

Lecture Notes Spring 2006

1 Origins

The Regularity lemma originally proved the following theorem [1]:

Theorem 1 (Szemerédi, 1975) *Let $A \subset \mathbb{Z}$ with positive upper density, then A contains arbitrarily long arithmetic progressions.*

This generalizes an old theorem by van der Waerden [3]:

Theorem 2 (van der Waerden, 1927) *If \mathbb{Z} is colored by $k \geq 1$ colors, then there exists a monochromatic arithmetic progression of arbitrary length.*

So, Szemerédi shows that the monochromatic arithmetic progression occurs in each color class that is not trivially sparse (we will not bother to define positive upper density).

From this theorem came the following lemma, by Szemerédi, written separately in 1976 [2].

Theorem 3 (RegLem) *For every $\epsilon > 0$ and positive integer m , there exist two integers $M(\epsilon, m)$ and $N(\epsilon, m)$ with the property that, for every graph G with $n \geq N(\epsilon, m)$ vertices, there exists a partition of the vertex set into $k+1$ classes $V = V_0 + V_1 + \dots + V_\ell$ such that*

- $m \leq \ell \leq M(\epsilon, m)$
- $|V_0| < \epsilon n$
- $|V_1| = |V_2| = \dots = |V_\ell|$
- For distinct $i, j \neq 0$, all but at most $\epsilon \ell^2$ of the pairs (V_i, V_j) are ϵ -regular of some density.

It almost seems as if this is totally vacuous. But it is very very useful. Note that the numbers M and N only depend on ϵ and m . In most applications, $m = 1$ and we see that the number of vertices is just larger than some constant depending only on ϵ .

2 Preliminaries to the proof of RegLem

There are two preliminary lemmas which we will use:

Lemma 1 (Cauchy-Schwarz inequality, defect form) *For some $m < n$, if*

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

then

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

Lemma 2 (Continuity of density) *For $0 \leq \delta \leq 0.3298$, if $X' \subseteq X$ and $Y' \subseteq Y$ with $|X'| \geq (1 - \delta)|X|$ and $|Y'| \geq (1 - \delta)|Y|$, then*

- $|d(X^*, Y^*) - d(X, Y)| < 6\delta$
- $|d(X^*, Y^*)^2 - d(X, Y)^2| < 12\delta$

We prove them first because they are of independent interest.

Proof of Lemma 1, the defect form of the Cauchy-Schwarz inequality.

Let $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$ and let $Y_k = X_k - \bar{X}$.
Observe that

$$\sum_{k=1}^n Y_k = \sum_{k=1}^n (X_k - \bar{X}) = n\bar{X} - n\bar{X} = 0,$$

but if we just count the first m terms,

$$\sum_{k=1}^m Y_k = \sum_{k=1}^m (X_k - m\bar{X}) = m\bar{X} + \delta - m\bar{X} = \delta.$$

Summarizing,

$$\sum_{k=1}^m Y_k = \delta \tag{1}$$

$$\sum_{k=m+1}^n Y_k = -\delta. \tag{2}$$

In addition,

$$\begin{aligned}\sum_{k=1}^n Y_k^2 &= \sum_{k=1}^n (X_k^2 - 2X_k\bar{X} + (\bar{X})^2) \\ &= \sum_{k=1}^n X_k^2 - n\bar{X}^2\end{aligned}\tag{3}$$

Let $S \subseteq \{1, \dots, n\}$, then Cauchy-Schwarz (or Jensen's inequality for convex functions) gives that

$$\left| \sum_S Y_k \cdot \mathbf{1} \right|^2 = |\langle \mathbf{y}, \mathbf{1} \rangle|^2 \leq \|\mathbf{y}\|^2 \|\mathbf{1}\|^2 = \left(\sum_S Y_k^2 \right) \left(\sum_S 1^2 \right) = |S| \sum_S Y_k^2.$$

So, using (1) and (2), respectively,

$$\frac{1}{m} \sum_{k=1}^m Y_k^2 \geq \left(\frac{1}{m} \sum_{k=1}^m Y_k \right)^2 = \left(\frac{\delta}{m} \right)^2\tag{4}$$

$$\frac{1}{n-m} \sum_{k=m+1}^n Y_k^2 \geq \left(\frac{1}{n-m} \sum_{k=m+1}^n Y_k \right)^2 = \left(\frac{\delta}{n-m} \right)^2\tag{5}$$

Using (3) as well as (4) and (5), we have

$$\sum_{k=1}^n X_k^2 - n\bar{X}^2 = \sum_{k=1}^n Y_k^2 \geq m \left(\frac{\delta}{m} \right)^2 + (n-m) \left(\frac{\delta}{n-m} \right)^2.$$

Therefore,

$$\sum_{k=1}^n X_k^2 \geq n\bar{X}^2 + \frac{\delta^2 n}{m(n-m)}.$$

□

Proof of Lemma 2, continuity of density.

First an upper bound:

$$\begin{aligned}e(X^*, Y^*) &\leq e(X, Y) \\ d(X^*, Y^*) &\leq d(X, Y) \frac{|X||Y|}{|X^*||Y^*|} \\ &\leq \frac{d(X, Y)}{(1-\delta)^2}\end{aligned}\tag{6}$$

Now a lower bound:

$$\begin{aligned}
e(X^*, Y^*) &\geq e(X, Y) - (|X^*||Y \setminus Y^*| + |Y^*||X \setminus X^*| + |X \setminus X^*||Y \setminus Y^*|) \\
d(X^*, Y^*) &\geq d(X, Y) \frac{|X||Y|}{|X^*||Y^*|} - \frac{|X \setminus X^*|}{|X^*|} - \frac{|Y \setminus Y^*|}{|Y^*|} - \frac{|X \setminus X^*||Y \setminus Y^*|}{|X^*||Y^*|} \\
&= d(X, Y) \frac{|X||Y|}{|X^*||Y^*|} - \frac{|X||Y| - |X^*||Y^*|}{|X^*||Y^*|} \\
&= 1 - (1 - d(X, Y)) \frac{|X||Y|}{|X^*||Y^*|} \\
&\geq 1 - \frac{1 - d(X, Y)}{(1 - \delta)^2} \tag{7}
\end{aligned}$$

As a result,

$$\begin{aligned}
1 - \frac{1 - d(X, Y)}{(1 - \delta)^2} &\leq d(X^*, Y^*) \leq \frac{d(X, Y)}{(1 - \delta)^2} \\
(d(X, Y) - 1) \left(\frac{1}{(1 - \delta)^2} - 1 \right) &\leq d(X^*, Y^*) - d(X, Y) \leq d(X, Y) \left(\frac{1}{(1 - \delta)^2} - 1 \right)
\end{aligned}$$

For $0 \leq \delta \leq 1/2$, we have

$$\frac{1}{(1 - \delta)^2} - 1 \leq 6\delta.$$

Now to the second part. As we have seen before,

$$\begin{aligned}
d(X^*, Y^*)^2 - d(X, Y)^2 &\leq d(X, Y)^2 \left(\frac{1}{(1 - \delta)^4} - 1 \right) \\
&\leq \left(\frac{1}{(1 - \delta)^4} - 1 \right)
\end{aligned}$$

As to the upper bound,

$$\begin{aligned}
d(X^*, Y^*)^2 - d(X, Y)^2 &\geq \left(1 - \frac{1 - d(X, Y)}{(1 - \delta)^2} \right)^2 - d(X, Y)^2 \\
&\geq \left(\frac{1}{(1 - \delta)^2} - 1 \right)^2 - \frac{2d(X, Y)}{(1 - \delta)^2} \left(\frac{1}{(1 - \delta)^2} - 1 \right) \tag{8} \\
&\quad + d(X, Y)^2 \left(\frac{1}{(1 - \delta)^4} - 1 \right) \tag{9}
\end{aligned}$$

The least value, according to the first derivative test, occurs when $d(X, Y) = 0$ (the value is $-(1/(1-\delta)^2 - 1)^2$), $d(X, Y) = 1$ (the value is 0) or

$$d(X, Y) = \frac{\frac{1}{(1-\delta)^2} \left(\frac{1}{(1-\delta)^2} - 1 \right)}{\frac{1}{(1-\delta)^4} - 1}$$

Plug this in and the value is

$$\begin{aligned} & \left(\frac{1}{(1-\delta)^2} - 1 \right)^2 - \frac{\frac{2}{(1-\delta)^4}}{\frac{1}{(1-\delta)^4} - 1} \left(\frac{1}{(1-\delta)^2} - 1 \right)^2 + \frac{\frac{1}{(1-\delta)^4} \left(\frac{1}{(1-\delta)^2} - 1 \right)^2}{\frac{1}{(1-\delta)^4} - 1} \\ &= -\frac{\left(\frac{1}{(1-\delta)^2} - 1 \right)^2}{\frac{1}{(1-\delta)^4} - 1} = -\frac{\frac{1}{(1-\delta)^2} - 1}{\frac{1}{(1-\delta)^2} + 1} \geq -\left(\frac{1}{(1-\delta)^4} - 1 \right). \end{aligned}$$

Finally, we can see numerically that

$$\frac{1}{(1-\delta)^4} - 1 \leq 12\delta$$

whenever $\delta < 0.3298$. □

3 Proving RegLem

Before we prove RegLem itself, let us use some definitions:

Definition 1

- The sets V_1, \dots, V_ℓ are called **clusters**
- The set V_0 is called the **exceptional set** or **leftover set**.
- If a pair is not ϵ -regular, we often refer to it as **ϵ -irregular**.
- We call a partition $V = V_0 + V_1 + \dots + V_\ell$ an **equitable partition** if $|V_1| = \dots = |V_\ell|$.
- Let V have an equitable partition, P , labelled $V = V_0 + V_1 + \dots + V_\ell$. The **index of P** is defined by

$$\text{index}(P) = \frac{1}{\ell^2} \sum_{s=1}^{\ell} \sum_{t=s+1}^{\ell} d(V_s, V_t)^2$$

Observe that $0 \leq \text{index}(P) < 1/2$.

We use the following:

Lemma 3 (Main Lemma) *Let $G = (V, E)$ be a graph on n vertices. Let P be an equitable partition of $V = V_0 + V_1 + \dots + V_\ell$, with exceptional class V_0 . Let ℓ be a positive integer with $4^\ell > 600\epsilon^{-5}$.*

If more than $\epsilon\ell^2$ pairs (V_s, V_t) with $1 \leq s < t \leq \ell$ are ϵ -irregular, then there exists an equitable partition Q of V into $1 + \ell 4^\ell$ classes with $|Q_0| \leq |V_0| + n/4^\ell$ such that $\text{index}(Q) \geq \text{index}(P) + \epsilon^5/20$.

Furthermore, Q is a refinement of P , which is to say each nonexceptional cluster of Q lies within a cluster of P .

Using this, let us prove RegLem:

Proof of the Regularity Lemma. Let s be the smallest positive integer such that $4^s > 600\epsilon^{-5}$, $s \geq m$ and $s \geq 2/\epsilon$.

Define $f(t)$ by $f(t) = \begin{cases} s, & \text{if } t = 0; \\ f(t-1)4^{t-1}, & \text{otherwise.} \end{cases}$

Let t be the largest integer such that there exists an equitable partition P of V into $1 + f(t)$ clusters such that $\text{index}(P) \geq t\epsilon^5/20$ and the size of the exceptional class is at most $\epsilon n (1 - 2^{-(t+1)})$.

We now claim that P is ϵ -regular (i.e., all but $\epsilon\ell^2$ pairs are ϵ -regular). Otherwise, the Main Lemma (Lemma 3) implies there exists a Q with index at least $(t+1)\epsilon^5/20$ with at least $1 + f(t)4^{f(t)} = 1 + f(t+1)$ clusters (including the exceptional set), and with exceptional set size at most $\epsilon n (1 - 2^{-(t+1)}) + \frac{n}{4^{f(t)}}$. We claim that this is at most $\epsilon n (1 - 2^{-(t+2)})$, which would contradict the maximality of t .

To verify this, we need $4^{f(t)} \geq 2^{t+2}/\epsilon$. We leave it to the reader to verify, by induction on t , that, with $s \geq 2/\epsilon \geq 1$ and $f(t)$ defined as above, that

$$4^{f(t)} \geq s2^{t+1} \geq 2^{t+2} \frac{1}{\epsilon}.$$

□

The final piece is the Main Lemma:

Proof of Lemma 3, the Main Lemma.

Partitioning V_i into atoms

Consider any ϵ -irregular pair (V_s, V_t) , $1 \leq s < t \leq \ell$. Then we can choose $X = X(s, t)$ and $Y = Y(s, t)$ such that $X \subseteq V_s$, $Y \subseteq V_t$, $|X| = |Y| =$

$\lceil \epsilon |V_s| \rceil = \lceil \epsilon |V_t| \rceil$ ¹ and

$$|d(X, Y) - d(V_s, V_t)| > \epsilon.$$

For every i , $1 \leq i \leq \ell$, we define an equivalence relation over V_i as follows:

$$\forall x, y \in V_i, \quad x \equiv y \iff x \in X(i, j) \text{ whenever } y \in X(i, j), \forall i \neq j$$

and $X(i, j) = X(j, i)$ for $j < i$. The equivalence classes are called *atoms*. Each cluster V_s has at most $2^{\ell-1}$ atoms. Let

$$L = \left\lfloor \frac{|V_i|}{4^\ell} \right\rfloor, \quad 1 \leq i \leq \ell.$$

Partition Q so that

1. each atom A has exactly $\lfloor |A|/L \rfloor$ members of Q ,
2. each cluster V_i , $1 \leq i \leq \ell$ has exactly $\lfloor |V_i|/L \rfloor$ members of Q ,
3. each member of Q has cardinality L .

This is easy to do by arbitrarily partitioning each A into pieces of size L . Combine all the (at most $L - 1$) vertices from each A and partition that set into pieces of size L . This leaves at most $L - 1$ leftover vertices. Add these to the exceptional set.

Note that $\lfloor |V_i|/L \rfloor = 4^\ell$, $1 \leq i \leq \ell$, so each V_i contains exactly 4^ℓ elements of Q and Q contains exactly $\ell 4^\ell$ members.

Furthermore, there are ℓ clusters and at most $L - 1$ are placed from each into the exceptional set, the new exceptional set is now of size at most

$$|V_0| + \ell L \leq |V_0| + \ell \left\lfloor \frac{n/\ell}{4^\ell} \right\rfloor \leq |V_0| + \frac{n}{4^\ell}.$$

Let $q = 4^\ell$ and each member of Q in V_s ($1 \leq s \leq \ell$) be

$$V_s(i), \quad i = 1, \dots, q = 4^\ell$$

and

$$V_s^* = \bigcup_{i=1}^q V_s(i).$$

¹We will prove later that if (V_s, V_t) is ϵ -irregular, then there are sets that witness this of size as small as possible.

Then,

$$\begin{aligned}
|V_s^*| \geq |V_s| - (L - 1) &= |V_s| - \left\lfloor \frac{|V_s|}{4^\ell} \right\rfloor - 1 \\
&\geq |V_s| \left(1 - \frac{1}{4^\ell}\right) - 1 \\
&\geq |V_s| \left(1 - \frac{\epsilon^5}{600}\right)
\end{aligned}$$

Index change for all pairs

By continuity of density, for $1 \leq s < t \leq \ell$,

$$|d^2(V_s^*, V_t^*) - d^2(V_s, V_t)| < 12 \frac{\epsilon^5}{600} = \frac{\epsilon^5}{50}. \quad (10)$$

So, given $1 \leq s < t \leq \ell$,

$$\frac{1}{q^2} \sum_{i=1}^q \sum_{j=1}^q d^2(V_s(i), V_t(j)) \geq \left[\frac{1}{q^2} \sum_{i=1}^q \sum_{j=1}^q d(V_s(i), V_t(j)) \right]^2 \quad (11)$$

$$\begin{aligned}
&= \left[\frac{1}{q^2 L^2} \sum_{i=1}^q \sum_{j=1}^q e(V_s(i), V_t(j)) \right]^2 \\
&= \left[\frac{1}{q^2 L^2} e(V_s^*, V_t^*) \right]^2 \\
&= [d(V_s^*, V_t^*)]^2 \\
&> d^2(V_s, V_t) - \frac{\epsilon^5}{50} \quad (12)
\end{aligned}$$

Inequality (11) comes from the (non-defect version of) Cauchy-Schwarz. Inequality (12) comes from inequality (10).

Index change for ϵ -irregular pairs

But now, consider an ϵ -irregular pair (V_s, V_t) with $1 \leq s < t \leq \ell$. Let $X = X(s, t)$ and $Y = Y(s, t)$. Recall $V_s^* = \bigcup_{i=1}^q V_s(i)$, $V_t^* = \bigcup_{j=1}^q V_t(j)$. Without loss of generality, we may assume that the members of Q in X and Y are, respectively,

$$\begin{aligned}
V_s(i) &\subseteq X, & i &= 1, \dots, r_X \\
V_t(j) &\subseteq Y, & j &= 1, \dots, r_Y
\end{aligned}$$

As a result,

$$\left| \bigcup_{i=1}^{r_X} V_s(i) \right| > |X| - L2^{\ell-1} \quad (13)$$

$$\begin{aligned} &> |X| - 2^\ell L \\ &\geq |X| - \frac{|V_s|}{2^\ell} \\ &\geq |X| - \frac{|X|}{\epsilon 2^\ell} \\ &\geq |X| - \frac{|X|}{\epsilon \sqrt{600} \epsilon^{-5}} \geq |X| \left(1 - \frac{\epsilon^{3/2}}{10\sqrt{6}} \right) \\ &\geq |X| \left(1 - \frac{\epsilon}{100} \right), \end{aligned} \quad (14)$$

as long as $\epsilon \leq 0.06$.

Inequality (13) comes from the fact that X splits into at most $2^{\ell-1}$ atoms and for each atom, there are less than L vertices not in one such atom.

Inequality (14) comes from $L = \lfloor |V_s|/4^\ell \rfloor$.

Therefore,

$$r_X = \frac{|\bigcup_{i=1}^{r_X} V_s(i)|}{L} > \frac{|X|(1 - \epsilon/100)}{\lfloor |V_s|/q \rfloor} \geq \frac{|X|}{|V_s|} q \left(1 - \frac{\epsilon}{100} \right) \geq \epsilon q \left(1 - \frac{\epsilon}{100} \right).$$

Analogously, $\left| \bigcup_{j=1}^{r_Y} V_t(j) \right| > |Y|(1 - \epsilon/100)$ and $r_Y > \epsilon q(1 - \epsilon/100)$. Let $r = \min\{r_X, r_Y\}$, then

$$r > \epsilon q \left(1 - \frac{\epsilon}{100} \right). \quad (15)$$

If you define

$$X^* = \bigcup_{i=1}^r V_s(i) \subseteq X \quad \text{and} \quad Y^* = \bigcup_{j=1}^r V_t(j) \subseteq Y,$$

then

$$|X^*| > |X| \left(1 - \frac{\epsilon}{100} \right), \quad \text{and} \quad |Y^*| > |Y| \left(1 - \frac{\epsilon}{100} \right).$$

By continuity of density,

$$|d(X^*, Y^*) - d(X, Y)| < 6 \frac{\epsilon}{100} < \frac{\epsilon}{4} \quad (16)$$

and so the triangle inequality, (16) and (10),

$$|d(X^*, Y^*) - d(V_s^*, V_t^*)| > \epsilon - \frac{\epsilon}{4} - \frac{\epsilon^5}{50} > \frac{\epsilon}{2} \quad (17)$$

Now we have the following identity:

$$\underbrace{\sum_{i=1}^r \sum_{j=1}^r d(V_s(i), V_t(j))}_{r^2 d(X^*, Y^*)} = \underbrace{\frac{r^2}{q^2} \sum_{i=1}^r \sum_{j=1}^r d(V_s(i), V_t(j))}_{r^2 d(V_s^*, V_t^*)} + r^2 d(X^*, Y^*) - r^2 d(V_s^*, V_t^*) \quad (18)$$

So, applying Lemma 1, the defect form of Cauchy-Schwarz to (18) with $n = q^2$, $m = r^2$ and $\delta = r^2 d(X^*, Y^*) - r^2 d(V_s^*, V_t^*)$, we get

$$\begin{aligned} \sum_{i=1}^q \sum_{j=1}^q d^2(V_s(i), V_t(j)) &\geq \frac{1}{q^2} \left[\sum_{i=1}^q \sum_{j=1}^q d(V_s(i), V_t(j)) \right]^2 \\ &\quad + \frac{(r^2 d(X^*, Y^*) - r^2 d(V_s^*, V_t^*))^2 q^2}{r^2 (q^2 - r^2)} \\ &= \frac{1}{q^2} (q^2 d(V_s^*, V_t^*))^2 \\ &\quad + \frac{r^2 q^2}{q^2 - r^2} (d(X^*, Y^*) - d(V_s^*, V_t^*))^2. \end{aligned}$$

So,

$$\frac{1}{q^2} \sum_{i=1}^q \sum_{j=1}^q d^2(V_s(i), V_t(j)) \geq d^2(V_s^*, V_t^*) + \frac{r^2}{q^2 - r^2} \left(\frac{\epsilon}{2}\right)^2 \quad (19)$$

$$\begin{aligned} &\geq \left(d^2(V_s, V_t) - \frac{\epsilon^5}{50} \right) \\ &\quad + \frac{(\epsilon q (1 - \epsilon/100))^2}{q^2 - (\epsilon q (1 - \epsilon/100))^2} \frac{\epsilon^2}{4} \quad (20) \end{aligned}$$

$$\geq d^2(V_s, V_t) - \frac{\epsilon^5}{50} + \frac{\epsilon^4}{16} \quad (21)$$

Inequality (19) comes from (17), and inequality (20) comes from (10) and (15). Inequality (21) requires $\frac{\epsilon^2(1-\epsilon/100)^2}{1-\epsilon^2(1-\epsilon/100)^2} > \frac{1}{4}$, which holds if $\epsilon < 0.4492$.

Combining the results

Now we use the fact that there are at least $\epsilon\ell^2$ pairs for which (21) holds. And we also use (12) for all pairs.

$$\begin{aligned}
\text{index}(Q) &\geq \frac{1}{\binom{\ell}{2}} \sum_{s=1}^{\ell} \sum_{t=s+1}^{\ell} \left[\frac{1}{q^2} \sum_{i=1}^q \sum_{j=1}^q d^2(V_s(i), V_t(j)) \right] \\
&\geq \frac{1}{\binom{\ell}{2}} \sum_{s=1}^{\ell} \sum_{t=s+1}^{\ell} \left[d^2(V_s, V_t) - \frac{\epsilon^5}{50} \right] + \frac{1}{\binom{\ell}{2}} \left(\epsilon\ell^2 \frac{\epsilon^4}{16} \right) \\
&= \text{index}(P) - \frac{\epsilon^5}{50} + \frac{\ell^2}{\binom{\ell}{2}} \frac{\epsilon^5}{16} \\
&\geq \text{index}(P) - \frac{\epsilon^5}{50} + \frac{\epsilon^5}{8} \\
&\geq \text{index}(P) + \frac{\epsilon^5}{20}.
\end{aligned}$$

□

4 ϵ -irregularity is witnessed by smallest sets

Recall that a pair (A, B) is ϵ -irregular of density $d = d(A, B)$ if there exist $X \subseteq A$, $Y \subseteq B$ with $|X| \geq \epsilon|A|$ and $|Y| \geq \epsilon|B|$ and $|d(X, Y) - d| > \epsilon$.

Proposition 1 *Let (A, B) be an ϵ -irregular pair of density d . Then, there exist $X' \subseteq A$, $Y' \subseteq B$ with $|X'| = \lceil \epsilon|A| \rceil$ and $|Y'| = \lceil \epsilon|B| \rceil$ and $|d(X', Y') - d| > \epsilon$.*

Proof of Proposition 1. By ϵ -irregularity, there exist X and Y such that $X \subseteq A$, $Y \subseteq B$ with $|X| \geq \epsilon|A|$ and $|Y| \geq \epsilon|B|$ and $|d(X, Y) - d| > \epsilon$.

If we choose all possible subsets of X of size $\epsilon|A|$ and all possible subsets of Y of size $\epsilon|B|$, then we will count each edge exactly $\binom{|X|-1}{\lceil \epsilon|A| - 1} \binom{|Y|-1}{\lceil \epsilon|B| - 1}$ times.

So,

$$\begin{aligned}
\sum_{\substack{X' \subseteq A \\ |X'| = \lceil \epsilon|A| \rceil}} \sum_{\substack{Y' \subseteq B \\ |Y'| = \lceil \epsilon|B| \rceil}} e(X', Y') &= \binom{|X| - 1}{\lceil \epsilon|A| - 1} \binom{|Y| - 1}{\lceil \epsilon|B| - 1} e(X, Y) \\
\sum_{\substack{X' \subseteq A \\ |X'| = \lceil \epsilon|A| \rceil}} \sum_{\substack{Y' \subseteq B \\ |Y'| = \lceil \epsilon|B| \rceil}} e(X', Y') &= \frac{\lceil \epsilon|A| \rceil \lceil \epsilon|B| \rceil}{|X||Y|} \binom{|X|}{\lceil \epsilon|A| \rceil} \binom{|Y|}{\lceil \epsilon|B| \rceil} e(X, Y)
\end{aligned}$$

Now we divide by $\lceil \epsilon|A| \rceil \lceil \epsilon|B| \rceil \binom{|X|}{\lceil \epsilon|A|} \binom{|Y|}{\lceil \epsilon|B|}$.

$$\frac{1}{\binom{|X|}{\lceil \epsilon|A|} \binom{|Y|}{\lceil \epsilon|B|}} \sum_{\substack{X' \subseteq A \\ |X'| = \lceil \epsilon|A|}} \sum_{\substack{Y' \subseteq B \\ |Y'| = \lceil \epsilon|B|}} d(X', Y') = d(X, Y)$$

By averaging, there exists a pair (X', Y') , $X' \subseteq X$, $Y' \subseteq Y$ with $|X'| = \lceil \epsilon|A|$, $|Y'| = \lceil \epsilon|B|$ and density at least $d(X, Y)$. There is also one such pair with density at most $d(X, Y)$.

If $d(X, Y) > d + \epsilon$, then we choose a pair such that $d(X', Y') \geq d(X, Y) > d + \epsilon$. If $d(X, Y) < d - \epsilon$, then we choose a pair such that $d(X', Y') \leq d(X, Y) < d - \epsilon$. \square

References

- [1] Szemerédi, Endre, On sets of integers containing no k elements in arithmetic progression. Collection of articles in memory of Juriĭ Vladimirovič Linnik. *Acta Arith.* 27 (1975), 199–245.
- [2] Szemerédi, Endre, Regular partitions of graphs. *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, pp. 399–401, *Colloq. Internat. CNRS, 260, CNRS, Paris, 1978*.
- [3] Van der Waerden, B. L., “Beweis einer Baudetschen Vermutung.” *Nieuw Arch. Wiskunde* 15, 212–216, 1927.