Lecture 13. Geometric constructions, part 1

Hong Liu

3rd July 2021

We will go over some geometric constructions in extremal combinatorics.

1 3-AP-free sets and corner-free sets

We have seen previously that Roth's theorem states that any subset of $[n]$ without 3-term arithmetic progressions (3-AP-free) has size $o(n)$. The classical lower bound construction, due to Behrend, gives a 3-AP-free subset of $[n]$ of size $\frac{n}{2^{(c+o(1))\sqrt{\log_2 n}}}$, where $c = 2\sqrt{2}$.

Here, we consider a 2-dimension problem in which we want to avoid a 2-dimension configuration defined as follows. A *corner* in $[n]^2$ is a triple of points of the form $(x, y), (x+d, y), (x, y+d)$, where $d > 0$. Then as you might already guess, we call a subset of the grid $A \subseteq [n^2]$ *corner-free* if it does not contain any corner.

How large a corner-free subset of $[n]^2$ can be? We have similar bounds as in the 3-AP-free case.

1.1 Upper bound via triangle removal lemma

Theorem 1.1. *If $A \subseteq [n]^2$ is corner-free, then $|A| = o(n^2)$.*

Sketch of proof. Construct a tripartite graph G with vertices being horizontal, vertical lines and lines with slope -1. That is, $V(G) = X \cup Y \cup Z$, where $X = \{y = i, i \in [n]\}$, $Y = \{x = i, i \in [n]\}$ and $Z = \{x+y = i, i \in [2n]\}$. Two vertices in G are adjacent if their corresponding line intersect at a point in our corner-free set A. By definition, each point $a \in A$ give rises to a triangle T_a in G, and so $e(G) = 3|A|$. Note also that T_a, $a \in A$, are pairwise edge-disjoint triangles. Suppose that $|A| = \Omega(n^2)$, then G contains $\Omega(n^2)$ many edge-disjoint triangles. Consequently, by triangle removal lemma, G must contain in fact $\Omega(n^3)$ triangles in total. That means there is a triangle in G not corresponding to a point in A. A moment of thought shows that such triangle in G corresponds to a corner in A, a contradiction. □

1.2 Dense corner-free sets

Previously best known corner-free sets construction is coming from using dense 3-AP-free sets construction of Behrend. To see this, consider a 3-AP-free set $B \subseteq [n]$, and define $A = \{(x, y) : x - y \in B\}$. This yields $|A| = \Theta(|B|n) = \frac{n^2}{2^{(c+o(1))\sqrt{\log_2 n}}}$, where $c = 2\sqrt{2} \approx 2.828...$.

The constant c above was improved by Linial and Shraibman earlier this year, and further improved by Green. We shall present Green's construction which yields a corner-free set of size $\frac{n^2}{2^{(c+o(1))\sqrt{\log_2 n}}}$, where $c = 2\sqrt{2\log_2 \frac{4}{3}} \approx 1.822...$.

The geometric idea behind this construction is the same as Behrend's: any line can intersects a sphere on at most two points.

We will consider n of the form $n = q^d$, with q, d to be determined later. We make use of the 1-to-1 correspondence between $[q^d - 1]$ and the high-dimension grid/cube $\{0, 1, \dots, q-1\}^d$ via base q expansion. That is, consider the map $\pi\colon [q^d-1] \mapsto \{0,1,\dots,q-1\}^d$, where

$$\pi(x) = (x_0, x_1, \dots, x_{d-1}), \quad \text{s.t. } x = \sum_{i=0}^{d-1} x_i q^i,\ 0 \le x_i < q, \forall i.$$

Consider the set $A_r \subseteq [n]^2$ consisting of all the points (x, y) such that

$$\|\pi(x) - \pi(y)\|_2^2 = r, \tag{1}$$

$$\text{and} \quad \frac{q}{2} \le x_i + y_i < \frac{3q}{2}, \forall i. \tag{2}$$

We shall show that A_r is corner-free, and the desired bound follows from taking the densest slice A_r and optimising the parameters q, d.

Suppose for a contradiction that there is a corner $(x, y), (x+d, y), (x, y+d) \in A_r$.

Claim 1.2. $\pi(x+d) + \pi(y) = \pi(x) + \pi(y+d)$.

Proof of claim. Induct on $i = 0, 1, \dots$ to show that $(x+d)_i + y_i = x_i + (y+d)_i$. Suppose for $j \ge 0$, it holds for all $i < j$. Write $x_{\ge j} = \sum_{i \ge j} x_i q^i$. Then the induction hypothesis and $(x+d) + y = x + (y+d)$ imply that

$$(x+d)_{\ge j} + y_{\ge j} = x_{\ge j} + (y+d)_{\ge j},$$

which in turns implies that

$$(x+d)_j + y_j = x_j + (y+d)_j, \quad \text{mod } q.$$

But both sides above lie in $[\frac{q}{2}, \frac{3q}{2}]$ due to (2). Thus in fact $(x+d)_j + y_j = x_j + (y+d)_j$. ■

Now let $a = \pi(x) - \pi(y)$ and let $b = \pi(x+d) - \pi(x) = \pi(y+d) - \pi(y)$ due to the above claim. Then as $(x, y), (x+d, y), (x, y+d) \in A_r$, by (1), we have

$$\|a\|_2^2 = \|\pi(x) - \pi(y)\|_2^2 = r, \quad \|a+b\|_2^2 = \|\pi(x+d) - \pi(y)\|_2^2 = r, \quad \|a-b\|_2^2 = \|\pi(x) - \pi(y+d)\|_2^2 = r.$$

We obtain a contradiction as $a-b, a, a+b$ lie on a line and so cannot all be on the sphere of radius $\sqrt{r}$.

To estimate the size of A_r, note that there are $\left(\frac{3}{4}q^2 + O(q)\right)^d$ many points satisfying (2), and there are at most $d \cdot q^2$ choices of r. Thus, by Pigeonhole Principle, there exists r such that $|A_r| \ge \frac{1}{dq^2}\left(\frac{3}{4}q^2 + O(q)\right)^d$. Setting $q = (\frac{2}{\sqrt{3}})^d$ and $n = q^d$, we get the desired bound.

We leave as an exercise to work out the details of Behrend's construction of 3-AP-free sets. Start with again $n = q^d$ and a set B_r consisting of all $x \in [q^d - 1]$ such that

$$\|\pi(x)\|_2^2 = r,$$

$$\text{and} \quad 0 \le x_i < \frac{q}{2}, \forall i.$$

We remark that Green's improvement on corner-free set comes from using the 2-dimension strip in (2) rather than just the (1-dimension) half-size interval in Behrend's construction.

2 Erdős-Rothschild's question on book

A *book* is a set of triangles sharing one common edge, and the size of the book is the number of triangles it contains. Denote by $b(G)$ the size of the largest book in a graph G. It was proved by Edwards, and independently Khadžiivanov and Nikiforov, that any n-vertex graph with at least $n^2/4+1$ edges contain a book of size at least $n/6$. On the other hand, if an n-vertex graph has at most $n^2/4$ edges, it could just be a subgraph of $K_{n/2,n/2}$ and so has no triangle, not to mention a book. What if we require that every edge of the graph must lie in a triangle? This was a question of Erdős and Rothschild. Is it true that for any $c > 0$, there exists $c' > 0$ such that the following holds: given an n-vertex graph G with cn^2 edges, if every edge lies in at least one triangle, then $b(G) \geq n^{c'}$.

This was disproved by Fox and Loh. They gave a geometric construction of a graph with edge density almost 1/2 and every edge lying in a triangle, but having sub-polynomial book size.

Theorem 2.1 (Fox-Loh 12)**.** *There is an n-vertex graph with $\frac{n^2}{4}\left(1 - e^{-(\log n)^{1/6}}\right)$ edges and every edge lying in a triangle, such that $b(G) \leq n^{14/\log\log n}$.*

The idea of their construction is the following. By concentration of measure, a high dimensional cube has a nice property that almost all pair of points are at a typical distance apart. Take three copies of cubes, say A, B, C. Then add edges between A and B when points are at typical distance; so A, B are almost completely joined. Between A (same for B) and C, add edges when points are at about half of the typical distance. Then it is not hard to show that every edge lies in a triangle. The fact that there is no large book follows from the geometric fact that two not-so-close high dimension balls have intersection size exponentially smaller than their volumes. Finally, blow up A and B to boost the edge density to 1/2.

Let us now see the detailed construction.

Step 1. Fix integer $r > 2$, $d = r^5$, and $n = r^d$. Define

$$\mu = \frac{r^2 - 1}{6} \cdot d.$$

Let $A = B = C = [r]^d$. Vertices $a \in A$ and $b \in B$ are adjacent if and only if their distance satisfies $\|a - b\|_2^2 = \mu \pm d$. Vertices $a \in A$ (same for $b \in B$) and $c \in C$ are adjacent if and only if their distance satisfies $\|a - c\|_2^2 = \frac{1}{4}\mu \pm 2d$.

Step 2. Blow up $A \cup B$. More precisely, replace each vertex of $A \cup B$ with 2^d copies of itself. Two copies of vertices are adjacent if their original vertices were adjacent.

The following lemma says that most of the pairs of points in the cube $[r]^d$ are at a typical distance apart, and so A, B are almost complete to each other. The proof follows from a simple application of concentration inequality; we leave it as an exercise.

Lemma 2.2. *Let x, y be two uniformly chosen points in $[r]^d$. Then*

$$\mathsf{Pr}\left(\|x - y\|_2^2 = \mu \pm d\right) \geq 1 - 2e^{-\frac{d}{2r^4}}.$$

The following lemma implies that the construction above has the properties as in Theorem 2.1. For its proof, we need to estimate the volume of Euclidean balls: for even d and real $r > 0$, the volume of the d-dimensional Euclidean ball with radius r is

$$\frac{\pi^{d/2} r^d}{(d/2)!} < (2\pi e)^{d/2} \cdot \frac{r^d}{d^{d/2}}. \tag{3}$$

Lemma 2.3. *Before blowing up $A \cup B$, every edge is contained in between 2^{d-1} and 15^d triangles.*

Proof. Fix an edge ab, we first prove the lower bound that ab is in at least 2^{d-1} triangles. Let $m = (m_1, \ldots, m_d)$ be the midpoint of $a = (a_1, \ldots, a_d)$ and $b = (b_1, \ldots, b_d)$, and let $x_i = b_i - a_i$.

Then $m_i - a_i = \frac{x_i}{2}$. For each i, let $\delta_i = \frac{1}{2}$ if x_i is odd, and $\delta_i = 1$ otherwise. Consider points $c = (c_1, \ldots, c_d) \in C$ with $c_i = m_i + \delta_i \varepsilon_i$, where $\varepsilon_i \in \{\pm 1\}$. Then

$$\|c-a\|_2^2 = \sum_i \left(\frac{x_i}{2} + \delta_i\varepsilon_i\right)^2 = \frac{\|b-a\|_2^2}{4} + \sum_i \delta_i^2 + \sum_i x_i\delta_i\varepsilon_i,$$

and similarly $\|b-c\|_2^2 = \frac{\|b-a\|_2^2}{4} + \sum_i \delta_i^2 - \sum_i x_i\delta_i\varepsilon_i$. As $\|b-a\|_2^2 = \mu \pm d$ and $\sum_i \delta_i^2 \le d$, we see that each choice of $(\varepsilon_1, \ldots, \varepsilon_d)$ satisfying

$$\Big|\sum_i x_i\delta_i\varepsilon_i\Big| \le \frac{3}{4}d$$

gives rise to a point $c \in C$ forming a triangle with ab.

Consider now each ε_i being a Rademacher random variable (i.e. with equal probability to be 1 or -1), then the random variable $Z := \sum_i x_i\delta_i\varepsilon_i$ is $2r$-Lipschitz (as $|x_i| \le r$) and has mean 0. An application of Azuma inequality gives

$$\mathsf{Pr}\Big(|Z| > \frac{3}{4}d\Big) < 2\exp\Big(-\frac{(\frac{3}{4}d)^2}{2(2r)^2 d}\Big) < 2\exp\Big(-\frac{d}{15r^2}\Big).$$

Thus, the number of coneighbours c of a and b is at least $\mathsf{Pr}(|Z| \le \frac{3}{4}d) \cdot 2^d > 2^{d-1}$.

We now move to upper bound. Let again $x_i = b_i - a_i$ and suppose $c =$ forms a triangle with ab, thus $\|b-c\|_2^2 = \|a-c\|_2^2 = \frac{1}{4}\mu \pm 2d$, say $c_i = a_i + \frac{x_i}{2} + \frac{w_i}{2}$ for some $w_i \in \mathbb{Z}$. Then

$$\|c-a\|_2^2 = \sum_i \left(\frac{x_i}{2} + \frac{w_i}{2}\right)^2 = \frac{\|b-a\|_2^2}{4} + \frac{1}{4}\sum_i w_i^2 + \frac{1}{2}\sum_i x_i w_i,$$

and similarly $\|b-c\|_2^2 = \frac{\|b-a\|_2^2}{4} + \frac{1}{4}\sum_i w_i^2 - \frac{1}{2}\sum_i x_i w_i$. So

$$\frac{\mu}{2} \pm 4d = \|c-a\|_2^2 + \|b-c\|_2^2 = \frac{\|b-a\|_2^2}{2} + \frac{1}{2}\sum_i w_i^2 = \frac{\mu}{2} \pm \frac{d}{2} + \frac{1}{2}\sum_i w_i^2,$$

and consequently

$$\sum_i w_i^2 \le 9d.$$

We are left to bound the number of lattice points in a d-dimensional Euclidean ball of radius $3\sqrt{d}$, which is at most the volume of a radius-$3.5\sqrt{d}$ ball. The desired bound then follows from (3). $\square$