Partial Symmetries

Tristan Shin

31 Jan 2025–7 Feb 2025

1 Distinct distances and partial symmetries

Recall the distinct distances problem:

Question 1.1: Distinct distances problem

Let $\mathcal{P}$ be a set of N points. Let $d(\mathcal{P}) = \{d(p_1, p_2) : p_1, p_2 \in \mathcal{P}^2, p_1 \neq p_2\}$. What is the minimum possible size of $d(\mathcal{P})$?

Let $d(N)$ be the answer to the question. The conjecture is that $d(N) \gtrsim N/\sqrt{\log N}$, because when $\mathcal{P}$ is a $\sqrt{N} \times \sqrt{N}$ grid, we have $|d(\mathcal{P})| \asymp N/\sqrt{\log N}$. Previously, we got a lower bound of $N^{2/3}$ from the Spencer–Szemerédi–Trotter bound on the unit distance problem, and we mentioned a crossing lemma proof of a $N^{4/5}$ lower bound.

The key insight of Elekes and Sharir was to tackle this problem using partial symmetries of $\mathcal{P}$. Let G be the group of orientation-preserving rigid motions of $\mathbb{R}^2$. The **r-rich partial symmetries** of $\mathcal{P} \subset \mathbb{R}^2$ are the elements of the set

$$G_r(\mathcal{P}) \coloneqq \{g \in G : |g(\mathcal{P}) \cap \mathcal{P}| \geq r\}.$$

In a sense, this has the feel of an incidence problem, where "points" are elements of G and "lines" are mappings between two points of $\mathcal{P}$. We will see later that this intuition is correct, i.e. bounding $|G_r(\mathcal{P})|$ is equivalent to an incidence problem of lines in $\mathbb{R}^3$.

First for some examples:

- For a generic set $\mathcal{P}$ of N points, we have $|G_2(\mathcal{P})| = \binom{N}{2} + 1$ and $|G_r(\mathcal{P})| = 1$ for $r \geq 3$.
- If $\mathcal{P}$ is a $\sqrt{N} \times \sqrt{N}$ grid, then $|G_r(\mathcal{P})| \asymp N^3 r^{-2}$ for $2 \leq r \leq N/2$. Details are a bit tricky; see book.

We will soon see that partial symmetries and distinct distances are closely related. If this is true, it is natural to guess that the grid is an extremum of $|G_r(\mathcal{P})|$. That is, we guess that $|G_r(\mathcal{P})| \lesssim N^3 r^{-2}$. Let us assume this, and show that we can get a very strong lower bound of $d(N) \gtrsim N/\log N$, which nearly matches the best upper bound given by the grid.

1.1 Small distance set implies large distance "energy"

Let

$$Q(\mathcal{P}) \coloneqq \{(p_1, p_2, q_1, q_2) \in \mathcal{P}^4 \,:\, d(p_1, p_2) = d(q_1, q_2) \neq 0\}.$$

Then $|Q(\mathcal{P})|$ is somewhat of a combinatorial "energy" quantity. A general principal guiding these energy quantities is that if the object being measured has small count, then the energy must be big. Equivalently, if the energy is small, then the count of the object must be big. (This is used in the combinatorics of sumsets, where the object is instead elements of $A+A$.)

Lemma 1.2

$|d(\mathcal{P})| \cdot |Q(\mathcal{P})| \gtrsim N^4$

Proof. For every $d \in \mathbb{R}_{>0}$, let $r(d) = \#\{(p_1, p_2) \in \mathcal{P}^2 \,:\, d(p_1, p_2) = d\}$. Then

$$|Q(\mathcal{P})| = \sum_{d \in d(\mathcal{P})} r(d)^2.$$

So the Cauchy–Schwarz inequality implies that

$$|d(\mathcal{P})| \cdot |Q(\mathcal{P})| = \left(\sum_{d \in d(\mathcal{P})} 1\right)\left(\sum_{d \in d(\mathcal{P})} r(d)^2\right) \geq \left(\sum_{d \in d(\mathcal{P})} r(d)\right)^2 = (N^2 - N)^2. \qquad \blacksquare$$

So if we can get an upper bound on the energy, we would get a lower bound on $d(N)$.

1.2 Computing energy using partial symmetries

Actually, the distance energy is closely related to the partial symmetries.

First we discuss the structure of rigid motions if I tell you a particular mapping. For any $p, q \in \mathbb{R}^2$, define

$$S_{p \mapsto q} \coloneqq \{g \in G \,:\, g(p) = q\}.$$

These are the "lines" in the earlier analogy to incidence geometry.

Then $S_{p \mapsto q}$ is a coset of a subgroup of G. Indeed, any $g \in S_{p \mapsto q}$ must be the translation from p to q, followed by any rotation around q. So if $R_q \leq G$ is the subgroup of rotations around q and $\tau_{q-p} \in G$ is the translation from p to q, then $S_{p \mapsto q}$ is the right coset $R_q \circ \tau_{q-p}$.

Lemma 1.3

$Q(\mathcal{P}) = \{(p_1, p_2, q_1, q_2) \in \mathcal{P}^4 \,:\, |S_{p_1 \mapsto q_1} \cap S_{p_2 \mapsto q_2}| = 1\}$

Proof. Suppose $(p_1, p_2, q_1, q_2) \in Q(\mathcal{P})$. Let $g \in S_{p_1 \mapsto q_1} \cap S_{p_2 \mapsto q_2}$. Then there exists a rotation $\gamma \in R_{q_1}$ such that $g = \gamma \circ \tau_{q_1 - p_1}$. But $g(p_2) = q_2$, so we need $\gamma(p_2 + q_1 - p_1) = q_2$. But $d(q_1, q_2) = d(p_1, p_2) = d(q_1, p_2 + q_1 - p_1)$, so there is a unique γ sending $p_2 + q_1 - p_1$ to q_2.

The converse is a quick check using the fact that $d(p_1, p_2) = d(g(p_1), g(p_2))$. □

Using the lemma, define a map $E \colon Q(\mathcal{P}) \to G$ by letting $E(p_1, p_2, q_1, q_2)$ be the unique $g \in S_{p_1 \mapsto q_1} \cap S_{p_2 \mapsto q_2}$. The fibres of E are very well-controlled by partial symmetries. For shorthand, let

$$G_{=r}(\mathcal{P}) \coloneqq \{g \in G \,:\, |g(\mathcal{P}) \cap \mathcal{P}| = r\} = G_r(\mathcal{P}) \setminus G_{r+1}(\mathcal{P}).$$

Note that the $G_{=r}(\mathcal{P})$ for $r = 0, \ldots, N$ partition G.

Lemma 1.4

If $g \in G_{=r}(\mathcal{P})$, then $|E^{-1}(g)| = r^2 - r$.

Proof. Note that

$$E^{-1}(g) = \{(p_1, p_2, g(p_1), g(p_2)) \,:\, p_1, p_2 \in \mathcal{P}, p_1 \neq p_2\} \cap \mathcal{P}^4,$$

so

$$\begin{aligned} |E^{-1}(g)| &= \#\{(p_1, p_2) \in \mathcal{P}^2 \,:\, (g(p_1), g(p_2)) \in g(\mathcal{P}) \cap \mathcal{P}, p_1 \neq p_2\} \\ &= \#\{(p_3, p_4) \in (g(\mathcal{P}) \cap \mathcal{P})^2 \,:\, p_3 \neq p_4\} \\ &= r^2 - r \end{aligned}$$

as claimed. □

So by counting fibres, we get that

$$\begin{aligned} |Q(\mathcal{P})| &= \sum_{r=1}^{N} |G_{=r}(\mathcal{P})| \cdot (r^2 - r) \\ &\overset{\text{Abel}}{=} |G_1(\mathcal{P})| \cdot 0 - \sum_{r=2}^{N} |G_r(\mathcal{P})| \cdot ((r-1)^2 - (r-1) - (r^2 - r)) \\ &= \sum_{r=2}^{N} |G_r(\mathcal{P})| \cdot (2r - 2). \end{aligned}$$

So we deduce that:

Lemma 1.5

$$|Q(\mathcal{P})| \asymp \sum_{r=2}^{N} |G_r(\mathcal{P})| \cdot r$$

1.3 Bounding $d(N)$

Now we can put together our tools to bound $d(N)$.

Proposition 1.6

Let $\epsilon \geq 0$. If $|G_r(\mathcal{P})| \lesssim_\epsilon N^{3+\epsilon} r^{-2}$ whenever $|\mathcal{P}| = N$ and $r \geq 2$, then $d(N) \gtrsim_\epsilon N^{1-\epsilon}/\log N$.

Proof. We have that

$$|Q(\mathcal{P})| \asymp \sum_{r=2}^{N} |G_r(\mathcal{P})| \cdot r \lesssim_\epsilon \sum_{r=2}^{N} N^{3+\epsilon} r^{-1} \asymp N^{3+\epsilon} \log N.$$

Then

$$d(N) \gtrsim \frac{N^4}{|Q(\mathcal{P})|} \gtrsim_\epsilon \frac{N^{1-\epsilon}}{\log N}$$

as desired. ■

In particular, if we can show the assumption with $\epsilon = 0$, we get the current best lower bound of $d(N) \gtrsim N/\log N$ by Guth and Katz. We will not manage to get there with the tools we currently have, but we will be able to get any $\epsilon > 0$, which in turn implies that $d(N) \gtrsim_\epsilon N^{1-\epsilon}$ for any $\epsilon > 0$ (by absorbing the log term in to the N^ϵ term).

Remark. This argument, even for $\epsilon = 0$, falls short of the $N/\sqrt{\log N}$ upper bound. Where did we lose ground in the grid example? When we applied Cauchy–Schwarz to relate the distance set and energy, we only have equality when $r(d)$ is around constant on all of $d(\mathcal{P})$. But this is far from the case in the grid example, where different distances have widely varying counts.

2 Incidence geometry of the $S_{p \mapsto q}$

The goal now is to show that $|G_r(\mathcal{P})| \lesssim_\epsilon N^{3+\epsilon} r^{-2}$ for any $\epsilon > 0$.

As promised earlier, the problem of bounding $|G_r(\mathcal{P})|$ can be translated into an incidence geometry problem about lines in $\mathbb{R}^3$. Recall our sets $S_{p \mapsto q}$ from earlier, the set of $g \in G$ that send $p \mapsto q$. Let $\mathcal{S} = \{S_{p \mapsto q}\}_{p,q \in P}$. Then one can check that

$$G_r(\mathcal{P}) = \{g \in G : g \text{ on} \geq r \text{ curves of } \mathcal{S}\}.$$

2.1 Taking care of translations

First, we can count the number of elements in $G_r(\mathcal{P})$ that are actually translations. Let $G^{\text{trans}} \leq G$ be the subgroup of translations.

Proposition 2.1

For $r \geq 2$, we have that $|G_r(\mathcal{P}) \cap G^{\text{trans}}| \lesssim N^3 r^{-2}$.

Proof. Let $g \in G_r(\mathcal{P}) \cap G^{\text{trans}}$. By Lemma 1.4, we have that $|E^{-1}(g)| \geq r^2 - r$. Since $g \in G^{\text{trans}}$, we have that $E^{-1}(g) \subseteq E^{-1}(G^{\text{trans}})$. Furthermore, the $E^{-1}(g)$ are disjoint for distinct g. So

$$|E^{-1}(G^{\text{trans}})| \geq \sum_{g \in G_r(\mathcal{P}) \cap G^{\text{trans}}} |E^{-1}(g)| \geq |G_r(\mathcal{P}) \cap G^{\text{trans}}| \cdot (r^2 - r).$$

Now it suffices to bound $|E^{-1}(G^{\text{trans}})|$. Every $(p_1, p_2, q_1, q_2) \in E^{-1}(G^{\text{trans}})$ satisfies $g(p_1) = q_1$ and $g(p_2) = q_2$ for some translation g. But then we need $q_1 - p_1 = q_2 - p_2$. There are at most N^3 solutions to this in $\mathcal{P}^4$, so $|E^{-1}(G^{\text{trans}})| \leq N^3$ and thus $|G_r(\mathcal{P}) \cap G^{\text{trans}}| \leq N^3/(r^2 - r)$. ■

2.2 Turning curves into lines

Let $G' = G \setminus G^{\text{trans}}$. It suffices to bound $|G_r(\mathcal{P}) \cap G'|$.

Observe that any element of G' is actually a (nontrivial) rotation. Indeed, a translation followed by a rotation is actually still a rotation. For $g \in G'$, let g be the rotation around (x, y) counterclockwise by $\theta \in (0, 2\pi)$. Define $\rho \colon G' \to \mathbb{R}^3$ by

$$\rho(g) = (x, y, \cot(\theta/2)).$$

We can check that this is a bijection, because $\cot \colon (0, \pi) \to \mathbb{R}$ is a bijection. The key fact is that ρ maps $S_{p \mapsto q} \cap G'$ to lines in $\mathbb{R}^3$. More precisely:

Lemma 2.2

Let $p, q \in \mathbb{R}^2$. Then $\rho(S_{p \mapsto q} \cap G')$ is a straight line $\ell_{p,q}$ in $\mathbb{R}^3$.

Let $p = (x_p, y_p)$ and $q = (x_q, y_q)$. Then $\ell_{p,q}$ is parametrised by

$$t \mapsto \left(\frac{x_p + x_q}{2}, \frac{y_p + y_q}{2}, 0 \right) + t \left(\frac{y_p - y_q}{2}, -\frac{x_p - x_q}{2}, 1 \right).$$

Furthermore, the lines $\ell_{p,q}$ for $(p, q) \in \mathbb{R}^2 \times \mathbb{R}^2$ are distinct.

For proof, see the book. It's just some Euclidean geometry and vectors. Distinctness follows from the parametrisation.

Corollary 2.3

Let $p = (x_p, y_p) \in \mathbb{R}^2$. Then $\{\ell_{p,q} : q \in \mathbb{R}^2\}$ partitions $\mathbb{R}^3$.

Furthermore, there is a nonvanishing vector field $V_p(x, y, z)$ such that:

- for all $(x, y, z) \in \mathbb{R}^2$, we have that $V_p(x, y, z)$ is tangent to the unique line $\ell_{p,q}$ through (x, y, z) (i.e. the integral curves of V_p are the $\ell_{p,q}$); and
- The entries of $V_p(x, y, z)$ are polynomials in x_p, y_p, x, y, z with $\deg_p \leq 1$ and $\deg_{x,y,z} \leq 2$.

Proof sketch. The first part is just a computation using the parametrisation of $\ell_{p,q}$. The result is that the unique $q \in \mathbb{R}^2$ such that $(x, y, z) \in \ell_{p,q}$ is

$$q = (x_q, y_q) = \left(\frac{2x - x_p - zy_p + 2yz + z^2x_p - y_p}{z^2 + 1}, \frac{2y + zx_p - y_p - 2xz + zx_p + z^2y_p}{z^2 + 1}\right).$$

Next, using the parametrisation of $\ell_{p,q}$, we see that $V_p(x, y, z) \propto \left(\frac{y_p - y_q}{2}, -\frac{x_p - x_q}{2}, 1\right)$. We can clear denominators by multiplying by $z^2 + 1$, and this has the desired properties. ■

Now let us translate from partial symmetries to the incidence geometry of lines. Let $\mathcal{L}(\mathcal{P}) = \{\ell_{p,q} : (p, q) \in \mathcal{P}^2\}$. Then $|\mathcal{L}(\mathcal{P})| = N^2$, and $|G_r(\mathcal{P}) \cap G'| = |\mathcal{P}_r(\mathcal{L}(\mathcal{P}))|$.

Recall the following result from polynomial partitioning that Tik showed in December:

Proposition 2.4

Let $\epsilon > 0$. If $\mathcal{L}$ is a set of L lines in $\mathbb{R}^3$ such that any algebraic surface of degree $O_\epsilon(1)$ contains at most $L^{1/2+\epsilon}$ lines of $\mathcal{L}$, then

$$|\mathcal{P}_r(\mathcal{L})| \lesssim_\epsilon L^{3/2+\epsilon} r^{-2} + Lr^{-1}.$$

We would like to apply this to $\mathcal{L} = \mathcal{L}(\mathcal{P})$. Noting that $r \leq N$, this tells us that if any algebraic surface of degree $O_\epsilon(1)$ contains at most $N^{1+\epsilon}$ lines of $\mathcal{L}(\mathcal{P})$, then $|G_r(\mathcal{P}) \cap G'| \lesssim_\epsilon N^{3+\epsilon} r^{-2}$, which is exactly what we want for the $d(N) \gtrsim N^{1-\epsilon}$ bound.

2.3 Lines don't cluster on low degree surfaces

The goal now is the verify the hypothesis of the polynomial partitioning result. It suffices to prove the following:

Proposition 2.5

Let $D \geq 1$. Then any algebraic surface of degree D contains $O_D(N)$ lines of $\mathcal{L}(\mathcal{P})$.

Indeed, by summing this implies that at most $O_\epsilon(N)$ lines of $\mathcal{L}(\mathcal{P})$ are contained in any algebraic surface of degree $O_\epsilon(1)$, and we can adjust the constant in the bound to account for the constant (in terms of ϵ) number of N such that $O_\epsilon(N) > N^{1+\epsilon}$.

We first prove this for $D = 1$, where degree 1 surfaces are planes.

Lemma 2.6

Any plane contains at most N lines of $\mathcal{L}(\mathcal{P})$.

Proof. The key claim is that any plane contains at most one of the lines $\ell_{p,q}$ for fixed p. To show the claim, it suffices to show that $\ell_{p,q}$ and $\ell_{p,q'}$ are skew for any q, q'. Note that they are disjoint, because $S_{p\mapsto q}$ and $S_{p\mapsto q'}$ are disjoint (we cannot have g mapping $p \mapsto q$ and $p \mapsto q'$). So we need to check that they are not parallel. To do this, just check the parametrisation and note that the velocity vectors are not parallel.

So any plane contains at most one line from each of $\{\ell_{p,q} : q \in \mathcal{P}\}$, which means that it can only contain at most N lines from all of $\mathcal{L}(\mathcal{P})$. ■

Now we turn to higher degrees. To prove Proposition 2.5, it suffices to show that for any irreducible polynomial f with $1 < \deg f \leq D$, the variety $Z(f)$ contains at most $(2D^2 + 1)N$ lines of $\mathcal{L}(\mathcal{P})$. The following lemma does the job:

Lemma 2.7

There is at most one point $p \in \mathbb{R}^2$ such that $Z(f)$ contains at least $2D^2$ lines of $\{\ell_{p,q} : q \in \mathbb{R}^2\}$.

For additional context about why this is a reasonable approach, see Section 9.7 of the book.

Assuming this lemma, we have that:

- for $N - 1$ points $p \in \mathcal{P}$, $Z(f)$ contains at most $2D^2$ lines of $\{\ell_{p,q} : q \in \mathcal{P}\}$; and
- for the last point $p \in \mathcal{P}$, $Z(f)$ contains at most N lines of $\{\ell_{p,q} : q \in \mathcal{P}\}$.

So collectively, $Z(f)$ contains at most $(N-1) \cdot 2D^2 + 1 \cdot N \leq (2D^2 + 1)N$ lines of $\mathcal{L}(\mathcal{P})$.

Now let us prove the lemma.

Proof. Fix $p \in \mathbb{R}^2$, and suppose $Z(f)$ contains at least $2D^2$ lines of $\{\ell_{p,q} : q \in \mathbb{R}^2\}$. The vector field V_p is tangent to each of those lines, and f vanishes on each such line, so $h_p := V_p \cdot \nabla f$ vanishes on each such line. In other words, $Z(f, h_p)$ contains at least $2D^2$ lines of $\{\ell_{p,q} : q \in \mathbb{R}^2\}$.

Now, observe that ∇f is polynomial in (x, y, z) with degree at most $D - 1$, and V_p is polynomial with degree at most 2, so h_p is polynomial with degree at most $D + 1$. If f

and h_p are relatively prime, then Bezout's theorem implies that $Z(f, h_p)$ contains at most $D^2 + D$ lines, contradiction. So by irreducibility of f, we see that f divides h_p. In particular, $Z(f) \subseteq Z(h_p)$.

Now suppose there are two points $p_1, p_2 \in \mathbb{R}^2$ such that $Z(f)$ contains at least $2D^2$ lines of each of $\{\ell_{p_1,q} : q \in \mathbb{R}^2\}$ and $\{\ell_{p_2,q} : q \in \mathbb{R}^2\}$. Then $Z(f) \subseteq Z(h_{p_1}) \cap Z(h_{p_2})$. Note that for fixed (x, y, z) and variable p, we have that $h_p(x, y, z)$ is a polynomial in p with degree at most 1, so the map $p \mapsto h_p(x, y, z)$ is affine. We can exploit this if $(x, y, z) \in Z(h_{p_1}) \cap Z(h_{p_2})$. Let $\text{line}(p_1, p_2) = \{\lambda p_1 + (1 - \lambda)p_2 : \lambda \in \mathbb{R}\}$ be the line passing through p_1 and p_2. Then for all $p = \lambda p_1 + (1 - \lambda)p_2 \in \text{line}(p_1, p_2)$, we have that $h_p(x, y, z) = 0$. So h_p vanishes on $Z(h_{p_1}) \cap Z(h_{p_2})$, and so $Z(f)$, for any $p \in \text{line}(p_1, p_2)$.

Suppose $Z(f)$ has a nonsingular point (x, y, z), i.e. $\nabla f(x, y, z) \neq 0$. Then there is a neighborhood $U \subseteq Z(f)$ around (x, y, z) on which ∇f is nonzero. For any $p \in \text{line}(p_1, p_2)$, we have that $h_p = V_p \cdot \nabla f$ vanishes on U, so V_p is a vector field on U, so its integral curves lie in U. But these integral curves are $\{\ell_{p,q} : q \in \mathbb{R}^2\}$, so the unique line in $\{\ell_{p,q} : q \in \mathbb{R}^2\}$ passing through (x, y, z) lies in $Z(f)$ and hence also the tangent plane to $T_{(x,y,z)}Z(f)$. But there are infinitely many p we can take, each leading to a different line, so $Z(f)$ contains infinitely many lines in the plane $T_{(x,y,z)}Z(f)$. By Bezout's theorem, this implies that $Z(f)$ is a plane, contradicting the fact that $\deg f > 1$.

So every point of $Z(f)$ is singular. Then ∇f vanishes on $Z(f)$, so in particular $\partial_x f$ vanishes on $Z(f)$. So the number of lines in $Z(f, \partial_x f)$ is at least the number of lines in $Z(f)$. If f and $\partial_x f$ are relatively prime, then Bezout's theorem implies that $Z(f, \partial_x f)$ contains at most $D(D - 1)$ lines, so $Z(f)$ contains at most $D^2 - D$ lines, so there are no points p such that $Z(f)$ contains $2D^2$ lines of $\{\ell_{p,q} : q \in \mathbb{R}^2\}$. Similarly if f and $\partial_y f$ are relatively prime, or f and $\partial_z f$ are relatively prime. Otherwise, by irreducibility of f, we see that f divides each of $\partial_x f, \partial_y f, \partial_z f$. But by degree counting, this means that $\partial_x f \equiv \partial_y f \equiv \partial_z f \equiv 0$, so f is constant. But f is not the zero polynomial, so $Z(f)$ is empty, so again there are no points p with the desired property.

So in any case, we get a contradiction, so there is at most one point p with the desired property. ■

This concludes the proof of Proposition 2.5, which can be plugged in to the polynomial partitioning result to deduce:

Proposition 2.8

Let $\epsilon > 0$. For all $r \geq 2$, we have that $|G_r(\mathcal{P})| \lesssim_\epsilon N^{3+\epsilon} r^{-2}$.

This in turn can be plugged in to our combinatorial estimate to reach our desired conclusion:

Theorem 2.9

$d(N) \gtrsim_\epsilon N^{1-\epsilon}$ for all $\epsilon > 0$