# *Szemerédi Regularity Lemma*

*Aakash Ghosh(19MS129)*

*August 16, 2022*

Let $G$ be a graph and let $A$ and $B$ be disjoint set of vertices of $G$. Let $e(A, B)$ be the number of edges from $A$ to $B$. Define $d(A, B) = \frac{e(A,B)}{|A||B|}$ [1]

[1] Note that $d(A, B)$ is a somewhat normalised measure of how well two vertex set is connected with each other. The value of $d$ will range from 0(i.e no connection) to 1(i.e all possible edges are present and very well connected)

**Lemma 1.** *If $y \in [a, b]$ then*

$$|x - y| \leq c \Leftrightarrow |x - a| \leq c, |x - b| \leq c \tag{1}$$

*Proof.* There are three possible cases as shown in Fig 1. The blue line shows the valuesw $y$ takes.

$\square$

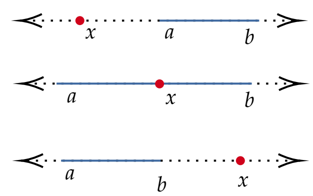


Figure 1: The three cases required to prove lemma 1

**Theorem 2.** *If $X \subset A$ and $Y \subset B$ with $\frac{|X|}{|A|} \geq 1 - \delta$ and $\frac{|Y|}{|B|} \geq 1 - \delta$ where $0 \leq \delta \leq \frac{1}{2}$. Then:*

$$|d(X, Y) - d(A, B)| \leq 2\delta \tag{2}$$

$$|d^2(X, Y) - d^2(A, B)| \leq 4\delta \tag{3}$$

***Proof outline:*** We try to go the other way around: we fix edges of $X, Y$ and try to build edges of $A, B$ around them. We note that since $d(A, B)$ increases with the number of edges added in the extension of the old graph, we just need to make sure that inequality holds for the maximum and minimum values only(lemma 1). We find bounds for this maximum and minimum values to complete the proof. eqn3 follows as a consequence of eqn3 qand bounds on $d(A, B), d(X, Y)$.

*Proof.* Let $d(X, Y) = d'$. Note that:

$$e(A, B) = e(X, Y) + e(X, B \setminus Y) + e(A \setminus X, Y) + e(A \setminus X, B \setminus Y) \tag{4}$$

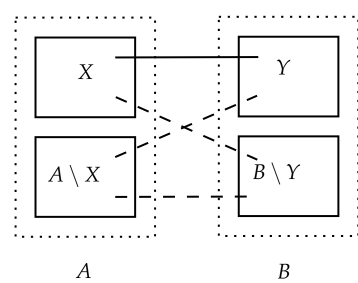


Figure 2: Each line represents all edges between the sets it joins. Edges represented by the solid line is given. Maximum is attained when all edges represented by the dotted line is present. Minimum is attained when no edge of represented by dotted lines are given.

We shall first try to prove the inequality when $d(A, B) = d_1$ is maximum. It is easy to see this occurs when each term of eqn4 attains their maximum value. Note only $e(X, Y)$ is fixed. For the other terms we have:

$$e(X, B \setminus Y) \leq (|X|)(|B| - |Y|) \tag{5}$$

$$e(A \setminus X, Y) \leq (|A| - |X|)(|Y|) \tag{6}$$

$$e(A \setminus X, B \setminus Y) \leq (|A| - |X|)(|B| - |Y|) \tag{7}$$

Where maximum is attained when all the respective edges are drawn. Putting in eqn4 and simplifying we get:

$$d_1|A||B| = e(X,Y) + |A||B| - |X||Y| \tag{8}$$

$$\Rightarrow d_1|A||B| = d'|X||Y| + |A||B| - |X||Y| = |A||B| - (1-d')|X||Y| \tag{9}$$

$$d_1 = 1 - (1-d')\left(\frac{|X|}{|A|}\right)\left(\frac{|Y|}{|B|}\right) \tag{10}$$

We know $1 \geq \left(\frac{|X|}{|A|}\right), \left(\frac{|Y|}{|B|}\right) \geq 1-\delta$. From this we get $d_1 \geq d'$. Therefore,

$$\begin{aligned} |d_1 - d'| &= d_1 - d' \\ &\leq 1 - (1-d')\left(\frac{|X|}{|A|}\right)\left(\frac{|Y|}{|B|}\right) - d' \\ &\leq 1 - (1-d')(1-\delta)^2 - d' \\ &\leq (1-d')\left(1-(1-\delta)^2\right) \\ &\leq \left(1-(1-\delta)^2\right) \\ &\leq 2\delta - \delta^2 \leq 2\delta \end{aligned} \tag{11}$$

Now we prove it when $d(A,B) = d_2$ is minimum. This occurs when each term of eqn4 attains minimum value. $e(X,Y)$ is fixed. For the other terms the mnimum value attained is 0. Therefore,

$$d_2|A||B| = e(X,Y) = d'|X||Y| \tag{12}$$

$$\Rightarrow d_2 = d'\left(\frac{|X|}{|A|}\right)\left(\frac{|Y|}{|B|}\right) \tag{13}$$

From the bounds of $\left(\frac{|X|}{|A|}\right), \left(\frac{|Y|}{|B|}\right)$ we get $d_2 \leq d'$. Therefore,

$$\begin{aligned} |d_2 - d'| &= d' - d_2 \\ &= d' - d'\left(\frac{|X|}{|A|}\right)\left(\frac{|Y|}{|B|}\right) \\ &= d'\left(1 - \left(\frac{|X|}{|A|}\right)\left(\frac{|Y|}{|B|}\right)\right) \\ &\leq d'\left(1-(1-\delta)^2\right) \\ &\leq 2\delta - \delta^2 \leq 2\delta \end{aligned} \tag{14}$$

Since the inequality is proved for maximum and minimum values of $d(A,B)$, we can use lemma 1 to conclude

Theorem 2 essentially tells that large subsets of two graph mimics the connectivity of the parent graph

$$|d(A,B) - d(X,Y)| \leq 2\delta$$

And from thsi it follows that:

$$|d(A,B)^2 - d(X,Y)^2| \leq |d(A,B) - d(X,Y)||d(A,B) + d(X,Y)| \leq 2\delta(1+1) = 4\delta$$

$\square$

**Definition 1** ($\epsilon$-Regularity). *We say $A, B$ are epsilon regular if for $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \epsilon|A|$ and $|Y| \geq \epsilon|B|$ we have $|d(A, B) - d(X, Y)| \leq \epsilon$*

## *References*