 + O(m)
\end{aligned}$$

Now we can choose ϵ and n_0 such that $n > n_0$ implies that $m > m_0$ since both the left hand side of this equation and the right hand side of this equation are increasing for increasing m and n if $\left(1 - \frac{r-2}{r-1} - \frac{\epsilon}{2}\right) > 0$.

All that is needed to conclude this proof, is the proof of lemma 2.

Proof of the lemma. We will prove this lemma using induction on r. Take $G = (V, E)$ to be a graph on m vertices and with the minimum degree of the vertices as $\left(\frac{r-1}{r} + \epsilon\right) * m$. For $r = 2$, we have seen this problem in section 4 in Kovari-Sos-Turan's theorem. This theorem said that there is a copy of $T_{2t,2}$ in G if $\binom{\epsilon m}{t} * m > (t-1)\binom{m}{t}$, or $\frac{\binom{\epsilon m}{t} * m}{\binom{m}{t}} > (t-1)$, so there definitely is some m_0 such that for $m \geq m_0$ there is a copy of $T_{2t,2}$ in G.

Assume now that $r \geq 2$. Let $s = \lceil \frac{t}{\epsilon} \rceil$. We will show that there exists a copy of $T_{(r+1)t,r+1}$ in G. The induction hypothesis now tells us that, for a large enough m, there is a copy of $T_{rs,r}$ in G. Let $A_1, A_2, ... A_r$ be the r parts of the r-partite graph $T_{rs,r}$, each of which has size s. Let U be the vertices in G that are not in the copy of $T_{rs,r}$ inside G, $U = V \setminus (A_1 \cup A_2 \cup ... \cup A_R)$. We are now going to look at the edges between A_i and U and show that we can find a group of t vertices that are connected to t vertices in A_i for all $1 \leq i \leq r$, and thus we can extend it to a $T_{(r+1)t,r}$.

Let W be the set of all vertices w in U such that $|\mathcal{N}(w) \cap A_i| > t$ for all i. Now the number of edges that are missing between U and $\cup_i A_i$ is at least $(|U| - |W|)(s - t)$ (the minimum amount of edges missing between the vertices not in W and $\cup_i A_i$) and at most $rs(\frac{1}{r} - \epsilon)m$ (since the degrees of all vertices in are at least $\left(\frac{r-2}{r-1} + \epsilon\right) * m$ and the degrees of all vertices in the subgraph induced by $\cup_i A_i$ are $(r-1)s$). This leads to the following inequality:

$$
\begin{aligned}
(|U| - |W|)\,(s-t) &\leq rs(\frac{1}{r} - \epsilon)m \\
(m - rs - |W|)\,(s-t) &\leq rs(\frac{1}{r} - \epsilon)m \\
m - rs - |W| &\leq \frac{rs}{s-t}(\frac{1}{r} - \epsilon)m \\
|W| &\geq m - \frac{rs}{s-t}(\frac{1}{r} - \epsilon)m - rs \\
|W| &\geq \left(1 - \frac{rs}{s-t}\left(\frac{1-r\epsilon}{r}\right)\right)m - rs \\
|W| &\geq \left(1 - \frac{s(1-r\epsilon)}{s-t}\right)m - rs \\
|W| &\geq \left(1 - \frac{s(1-r\epsilon)}{s(1-t/\lceil t/\epsilon\rceil)}\right)m - rs \\
|W| &\geq \left(1 - \frac{1-r\epsilon}{1-t/\lceil t/\epsilon\rceil}\right)m - rs \\
|W| &\geq \left(1 - \frac{1-r\epsilon}{1-\epsilon}\right)m - rs \\
|W| &\geq \frac{r\epsilon - \epsilon}{1-\epsilon}m - rs \\
|W| &\geq \frac{(r-1)\epsilon}{1-\epsilon}m - rs
\end{aligned}
$$

For $\epsilon < 1$, $\frac{(r-1)\epsilon}{1-\epsilon}m - rs$ is increasing linearly with m. Therefore, we can choose m_0 such that, if $m \geq m_0$, $|W| > \binom{s}{t}^r(t-1)$ and thus we can extend the $T_{rs,r}$ to a $T_{(r+1)t,r}$. This concludes the proof of the lemma. □

□

6 Bibliography

References

[1] S. Kopparty,
Graph theory,
Lecture notes,
http://www.math.rutgers.edu/ sk1233/courses/graphtheory-F11/,
Rutgers University,
fall 2011.

[2] Multiple authors,
Big O notation,

https://en.wikipedia.org/wiki/Big_O_notation,
June 17, 2016.

[3] R. Morris,
Extremal and Probabilistic Combinatorics,
Videorecording,
http://video.impa.br/index.php?page=programa-de-mestrado-2012-extremal-and-probabilistic-combinatorics,
Instituto Nacional de Matemática Pura e Aplicada,
Januari 13, 2012.

[4] P. Erdős, A. Rényi, V. T. Sós,
On a problem of graph theory,
Studia Scientiarum Mathematicarum Hungaria I,
pages 215-235,
1966.

[5] T. Kővári, V. T. Sós, P. Turán,
On a problem of K. Zarankieqicz,
Colloquium math. 3,
pages 50-57,
1954.