Forbidden Configurations in Discrete Geometry

David Eppstein

29th Canadian Conference on Computational Geometry
Ottawa, ON, July 2017

I: A happy ending

Quadrilaterals in five-point sets

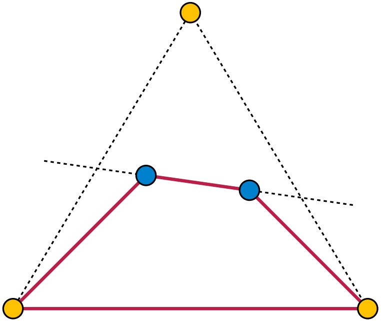
Esther Klein, early 1930s:

Five points in general position (no three in a line) contain a convex quadrilateral

Proof:

If convex hull has 4 or 5 vertices, obvious

Otherwise, line through inner points misses a hull edge; use that edge + inner points



Erdős and Szekeres, 1935

Generalization to larger convex polygons

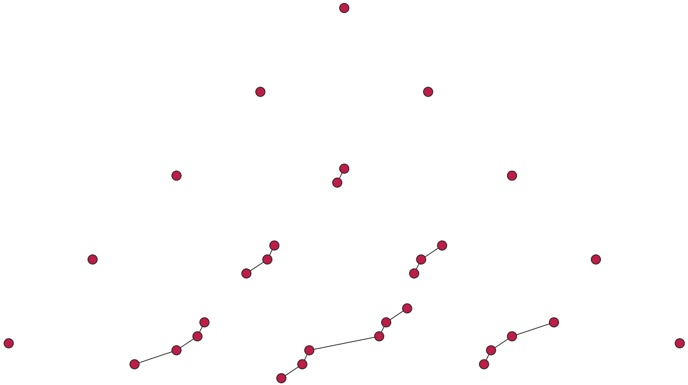
Every $\binom{2k-4}{k-2} + 1 < 4^k$ points in general position
contain a convex k-gon

Conjecture: $\#$ points needed to ensure a convex k-gon is $2^{n-2} + 1$

Still open, \$500 prize for solution

Klein and Szekeres marry, commemorated in the theorem name

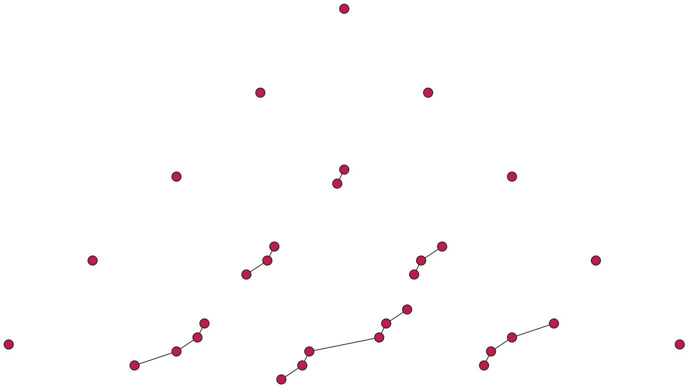
A Pascal's triangle of point sets



Each set of $\binom{n}{i}$ points combines above-left $\binom{n-1}{i-1}$ and right $\binom{n-1}{i}$

Separated far enough so lines through two left points pass above right points, lines through two right points pass below left points

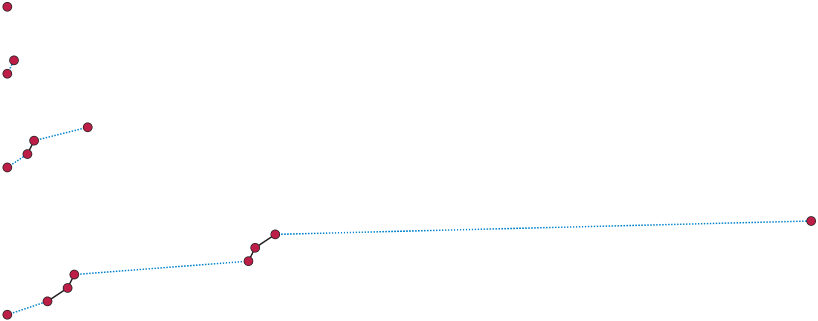
Cups and caps



For the set in position $\binom{n}{i}$, largest cup (subset of points on the graph of a convex function) has size $n - i + 1$ and largest cap (subset on the graph of a concave function) has size $i + 1$

Erdős and Szekeres, 1960

One side of the 1935 conjecture:
there exist sets of 2^{k-2} points with no convex k-gon



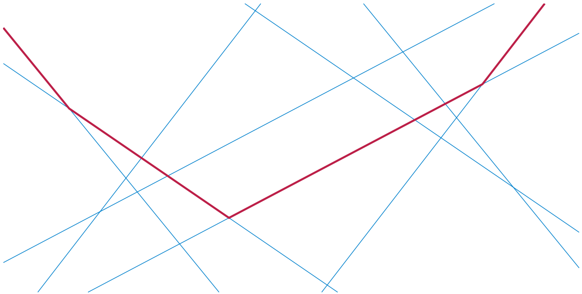
Construction: Glue together the sets on one row of Pascal's triangle in the same way the individual sets are glued together

Convex subsets either stay in one subset, or have the form
cap – point – point – $\cdots$ – cup

Finding the largest convex subset

Chvátal and Klincsek 1980; Edelsbrunner and Guibas 1989

Biggest polygon with fixed bottom vertex is dual to longest convex chain in line arrangement, found by sweeping the arrangement

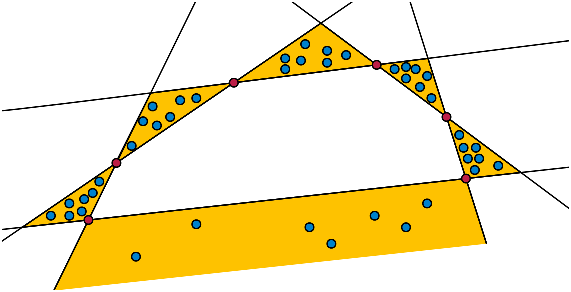


Test all choices of bottom vertex: $O(n^3)$ time, $O(n)$ space

(Related algorithms for finding convex subsets with other optimization criteria: E, Overmars, Rote, Woeginger, 1992; E, Erickson, 1994)

Suk 2016

Every $2^{k+O(k^{2/3}\log k)}$ points in general position have a convex k-gon



Proof strategy:

Use E–S 1935 to find a big cup or cap (red)
such that each yellow region has many points

Use Dilworth to find large chains or antichains in each yellow region
for partial ordering by triangle containment

Apply case analysis to glue together yellow regions

Key properties of largest convex polygon

The function that maps point sets to the size of their largest convex polygon has two key properties.

Monotone:
Removing points can only reduce largest convex subset

Invariant:
Depends only on relative orientations of triples of points, not on their exact locations

Motivation for monotone invariant properties

Hereditary properties (monotonic under vertex removal) are central to graph theory:

- Cliques, independent sets, and coloring
- Perfect graphs
- Claw-free and triangle-free graphs
- Chordal graphs and other intersection classes of graphs
- Etc.

We should build an analogous theory in discrete geometry!

. . . or maybe we already have and we just didn't realize it?

II: A menagerie of monotone invariant problems

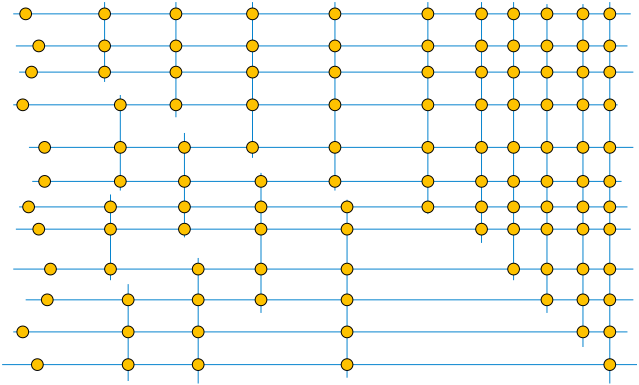
Projective clustering

How many lines are needed to cover all points of a point set?

NP-complete [Megiddo and Tamir 1982]

Fixed parameter tractable [Langerman and Morin 2005]

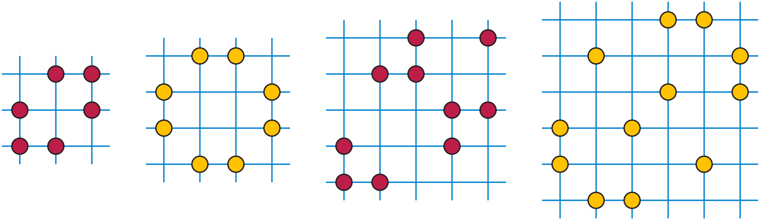
Greedy cover gives only log-approximation



Open: Can we do better?

The no-three-in-line problem

Dudeney 1917: Largest general-position subset of $n \times n$ grid



Erdős 1951: at least $n(1 - o(1))$

Hall et al. 1975: at least $n(1.5 - o(1))$

Guy 2005: conjectures at most

$$\frac{\pi n}{\sqrt{3}} + o(n) \approx 1.814n + o(n).$$

Open: Any upper bound better than $2n$

Trading line length for general-position subsets

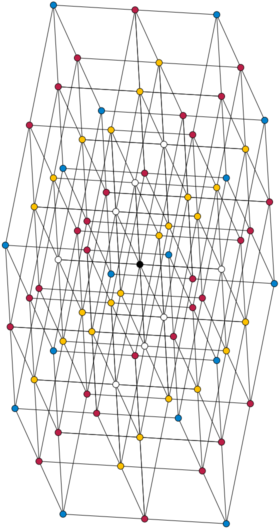
Erdős 1986, 1988: if not many points are on any line, must there be a large general-position subset?

Square grid: Longest line and largest general-position subset both $O(\sqrt{n})$

Payne and Wood: Whenever longest line is $\ell = O(\sqrt{n})$, largest general-position subset is $\Omega(\sqrt{n/\log \ell})$

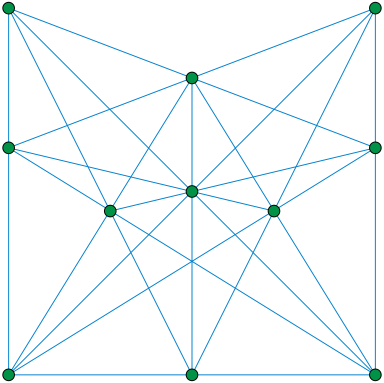
Balogh and Solymosi 2017: Some point sets with no four on a line have largest general-position subset $O(n^{5/6+\epsilon})$

Open: Close gap between 1/2 and 5/6



Orchard planting

Jackson 1821, Lloyd 1914:
How many three-point lines can we form from n points?



Burr et al. 1974: at least $\left\lfloor \frac{n(n-3)}{6} \right\rfloor + 1$

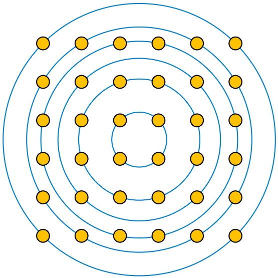
Green and Tao 2013: Burr et al. is optimal for large n

Onion layers

Repeatedly remove convex hull vertices

How many layers do you get?

Chazelle 1985: $O(n \log n)$-time algorithm



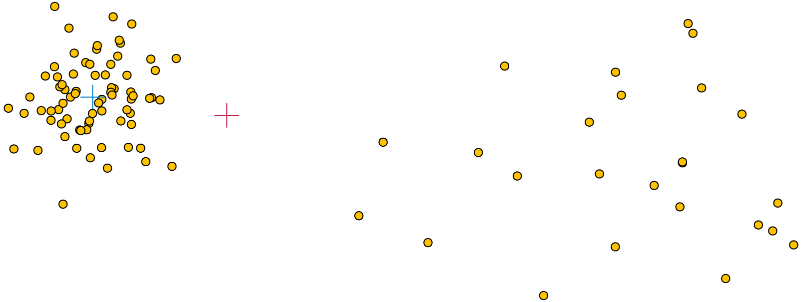
Har-Peled and Lidický 2013: $n \times n$ grid has $\Theta(n^{4/3})$ layers

Conjecture (E, Har-Peled, Nivasch 2017): Onion-peeling convex subsets of grids approximates the *affine curve-shortening flow*

Tukey depth

Depth(q) = minimum # points in a halfplane containing q

Deepest point = estimate of central location, more robust to outliers than the centroid (Tukey 1975)



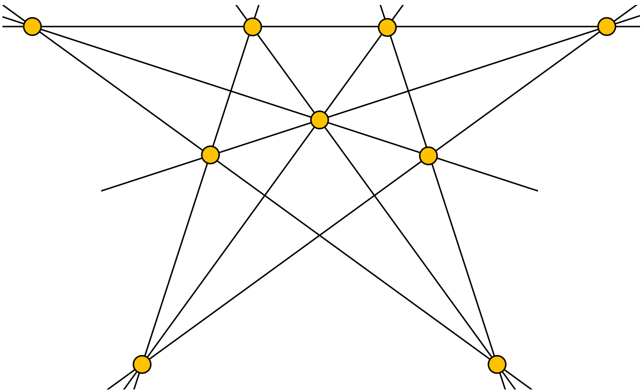
Deepest point in the plane: not invariant

Deepest from a given point set: invariant and monotone

Integer coordinates

The Perles configuration (Perles, 1960s)

No combinatorially-equivalent set of points has integer coordinates

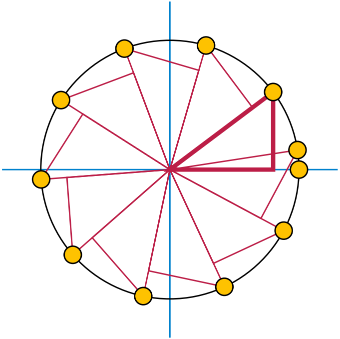


Open (Grünbaum 2003): Smallest non-integer configuration?

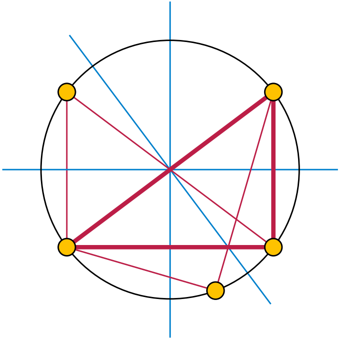
Open: Computational complexity of integer realization?

Rational distances

Euler 1862: For every k there exists a convex polygon with all pairwise distances rational



Rotate unit vector by the angle of a Pythagorean triangle



Reflect integer-sided triangle across perpendicular bisectors

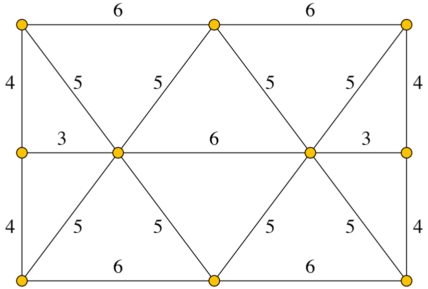
(Both constructions listed by Harborth 1998)

Harborth, Erdős, and Ulam

Erdős–Ulam conjecture: there exists a dense subset of the plane for which all distances are rational

would imply

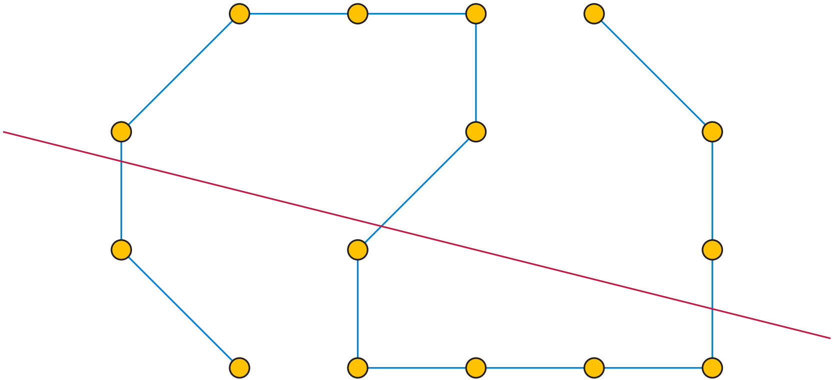
Harborth's conjecture: every planar graph can be drawn with all edges as integer-length line segments



Open, would answer one of the two conjectures:
Is there a general-position set of points with no combinatorially equivalent rational-distance realization?

Paths with low stabbing number

Any point set has a path with stabbing number $O(\sqrt{n})$, used in range searching (Agarwal, Matoušek, Welzl 1991–2)



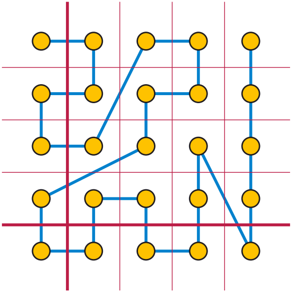
Stabbing number has logarithmic approximation (Har-Peled 2009)

What is its computational complexity?

Shattering number

How many lines needed to separate all points?

Path stabbing number $\geq (n-1)/L$



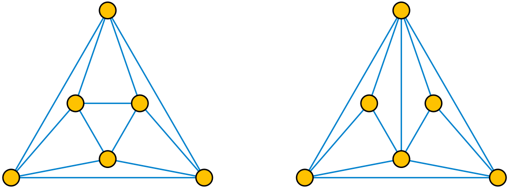
For grid, $L \approx \sqrt{n}$, but for random points, $L \approx n^{2/3} \log^{O(1)} n$
(Har-Peled and Jones 2017)

Universality

Max k s.t. we can draw all k-vertex planar graphs on given points?

The six points below can be used for all six-vertex planar graphs, but for larger k we need $n > k$

(Chrobak and Karloff 1989; Kurowski 2004; Cardinal et al. 2015)



Smallest known k-universal set: $k^2/4 - O(n)$

(Bannister, Cheng, Devanny, E, 2014)

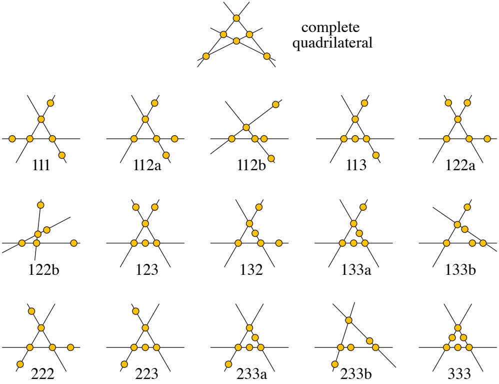
Open: Close gap between $\Omega(k)$ and $O(k^2)$

III: New perspectives from monotonicity

Obstacles

Every monotone property (preserved on deletion) can be characterized by minimal point sets that violate it ("obstacles")

Every monotone parameter can be characterized by minimal point sets forcing parameter $\geq k$ (obstacles to bounded parameter value)



(Some obstacles to covering with two lines, being two points away from collinear, and being one point away from general position)

Parameters from obstacles

Given obstacles, we can define monotone parameters

Largest obstacle-avoiding set
(e.g. no-three-in-line:
largest 3-line-avoiding subset)

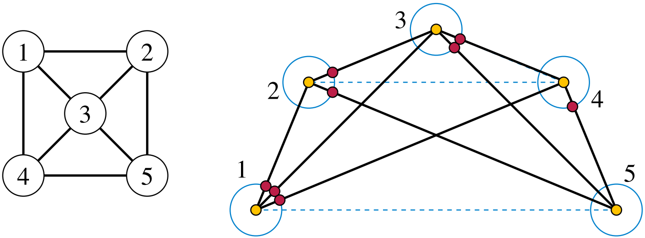
Min points to delete
to eliminate all obstacles

Fewest subsets in partition
into obstacle-avoiding subsets
(projective clustering: fewest
triangle-avoiding subsets)

How hard is finding an obstacle?

Theorem: Testing whether n points include a given k-point obstacle is NP-complete, W[1]-hard, and not solvable in time $n^{o(\sqrt{k})}$ unless the exponential time hypothesis fails

Reduction from clique-finding:



…nevertheless, it's (obviously) polynomial for $k = O(1)$.

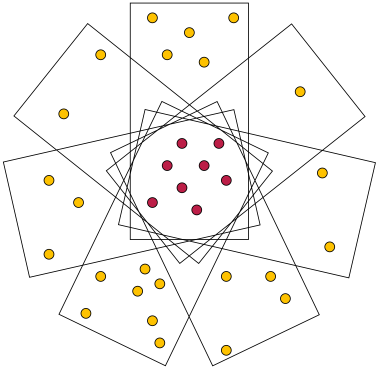
Parameterized complexity

Distance to an obstacle-free subset is always fixed-parameter tractable

(Sunflower-based kernelization for hitting sets of bounded size, Flum and Grohe 2006)

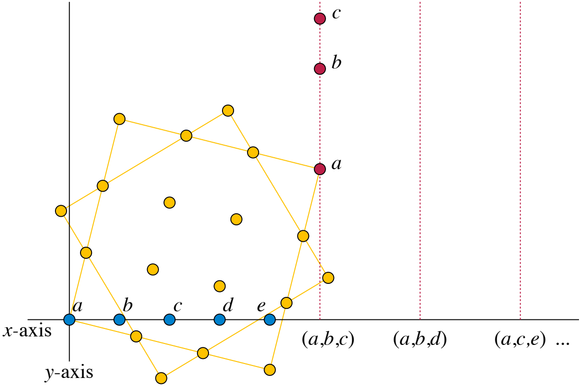
Open: Is there a finite obstacle set for which largest obstacle-avoiding subset is *not* FPT?

(both Σ_2^P-complete for variable sets of obstacles)



New parameters for hard problems

Partition into general-position subsets is NP-complete for its natural parameter



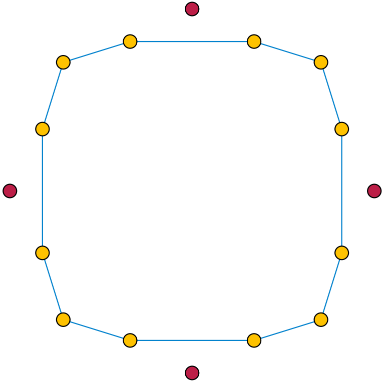
. . . but FPT parameterized by min # lines to cover all points

Property testing

Theorem: For every property with finitely many obstacles, sampling n^c points for some $c < 1$ distinguishes sets with the property from sets far from the property, w.h.p.

Sometimes $c > 0$ is necessary (e.g. convexity)

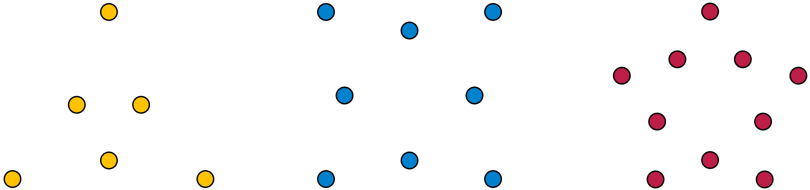
Sometimes even $O(1)$ points suffice (e.g. collinearity)



A set that is $n/4$-far from convex whose samples of $o(n^{2/3})$ points are w.h.p. convex

Are there always $O(1)$ obstacles?

No! There exist infinite antichains of point sets



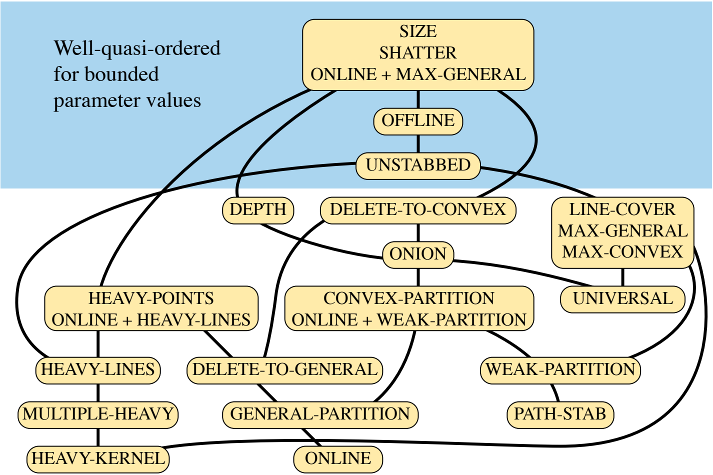
Any subset of the antichain can be an obstacle set

But many natural properties have finite obstacles
(including largest obstacle-avoiding subset, projective clustering,
onion layers, Tukey depth, universality)

Properties of some restricted classes of points (e.g. weakly convex,
or covered by two lines) always have $O(1)$ obstacles

Comparison of parameters

Which parameters are bounded by a function of which other parameters?



Conclusions

Monotonicity provides a unifying framework for many famous problems in discrete geometry

Characterization by obstacles leads to algorithms of many types (exact, parameterized, approximation, property testing)

Much more remains to be done!