Mathematics 272 Lecture 5 Notes

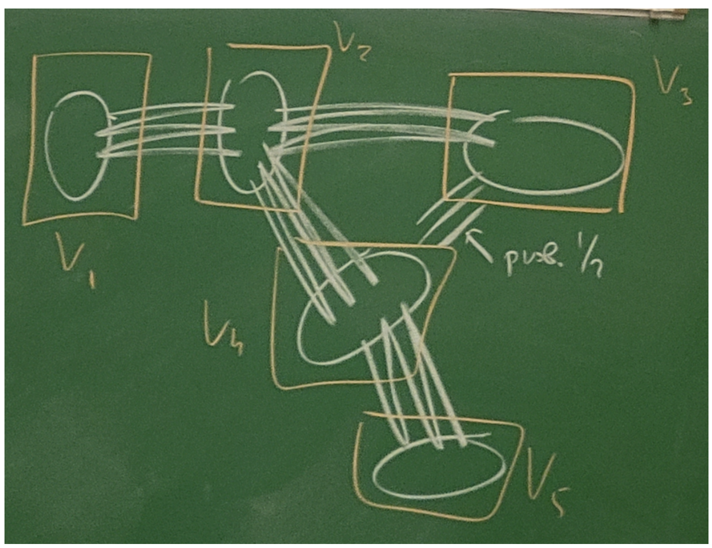
Professor: Dan Král
Scribe: Daniel Raban

February 4, 2025

1 Szemerédis Regularity Lemma and the Removal Lemma

1.1 Szemerédis regularity lemma

The idea behind Szemerédi's regularity lemma is that if we have a large graph with a number of separate pieces, then the most efficient partition is to pick each piece of the graph.



Let $e(X, Y)$ denote the number of edges between the sets of vertices X and Y.

Lemma 1.1 (Szemerédi's regularity lemma, 1978)**.** *For every $\varepsilon_R > 0$ and $k_0 \in \mathbb{N}$, there exists a K_0 such that for every graph G, we can find a partition $V(G) = V_0 \sqcup V_1 \sqcup \cdots \sqcup V_k$ with $k_0 \leq k \leq K_0$ satisfying the following properties:*

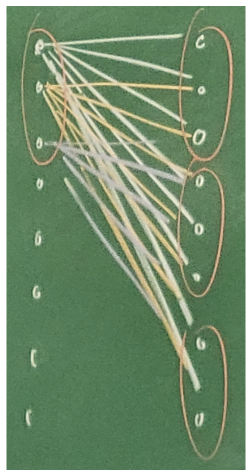
1. $|V_1| = \cdots = |V_k|$,
2. $|V_0| \leq \varepsilon_R |V(G)|$,

3. *All pairs $1 \le i < j \le k$ except $\le \varepsilon_R k^2$ pairs are such that for all subsets $A \subseteq V_i$, $B \subseteq V_j$ with $|A| \ge \varepsilon_R |V_i|$ and $|B| \ge \varepsilon_R |V_j|$,*

$$\left| \frac{e(A,B)}{|A||B|} - \frac{e(V_i,V_j)}{|V_i||V_j|} \right| \le \varepsilon_R.$$

*Here, V_i, V_j are said to be ε_R-**regular**.*

Example 1.1. Consider the following bipartite graph where the vertices on the left are connected to all vertices at the same level or below.



If we split up the vertices as above, then the edge densities between the different pieces of the graph are regular within each pair of pieces.

1.2 The removal lemma

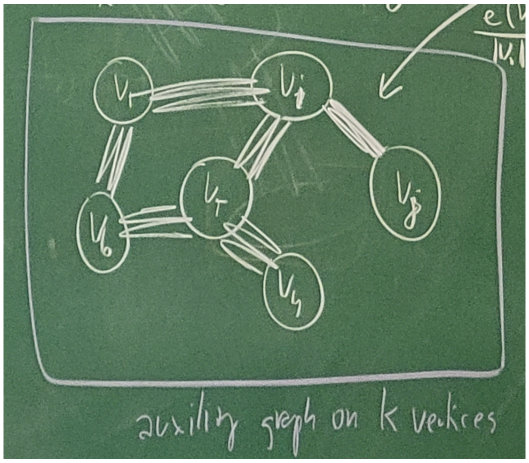
We want to derive the following statement, which is morally equivalent to counting the number of subgraphs with certain properties. The regularity lemma gives us control of various things, one of which is the density of subgraphs.

Lemma 1.2 (Removal lemma)**.** *For every $\varepsilon > 0$ and graph H, there exists a $\delta > 0$ such that every n-vertex graph G satisfies one of the following:*

1. *G has at least $\delta n^{|V(H)|}$ copies of H.*

2. *There exists a small set of edges $F \subseteq E(G)$ with $|F| \le \varepsilon n^2$ such that $G \setminus F$ is H-free.*

Proof. We will give the proof for $H = K_3$; the general case is left as an exercise (try proving it for $H = K_n$ first). We want to apply the regularity lemma, so we choose $\varepsilon_R = \frac{\varepsilon}{100}$ and $k_0 = \lceil \frac{100}{\varepsilon} \rceil$. Szemerédi's regularity lemma spits out some K_0, and based on this value, we will later specify δ as a function of K_0 and ε_R. Given an n-vertex graph G, the lemma gives us the partition $V_0 \sqcup V_1 \sqcup \cdots \sqcup V_k$ with regular pairs (V_i, V_j). We then define an auxilliary graph with the V_i as the vertices, connecting V_i, V_j by an edge if $\frac{e(V_i,V_j)}{|V_i||V_j|} \geq \frac{\varepsilon}{10}$.



Case 1: If the auxiliary graph does not contains any copies of K_3, then let F contain the following edges:

- edges between pairs of parts that are not ε_R-regular

$$\leq \varepsilon_R k^2 |V_1|^2 \leq \varepsilon_R k^2 \left(\frac{n}{k}\right)^2 = \varepsilon_R n^2 = \frac{\varepsilon n^2}{100} \text{ many of these}$$

- edges between pairs V_i, V_j that have $\frac{e(V_i,V_j)}{|V_i||V_j|} < \frac{\varepsilon}{10}$

$$\leq \binom{k}{2} \frac{\varepsilon}{10} |V_1|^2 \leq k^2 \frac{\varepsilon}{10} \left(\frac{n}{k}\right)^2 = \frac{\varepsilon n^2}{10} \text{ many of these}$$

- edges within the same part V_i

$$\leq k \binom{|V_1|}{2} \leq k \left(\frac{n}{k}\right)^2 \leq \frac{n^2}{k_0} \leq \frac{\varepsilon n^2}{100} \text{ many of these}$$

- edges incident to a vertex in V_0

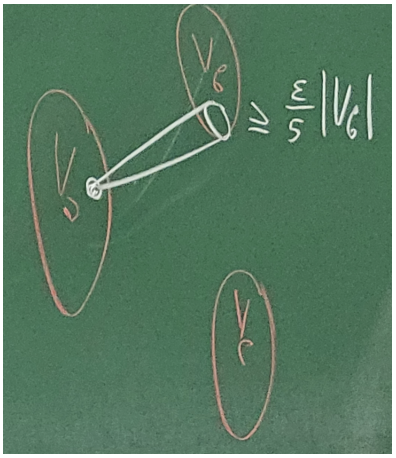
$$\leq |V_0| n \leq \varepsilon_R n^2 \leq \frac{\varepsilon n^2}{100} \text{ many of these}$$

If we remove all these edges and there are no triangles in the auxiliary graph, then there can be no triangles left in the graph.

Case 2: If the auxiliary graph does contain a copy of K_3, then there exists V_a, V_b, V_c with $a < b < c$ such that every pair of them is ε_R-regular and has density $\geq \frac{\varepsilon}{10}$. The number of vertices is a constant fraction of n:

$$|V_a| = |V_b| = |V_c| \geq (1 - \varepsilon_R)\frac{n}{K_0} \geq \frac{n}{2K_0}.$$

We claim that at least $(1 - \varepsilon_k)|V_a|$ vertices of V_a have at least $\frac{\varepsilon}{50}|V_b|$ neighbors in V_b.



Suppose not. Then let $A = \{w \in V_a \text{ with } \leq \frac{\varepsilon}{50}|V_b| \text{ neighbors in } V_b\}$ and $B = V_b$. Then

$$e(A, B) \leq \frac{\varepsilon}{50}|A||B|,$$

which gives

$$\frac{e(A, B)}{|A||B|} \leq \frac{\varepsilon}{50}.$$

On the other hand, we already know that

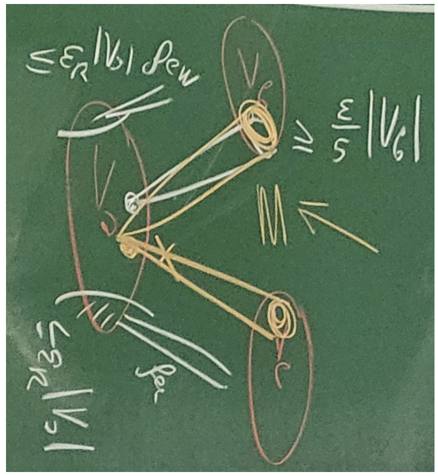
$$\frac{e(V_a, V_b)}{|V_a||V_b|} \geq \frac{\varepsilon}{10},$$

which gives a contradiction.

Similarly, we also have at least $(1 - \varepsilon_R)|V_a|$ vertices of V_a with at least $\frac{\varepsilon}{50}|V_c|$ neighbors in V_c. Putting these two statements together, we have at least $(1 - 2\varepsilon_k)|V_a|$ vertices of V_a with at least $\frac{\varepsilon}{50}|V_b|$ neighbors in V_b and $\frac{\varepsilon}{50}|V_c|$ neighbors in V_c.

To make a triangle, the number of choices for a vertex x in V_a is $\geq (1 - 2\varepsilon_R)|V_a|$. Let

X_b be the neighbors of x in V_b, and let X_c be the neighbors of x in V_c.



The number of edges between X_b and X_c is $\geq (\frac{\varepsilon}{10} - \varepsilon_R)|X_b||X_c|$.

$$\begin{aligned}
\#\text{ triangles} &\geq (1 - 2\varepsilon_R)|V_a| \left(\frac{\varepsilon}{10} - \varepsilon_R\right) |X_b||X_c| \\
&\geq \frac{1}{2} \frac{n}{2K_0} \frac{\varepsilon}{20} \left(\frac{\varepsilon}{50} \frac{n}{K_0}\right)^2 \\
&= \frac{\varepsilon^3}{200000K_0^3} n^3.
\end{aligned}$$

so we get the result with $\delta = \frac{\varepsilon^3}{200000K_0^3}$. $\square$

Remark 1.1. The same proof gives that

$$\#\text{ triangles} \approx \frac{e(V_a, V_b)}{|V_a||V_b|} \frac{e(V_a, V_c)}{|V_a||V_c|} \frac{e(V_b, V_c)}{|V_b||V_c|} |V_a||V_b||V_c|.$$

This is what we mean by the removal lemma being morally equivalent to counting subgraphs.

Next time, we with prove Roth's theorem.