Regularity Lemma and its applications

Miklós Simonovits

Moscow, 2015
The PLAN (?)

Generalized Random Graph Sequences
- \( \varepsilon \)-regular pairs
- Generalized Quasi-Random Sequences

The Szemerédi Regularity Lemma
Why do we like Szemerédi Regularity Lemma?
- Cluster graph

Applications: \( \text{RT}(n, K_4) \leq \frac{1}{8} n^2 + o(n^2) \)
- Removal Lemma

Ruzsa-Szemerédi Theorem, and its importance

The plan was to prove \( r_k(n) = o(n^2) \), using Extremal Hypergraph Theory
but Tim Gowers cam!
Extensions

- Sparse Regularity Lemmas: Kohayakawa-Rödl
- Weak Hypergraph Regularity Lemmas: Frankl-Rödl
- Strong Hypergraph Regularity Lemmas: Rödl-Nagle-Skokan-Schacht
- and some newer ones, Tao ...
Skipping, among others:

- Algorithmic aspects
- Connections to Property testing
- Weak Regularity Lemma Frieze-Kannan, …
Szemerédi Regularity Lemma

- Origins/connections to the existence of arithmetic progressions in dense sequences
- Connection to the quantitative Erdős-Stone theorem
- First graph theoretic applications
  (Ruzsa-Szemerédi theorem, Ramsey-Turán problems)
- Counting lemma, removal lemma, coloured regularity lemma
Given $G$, with $X$ and $Y$, the edge-density between $X$ and $Y$ is

$$d(X, Y) := \frac{e(X, Y)}{|X||Y|}.$$
Regular pairs are highly uniform bipartite graphs, namely ones in which the density of any reasonably sized subgraph is about the same as the overall density of the graph.

**Definition (ε-regular set-pairs)**

Let $\varepsilon > 0$. Given a graph $G$ and two disjoint vertex sets $A \subset V$, $B \subset V$, we say that the pair $(A, B)$ is $\varepsilon$-regular if for every $X \subset A$ and $Y \subset B$ satisfying

$$|X| > \varepsilon|A| \text{ and } |Y| > \varepsilon|B|$$

we have

$$|d(X, Y) - d(A, B)| < \varepsilon.$$
Generalized random graphs

Given a probability matrix $A := (p_{ij})_{r \times r}$ and integer $n_1, \ldots, n_r$.

- We choose the subsets $U_1, \ldots, U_r$ and join $x \in U_i$ to $y \in U_j$ with probability $p_{ij}$ independently.

- Regularity Lemma: the graphs can be approximated by generalized random graphs well.
The Regularity Lemma says that

- every dense graph can be partitioned into a small number of regular pairs and a few leftover edges.

- Since regular pairs behave as random bipartite graphs in many ways, the R.L. provides us with an approximation of an arbitrary dense graph with the union of a constant number of random-looking bipartite graphs.
Theorem (Szemerédi, 1978)

For every $\varepsilon > 0$ and $m$ there are $M(\varepsilon, m)$ and $N(\varepsilon, m)$ with the following property: for every graph $G$ with $n \geq N(\varepsilon, m)$ vertices there is a partition of the vertex set into $k$ classes

$$V = V_1 + V_2 + \ldots + V_k$$

such that

- $m \leq k \leq M(\varepsilon, m)$,
- $||V_i| - |V_j|| < 1$, $(1 \leq i < j \leq k)$
- all but at most $\varepsilon k^2$, of the pairs $(V_i, V_j)$ are $\varepsilon$-regular.

See → SzemRegu, KomSim
The role of $m$

is to make the classes $V_i$ sufficiently small, so that the number of edges inside those classes are negligible. Hence, the following is an alternative form of the R.L.

**Theorem (Regularity Lemma – alternative form)**

*For every $\varepsilon > 0$ there exists an $M(\varepsilon)$ such that the vertex set of any $n$-graph $G$ can be partitioned into $k$ sets $V_1, \ldots, V_k$, for some $k \leq M(\varepsilon)$, so that

- $|V_i| \leq \lceil \varepsilon n \rceil$ for every $i$,
- $||V_i| - |V_j|| \leq 1$ for all $i, j$,
- $(V_i, V_j)$ is $\varepsilon$-regular in $G$ for all but at most $\varepsilon k^2$ pairs $(i, j)$.*

For $e(G_n) = o(n^2)$, the Regularity Lemma becomes trivial.
How to prove Regularity Lemma?

- Use the Defect form of Cauchy-Schwarz.

Index:

$$I(\mathcal{P}) = \frac{1}{k^2} \sum d(V_i, V_j)^2 < \frac{1}{2}.$$
Lemma (Improved Cauchy-Schwarz inequality)

If for the integers $0 < m < n$,

$$
\sum_{k=1}^{m} X_k = \frac{m}{n} \sum_{k=1}^{n} X_k + \delta,
$$

then

$$
\sum_{k=1}^{n} X_k^2 \geq \frac{1}{n} \left( \sum_{k=1}^{n} X_k \right)^2 + \frac{\delta^2 n}{m(n - m)}.
$$
Coloured Regularity Lemma

If we have several colours, say, Black, Blue, Red, then we have a Szemerédi partition good for each colour simultaneously.

How to apply this?
**Counting Lemma**

Through a simplified example:

- If the Generalized Random Graph corresponding to $G_n$ contains many copies of $L$, then $G_n$ also contains many (approximately the same number of copies of $L$)
- If the reduced graph contains an $L$ then $G_n$ contains at least $cn^{\nu(L)}$ copies of $L$. 
The classes $V_i$ will be called groups or clusters. Given an arbitrary graph $G = (V, E)$, a partition $P$ of the vertex-set $V$ into $V_1, \ldots, V_k$, and two parameters $\varepsilon, d$, we define the Reduced Graph (or Cluster Graph) $R$ as follows: its vertices are the clusters $V_1, \ldots, V_k$ and $V_i$ is joined to $V_j$ if $(V_i, V_j)$ is $\varepsilon$-regular with density more than $d$.

Most applications of the Regularity Lemma use Reduced Graphs, and they depend upon the fact that many properties of $R$ are inherited by $G$. 
**Inheritance**

\( G_n \) inherits the properties of the cluster graph \( H_k \).
- sometimes in an improved form!

Through a simplified example:

- If \( H_k \) contains a \( C_7 \) then \( G_n \) contains many: \( cn^7 \).
## Ramsey-Turán problems

### Theorem (Szemerédi)

<table>
<thead>
<tr>
<th>→ $\text{SzemRT}$</th>
<th>If $G_n$ does not contain $K_4$ and $\alpha(G_n) = o(n)$ then</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$e(G_n) = \frac{n^2}{8} + o(n^2)$.</td>
</tr>
</tbody>
</table>

How to prove this?
- Use Regularity Lemma
- Show that the reduced graph does not contain $K_3$.
- Show that the reduced graph does not contain

$$d(V_i, V_j) > \frac{1}{2} + \varepsilon$$
Szemerédi–Ruzsa

\( f(n, 6, 3) \)
Removal Lemma

Originally for $K_3$, Ruzsa-Szemerédi

Generaly: through a simplified example:

For every $\varepsilon > 0$ there is a $\delta = \delta(\varepsilon) \to 0$, $\delta > 0$:
If a $G_n$ does not contain $\delta n^{10}$ copies of the Petersen graph, then we can delete $\varepsilon n^2$ edges to destroy all the Petersen subgraphs.

something similar is applicable in Property testing.
The Cluster graph, illustrated:
The Cluster graph, illustrated:
The Cluster graph, illustrated:
The Cluster graph, illustrated:
The Cluster graph, illustrated:
Encoding? Logarithm? Generating function?

- Original Graph, satisfying some $\mathcal{P}$.
- Cluster graph $H_k$ satisfying some $\mathcal{P}'$ and having proportionally many edges.
- Solving the corresponding problem for $H_k$.
- Translating the result for $G_n$. 
How to prove Erdős-Stone?

- No $K_{p+1}$ in the Reduced graph $H_k$
- Apply Turán’s theorem
- Estimate the edges of the original graph:

$$e(G_n) \leq e(H_k)m^2 + 3\varepsilon n^2.$$
How to prove Stability?

- No $K_{p+1}$ in the Reduced graph $H_k$
- Apply Turán’s theorem with stability (Füredi)
- Estimate the edges of the original graph
The largest $K_3$ subgraph of a Random graph is its largest bipartite subgraph

*de Marco–Jeff Kahn*
Various Regularity Lemmas

- Original Ugly
- Original Nice
- Weak Regularity Lemma
  - Frieze-Kannan
  - Connections to Statistical approach
- Weak Hypergraph Regularity
- Good Hypergraph Regularity: Rödl, ... Schacht, Gowers
How to get rid of Regularity Lemma?

and why????

The thresholds are too large
But Regularity Lemma often makes the things transparent

— See Luczak: Odd cycle Ramsey
Blowup Lemma

Komlós, G. Sárkőzy, Szemerédi:

Good to prove the existence of spanning subgraphs

Pósa-Seymour conjecture, . . .

$(A, B)$ is $(\varepsilon, \delta)$-super-regular if for every $X \subset A$ and $Y \subset B$ satisfying

$$|X| > \varepsilon|A| \text{ and } |Y| > \varepsilon|B|$$

we have

$$e(X, Y) > \delta|X||Y|,$$

and

$$\deg(a) > \delta|B| \text{ for all } a \in A,$$

and $\deg(b) > \delta|A| \text{ for all } b \in B.$
Theorem

Given a graph $R_r$ and $\delta, \Delta > 0$, there exists an $\varepsilon > 0$ such that the following holds. $N$ = arbitrary positive integer,

- replace the vertices of $R$ with pairwise disjoint $N$-sets $V_1, V_2, \ldots, V_r$.

- Construct two graphs on the same $V = \bigcup V_i$. $R(N)$ is obtained by replacing all edges of $R$ with copies of $K_{N,N}$, and a sparser graph $G$ is constructed by replacing the edges of $R$ with $(\varepsilon, \delta)$-super-regular pairs.

If $H$ with $\Delta(H) \leq \Delta$ is embeddable into $R(N)$ then it is already embeddable into $G$. 
**Other Regularity Lemmas**

- **Frieze-Kannan**
  Background in statistics, more applicable in algorithms

- **Lovász-B. Szegedy**: Limit objects, continuous version

- **Alon-Fischer-Krivelevich-M. Szegedy**: Used for property testing

- **Alon-Shapira**: Property testing is equivalent to using Regularity Lemma
Szemerédi’s Lemma for the Analyst

This is the title of a paper of L. Lovász and B. Szegedy

Hilbert spaces, compactness, covering
## Hypergraph regularity lemmas

- Frankl-Rödl
- Frankl-Rödl 2.
- F. Chung
- A. Steger
- Rödl, Skokan, Nagle, Schacht,…
- Gowers, Tao,…
Many thanks for your attention.