USAMO 2020/5

Evan Chen

Twitch Solves ISL

Episode 16

Problem

A finite set S of points in the coordinate plane is called *overdetermined* if $|S| \geq 2$ and there exists a nonzero polynomial $P(t)$, with real coefficients and of degree at most $|S| - 2$, satisfying $P(x) = y$ for every point $(x, y) \in S$.

For each integer $n \geq 2$, find the largest integer k (in terms of n) such that there exists a set of n distinct points that is *not* overdetermined, but has k overdetermined subsets.

Video

https://youtu.be/r7j0oRtpErA

Solution

We claim the answer is $k = 2^{n-1} - n$. We denote the n points by A.

Throughout the solution we will repeatedly use the following fact:

Lemma. *If S is a finite set of points in the plane there is at most one polynomial with real coefficients and of degree at most $|S| - 1$ whose graph passes through all points of S.*

Proof. If any two of the points have the same x-coordinate then obviously no such polynomial may exist at all.

Otherwise, suppose f and g are two such polynomials. Then $f - g$ has degree at most $|S| - 1$, but it has $|S|$ roots, so is the zero polynomial. □

Remark. Actually Lagrange interpolation implies that such a polynomial exists as long as all the x-coordinates are different!

Construction: Consider the set $A = \{(1, a), (2, b), (3, b), (4, b), \dots, (n, b)\}$, where a and b are two distinct nonzero real numbers. Any subset of the latter $n - 1$ points with at least one element is overdetermined, and there are $2^{n-1} - n$ such sets.

Bound: Say a subset S of A is *flooded* if it is not overdetermined. For brevity, an *m-set* refers simply to a subset of A with m elements.

Claim. If S is an flooded m-set for $m \ge 3$, then at most one $(m-1)$-subset of S is not flooded.

Proof. Let $S = \{p_1, \dots, p_m\}$ be flooded. Assume for contradiction that $S - \{p_i\}$ and $S - \{p_j\}$ are both overdetermined. Then we can find polynomials f and g of degree at most $m - 3$ passing through $S - \{p_i\}$ and $S - \{p_j\}$, respectively.

Since f and g agree on $S - \{p_i, p_j\}$, which has $m - 2$ elements, by the lemma we have $f = g$. Thus this common polynomial (actually of degree at most $m - 3$) witnesses that S is overdetermined, which is a contradiction. □

Claim. For all $m = 2, 3, \dots, n$ there are at least $\binom{n-1}{m-1}$ flooded m-sets of A.

Proof. The proof is by downwards induction on m. The base case $m = n$ is by assumption.

For the inductive step, suppose it's true for m. Each of the $\binom{n-1}{m-1}$ flooded m-sets has at least $m - 1$ flooded $(m-1)$-subsets. Meanwhile, each $(m-1)$-set has exactly $n - (m-1)$ parent m-sets. We conclude the number of flooded sets of size $m - 1$ is at least

$$\frac{m-1}{n-(m-1)} \binom{n-1}{m-1} = \binom{n-1}{m-2}$$

as desired. □

This final claim completes the proof, since it shows there are at most

$$\sum_{m=2}^{n} \left(\binom{n}{m} - \binom{n-1}{m-1} \right) = 2^{n-1} - n$$

overdetermined sets, as desired.

Remark (On repeated x-coordinates)**.** Suppose A has two points p and q with repeated x-coordinates. We can argue directly that A satisfies the bound. Indeed, any overdetermined set contains at most one of p and q. Moreover, given $S \subseteq A - \{p, q\}$, if $S \cup \{p\}$ is overdetermined then $S \cup \{q\}$ is not, and vice-versa. (Recall that overdetermined sets always have distinct x-coordinates.) This gives a bound $\left[2^{n-2} - (n-2) - 1\right] + \left[2^{n-2} - 1\right] = 2^{n-1} - n$ already.

Remark (Alex Zhai)**.** An alternative approach to the double-counting argument is to show that any overdetermined m-set has an flooded m-superset. Together with the first claim, this lets us pair overdetermined sets in a way that implies the bound.