We begin by proving Lemma 3.4 on the first handout.

Proof of Lemma 3.4.

$\square$

Lemma 5.2 (Equivalent definitions of additive energy). *Let A and B be finite non-empty sets in a commutative group. Then*

$$\begin{aligned} E(A,B) &= \sum_{x+y=z+w} 1_A(x)1_B(y)1_A(z)1_B(w) \\ &= \sum_{x-w=z-y} 1_A(x)1_B(y)1_A(z)1_B(w) \\ &= \sum_{x\in A-B} r_{A-B}(x)^2 \\ &= \sum_{a\in A, b\in B} |(a+B)\cap(A+b)| \\ &= \sum_{a\in A, b\in B} |(a-B)\cap(A-b)|. \end{aligned}$$

Proof.

□

The last two identities combined with the Cauchy–Schwarz lower bound and averaging arguments allow one to prove variants of Lemma 3.4. On of the four possible statements is.

Lemma 5.3. *Let $\alpha\in\mathbb{R}$ and A,B be finite non-empty sets in a commutative group. Suppose that $|A+B|\le\alpha|A|$. There exists $S\subseteq B-A$, $|S|\le\lceil\alpha\log(|B|)\rceil$ such that $B\subseteq S+A$.*

6 The Balog–Szemerédi–Gowers theorem

This celebrated theorem is a converse of sorts to the statement 'small doubling implies large additive energy'.

We have seen that 'small doubling' means a doubling constant not far off from the absolute minimum that is 1. So 'large additive energy' must mean additive energy not far off the absolute maximum. What is this absolute maximum?

Lemma 6.1. *Let A and B be finite sets in a commutative group. Their additive energy is bounded by each of the quantities $|A|^2|B|$, $|A||B|^2$ and $|A|^{3/2}|B|^{3/2}$.*

In particular $E(A, A) \le |A|^3$.

Proof.

$\square$

Let us now check what can we say about some sets with large additive energy.

Real time exercise. For each of the following sets estimate: the additive energy, the doubling constant, the largest cardinality of a subset that has small doubling – here take small to mean a number smaller than any power of the cardinality of the subset.

(i) $A = \mathbb{Z}_n$.

(ii) $A = \mathbb{Z}_n \times \{0\} \cup \{0\} \times \mathbb{Z}_n \subset \mathbb{Z}_n^2$. A can be written in a convenient if sloppy way as $(\mathbb{Z}_n, 0) \cup (0, \mathbb{Z}_n)$.

(iii) $A = (\mathbb{Z}_n, 0, \ldots, 0) \cup (0, \mathbb{Z}_n, 0, \ldots, 0) \cup \cdots \cup (0, \ldots, 0, \mathbb{Z}_n) \subseteq \mathbb{Z}_n^d$.

Solution.

The conclusions of the last example and Lemma 5.1 is that we cannot hope to do much better than the following statement: if $E(A,A) \geq \delta|A|^3$, then A must contain a subset A' of relative density at lest δ and doubling at most δ^{-1}.

The Balog–Szemerdi–Gowers is a statement like the above, where δ is replaced by some power. Balog and Szemerédi had proved the theorem for much worse bounds.

We will be a little sloppy in the statement and proof of the theorem by not keeping track of constants. We write $P \ll Q$ if there exists a constant C such that $P \leq CQ$ and $P \gg Q$ if there exists a constant C such that $P \geq CQ$.

Theorem 6.2 (Balog–Szemerédi–Gowers)**.** *Let $\delta > 0$ be a real number and A be a finite set in a commutative group. Suppose that $E(A,A) \geq \delta|A|^3$.*

There exists a subset $A' \subseteq A$ of cardinality at least

$$|A'| \gg \delta^{10}|A|$$

and difference set cardinality at most

$$|A' - A'| \ll \delta^{-32}|A'|.$$

Remark. There are more efficient versions of this result. The important fact is that in the conclusion only powers of δ or δ^{-1} appear.

Idea of the proof:

The first part described above is accomplished by a random selection process.

Lemma 6.3 (Gowers)**.** *Let m and n be positive integers and $\varepsilon > 0$ a positive real number. Suppose $A_1, \ldots, A_m$ are sets in $\{1, \ldots, n\}$ such that* $\sum_{i=1}^{m} |A_i| \geq \varepsilon nm$.

There exists a subset $B \subseteq \{1, \ldots, m\}$ of size at least $|B| \geq \frac{\varepsilon^5}{2} m$ such that for at least 90% of pairs $(i, j) \in B \times B$, $|A_i \cap A_j| \geq \frac{\varepsilon^2}{2} n$.

Proof.

$\square$

Remarks. The method is called dependent random choice, The number of random points x_i is determined by the desired degree of accuracy – 5 corresponds to 15/16.

Proof of Balog–Szemerédi–Gowers. Reference list:

- popular difference d: $r_{A-A}(d) \geq \delta|A|/2$.

- G graph with vertex set A and ab an edge iff ab is a popular difference.

- $B \subseteq A$ where at least 90% of pairs (b, b') satisfy $|\Gamma_G(b) \cap \Gamma_G(b')| \geq \delta^4|A|/32$. $|B| \ll \alpha^{10}|A|$.

- H graph with vertex set B and bb' an edge iff $|\Gamma_G(b) \cap \Gamma_G(b')| \geq \delta^4|A|/32$.

- $A' \subseteq B$ is determined by $a' \in A'$ if $|\Gamma_H(a')| \geq 4|B|/5$. $|A'| \geq |B|/2 \gg \alpha^{10}|A|$.

$\square$