Graph removal lemma

Bartosz Wodziński

Jagiellonian University

January 20, 2021

Motivation - number theory problems

Van der Waerden's theoem (1927)

In any partition $\mathbb{N} = C_1 \cup C_2 \cup \ldots \cup C_r$, some C_i contains arbitrarily long arithmetic progressions.

Definition

A subset A of the natural numbers is said to have positive upper density if $\limsup_{n\to\infty} \frac{|A \cap \{1, 2, \ldots, n\}|}{n} > 0$.

Roth's theorem (1953)

A subset of $\mathbb{N}$ with positive upper density contains a 3-term arithmetic progression.

Szeremédi's theorem (1975)

A subset of $\mathbb{N}$ with positive upper density contains a k-term arithmetic progression for every k.

Motivation number theory problems

Green–Tao theorem (2004)

Sequence of prime numbers contains arbitrarily long arithmetic progressions.

Proof uses Szeremédi's theorem.

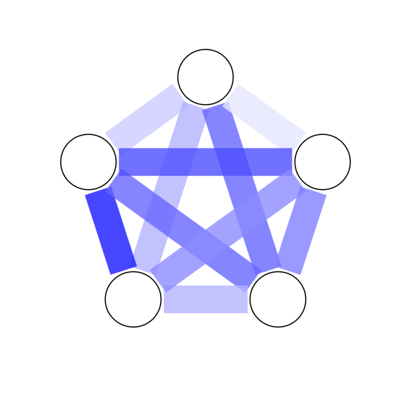
But how to proof Szeremédi's theorem in the first place?

The original proof used so-called *Szemerédi's Regularity Lemma* - very powerful tool in graph theory.

Szemerédi's Regularity Lemma

Informal statement

Vertices of a large enough graph can be partitioned into a bounded number of roughly equally-sized parts so that the edges between different parts behave almost randomly.



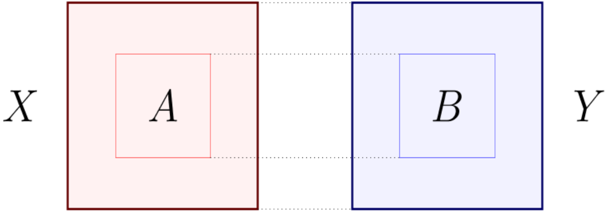
What do we mean by "almost randomly"?

Edge density

For disjoint vertex sets U and V, let $e(U, V)$ denote the number of edges with one endpoint in U and the other in V and let $d(U, V) := \frac{e(U,V)}{|U||V|}$ denote *edge density* between U and V.

ϵ - regularity

A bipartite graph between X and Y is said to be ϵ - regular if for all $A \subset X$ and $B \subset Y$ with $|A| \geq \epsilon|X|$ and $|B| \geq \epsilon|Y|$ we have $|d(A, B) - d(X, Y)| < \epsilon$.



X
A
B
Y

What do we mean by "almost randomly"?

ϵ - regular partition

A partition of vertices of a graph into k equal parts (± 1) is said to be ϵ - regular if all but ϵk^2 pairs of parts induce ϵ - regular bipartite graphs.

Szeremédi's Regularity Lemma

For every $\epsilon > 0$ there is some M such that for **every** graph has an ϵ-regular partition into at most M parts.

Intuition

Regularity lemma partitions the graph into some number of parts so that the graph behaves as if it was a random graph with prescribed densities between the parts.
For instance, if we would like to count the number of triangles between some 3 subsets X, Y, Z induced by lemma, we would expect that it should be close to $d(X, Y)d(Y, Z)d(Z, X)|X||Y||Z|$.

Counting triangles

Notation

Let d_{XY} denote $d(X, Y)$.

Triangle counting lemma

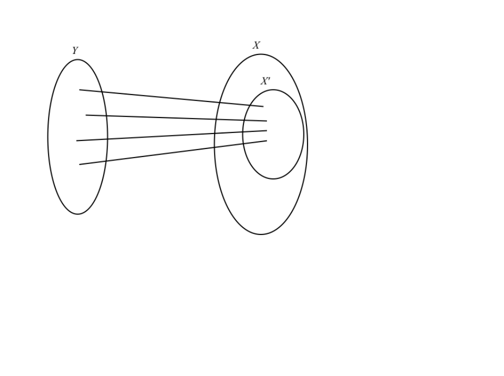
Suppose subsets X, Y, Z are pairwise ϵ - regular and d_{XY}, d_{YZ}, $d_{ZX} > 2\epsilon$. Then the number of triangles between subsets X, Y, Z is at least $(1-2\epsilon)(d_{XY}-\epsilon)(d_{YZ}-\epsilon)(d_{ZX}-\epsilon)|X||Y||Z|$.

Proof

There are less than $\epsilon|X|$ vertices in X that have $< (d_{XY}-\epsilon)|Y|$ neighbours in Y.

Proof

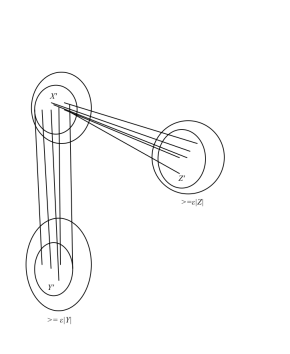
Suppose that it is not the case. Then There exists $X' \subset X$ such that $d_{X'Y} < \frac{(d_{XY}-\epsilon)|Y||X'|}{|X'||Y|} = d_{XY} - \epsilon$ - contradiction with ϵ - regularity of X and Y.



Counting triangles

Proof

From X, we remove $\epsilon|X|$ veritces with smallest degree to Y (Z). Now all have degree_to_Y $\geq (d_{XY} - \epsilon)|Y|$.
Since $d_{XY} > 2\epsilon$, $(d_{XY} - \epsilon)|Y| \geq \epsilon|Y|$. So $|Y'| \geq \epsilon|Y|$, $|Z'| \geq \epsilon|Z|$.
By regularity: #of $Y'Z'$ edges $\geq (d_{XY} - \epsilon)|Y|(d_{XZ} - \epsilon)|Z|(d_{YZ} - \epsilon)$.
Total # of Δ's $\geq (1 - 2\epsilon)|X|(d_{XY} - \epsilon)|Y|(d_{XZ} - \epsilon)|Z|(d_{YZ} - \epsilon)$.



Towards graph removal lemma

Triangle removal lemma

For every $\epsilon > 0$ there exists δ such that every graph with δn^3 triangles can be made triangle-free by removing $\leq \epsilon n^2$ edges.
So, if graph has $o(n^3)$ triangles then it can be made triangle-free by deleting $o(n^2)$ edges.

This lemma can be used to prove Roth's theorem about 3 - term arithmetic progressions.

Towards graph removal lemma

Proof

Take $\frac{\epsilon}{4}$ - regular partition $V_1 \cup ... \cup V_M$ (exists by Szemerédi Regularity Lemma).
Remove all edges between V_i, V_j if:

- (V_i, V_j) is not $\frac{\epsilon}{4}$ - regular. $(\leq \frac{\epsilon}{4}n^2)$
- $d_{V_iV_j} < \frac{\epsilon}{2}$ $(\leq \frac{\epsilon}{2}n^2)$
- V_i or V_j has at most $\frac{\epsilon n}{4M}$ vertices $(\frac{\epsilon n}{4M}Mn \leq \frac{\epsilon}{4}n^2)$

Total number of edges deleted: $\leq \epsilon n^2$.
Suppose some triangle remains. Take one with vertices in V_i, V_j, V_k.
Each pair in (V_i, V_j, V_k) is $\frac{\epsilon}{4}$- regular and has edge density $\geq \frac{\epsilon}{2}$.
So, we have Triangle Counting Lemma assumptions, so by applying it, we have that # of triangles is at least:
$\frac{1}{6}(1-\frac{\epsilon}{2})(\frac{\epsilon}{4})^3|V_i||V_j||V_k| = \frac{1}{6}(1-\frac{\epsilon}{2})(\frac{\epsilon}{4})^3(\frac{\epsilon n}{4M})^3 = an^3$. If we take $\delta < a$, we get a contradiction (there are more than δn^3 triangles, but we assumed there are $\leq \delta n^3$).

Graph removal lemma

Graph Removal Lemma

Let G be a graph with $o(n^h)$ subgraphs isomorphic to H (graph with h vertices). It is possible to eliminate all copies of H by removing $o(n^2)$ edges from G.

Additional information

- Regularity Lemma assures that we can partition the graph into at most M ϵ - regular parts. M depends on ϵ only, but turns out to be **huge**, i.e. $2^{2^{\cdots^{2}}}$, where the 2's tower has height $O(\epsilon^{-5})$.
- In practice, the weaker version of lemma is used to obtain $2^{O(\epsilon^{-2})}$ bound on M.
- Graph Removal Lemma can be used to graph property testing: either the graph is near H - free (we have to delete less than ϵn^2 edges to eliminate all copies of G), or it contains a lot of copies of H, that can be easily found by random sampling.

Bonus - Roth's Theorem proof

Lemma

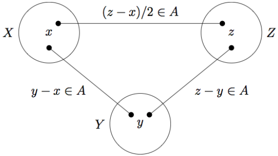
If G is a graph such that each edge in G belongs to exactly 1 triangle, then G has $o(n^2)$ edges.

Proof

Number of triangles in G is at most n^2 (maximum number of edges). Then by TRL ($n^2 \in o(n^3)$), it suffices to remove $o(n^2)$ edges from G to make it Δ-free. Since $\#\_of\_\Delta\_in\_G = \frac{|E|}{3}$, then $|E| \in o(n^2)$.

Bonus - Roth's Theorem proof

Take $A \subset [N] : |A| \geq \epsilon N$ and suppose it doesn't have 3-AP. View A as subset of $\mathbb{Z}_M$ group, where $M = 2N + 1$.
Create following graph G:



$|X|, |Y|, |Z| = M$. Notice that $y - x$, $\frac{z-x}{2}$, $z - y$ form a 3-AP.
If $\Delta \in G$, then A is not 3-AP-free, unless $y - x = \frac{z-x}{2} = z - y$. But then, it follows that $y = \frac{x+z}{2}$, so (x, y, z) forms 3-AP. Additionally, fixing 2 of them determines the third. So every edge in G lies in exactly 1 triangle.
Now, $E(G) \leq 3M|A|$ and from previous lemma $E(G) = o(M^2)$. It follows that $|A| \in o(N)$ - contradiction.