Extremal graph theory.

Let G be a graph with n vertices. The main question we would like to study is the following one: for a graph property P, which is the value $F_P(n)$ such that if $|E(G)| > F_P(n)$ then G satisfies property P?

We will emphasize our study over the property "H is a subgraph of G".

Obs/ This general idea admits multiple variations: for instance we can change the parameter "number of edges" by the parameter "minimum degree".

Obs/ In many situations we won't be able to get the exact value for $F_P(n)$ (only asymptotic estimates)

Some motivational examples

Before going to the study of the subgraph property, we will motivate the previous ideas with examples.

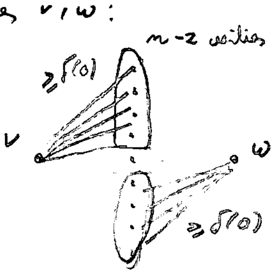
Prop/ All graphs with n vertices and less than $n-1$ edges are disconnected.

Proof/ We apply the handshaking lemma: $\sum_{v \in V} d(v) = 2|E(G)| \leq 2n-4$. If there is a vertex of degree 0, we are done. So, assume that G is connected: it must have then at least one vertex of degree 1. Deleting it (and the corresponding edge) and applying induction we get the result as claimed.

On the other side, all trees have $n-1$ edges and are connected. So this is the best result we can obtain in this direction. We have

Prop/ Let G be a n-vertex graph. Then, if $\delta(G) \geq \lfloor \frac{n}{2} \rfloor$, the graph is connected.

Proof/ Observe that $\lfloor \frac{n}{2} \rfloor = n/2$ if $n \equiv 0 (2)$ and $n-1/2$ if $n \equiv 1 (2)$. Consequently, for every two vertices v, w:



v and w must have at least one common neighbour, otherwise we would have that:

$$n-2 \geq d(v) + d(w) \geq 2\delta(G) \geq \begin{cases} n & n \equiv 0 (2) \\ n-1 & n \equiv 1 (2) \end{cases} \quad !!$$

Again, this is the best possible: take two disjoint copies of $K_{\lfloor \frac{n}{2} \rfloor}$. Then $\delta(G) = \lfloor \frac{n}{2} \rfloor - 1$ and the graph is NOT connected!

We can exploit this result in the context of hamiltonian cycle:

Theorem / (Dirac, 53) Let G be an n-vertex graph with $\delta(G) \geq \frac{n}{2}$ (and $n \geq 3$). Then G has a Hamiltonian cycle.

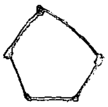
Proof/ As $\frac{n}{2} \geq \lfloor \frac{n}{2} \rfloor$, G is connected. Let P be the longest path in G; call x_0 and x_k the starting and the finishing vertices of this path:

Observe now that all neighbours of x_0 and x_k must belong to the path, otherwise we could extend the path. Hence, $n/2$ of the vertices $x_0, \ldots, x_{k-1}$ are adjacent with x_k, and $\frac{n}{2}$ of the vertices $x_1, \ldots, x_k$ are adjacent with x_0. By the pigeonhole principle there is an i such that $\overline{x_0 x_{i+1}} \in E(G)$, $\overline{x_i x_k} \in E(G)$. This defines a cycle. Finally, observe that all vertices must be in the initial path, because the graph is connected and we would be able to extend the path.

Containing a certain subgraph.

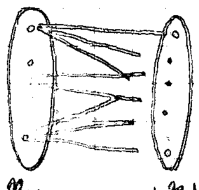
We move to study the following question: for a graph G with n vertices, which is the maximum number of edges (if exist) in order to assure that G has not the graph H as a subgraph? We will denote this number by $ex(n,H)$, and we call it <u>Turan number</u> for H.

The base case is to study $ex(n,K_3)$. For instance, we can do the following construction:



5 edges

6 edges

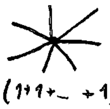


$\lfloor \frac{n}{2} \rfloor$ $\lfloor \frac{n}{2} \rfloor$

If $n \equiv 0(2) \to \frac{n}{2} \times \frac{n}{2}$ edges $\equiv \frac{n^2}{4}$ edges. If $n \equiv 1(2)$, we take $n-1/2$ and $n+1/2$ vertices $\Rightarrow \frac{n^2}{4} - \frac{1}{4}$ edges. In fact, this construction is the best possible

Theorem/ (Mantel, 1907) $ex(n,K_3) = \left\lfloor \frac{n^2}{4} \right\rfloor$.

Proof/ Assume that G has n vertices and $|E(G)| > \frac{n^2}{4}$ edges. Assume also that G has no triangles. Then adjacent vertices cannot have common neighbours. Hence, if $\overline{xy} \in E(G)$, necessarily $d(x) + d(y) \le n$. Hence,



$d(x)^2 = d(x)(1+1+\_ +1)$

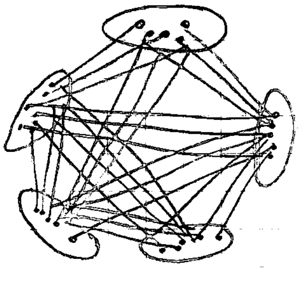
$$\sum_{x \in V(G)} d(x)^2 = \sum_{\overline{xy} \in E(G)} d(x) + d(y) \le n\,|E(G)|$$

On the other hand, using Cauchy-Schwarz inequality we get that

$$\sum_{x \in V(G)} d(x)^2 \ge \frac{1}{n} \Big(\sum_{x \in V(G)} d(x) \Big)^2 = \frac{4}{n} |E(G)|^2$$

Finally, $|E(G)| \le \frac{n^2}{4}$, which contradicts the hypothesis.

The next step is to try to generalize the previous construction in order to deal with $ex(n, K_{r+1})$. In order to do so, we can consider the so-called <u>Turan graph</u> ($T_{r,n}$)



we take $n(r)$ subsets of size $\lceil \frac{n}{r} \rceil$, and $r - n(r)$ subsets of size $\lfloor \frac{n}{r} \rfloor$, and we consider all the edges between the partite sets. The number of edges is equal to:

$$\underbrace{\binom{n}{2}}_{\approx \frac{n^2}{2}} - \underbrace{\binom{\lceil \frac{n}{r} \rceil}{2}}_{\frac{1}{2} \lceil \frac{n}{r} \rceil^2} \{ n(\mathrm{mod}\, r) \} - \underbrace{\binom{\lfloor \frac{n}{r} \rfloor}{2}}_{\frac{1}{2} \lfloor \frac{n}{r} \rfloor^2} \{ r - n(\mathrm{mod}\, r) \} \lesssim \left(1 - \frac{1}{r}\right) \frac{n^2}{2}$$

$\Rightarrow$ If $r | n \Rightarrow = \left(1 - \frac{1}{r}\right) \frac{n^2}{2}$

This is again the best possible bound:

Theorem/ (Turan, 1941) A graph G on n vertices without a K_{r+1} as a subgraph has $|E(G)| \le \left(1 - \frac{1}{r}\right) \frac{n^2}{2}$ edges. Additionally, the maximum value for $|E(G)|$ is reached when G is a Turan graph.

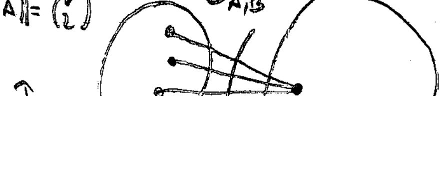
Proof/ We apply induction on n (fixing the value of r). The theorem is trivially true for $n = 1, \dots, r$. ($\binom{n}{2} \le (1 - 1/r)\, n^2/2$). So let us assume that the result is true for values $< n$, and let us prove for $|V(G)| = n$. Hence, let G be a graph with n vertices, and without K_{r+1} as a subgraph with maximum of edges. Observe that it must have some copy of K_r, otherwise we could add some edge without creating K_{r+1}, and this is not possible by maximality.

Let $A \subseteq V(G)$ a set of vertices define a K_r, and $B = V(G) - A$. Then $|B| = n - r$, $|A| = r$:

$|V(A)| = r$

$|E(A)| = \binom{r}{2}$

$e_{A,B}$

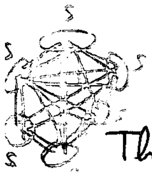


$|V(B)| = n - r$

$|E(B)| \le \left(1 - \frac{1}{r}\right) \frac{(n-r)^2}{2}$ $\Rightarrow |E(G)| = |E(A)| + |E(B)| + e_{A,B} \lesssim$

We show now that extremal graphs are precisely Turan graphs. For this, observe that the previous inequality must be an equality, and so each vertex in B is connected with exactly $r-1$ vertices in A. Writing $V(A)=\{x_1,\dots,x_r\}$, we denote by $V_i=\{v\in V(G): vx_i\notin E(G)\}$. Observe that each V_i is an independent set, and $\bigcup V_i = V(G)$. Hence, G is r-partite. And then G must be a Turan graph in order to maximise the number of edges.

This result gives us that $ex(n,K_{r+1}) = \left(1-\frac{1}{r}\right)\binom{n}{2}+O(n)$, which is the number of edges in the Turan graph. So, in this case, Turan graphs are dense graphs. This is what happen in most of the cases.



To show this fact, we start with the base case $K_s^{r+1} = K_{s,s,\dots,s}$, which is the base case towards the general result:

Theorem / (Erdős–Stone, 1946) For fixed r and s, $ex(n,K_s^{r+1}) = \left(1-\frac{1}{r}\right)\binom{n}{2}+o(n^2)$.

Equivalently, this theorem says that for all $\varepsilon>0$, and for fixed r and s, there is an integer n_0, such that for every graph with $n\geq n_0$ vertices and at least $ex(n,K_{r+1})+\varepsilon n^2$ edges contain K_s^{r+1} as a subgraph. We will prove this result as an application of the Regularity Lemma.

We now exploit this theorem to show the following corollary:

Corollary / (Erdős–Simonovits, 1966) For any graph H, $ex(n,H) = \left(1-\frac{1}{\chi(H)-1}\right)\binom{n}{2}+o(n^2)$

Observe that this corollary covers the previous results.

Proof / Write $r=\chi(H)$. As H cannot be coloured with $r-1$ colours, H is not a subgraph of $T_{r-1,n}$ for all $n\in\mathbb{N}$, hence $|E(T_{r-1,n})|\leq ex(n,H)$. On the other side, H is a subgraph of $K_{s,\overset{r}{\dots},s}$, for s fixed (but large enough). Hence,

$$|E(T_{r-1,n})|\leq ex(n,H)\leq ex(n,K_s^r)$$

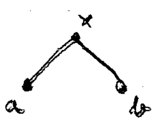
We will use now a "sandwich" argument, concluding that $ex(n,H)=\left(1-\frac{1}{r-1}\right)\binom{n}{2}+o(n^2)$, as we wanted to prove.

This result is very general but it has a big obstruction: it does not give anything precise when $\chi(H)=2$: this is because all constructions we have obtained so far contain bipartite subgraphs.

Hence, we need a "case by case" analysis when dealing with bipartite graphs. We start with the easier case:

Theorem / (Erdős, 1938) For the cycle C_4 (or $K_{2,2}$) we have $ex(n,C_4)=O(n^{3/2})$.

Proof / We double count the number of copies of $K_{1,2}$ (namely, paths of lenght 3):



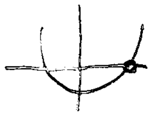
$A=\{(\{a,b\},x): a,b,x\in V(G),\ x$ is adjacent with a and $b\}$ $\Rightarrow$ We apply double counting on A:

a) Fix $\{a,b\}$ $\Rightarrow$ We have a unique choice for x $\Rightarrow$ $|A|\leq\binom{|V(G)|}{2}$.

b) Fix x $\Rightarrow$ We have exactly $\binom{d(x)}{2}$ copies of $K_{1,2}$. $|A|=\sum_{x\in V(G)}\binom{d(x)}{2}$. Hence (writing $|V(G)|=n$)

$$\binom{n}{2}\geq\sum_{x\in V(G)}\binom{d(x)}{2}=\frac{1}{2}\sum_{x\in V(G)}d^2(x)-\frac{1}{2}\sum_{x\in V(G)}d(x)=\frac{1}{2}\sum_{x\in V(G)}d^2(x)-|E(G)|$$

Applying now CS $\Rightarrow \binom{n}{2} \geq \frac{2}{n}|E(G)|^2 - |E(G)|$, hence:



$$0 \geq \frac{2}{n}|E(G)|^2 - |E(G)| - \binom{n}{2} \Rightarrow |E(G)| \leq \frac{n}{4}\left(1+\left(1+\frac{8}{n}\binom{n}{2}\right)^{1/2}\right) = \frac{1}{2}n^{3/2}+o(n^{3/2})$$

On the other hand, we have explicit constructions reaching this value:

Theorem/ (Klein, 1938) $ex(n,C_4) = \Theta(n^{3/2})$.

One can generalise this argument in order to deal with the case $K_{t,s}$:

Theorem/ (Kővari-Sós-Turán, 1954) For $s \geq t \geq 1$, $ex(n,K_{t,s}) = O(n^{2-1/t})$

Proof/ Homework.

In the rest of the cases many things are known, but many more are unknown!

Graph	Who?	Result / Conjecture
$K_{s,t}$, $s \leq t$, $s \in \{2,3\}$ or $t > (s-1)!$	Klein, Brown, Kollár-Rónyai-Szabó	It is proved that $ex(n,K_{s,t}) = \Theta(n^{2-1/s})$
T, T is a tree, $\lvert V(T)\rvert = k$	Erdős-Sós, 63	It is conjectured that if $e(G) > \frac{1}{2}(k-2)\lvert G\rvert$, then G contain all trees with k vertices (True for k large enough)
C_{2k}	Bondy-Simonovits 74	$ex(n,C_{2k}) = O(n^{1+1/k})$. This is sharp for $k=2,3,5$.
C_8		$\Omega(n^{9/8}) \leq ex(n,C_8) \leq O(n^{5/4})$
Q_3 (cube)		$\Omega(n^{3/2}) \leq ex(n,Q_3) \leq O(n^{8/5})$

In fact, from Kővari-Sós-Turán theorem we have:

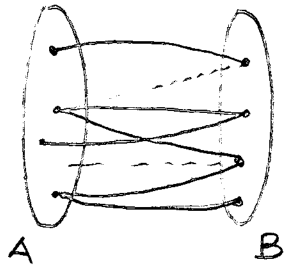
$$\Omega(n^{3/2}) \leq ex(n,K_{2,2}) \leq ex(n,K_{2,3}) \leq \dots \leq O(n^{3/2}) \Rightarrow \Theta(n^{3/2})$$

$$\Omega(n^{5/3}) \leq ex(n,K_{3,3}) \leq ex(n,K_{3,4}) \leq \dots \leq O(n^{5/3}) \Rightarrow \Theta(n^{5/3})$$

$$\Omega(n^{5/3}) \leq ex(n,K_{4,4}) \leq ex(n,K_{4,5}) \leq ex(n,K_{4,6}) \leq ex(n,K_{4,7}) \leq \Theta(n^{7/4})$$

Szemerédi's Regularity Lemma

We introduce a central tool in extremal combinatorics discovered by Endré Szemerédi in the course of his proof of a conjecture of Erdős and Turan (we will talk about this later). The first basic idea we need is the one related to a random model for bipartite graphs:



We choose each possible edge independently with a probability p. Denoting by X the random variable which counts the number edges, we have that:

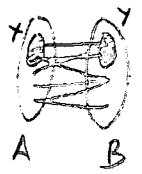
$$\mathbb{E}[X] = p|A|\cdot|B| \Rightarrow d(A,B) := \frac{\mathbb{E}[X]}{|A||B|} = p$$

This is valid for all $X \subseteq A$, $Y \subseteq B$: $d(X,Y) = p$.

So, this random model satisfies a regularity condition: the "density" of edges does not depend on X and Y.

Of course this is the case when we consider concrete graphs. But the idea is that large graphs have some kind of random behaviour, which implies some regularity properties.

Def/ Let (A,B) be a set of disjoint vertices. For $X \subseteq A$ and $Y \subseteq B$ denote by $e(X,Y)$ the number of edges with endvertices in X and Y, and denote by $d(X,Y) = e(X,Y) \cdot |X|^{-1} \cdot |Y|^{-1}$.



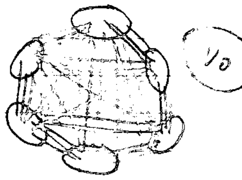
For $\varepsilon > 0$, we say that the pair (A,B) is <u>ε-regular</u> if for all $X \subseteq A, Y \subseteq B, |X| \geq \varepsilon |A|$, $|Y| \geq \varepsilon |B|$, we have that

$$|d(A,B) - d(X,Y)| < \varepsilon \qquad \text{(or } \underline{\varepsilon\text{-pair}}\text{)}$$

We now generalize the notion of ε-pair to more sets.

Def/ Let G be a graph, and let $\{V_0, V_1, \dots, V_k\}$ a partition of $V(G)$. Let $\varepsilon > 0$. We say that the previous partition is <u>ε-regular</u> if:

i) $|V_0| \leq \varepsilon |V(G)|$ (<u>exceptional set</u>)
ii) $|V_1| = \dots = |V_k|$
iii) All but εk^2 of the pairs (V_i, V_j), $i<j$ are ε-regular.



We have an equivalent definition for ε-regularity:

Not covered

Def/ Let G be a graph, and let $V_1 \cup \dots \cup V_k = V(G)$ be a partition of vertices of $V(G)$. We say that the previous partition is <u>ε-regular (v2)</u> if for all $i \neq j$, $||V_i| - |V_j|| \leq 1$ and all but at most εk^2 of the pairs (V_i, V_j) are ε-regular.

The Szemeredi's regularity lemma states that such a partition always exists:

Lemma/ (Szemeredi's regularity lemma) For every $\varepsilon > 0$ and every integer $m \geq 1$ there exists an integer $M := M(\varepsilon, m)$ such that every graph with more than m vertices has an ε-regular partition $\{V_0, \dots, V_k\}$ with $k \leq M$.

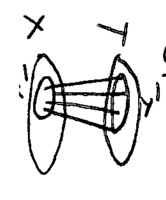
Obs/ 1) M just depend on ε and m, and not on the number of vertices of my graph (this number must be $\gg M$). The dependence is very bad: $|V(G)|$ must be a tower of height $(1/\varepsilon)^5$ of m, and this is tight (Gowers, 97).

2) This lemma is useful when the graph is <u>dense</u> (namely, when $|E(G)| = \Omega(n^2)$). Otherwise, our graph is <u>sparse</u> ($|E(G)| = o(n^2)$): recall that $d(A,B) = e(A,B) \cdot |A|^{-1} |B|^{-1}$. Then, if $e(A,B) = o(n^2)$ all densities tend to 0, and all partitions are ε-regular.

This lemma is usually applied in two steps: preparation of the graph + counting. Let us see an example

Lemma/ (Counting lemma for triangles) Let G be a graph and $\{X, Y, Z\}$ a partition of the vertex set $V(G)$. Assume that $d(X,Y) = \alpha$, $d(Y,Z) = \beta$ and $d(Z,X) = \gamma$. If there exists an $\varepsilon > 0$ such that $\min\{\alpha, \beta, \gamma\} \geq 2\varepsilon$ and the pairs (X,Y), (Y,Z) and (Z,X) are ε-regular, then the number of triangles xyz with $x \in X$, $y \in Y$ and $z \in Z$ is at least $(1-2\varepsilon)(\alpha-\varepsilon)(\beta-\varepsilon)(\gamma-\varepsilon)|X||Y||Z|$

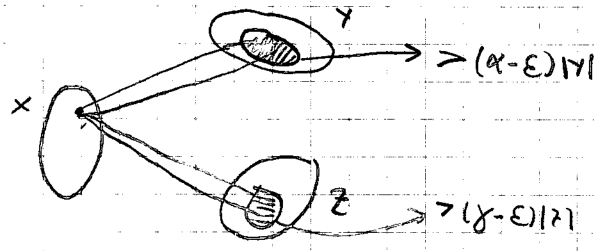
Proof/ For each $x \in X$, write $d_Y(x)$ and $d_Z(x)$ the number of neighbours of x in Y and Z, respectively.



<u>Claim</u>: $\# x \in X$ such that $d_Y(x) < (\alpha - \varepsilon)|Y|$ is $\leq \varepsilon |X|$: assume the contrary, so take this set of vertices (call it X') and the neighbours in Y. Then, $|X'| > \varepsilon |X|$, and the density of edges $d(X', Y) < \alpha - \varepsilon$. Then $d(X',Y) - d(X,Y) < \alpha - \varepsilon - \alpha = -\varepsilon \Rightarrow$ The pair cannot be ε-reg.

The same result holds for the pair (X,Z) (namely, $\#\, x\in X$ such that $d_Z(x)<(\gamma-\varepsilon)|Z|$ is $\leq \varepsilon|X|$).

Let us point our attention over $x\in X$ such that $d_Y(x)>(\alpha-\varepsilon)|Y|$ and $d_Z(x)>(\gamma-\varepsilon)|Z|$. Then we consider the sets $N(x)\cap Y$ and $N(x)\cap Z$:



As now the pair (Y,Z) is ε-regular, we can estimate the number of edges between $N(x)\cap Y$ and $N(x)\cap Z$:

$$|d(Y,Z) - d(N(x)\cap Y, N(x)\cap Z)|<\varepsilon \Rightarrow$$

$$\Rightarrow e(N(x)\cap Y, N(x)\cap Z) > (\alpha-\varepsilon)(\beta-\varepsilon)(\gamma-\varepsilon)|Y||Z|$$

Finally, we sum over all x such that $d_Y(x)>(\alpha-\varepsilon)|Y|$, $d_Z(x)>(\gamma-\varepsilon)|Z|$, which is at least $(1-2\varepsilon)|X|$, getting the result as claimed.

The next step is a prototype of the type of results we obtain when applying the regularity lemma:

Theorem/ (Triangle Removal lemma) For every $\varepsilon>0$ there exist $\delta>0$ such that for every graph G on n vertices at at most δn^3 triangles, it can made triangle-free by removing at most εn^2 edges.

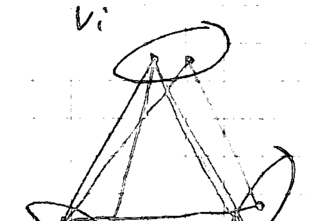
This tells us that we can remove not all the edges of the graph in order to erase all triangles.

Proof/ We prove that if we need more than εn^2 edges to eliminate all triangles, then $\#\text{triangles} > \delta n^3$, for $\delta=\delta(\varepsilon)$.

Let $V_0\cup V_1\cup\ldots\cup V_k$ be a $\frac{\varepsilon}{4}$-regular partition of the vertices of G. We will remove the following edges (we write $|V_1|=\ldots=|V_k|=c$, and then $k\cdot c<n$), taking $m:\lfloor\frac{4}{\varepsilon}\rfloor$:

- All edges incident with V_0: bounded by $\frac{\varepsilon}{4}\cdot n\cdot n = \frac{\varepsilon}{4}n^2$.
- All edges inside $V_1,\ldots,V_k$: $\leq k\cdot c^2 < n^2/k < \frac{\varepsilon}{4}n^2$ $\quad (k \gg (\frac{\varepsilon}{4})^{-1})$
- All edges defined by a non $\frac{\varepsilon}{4}$-regular pair: $<(\frac{\varepsilon}{4}\cdot k^2)\cdot c^2 < \frac{\varepsilon}{4}n^2$.
- All edges lying between a pair V_i, V_j, which define a $\frac{\varepsilon}{4}$-regular pair, whose density is less than $\frac{\varepsilon}{2}$: $\binom{k}{2}\cdot\frac{\varepsilon}{2}c^2 < \frac{\varepsilon}{4}n^2$.

At this point we have deleted $<\varepsilon n^2$ edges. If there are no triangles, we are done; otherwise,



Consider the triple (V_i,V_j,V_k) such that every pair is $\frac{\varepsilon}{4}$-regular and $d(V_r,V_s)\geq\frac{\varepsilon}{2}$. Then, by the counting lemma we have that for a triplet (V_i,V_j,V_k):

$$\#\text{triangle} > \left(1-\frac{\varepsilon}{2}\right)\left(\frac{\varepsilon}{4}\right)^3 c^3 > \delta(\varepsilon)\, n^3$$

And this finish the proof.

Obs/ 1) The triangle removal lemma has a generalization to general subgraphs: this is what is called the graph removal lemma. In this case, if H is the subgraph, we need to change δn^3 by $\delta n^{|V(H)|}$, but the εn^2 is the same.

2) Jacob Fox has recently proved the graph removal lemma without the use of the regularity lemma. This gives better bounds.

3) There are similar versions of this result for induced subgraphs; but the proof is even more involved.

Roth's and Szemerédi's Theorem

We will show how to apply the triangle removal lemma in the context of additive number theory.

Def/ A **k-AP** in $\mathbb{Z}$ is a set of the form $a, a+d, \dots, a+(k-1)d$, where $d \neq 0$.

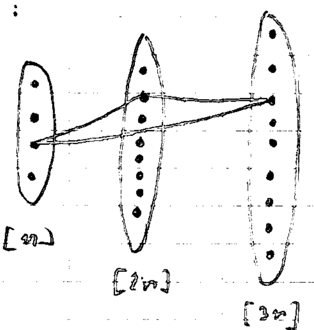
The question is the following: write $r_k(n)$ for the maximal number of elements of a subset of $\{1,2,\dots,n\} = [n]$ without k-AP. Which is the value of $r_k(n)$?

Theorem/ (Erdős–Turan Conjecture $\to$ Szemerédi's Theorem) For every k, $r_k(n) = o(n)$.

We will prove this theorem in the particular case $k=3$, which is known as Roth's Thm:

Theorem/ (Roth's theorem, 1953) $r_3(n) = o(n)$.

Proof/ We use the triangle removal lemma. Given a set $S \subseteq [n]$, we create a tripartite graph $H(S)$ with vertex set $\{(i,1) : i \in [n]\} \cup \{(j,2) : j \in [2n]\} \cup \{(k,3) : k \in [3n]\}$, and:



1) $(i,1)$ and $(j,2)$ are adjacent if $j-i \in S$
2) $(j,2)$ and $(k,3)$ are adjacent if $k-j \in S$
3) $(i,1)$ and $(k,3)$ are adjacent if $k-i \in 2\cdot S = \{2\cdot s : s \in S\}$

Then, if $(i,1)$, $(j,2)$ and $(k,3)$ form a triangle, writing $j-i = a_1$, $k-j = a_3$ and $k-i/2 = a_2$ then (a_1, a_2, a_3) is a 3-AP. In particular, we always have the triangles defined by $(i,1), (i+a,2), (i+2a,3)$ (a total of $|S|\cdot n$). Additionally, all this triangles are edge disjoint. Now, if $|S| > \varepsilon n$ for all $n \geq n_0$ and S is 3-AP free, then $H(S)$ has more than εn^2 trivial triangles (and less than, say, $2n^2$)

Then, by disjointness, we need to delete more than εn^2 edges in order to make this graph triangle free. But then by the triangle removal lemma, it has at least δn^3 triangles, such that at least $\delta n^3 - 2n^2$ of them are <u>NOT</u> trivial. Now, choosing $n > 2/\delta$ we can ensure that there is some non trivial triangle, and the proof is done.

What about the history of the problem?

- Roth: $k=3$, analytic method (1953)
- Szemerédi: first $k=4$, then general, combinatorial (1975)
- Furstenberg: ergodic theory (1977)
- Gowers: functional analysis (2001)
- Rödl–Schacht–Gowers: hypergraphs (2007)
- Elek–Szegedy: non-standard analysis (2007+)

For $r_3(n)$:

- Roth (53): $r_3(n) = O\left(\frac{n}{\log\log n}\right)$
- Heath-Brown (1987): $O(n/(\log n)^c)$
- Szemerédi (1990): "
- Bourgain (1999): $O\left(n\left(\frac{\log\log n}{\log n}\right)^{1/2}\right)$
- Bourgain (2008): $O(n(\log\log n)^2/\log n^{2/3})$
- Sanders (2012): $O(n(\log\log n)^5/\log n)$

What is known in the opposite direction? $\Rightarrow$ Behrend's construction:

Theorem/ (Behrend, 1949) For n large enough, $r_3(n) \geq n e^{-O(\sqrt{\log n})}$.

Proof/ We exploit the following easy idea: a line could cut an sphere in at most **2** points. Let then m and M be large integers (we will determine them later) and write

$$S(r) = \{\bar{x} = (x_1, \dots, x_m) \in [M]^m : x_1^2 + \dots + x_m^2 = r^2\}$$

Observe that $m \leq r^2 \leq mM^2$, hence $\bigcup_r S(r) \supseteq [M]^m$. Hence, by the pigeonhole principle, there is a radius r_0 such that

$$|S(r_0)| \geq \frac{M^m}{mM^2 - m} > \frac{M^{m-2}}{m}$$

We will project these vertices into $[n]$ using the mapping $P(x_1, \dots, x_m) = \frac{1}{2M} \sum_{i=1}^{m} x_i (2M)^{i}$. This mapping has the following properties:

i) P is injective ($\equiv$ unique representation in base $2M$)
ii) $\bar{x} + \bar{y} = 2\bar{z}$ iff $P(\bar{x}) + P(\bar{y}) = P(2\bar{z})$
iii) $\max_{\bar{x} \in S(r_0)} P(\bar{x}) \leq (2M+1)^M$

Then taking $(2M+1)^m \simeq n \Rightarrow M \approx n^{1/m}/2$. So, take $M = \lceil n^{1/m}/2 \rceil$. Then all points $P(S(r_0))$ belong to $[n]$ and it does not have a 3-AP. Then:

$$|P(S(r_0))| = |S(r_0)| \geq \frac{M^{m-2}}{m} \geq \frac{n^{1-2/m}}{m 2^m} \underset{\substack{\text{optimizing} \\ m := \sqrt{\log_2 n}}}{\geq} n \exp(-c\sqrt{\log n})$$

So the real value for $r_3(n)$ is between $n \exp(-O(\sqrt{\log n}))$ and $O(n \frac{(\log \log n)^5}{\log n})$. In fact, one would wonder to eliminate the $(\log \log n)^5$ in the upper bound. This would have important consequences.

Another application: Erdős-Stone revisited

We will use the previous ideas related to the Szemerédi's Regularity lemma in order to get a proof of Erdős-Stone Thm. Recall that this theorem states the following.

Theorem / (Erdős-Stone, 46) For fixed r and s, $ex(n, K_{s,s,\overset{r}{\dots},s}) = \left(1 - \frac{1}{r-1}\right)\binom{n}{2} + o(n^2) = \left(1 - \frac{1}{r-1}\right)\frac{n^2}{2} + o(n^2)$

Or equivalently, that for all $\gamma > 0$ and fixed r and s, there is an integer N such that all graphs with $n \geq N$ vertices and at least $(1 - \frac{1}{r-1})\binom{n}{2} + \gamma n^2$ edges contains a copy of $K_{s,s,\overset{r}{\dots},s}$. To do so, we use similar ideas as the ones used in previous applications of the SRL.

Def / Let G be a graph with ε-regular partition $\{V_0, \dots, V_k\}$. Given $d \in [0,1]$, the <u>regularity graph</u> R_d of the partition $\{V_0, \dots, V_k\}$ is a graph with vertex set $V(R) = \{V_1, \dots, V_k\}$ and two vertices V_i, V_j are adjacent iff they form an ε-regular pair such that $d(V_i, V_j) \geq d$.

We will use a lot the following property that can be proved using the arguments used in the proof of counting lemma for triangles:

Lemma / Let (A, B) be an ε-regular pair with $d(A,B) \geq d$. Let $Y \subseteq B$ such that $|Y| \geq \varepsilon |B|$. Then the number of vertices of A with small degree into Y:

$$|\{v \in A : d_Y(v) < (d - \varepsilon)|Y|\}| < \varepsilon |A|$$

Proof / Assume the contrary ($\geq \varepsilon|A|$) and conclude a contradiction with ε-regularity.

Now we can start with the proof of Erdős-Stone. The strategy will be the following one:

1) Consider a ε-regular partition, and study the regularity graph (for some d)
2) Show that K_r is a subgraph of R.
3) Show that $K_r \subseteq R_d \Rightarrow K_{s,\overset{r}{\dots},s}$ is a subgraph of G (*)

Proof / Assume that G has $\left(1-\frac{1}{r-1}\right)\frac{n^2}{2}+\gamma n^2$, where γ is fixed and >0 (but arbitrary). We consider an ε-regular partition of G and the corresponding R_d, for some value of d. Both ε and d will be determined later.

with m fixed

Let us start estimating the number of edges we are not interested in.

a) Edges incident with V_0: $\leq \varepsilon \cdot n \cdot n = \varepsilon n^2$
b) Edges inside each V_i: $< k\binom{n/k}{2} < n^2/k < n^2/m$
c) Edges between non ε-regular pairs: $< \varepsilon k^2 \left(\frac{n}{k}\right)^2 < \varepsilon n^2$
d) Edges between ε-regular pairs with density $< d$: $< \binom{k}{2} d \left(\frac{n}{k}\right)^2 \leq dn^2$.

This is essentially $< (d+2\varepsilon+\frac{1}{m})n^2$. Recall that our graph has $\left(1-\frac{1}{r-1}\right)\frac{n^2}{2}+\gamma n^2$ edges, hence, if

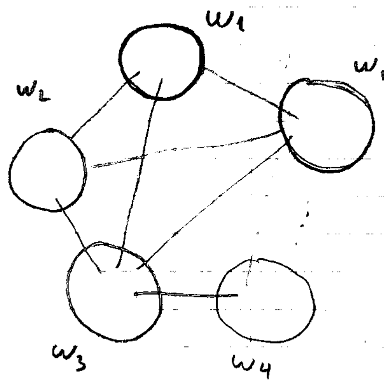
$$\boxed{d+2\varepsilon+\frac{1}{m}<\frac{\gamma}{2}} \Rightarrow \text{at least } \left(1-\frac{1}{r-1}\right)\frac{n^2}{2}+\frac{\gamma}{2}n^2 \text{ edges of } G \text{ go between sets } V_i, V_j$$

with $V_iV_j \in E(R_d)$.

On the other hand, the number of edges can be bounded by $|E(R_d)|\left(\frac{n}{k}\right)^2$. Hence,

$$\left(1-\frac{1}{r-1}\right)\frac{n^2}{2}+\frac{\gamma}{2}n^2 \leq |E(R_d)|\left(\frac{n}{k}\right)^2 \Rightarrow \left(1-\frac{1}{r-1}\right)\frac{k^2}{2}+\frac{\gamma}{2}k^2 \leq |E(R)|.$$

Now, as $m \leq k$, choosing $m := m(\gamma)$ such that $ex(m, K_r) \leq \left(1-\frac{1}{r-1}\right)\binom{m}{2}+\frac{\gamma}{2}m^2$ we assure that R_d contains K_r.

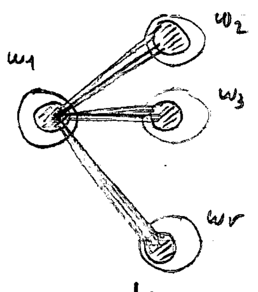


The next step now is to construct a copy of $K_{s,\ldots,s}$. We call the V_i involved in the copy of K_r $W_1, \ldots, W_r$, respectively.

We will find, one by one and inductively vertices $v_{1,1}, \ldots, v_{1,s} \in W_1$, $v_{2,1}, \ldots, v_{2,s} \in W_2, \ldots, v_{r,1}, \ldots, v_{r,s} \in W_r$ which define a $K_{s,\ldots,s}$.

Def / For $S \subseteq V(G)$, write $\Gamma_{W_i}(S) = \bigcap_{v \in S} N_{W_i}(v)$ (set of common neighbours of vertices of S in W_i)

• $v_{1,1}$: we want to choose $v_{1,1}$ in W_1 with big neighbourhoods in $W_2, \ldots, W_r$. We have that:



$$\left.\begin{array}{l} |\{v \in W_1 : d_{W_2}(v) < (d-\varepsilon)|W_2|\}| < \varepsilon |W_1| \\ |\{v \in W_1 : d_{W_3}(v) < (d-\varepsilon)|W_3|\}| < \varepsilon |W_1| \\ \vdots \\ |\{v \in W_1 : d_{W_r}(v) < (d-\varepsilon)|W_r|\}| < \varepsilon |W_1| \end{array}\right\} \Rightarrow \begin{array}{l}\text{There are at least } |W_1|-(r-1)\varepsilon|W_1| \\ \text{vertices in } W_1 \text{ with degree at least} \\ (d-\varepsilon)|W_i| \text{ in each } W_i.\end{array}$$

$$\Rightarrow \boxed{\varepsilon < \frac{1}{r-1}}.$$

• $v_{1,2}$: now we proceed similarly, but instead of taking $W_2, \ldots, W_r$ we must restrict ourselves to $\Gamma_{W_2}(\{v_{1,1}\}), \Gamma_{W_3}(\{v_{1,1}\}), \ldots, \Gamma_r(\{v_{1,1}\})$. We know that each $|\Gamma_{W_i}(\{v_{1,1}\})| \geq (d-\varepsilon)|W_1|$, but in order to apply the degree lemma, we need $|\Gamma_{W_i}(\{v_{1,1}\})| \geq \varepsilon |W_i|$. Hence, we need that $d-\varepsilon > \varepsilon$. In this case:

$$\left.\begin{array}{l} |\{v \in W_1 : d_{\Gamma_{W_2}(\{v_{1,1}\})}(v) < (d-\varepsilon)|\Gamma_{W_2}(\{v_{1,1}\})|\}| < \varepsilon |W_1| \\ \vdots \\ |\{v \in W_1 : d_{\Gamma_{W_r}(\{v_{1,1}\})}(v) < (d-\varepsilon)|\Gamma_{W_r}(\{v_{1,1}\})|\}| < \varepsilon |W_1| \end{array}\right\} \Rightarrow \begin{array}{l}\text{There are at least} \\ |W_1|-(r-1)\varepsilon|W_1| \text{ vertices} \\ \text{in } W_1 \text{ with degree at least} \\ (d-\varepsilon)|\Gamma_{W_i}(\{v_{1,1}\})| < \\ (d-\varepsilon)^2|W_i| \text{ in each } W_i.\end{array}$$

• $v_{1,j}$: we continue inductively up to the case in which we have choosen $v_{1,1},\dots,v_{1,j-1}$. Then, we study $\Gamma_{W_i}(\{v_{1,1},\dots,v_{1,j-1}\})$, and by the same means:

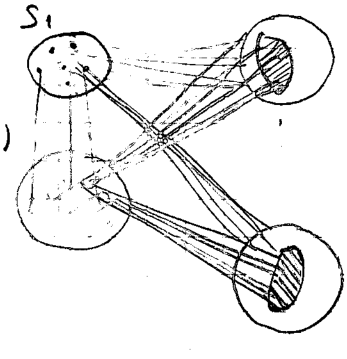
$$i=2,\dots,r \quad |\{v\in W_1 : d_{\Gamma_{W_i}(\{v_{1,1},\dots,v_{1,j-1}\})}(v) < (d-\varepsilon)\,|\Gamma_{W_i}(\{v_{1,1},\dots,v_{1,j-1}\})|\}| < \varepsilon|W_1|$$

We finally conclude that:

i) We can always pick $v_{1,i}$, if $\boxed{|W_1|-(r-1)\varepsilon|W_1|>s-1}$ $\Rightarrow$

ii) The sizes of the Γ_{W_i} are, in each step $>(d-\varepsilon)^{j-1}|W_1|>\varepsilon|W_1| \Rightarrow \boxed{(d-\varepsilon)^{s-1}>\varepsilon}$

We write $\{v_{1,1},\dots,v_{1,s}\}=S_1$; all of them have a big neighbourhood in $W_2,\dots,W_r$:



We proceed in the same way for W_2. Now, when we get the vertex $v_{2,s}$, we get the conditions:

a) $\boxed{(d-\varepsilon)^{t}|W_1|-(r-2)\varepsilon|W_1|>s-1}$

b) $\boxed{(d-\varepsilon)^{2s-1}>\varepsilon}$

So continuing the procedure, we get the last vertex $v_{r,s}$, and then we have the conditions:

$\boxed{(d-\varepsilon)^{(r-1)s}|W_1|>s-1}$; $\boxed{(d-\varepsilon)^{(r-1)s-1}>\varepsilon}$

So, now we need to see if all conditions we have found can be satisfied at the same time (recall that ε, d and m are not defined yet!)

1) $d+2\varepsilon+\frac{1}{m}<\frac{\gamma}{2}$ (for the condition in the begining)

2) $\varepsilon<(d-\varepsilon)<(d-\varepsilon)^2<\dots<(d-\varepsilon)^{(r-1)s-1}$ (Conditions on the sizes)

3) $(d-\varepsilon)^{(r-1)s}|W_1|-(r-1)\varepsilon|W_1|>s-1$. (This is stronger!)

So now take $d=\frac{\gamma}{6}$, $m=\frac{6}{\gamma}$ and $\varepsilon=\frac{1}{r}\left(\frac{d}{2}\right)^{s(r-1)}$, and one can check that 1)–3) is satisfied. So, fixed now ε, m and d, we have the parameter $M(m,\varepsilon)$ for the partition $\Rightarrow$ we have the value of vertices which is necessary to assure the condition.