MTH 394A
Additive Combinatorics and Graph Theory: The Szemerédi Regularity Lemma

Anurag Sahay
11141

13th April, 2014

About the Project

This project, done under the supervision of Prof. Rajat Mittal, Dept. of CSE is an exploration of the field of Additive Combinatorics. Additive Combinatorics is a relatively new field of mathematics which has deep connections to fields such as number theory, graph theory, fourier analysis and ergodic theory. This presentation is focused in particular at one part of the project, which is the tool known as Szemerédi's Regularity Lemma. The project was done jointly with Mr. Vijay Keswani, an undergraduate in the Dept. of CSE. The department internal supervisor was Prof. Shobha Madan.

The Field of Additive Combinatorics

The deep connections between Additive Combinatorics and its allied fields arise from the similarity of the arguments that are used in the fields. Despite the fact that the fields are as varied as ergodic theory, number theory and graph theory, the proof methods have surprising similarities. A major similarity is that the notion of "randomness" plays a crucial role in all forms of the subject. What notion of randomness is being used largely depends on the context.

For example in the context of the Regularity Lemma, we will be defining the notion of ϵ-regularity for graphs.

Some motivating theorems

One of the problems that motivated my interest in the field of Additive Combinatorics is Szemerédi's theorem, a theorem which has many proofs, including ones that are fourier-analytic in nature, combinatorial in nature and ergodic-theoretic in nature. We briefly state a toy version of this theorem:

Theorem (Roth's Theorem)

For every $1 \geq \delta > 0$, there exists an N_0 such that for every $N \geq N_0$ and $A \subset [1, N]$ with $|A| = \delta N$, we must have that A contains a non-trivial arithmetic progression of length 3. (3-AP)

Replacing 3 here with an arbitrary k gives the general case of Szemerédi's theorem.

Some motivating theorems

The original Regularity Lemma was proven to establish the general case of the previous theorem.

We shall now set up the basic requirements to state the Regularity Lemma.

Notation and Definitions

First we fix some notation. For us, a graph $G = (V, E)$ is simple and undirected, with vertex set V and edge set E. For two disjoint subsets A and B of V, we define $e(A, B)$ to be the number of edges across the two sets.

Definition (Edge Density)

For disjoint subsets A and B as above, we define the edge density as follows:

$$d(A, B) = \frac{e(A, B)}{|A||B|}$$

Definition (ϵ-regular pair (A, B))

The above pair (A, B) is called ϵ-regular if for any $X \subset A$ and $Y \subset B$ with $|X| \geq \epsilon|A|$ and $|Y| \geq \epsilon|B|$, we have that

$$|d(A, B) - d(X, Y)| < \epsilon$$

Notation and Definitions

Definition (ϵ-regular partition)

A partition $V = V_0 \cup V_1 \cup ..... \cup V_k$ is called ϵ-regular if

- $|V_0| \leq \epsilon |V|$
- $|V_1| = |V_2| = \cdots = |V_k|$
- At most ϵk^2 pairs of (V_i, V_j) are not ϵ-regular

Why does this give a suitable notion of a random graph?

Statement of the Regularity Lemma

Theorem (Szemerédi's Regularity Lemma)

$\forall \epsilon > 0$ and any integer t, there exists integers $T(\epsilon, t)$ and $N(\epsilon, t)$, for which every graph with atleast $N(\epsilon, t)$ vertices has an ϵ-regular partition $(V_0, V_1, \ldots.., V_k)$, where $t \leq k \leq T(\epsilon, t)$.

The crucial aspect of this theorem is the fact the k given here is bounded above.

In other words, this theorem could be taken to mean that any large enough graph can "roughly" be decomposed into boundedly many equi-sized clusters, which "roughly" behave "randomly" with each other.

Proof Method

The details of the proof of the Regularity Lemma are long and arduous, however we can present the idea of the proof here,. This proof idea has other applications, and in fact comes up somewhere in the proof of the general Szemerédi's theorem in all known proofs.

The idea of the proof is as follows.

1. Define an appropriate notion and measure of randomness.
2. Define an appropriate density/energy function. Prove that this function has an upper bound.
3. Show that in the random case, whatever theorem we want to prove holds easily.
4. Show that if the situation is not random, a new situation can be constructed such that both the appropriate density/energy function increases and the amount of randomness increases.

Proof Method

We now give an intuitive idea for how this proof idea is put into action in this particular case.

1. Start with an arbitrary equipartition $V_1, V_2, \cdots, V_t$.
2. Define energy of a partition appropriately.
3. If the partition is not ϵ-regular, then there exists many pairs in which the density varies greatly for subsets.
4. Construct an appropriate refinement of this partition such that the deviation in density decreases, and the energy function increases.
5. Show that the energy function is bounded above.

Applications of the Regularity Lemma

We now state a few applications of the Regularity Lemma that we looked into.

Theorem (Triangle Removal Lemma)

For every $\epsilon > 0$, there is a $\delta > 0$ such that if G is a simple and undirect graph which can be made free of triangles (that is, a subgraph isomorphic to K_3) by making $\geq \epsilon n^2$ edge deletions, then G has $\geq \delta n^3$ triangles.

Theorem (Roth's Theorem)

For every $1 \geq \delta > 0$, there exists an N_0 such that for every $N \geq N_0$ and $A \subset [1, N]$ with $|A| = \delta N$, we must have that A contains a non-trivial arithmetic progression of length 3. (3-AP)

Y. Lima, *Szemeredi's Regularity Lemma*, http://www2.math.umd.edu/~yurilima/szemeredi_regularity_lemma.pdf

T. Tao, V. Hu, *Additive Combinatorics*, Cambridge Studies in Advanced Mathematics, **105**, Cambridge University Press.