# Additive Combinatorics and Szemerédi's Regularity Lemma

Vijay Keswani
Anurag Sahay

20th April, 2015

**Supervised by :** Dr. Rajat Mittal

# Contents

# 1 Introduction

This project, done under the supervision of Prof. Rajat Mittal, 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 project is focused in particular on two topics, Sumset Estimates and Szemerédi's Regularity Lemma. The project was done jointly with Mr. Anurag Sahay, an undergraduate in the Dept. of Mathematics.

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. In this project, we explore this notion in a graph-theoretic setting, while studying Szemerédi's Regularity Lemma.

Throughout this project, we referred to multiple surveys and expostions. For Sumset Estimates, the primary references were [1] [2]. For Szemerédi's Regularity Lemma, we mainly referred to [6].

## 2 Sum-set Estimates

In this section, given additive sets $A$ and $B$ which are subsets of some additive group, say $G$, we will analyze the sets $A + B$, $A - B$ and related objects. A sumset $A+B$ (and correspondingly, difference set) is defined as,

$$A + B := \{a + b \mid a \in A, b \in B\}$$

Our main objective is to define various measures which would help us estimate the additive properties of $A+B$ in a very general manner. For example, we would like to know under what conditions is $A + B$ small or big, given $A$ and $B$.

### 2.1 Size of sumset and Doubling constants

To this end, the first measure which one would like to analyze would be the size of the set $A + B$.
Trivially we can observe the following inequalities,

$$|A + x| = |A| \text{ where } x \in Z$$

$$max(|A|, |B|) \leq |A + B| \leq |A||B|$$

$$|A| \leq |A + A| \leq \frac{|A|(|A| + 1)}{2}$$

We will refer to the set $A + x$, for some $x \in G$, as a *translate* of the set $A$. In general one can observe that the sizes of $A+B$ or $A-B$, remains unaffected if one translates $A$ or $B$ by an arbitrary amount.

Also one can see that if $A$ is a subgroup than $A + A = A$.

To get an intuition into the sizes of the sumsets, we take an example which we will freely use throughout this section.

Consider an arithmetic progression $A$. The sumset $A+A$ will mostly contain elements of $A$ itself. Hence, we can see that $|A+A| \sim c|A|$, for some constant $c$. Next consider an geometric progression $B$. Here the sumset $B + B$ will contain relatively few elements elements from the set $B$. Infact, one can see that $|B + B| \sim \frac{|B|(|B|+1)}{2}$.

The analysis of the above the two examples of an A.P. and a G.P. show that they lie on two opposite of what we can call an *additive spectrum*.

The size of the sumset, even though captures to an extent how *additive* two sets are together, does not present the whole picture, and we will examples later as to why this is indeed not such a good measure.

The next measure that we define is the doubling constant.

**Definition 2.1** (Doubling Constant)**.** For an additive set $A$, the doubling constant $\sigma[A]$ is defined as the following quantity,

$$\sigma[A] := \frac{|A+A|}{|A|}$$

From the inequalities in the previous page, we can see that,

$$1 \leq \sigma[A] \leq \frac{|A|+1}{2}$$

The upper bound is easy to see from an example of a GP. Also $\sigma[A] = 1$, if and only if $A$ is coset of a finite subgroup of $Z$. In general, the doubling contant intuitively tell us how close the set is to being a group, ie, it behaves "approximately" as a group if the doubling constant is small.

Next we will look at a set of inequalities and measures which further analyze the additive structure of sets. This set of theorems, developed by Ruzsa, is commonly referred to as Ruzsa calculus. The following section on Ruzsa calculus has been included for the sake of completeness and have been taken verbatim from [8].

## 2.2 Rusza Calculus

The fundamental result in Rusza Calculus is the triangle inequality, viz.

**Theorem 2.1** (Rusza triangle inequality)**.** *Let $G$ be an abelian group, and $A, B, C \subset G$. Then we have the following inequality among cardinalities:*

$$|A||B-C| \leq |A+B||A+C|$$

*Proof.* For any $x \in B - C$, fix a representation $x = b - c = b(x) - c(x)$ with $b \in B$ and $c \in C$. Now define a map $f : A \times (B - C) \to (A + B) \times (A + C)$ as $f(a, x) = (a + b, a + c)$.

Now, suppose $f(a,x) = f(a',x')$. Thus, $x = b - c = (a+b) - (a+c) = (a'+b') - (a'+c') = b' - c' = x'$. Since we fixed a representation, this implies that $b = b'$ and $c = c'$, and hence $a = a'$.

Thus, $f$ is an injective map. Comparing the cardinalities of the domain and co-domain, the theorem follows.

$\square$

To see why this theorem is called a triangle inequality, we first define the following notion of distance between subsets of a group:

**Definition 2.2** (Rusza distance)**.** The *Rusza distance* $d(A,B)$ between two sets $A, B \subset G$ is defined as

$$d(A,B) = \log \frac{|A-B|}{|A|^{1/2}|B|^{1/2}}$$

It is easy to see that this distance is symmetric since $|A-B| = |B-A|$. Further, note that the triangle inequality

$$d(A,C) \leq d(A,B) + d(B,C)$$

can be rewritten, by taking exponentials both sides to

$$\frac{|A-C|}{|A|^{1/2}|C|^{1/2}} \leq \frac{|A-B|}{|A|^{1/2}|B|^{1/2}} \times \frac{|B-C|}{|B|^{1/2}|C|^{1/2}}$$

or, in other words,

$$|A-C||B| \leq |A-B||B-C|$$

This is equivalent to the previous theorem (which can be seen by replacing $(A,B,C)$ in the previous theorem by $(-B,A,C)$).

Thus, the previous theorem is actually equivalent to the statement that the Rusza distance defined above satisfies the triangle inequality.

The Rusza disance is, however, clearly not reflexive in the general case.

The Rusza distance is a very useful tool for proving general inequalities. In particular, it allows us to connect the notion of sets that grow slowly under addition and substraction. For example, we have the following theorem:

**Theorem 2.2.** *If $|A+A| \leq K|A|$ for some absolute constant $K$, then we have $|A-A| \leq K^2|A|$. Conversely, $|A-A| \leq K|A|$ implies $|A+A| \leq K^2|A|$.*

*Proof.* Note that

$$\frac{|A-A|}{|A|} = \exp(d(A,A)) \leq \exp(d(A,-A) + d(-A,A)) = \frac{|A+A|^2}{|A|^2}$$

and that

$$\frac{|A+A|}{|A|} = \exp(d(A,-A)) \leq \exp(d(A,A) + d(-A,-A)) = \frac{|A-A|^2}{|A|^2}$$

Both these inequalities together with the respective hypothesis give the desired conclusion.

□

For a given $K$, a set satisfying $|A+A| \leq K|A|$ is said to be a set of small doubling. The expectation is that if $A$ has small doubling, then in fact, all possible additions and substractions of $A$ with itself must be small (since there must be inherent structure in $A$ of some sort). The formal result is by Plünneke and Rusza:

**Theorem 2.3** (Plünneke-Rusza inequality)**.** *Let $G$ be an abelian group, and $A, B \subset G$ be sets of equal size satisfying $|A+B| \leq K|A|$. Then we have*

$$|kA - lA| \leq K^{k+l}|A|$$

## 2.3 Covering Lemmas

Covering lemmas, roughly speaking say that, if $A$ and $B$ have similar additive structure, then one can cover $A$ by a small number of translates of $B$.

**Lemma 2.1** (Ruzsa Covering lemma [1])**.** *For any additive sets $A, B$, there exists an additive set $X \subseteq B$ with*

$$B \subseteq A - A + X$$

$$|X| \leq \frac{|A+B|}{|A|}$$

$$|A + X| = |A||X|$$

*In particular, $B$ can be covered by $\frac{|A+B|}{|A|}$ translates of $A - A$.*

*Proof.* Consider the family $\{A + b : b \in B\}$ of translates of $A$. We take the **maximal disjoint** sub-family, say $\{A + x : x \in X\}$, for $X \subseteq B$ and prove that this $X$ satisfies all the above conditions.

Each $A + x$ has size $|A|$ and is a subset of $A + B$. Therefore, $|X| \leq \frac{|A+B|}{|A|}$.

Since all sets in $\{A + x : x \in X\}$ are disjoing, clearly, $|A + X| = |A||X|$.

Now for any element $b \in B$, $(A + b) \cap (A + x)$ must not be null for atleast one $x \in X$, since otherwise the condition for maximality would be violated. Therefore, $b$ must be present in the set $A - A + X$, for all $b \in B$. Hence, $B \subseteq A - A + X$.

□

We next give a modification of Ruzsa covering lemma, which is a stronger statement than the above lemma.

**Lemma 2.2** (Green-Ruzsa Covering lemma [1])**.** *For any additive sets $A, B$, there exists an additive set $X \subseteq B$ with $|X| \leq 2\frac{|A+B|}{|A|} - 1$ such that $A - A + X$ covers $B$ with multiplicity atleast $|A|/2$.*

Informally, the lemmas says that for each $y in B$, there are atleast $|A|/2$ triplets $(x, a, a')$, such that $y = x + a - a'$.

*Proof.* The proof for this lemma is constructive.
We initialise $X$ to be the empty set, and in each iteration we check if there exists $y \in B$ for which $|(y + A) \cap (X + A)| \leq |A|/2$. If such a $y$ exists, we add it to $X$, else we terminate the loop.

In the first step, we add $|A|$ elements to $X+A$. In all the other iterations, we add atleast $|A|/2$ new elements to the set $X+A$. Since $|X+A|$ is bounded by $|A+B|$,

$$|X| \leq 2\frac{|A+B|}{|A|} - 1$$

By our construction, for every $yinB$, $|(y+A) \cap (X+A)| \geq |A|/2$ and so we will have atleast $|A|/2$ triplets $(x, a, a')$, such that $y = x + a - a'$. □

Covering lemmas tell us a lot about the additive structure of $A+B$. For example, from the Ruzsa covering lemma,

$$A + B \subseteq A + X$$

for some $X \subseteq B$.

In particular, if $X$ is small, it shows that $A+B$ does not grow much under addition and at the same time we also get an alternate small representation for it, and so are useful in establishing sumset properties.

Before moving further we quantify our earlier notion of approximate groups, which we said, have a small doubling constant.

**Definition 2.3** (Approximate groups)**.** Let $K > 1$. An additive set $H$ is said to be $K$-approximate group if it is symmetric, i.e., $H = -H$, contains the origin and $H+H$ can be covered by atmost $K$ translates of $H$.

Observe that a 1-approximate group is a finite group and vice-versa.

Having introduced the notion of Ruzsa distance, covering lemmas and approxiamte groups, we look at the closely related measure called Additive Energy.

## 2.4 Additive Energy

**Definition 2.4** (Additive Energy)**.** For two additive sets $A$ and $B$, let $Q(A, B)$ be the following set,

$$Q(A, B) := \{(a, a', b, b') \in A \times A \times B \times B : a + b = a' + b'\}$$

Then we define the energy $E(A, B)$ as

$$E(A, B) = \frac{|A|^2|B|^2}{|Q(A, B)|}$$

Trivially, the energy can be bounded as,

$$max(|A|, |B|) \leq E(A, B) \leq |A + B|$$

To understand why the measure of additive energy better captures the additive structure of $A + B$ than cardinality, we look at our previous example of an AP and a GP.

Say $A$ is an AP of size $N$ and $B$ is a GP of size $N$. Then we saw earlier that $|A + A| \sim kN$ and $|B + B| \sim \frac{N(N-1)}{2}$.

Also $|Q(A, A)| \sim N^3$, since choosing any 3 elements of $A$ will give the fourth element, such that sum of two of these is equal to the sum of other two. This implies that $E(A, A) \sim N$.

$Q(B, B)$, however, will only contain quadruples of the form $(a, a', b, b')$, where $a = a'$ and $b = b'$. Hence $|Q(B, B)| = N^2$ and $E(B, B) = N^2$.

Now consider another set $C = A \cup B$. For this set,

$$|C + C| \geq |B + B|$$

$$\implies |C + C| = O(N^2)$$

However, $Q(C, C) \geq Q(A, A)$, since it has atleast as many valid quadruples as $Q(A, A)$. This implies that $E(C, C) \leq kN$, for some constant $k$.

Hence we see that $C$ is a set for which the size is large but the additive energy is small. Intuitively the reason for energy being small is that $C$ has a *large* subset which does not *grow much* under addition (that is the set $A$). This information is not captured by the cardinality measure.

Therefore, the measure of additive energy affectively gives us information about the additive structure of large subsets as well. Indeed this is what the Balog-Szemeredi-Gowers theorem tries to say in a general manner.

### 2.5 Balog-Szemeredi-Gowers Theorem

**Theorem 2.4** (BSG Theorem [4] [2])**.** *Let $A, B \subset G$ be sets of size $N$ in an abelian group $G$. Suppose that $E(A, B) \leq KN$. Then, there exists subsets $A' \subset A$ and $B' \subset B$ with $|A'|, |B'| \geq N/K^c$ such that $|A' + B'| \leq K^c N$. Here, $c > 0$ is an absolute constant.*

To the reader interested in the proof of the BSG theorem, we refer to [8], where we surveyed the proof theorem along with its applications to the proof of Szemeredi-Trotter theorem.
The proof uses a graph-theoretic and relates the number of three length paths in a specifically constructed bipartite graph to the size of some subset $|A' + B'|$.

In general, BSG theorem formalises our notion that a pair of sets with low energy must have a large subset which does not grow under addition.

We next state a corollary of the above BSG theorem, which would help us generalise the BSG theorem for many other cases.

**Corollary 2.1.** *Let $A, B \subset G$ be sets of size $N$ in an abelian group $G$. Suppose that $E(A, B) \leq KN$. Then, there exists a $K'$-approximate group $H$ and $x, y \in G$ such that*

$$|A \cap (H + x)|, |B \cap (H + y)| \geq \frac{1}{K'}|H|$$

*and $|A|, |B| \leq K'|H|$.*

Informally it says that $A, B$ can both be almost covered by a single translate of an approximate group $H$.

We end our discussion of sumset estimates with BSG Theorem, but for the readers interested in further studying this direction of Additive Combinatorics, we would recommend the survey [3].

# 3 Szemeredi's Regularity Lemma

Szemeredi's regularity lemma roughly says that the vertex set of any *large* graph can be partitioned into a small number of clusters of equal sizes and a few leftover vertices, such that between most pairs of clusters, the edges seem to be distributed randomly.

Before stating the formal statement of the lemma, we first define a few measures which would help us quantify our notion of *randomness*.

## 3.1 Regularity and Randomness

Consider a graph $G = (V, E)$. For disjoint sets $A, B \subset V$, we denote the edges across $A$ and $B$ as $e(A, B)$.
For this pair $A, B$, we define density $d(A, B)$ as,

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

**Definition 3.1.** $(A, B)$ is $\epsilon$-regular if $\forall X \subseteq A, \forall Y \subseteq B$ with $|X| \geq \epsilon|A|, |Y| \geq \epsilon|B|$, we have $|d(X, Y) - d(A, B)| < \epsilon$.

$\epsilon$-regularity is a way of formalising the notion of randomness in a graph. If the edges between $A$ and $B$ are chosen independently with probability $p$, then the expected number of edges between $A$ and $B$ would be $p|A||B|$.

So given that we have $e(A, B)$ edges, the expected probability of choosing an edge would be $\frac{e(A,B)}{|A||B|}$, which is $d(A, B)$. In case of a random graph, the probability of choosing an edge would be uniform across all edges.

Hence for any pair $A, B$, one would expect the density of the any subset of the pair to be almost equal to $d(A, B)$, ie, for any $X \subseteq A, Y \subseteq B$, $|d(X, Y) - d(A, B)|$ should be small. However this would not make much sense if the subsets are small, so we give a reasonable lower bound on its size.

An partition $V = V_0 \cup V_1 \cup ..... \cup V_k$ is $\epsilon$-regular if
1. $|V_0| \leq \epsilon|V|$
2. $|V_1| = |V_2| = \cdots = |V_k|$
3. atmost $\epsilon k^2$ pairs are $(V_i, V_j)$ are not $\epsilon$-regular.

We are now ready to state the formal statement of Szemeredi's regularity lemma. The proof presented here has been inspired from [6], though many parts have been modified and/or simplified as per our understanding of the same.

## 3.2 Statement and Proof

**Theorem 3.1** (Szemeredi'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 $\epsilon$-regular partition $(V_0, V_1, ....., V_k)$, where $t \leq k \leq T(\epsilon, t)$.*

Before proving the lemma, which we will do by construction, we will state and prove certain lemmas which will provide us with tools to get and $\epsilon$-regular partion out of a non $\epsilon$-regular partition.

First we recall the definition of equivalence relation. An equivalence relation, say $\sim$, is a binary relation which satisfies the properties of reflexivity, symmetry and transitivity.

An equivalence class of an element $a$ of a set $X$ is defined as the set $:= \{x \in X : a \sim x\}$. For any two elements $a, b \in X$, it is easy to see that the equivalence classes of $a$ and $b$ are either equal or disjoint. Hence the set of equivalence classes form a partition of the set $X$.

We say that a partition $P'$ is a *refinement* of the partion $P$ if every element in $P'$ is a subset of an element in $P$.

From the above definitions, we can make the following observation:

**Observation 3.1.** *If $P = (P_1, P_2, ...., P_k)$ and $P' = (P'_1, P'_2, ...., P'_l)$ are two partitions of the same set $X$, then the partition given by $((P_1 \cap P'_1), ..., (P_1 \cap P'_l), (P_2 \cap P'_1), ..., (P_2 \cap P'_l), ...., (P_k \cap P'_1), ..., (P_k \cap P'_l))$ is a refinement of both $P$ and $P'$.*

The above observation makes it easier to visualize the refinement we will use of the partitions.

Next we define a bounded energy function, which would in a way, measure the potential of a pair being $\epsilon$-regular.
The energy function for a disjoint pair of vertices $A, B$ is defined as

$$f(A, B) := \frac{|A||B|}{n^2} d^2(A, B)$$

For partitions $P_1, P_2$,

$$f(P_1, P_2) := \sum_{A \in P_1, B \in P_2} f(A, B)$$

For any partition $P$,

$$f(P) := \sum_{A,B \in P : A \neq B} f(A, B)$$

The above definitions and inequalities will be more clear if we consider and work with the following probabilistic interpretation.
Consider a pair $(U, W) \in V \times V$, and the partitions $P_1 = (U_1, U_2 ...)$ of $U$ and $P_2 = (W_1, W_2 ...)$ of $W$.

We define a random variable $Z$ on the elements of $U \times W$, in the following way :
Say $u$ belongs to some partition $U_i$ of $U$ and $w$ belongs to some partition $W_j$ of $W$. Then,

$$Z(u, w) = d(U_i, W_j)$$

The expected value of $Z$ is,

$$\mathbb{E}[Z] = \sum_{U' \in P_1, W' \in P_2} \frac{|U'||W'|}{|U||W|} d(U', W')$$

$$= \frac{1}{|U||W|} \sum_{U' \in P_1, W' \in P_2} |U'||W'| d(U', W')$$

$$= \frac{1}{|U||W|} \sum_{U' \in P_1, W' \in P_2} e(U', W')$$

$$= \frac{1}{|U||W|} e(U, W) = d(U, W)$$

The probabilistic interpretation will make it relatively easy to prove the bounds on the energy of refined partitions and also contributes to an intuitive understanding of the construction.

**Lemma 3.1.** *For any partition $P$, $f(P) \leq 1$.*

*Proof.* Since $d(A, B) \leq 1$, for any $A, B$, we can say that,

$$f(A, B) \leq \frac{1}{n^2} \sum_{A,B \in P} |A||B| \leq \frac{1}{n^2} \sum_{A \in P} |A| \sum_{B \in P} |B| \leq 1$$

□

**Lemma 3.2** (Refinment increases energy)**.** *If partition $(U', W')$ is a refinement of partition $(U, W)$, then $f(U', W') \geq f(U, W)$.*

*Proof.* The statement is easy to infer when we look at it from a probabilistic point of view.

We know,

$$\mathbb{E}[Z^2] \geq \mathbb{E}[Z]^2$$

$$\implies \sum_{A \in U', B \in W'} \frac{|A||B|}{|U||W|} d^2(A, B) \geq d^2(U, W)$$

$$\implies \sum_{A \in U', B \in W'} \frac{|A||B|}{n^2} d^2(A, B) \geq \frac{|U||W|}{n^2} d^2(U, W)$$

$$\implies f(U', W') \geq f(U, W)$$

□

Before we move further and describe the exact details of the construction of the $\epsilon$-regular partition, we give a general outline:

- Our goal is, given a non $\epsilon$-regular partition, to refine this partition in such a way that the new partition is more $\epsilon$-regular.
- Ensure that the energy of the new partition increases by atleast a fixed amount.
- Keep refining till the energy becomes 1. Since the energy cannot increase beyond that, we must get an $\epsilon$-regular partition till then.

To give an intuitive idea of second step, if there is a non-$\epsilon$-regular pair, then there will be a subset of the pair for which the density would be largely different from the density of the pair. Hence if we remove this subset from the original set we will get a significant contribution to the overall energy, which was not there earlier. So we would expect our energy to increase atleast by a certain amount.

We now formally state the ideas that we gave and prove the Regularity lemma.

**Lemma 3.3.** *If* $(U, W)$ *are disjoint subsets of* $V$ *such that the pair* $(U, W)$ *is not* $\epsilon$*-regular, then there exists partitions* $P_1 = (U_1, U_2)$ *of* $U$ *and* $P_2 = (W_1, W_2)$ *of* $W$*, such that*

$$f(P_1, P_2) \geq f(U, W) + \epsilon^4 . \frac{|U||W|}{n^2}$$

*Proof.* For a non $\epsilon$-regular pair $(U, W)$, we know that there exists $U' \subseteq U, W' \subseteq W$ with $|U'| \geq \epsilon|U|, |W'| \geq \epsilon|W|$, such that $|d(U, W) - d(U', W')| \geq \epsilon$.

From our earlier discussion, it would make sense to remove this large enough subsets from $(U, W)$ so as to accomplish two objectives, (1) Make the partition "more $\epsilon$-regular", and (2) Increase the value of energy function using the amount contributed by $(U', W')$ [both the objectives in a way are equivalent].

So we define our partitions to be $P_1 = (U', U \setminus U')$ and $P_2 = (W', W \setminus W')$.

We will shift to the probabilistic argument to get the lower bound on energy of the partition.
Consider the random variable $Z$ as defined earlier. We know, that whenever $(u, w) \in U' \times W'$, $Z(u, w)$ deviates atleast $\epsilon$ from $\mathbb{E}[Z]$ $(= d(U, W))$. Therefore,

$$Var[Z] = \mathbb{E}[(Z - \mathbb{E}[Z])^2] \geq \frac{|U'||W'|}{|U||W|}\epsilon^2 \geq \epsilon^4$$

But we know, $Var[Z] = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2$

$$\implies \mathbb{E}[Z^2] - \mathbb{E}[Z]^2 \geq \epsilon^4$$

$$\implies \sum_{U' \in P_1, W' \in P_2} \frac{|U'||W'|}{|U||W|} d^2(U', W') \geq d^2(U, W) + \epsilon^4$$

$$\implies \sum_{U' \in P_1, W' \in P_2} \frac{|U'||W'|}{n^2} d^2(U', W') \geq \frac{|U||W|}{n^2} d^2(U, W) + \frac{|U||W|}{n^2}\epsilon^4$$

$$\implies f(P_1, P_2) \geq f(U, W) + \epsilon^4 . \frac{|U||W|}{n^2}$$

□

We will now generalise the above lemma to get a refinement of a non $\epsilon$-regular partition.

**Lemma 3.4.** *Say $P = V_0, V_1, ...., V_k$ is a non $\epsilon$-regular partition of $V$. Then there exists a refinement $P' = V_0, V_1', ...., V_l'$ of $P$, such that*

$$f(P') \geq f(P) + \epsilon^5/4$$

*Proof.* Since $P$ is a non $\epsilon$-regular partition, there will be $\geq \epsilon k^2$ pairs which are non $\epsilon$-regular. For each such pair $(V_i, V_j)$, use Lemma 3.3 to get partitions $\mathcal{V}_{ij}$ of $V_i$ and $\mathcal{V}_{ji}$ of $V_j$.

For some $V_i$, consider all such partitions $\{\mathcal{V}_{ij} : j \neq i\}$. We say that two elements $v, v' \in V_i$ are equivalent if they occur in the same partition $\mathcal{V}_{ij}$, for all $j \neq i$. This is a valid equivalence relation since it satisfies the properties of reflexivity, transitivity and symmetry.

Say $\mathcal{V}_i$ is the set of all such equivalence classes of the element of $V_i$. Then $\mathcal{V}_i$ forms a partition of $V_i$. Given each such partition $\mathcal{V}_i$, from **Observation 3.1**, it can be seen that the *intersection* of partition $P$ with $\mathcal{V}_i$ gives a refinement of $P$. In other words, we are generalising Lemma 3.3 for general partitions.

Since $P' = (V_0, \mathcal{V}_1, ...., \mathcal{V}_k)$ is also a partition of $V$, extending the above argument to all $i$, we conclude that $P'$ is a refinement of $P$.

$$f(P') = \sum_{1 \leq i < j} f(\mathcal{V}_i, \mathcal{V}_j)$$

Since $\mathcal{V}_i, \mathcal{V}_j$ are a refinement of $\mathcal{V}_{ij}, \mathcal{V}_{ji}$,

$$f(P') \geq \sum_{1 \leq i < j} f(\mathcal{V}_{ij}, \mathcal{V}_{ji})$$

Now we know that there are atleast $\epsilon k^2$ irregular pairs and for each such $(V_i, V_j)$, energy increases by $\geq \epsilon^4 . \frac{|V_i||V_j|}{n^2}$. Therefore,

$$f(P') \geq f(P) + \epsilon k^2 (\epsilon^4 . \frac{|V_1|^2}{n^2})$$

$$f(P') \geq f(P) + \epsilon^5 (\frac{k|V_1|}{n})^2$$

But $k|V_1| = n - |V_0| \geq n/2$,

$$f(P') \geq f(P) + \epsilon^5/4$$

□

Even though the partition given by the above lemma is a refinement of the earlier, it is not an equi-partition. To make it an equipartition we split every set in $P'$ into disjoint sets of size $b = \lfloor \frac{|V_1|}{4^k} \rfloor$, and throw away leftover vertices in $V_0$.

However it is not immediately clear as to why the considerable number of sets in $P'$ would have size $\geq b$. To see this, note that by our construction, for any set $V_i$ there can be atmost $2^{k-1}$ equivalence classes.
Since the set of equivalence classes ($\mathcal{V}_i$) result in a partition, the average size of each partition $= \frac{|V_i|}{|\mathcal{V}_i|} \geq \frac{|V_i|}{2^{k-1}}$.

Note that $b << \frac{|V_i|}{2^{k-1}}$. So in expectation, there would many sets in $\mathcal{V}_i$ which would have size greater than $b$.

**Claim 3.1.** *Energy after making equipartitions does not decrease significantly.*

*Proof.* Since we put a few elements from the non-trivial sets in $P'$ into $V_0$ while constructing an equipartition, there should be some change in energy, in particular, it will decrease. We need to show that it does not decrease much.

To see that, say we remove atmost $b$ elements (say set $q$) from an equivalence class $Q$. Say we call the set after removing the elements as $Q'$.

The decrease in energy for the pair $(Q, B)$, for any other set $B$ in the partition $P$, will be

$$= f(Q, B) - f(Q', B)$$

$$= \frac{e^2(Q, B)}{n^2|Q||B|} - \frac{e^2(Q', B)}{n^2|Q'||B|}$$

$$\leq \frac{1}{n^2|Q'||B|}[e^2(Q, B) - e^2(Q', B)]$$

But $e(Q, B) = e(Q', B) + e(q, B)$. Therefore decrease in energy becomes

$$\leq \frac{1}{n^2|Q'||B|}[e^2(q, B) + 2e(Q', B)e(q, B)]$$

$$\leq \frac{1}{n^2|Q'||B|}[b^2|B|^2 + 2b|Q'||B|^2]$$

But $|B| = |V_1|$ and $|Q'| > b$.

$$\leq \frac{3b|V_1|}{n^2} \leq \frac{3|V_1|^2}{4^k n^2}$$

Looking at the total decrease for all pairs and for all equivalence classes, we get, that the decrease in energy

$$\leq \frac{3|V_1|^2}{4^k n^2}.k2^{k-1}$$

$$\leq \frac{2k|V_1|^2}{n^2 2^k}$$

Also, $k|V_1| \leq n$

$$\leq \frac{2}{k2^k} \leq \frac{2}{t2^t}$$

Since the decrease in energy is exponentially small compared to the number of partitions, we can still safely assume that there is a constant increase in energy every iteration [specifically, $\frac{\epsilon^5}{4} - \frac{2}{t2^t}$]

□

Hence, we have an equipartition, say $P'' = V_0'', V_1'', ...., V_m''$ which is considerably "more regular" than the previous partition. We only need to show that $V_0$ does not increase much and that the new numbers of partitions is bounded.

For each set $V_i$ there can be atmost $2^{k-1}$ equivalence classes and each equivalence class contributes atmost $b$ elements to $V_0$. Hence,

$$|V_0''| \leq |V_0| + k.2^{k-1}.b \leq |V_0| + kc.2^{k-1}/4^k \leq |V_0| + n/2^k$$

Also, because of the choice of $b$, each set $V_i$ can be divided into atmost $4^k$ parts. So the new number of partitions $m \leq k4^k$.

Now using the above constructions and lemmas, we can give a proof of Szemeredi's regularity lemma.

*Proof of Theorem 3.1.* We start with an arbitrary partition of the set $V$ into $t$ subsets, each of size $\lfloor \frac{n}{t} \rfloor$, with $|V_0| \leq \lfloor \frac{n}{t} \rfloor$.

Then we refine our partition using the above cosntruction $4\epsilon^{-5}$ times to obtain a equipartition of $V$. Since by then, the value of energy function will reach 1, we will get an $\epsilon$-regular partition at some point in the process.

The value of $T(\epsilon, t)$ will be the number obtained by iterating the map $x \rightarrow x.4^x$ $4\epsilon^{-5}$ times, starting from $t$. □

It may be interesting to note that the value of $T(\epsilon, t)$ obtained from the above construction is a tower of 4's of height proportional to $\epsilon^{-5}$. This makes this lemma useful only for large graphs.

However, Gowers [7] showed that $T(\epsilon, t)$ will indeed have to be very large, or in other words, a bound of the type $exp(\epsilon^{-c})$, for some $c > 0$, does not exist.

## References


[1] T. Tao, V. Hu, *Additive Combinatorics*, 69-78.

[2] Z. Dvir, *Incidence Theorems and Their Applications* (2013), `http://arxiv.org/pdf/1208.5073v2.pdf`

[3] S. Lovett, *Additive Combinatorics and its Applications in Theoretical Computer Science* (2013), `http://cseweb.ucsd.edu/~slovett/files/addcomb-survey.pdf`

[4] A. Balog and E. Szemeredi., *A statistical theorem of set addition*, Combinatorica (1994) 14(3):263-268.

[5] W. T. Gowers, *A new proof of Szemeredis theorem for arithmetic progressions of length four*, Geom. Funct. Anal. (1998), 17(2):230261.

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

[7] W.T. Gowers, *Lower bounds of tower type for Szemeredi's uniformity lemma*, `http://link.springer.com/article/10.1007%2FPL00001621`

[8] *Additive Combinatorics and Incidence Geometry : The Szemeredi-Trotter Theorem*, `http://home.iitk.ac.in/~vijaykes/ugp1.pdf`