13. Proof of the Balog-Szemeredi-Gowers Theorem

Thur Apr 3

In this lecture we'll prove the BSG theorem that we used in the proof of the BKT theorem. Here's the statement again:

Theorem 13.1 (Balog–Szemerédi–Gowers). *Let A and B be subsets of an abelian group and suppose $X \subset A \times B$. If $|A|, |B| \leq N$, $|X| \geq K^{-1}N^2$, and $|\pi_1(X)| \leq KN$, then there are $A' \subset A$ and $B' \subset B$ such that $|A' + B'| \leq K^{O(1)}N$ and $|X'| \geq K^{-O(1)}N^2$, where $X' = X \cap (A' \times B')$.*

Here $\pi_1(X) = \{a + b \ : \ (a, b) \in X\}$.

Example 13.2. *As subsets of $\mathbb{Z}$, let $A = B$ be the union of $[N]$ and some garbage. Let X be the union of any subset of $[N] \times [N]$ and a little garbage. Then $\pi_1(X)$ is small ($\lesssim N$), while $|A + B| \gtrsim N^2$ is large. We can take $A' = B' = [N]$.*

The theorem was originally proved by Balog and Szemerédi, but in the bounds $K^{O(1)}$ was instead $F(k)$ and $K^{-O(1)}$ was $\frac{1}{F(K)}$, and $F(K)$ was some function with crazy growth. The bounds in the version stated above are due to Gowers.

We will proceed by thinking of $X \subset A \times B$ as a bipartite graph with A on the left, B on the right, and an element $(a, b) \in X$ representing an edge from a to b. Let

$$P_K(a, b) = \#\{\text{paths of length } K \text{ in the graph } X \text{ from } a \text{ to } b\}.$$

Lemma 13.3. *If $A' \subset A$ and $B' \subset B$, and for any $a \in A'$, $b \in B'$, $P_3(a, b) \geq P$, then*

$$|A' + B'| \leq \frac{\pi_1(X)^3}{P}.$$

Proof. A path of length 3 from a to b goes from a to some $b_1 \in B$, then to some $a_1 \in A$, then to b. So $(a, b_1), (a_1, b_1), (a_1, b) \in X$ and hence

$$\underbrace{a + b_1}_{z_1}, \underbrace{a_1 + b_1}_{z_2}, \underbrace{a_1 + b}_{z_2} \in \pi_1(X).$$

We can write $a + b = z_1 - z_2 + z_3$. Therefore

$$\#\{(z_1, z_2, z_3) \in \pi_1(X)^3 \ : \ a + b = z_1 - z_2 + z_3\} \geq P_3(a, b) \geq P.$$

Summing over $A' + B'$ we get

$$|A' + B'| \cdot P \leq |\pi_1(X)|^3.$$

∎

From now on, everything we prove will be a statement about bipartite graphs, i.e. we don't need the addition law for anything that follows.

Lemma 13.4 (Key Lemma)**.** *If $X \subseteq A \times B$ and $|X| \geq K^{-1}|A||B|$, then there are $A' \subset A$ and $B' \subset B$ such that $|X'| \geq K^{-O(1)}|A||B|$ where $X' = X \cap (A' \times B')$ and for any $a \in A'$, $b \in B'$,*

$$P_3(a,b) \geq K^{-O(1)}|A||B|.$$

The BSG theorem is proved by combining Lemma 13.3 and the Key Lemma.

13.1. Simple Bounds About $P_K(a,b)$.

In this section, we have

$$\begin{aligned} \#\text{ edges } &= |X| =\geq K^{-1}|A||B|, \\ P_\ell &:= \#\text{paths of length } \ell \text{ starting in } A, \\ P_1 &= |X| \geq K^{-1}|A||B|. \end{aligned}$$

Definition 13.5. *For $a \in A$, the **neighborhood** of a is the set $N(a)$ of points that share an edge with a.*

We can average over $|A|$ to get

$$\operatorname{Avg}_{a\in A} P_1(a,\cdot) = \frac{|P_1|}{|A|} \geq K^{-1}|B|.$$

To get an estimate for the average of P_2, we use Cauchy-Schwarz to get

$$\begin{aligned} P_2 &= \sum_b |N(b)|^2 \\ &\geq \frac{\left(\sum_b |N(b)|\right)^2}{|B|} \\ &\geq \frac{(K^{-1}|A||B|)^2}{|B|} \\ &= K^{-2}|A|^2|B|. \end{aligned}$$

Averaging this get us

$$\operatorname{Avg}_{a_1,a_2} P_2(a_1,a_2) \geq K^{-2}|B|.$$

As an exercise, use similar methods to prove $|P_3| \geq K^{-3}|A|^2|B|^2$ and

$$\operatorname{Avg}_{a,b} P_3(a,b) \geq K^{-3}|A||B|.$$

The Key Lemma says that $P_3(a,b)$ is at least a small fraction of the average for *all* $a \in A'$, $b \in B'$.

Lemma 13.6 (Length 2)**.** *If $X \subset A \times B$, $|X| \geq K^{-1}|A||B|$, $\epsilon > 0$, then there is a subset $A' \subset A$ such that $|A'| \geq \frac{1}{2}K^{-1}|A|$ and $P_2(a_1,a_2) \geq \epsilon K^{-2}|B|$ for $(1-2\epsilon)|A'|^2$ choices of $(a_1,a_2) \in (A')^2$.*

Note that we cannot always take $A' = A$, because there are graphs X where only $\frac{1}{K}|A|$ vertices in A have an edge and there are also graphs with multiple connected components. What we will do is let $A' = N(b)$ for some $b \in B$.

Definition 13.7. *A pair (a_1, a_2) is **ϵ-bad** if $P_2(a_1, a_2) < \epsilon K^{-2}|B|$. Let*

$$BP_\epsilon(b) = \#\{(a_1, a_2) \in N(b)^2 \;:\; (a_1, a_2) \text{ is } \epsilon\text{-bad}\}.$$

Lemma 13.8 (P1)**.**

$$\mathbb{E}_b|BP_\epsilon(b)| \le \epsilon K^{-2}|A|^2.$$

Lemma 13.9 (P2)**.**

$$\mathbb{E}_b|N(b)|^2 \ge K^{-2}|A|^2.$$

This says there's only about an ϵ-fraction of bad pairs.

Proof of P2. By Cauchy-Schwarz,

$$\begin{aligned}\sum_b |N(b)|^2 &\ge \frac{\left(\sum_b |N(b)|\right)^2}{|B|} \\ &\ge \frac{(K^{-1}|A||B|)^2}{|B|} \\ &= K^{-2}|A|^2|B|.\end{aligned}$$

Divide by $|B|$. ☐

Proof of P1.

$$\begin{aligned}\sum_b |BP_\epsilon(b)| &= \#\{a_1, a_2, b \text{ such that } (a_1, b), (a_2, b) \in X \text{ and } P(a_1, a_2) \le \epsilon K^{-2}|B|\} \\ &\le |A|^2 \epsilon K^{-2}|B|.\end{aligned}$$

Divide by $|B|$. ☐

Proof of Length 2. Let $A' = N(b)$. Then by the previous two lemmas,

$$\mathbb{E}\left(|N(b)|^2 - \frac{1}{2\epsilon}|BP_\epsilon(b)|\right) \ge \frac{1}{2}K^{-2}|A|^2.$$

So we can pick b to satisfy

$$|N(b)|^2 - \frac{1}{2\epsilon}|BP_\epsilon(b)| \ge \frac{1}{2}K^{-2}|A|^2$$

and let $A' = N(b)$. Then $|BP_\epsilon(b)| \le 2\epsilon|N(b)|^2$. ☐

By discarding some a_1's, we can upgrade this.

Lemma 13.10 (2)**.** *If $X \subset A \times B$, $|X| \geq K^{-1}|A||B|$, and $\epsilon > 0$, then there exists $A_2 \subset A$ such that $|A_2| \geq \frac{1}{4}K^{-1}|A|$ and for every $a \in A_2$, there are at most $10\epsilon|A_2|$ choices for a_2 such that (a, a_2) is ϵ-bad.*

We won't prove this, but the idea is to let

$$A_2 = A' \setminus \{a \in A' \ : \ (a, a_2) \text{ is } \epsilon\text{-bad for many } a_2 \in A'\}.$$

The second part of the conclusion can be written as $A = B(a) \cup G(a)$, where $|B(a)| \leq 10\epsilon|A_2|$ and for any $a_2 \in G(a)$, $P_2(a, a_2) \geq \epsilon K^{-2}|B|$.

Proof of Key Lemma. First, let

$$A_1 = \{a \in A \ : \ |N(a)| \geq \frac{1}{10}K^{-1}|A|\}.$$

Let

$$X(A', B') := \{(a, b) \in (A' \times B') \cap X\} = (A' \times B') \cap X.$$

Choose $A' \subset A$ be the A_2 of Lemma 2. Let

$$B' = \{b \in B \ : \ |N(b) \cap A'| > 20\epsilon|A'|\}$$

so

$$|B(a)| \leq 10\epsilon|A'|.$$

For any $a \in A'$, $b \in B'$, we have

$$\begin{aligned} P_3(a, b) &\geq \epsilon K^{-2}|B|(|N(b) \cap G(a)|) \\ &\geq \epsilon K^{-2}|B|(|N(b) \cap A'| - |B(a)|) \\ &\gtrsim \epsilon^2 K^{-3}|A||B| \end{aligned}$$

using $|A'| \gtrsim K^{-1}|A|$. Now we just need to check $|X(A', B')| \geq K^{-O(1)}|X|$. Since $|A'| \gtrsim K^{-1}|A|$, $A' \subset A$, $N(a) \geq \frac{1}{10}K^{-1}|B|$ for $a \in A'$, and $|X(A', B)| \gtrsim K^{-2}|A||B|$, so

$$\begin{aligned} |X(A', B \setminus B')| &\leq 20\epsilon|A'||B| \\ &\leq 20\epsilon|A||B|. \end{aligned}$$

Let $\epsilon = \frac{1}{10^6}K^{-2}$, so $|X(A', B \setminus B')| \ll |X(A', B)|$. Hence

$$|X(A', B')| \sim |X(A', B)| \geq K^{-O(1)}|A||B|.$$

□

18.156 Projection Theory
Spring 2025