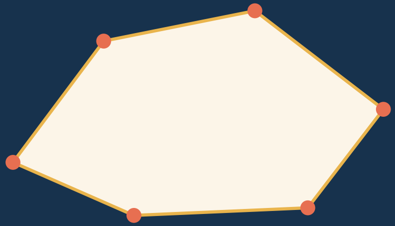
The Happy Ending Problem

Math Club, ISI Bangalore



Hrishik

Convexity and the convex hull

Convex set

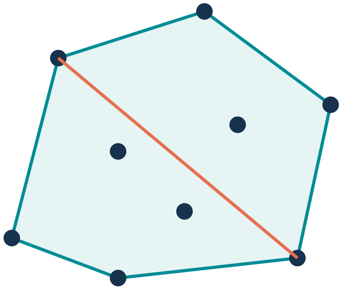
A set $C \subseteq \mathbb{R}^2$ is convex if, whenever $x, y \in C$, the whole segment $[x, y]$ lies in C.

Convex hull

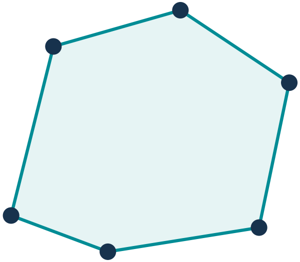
For $P = \{p_1, \dots, p_m\}$,

$$\operatorname{conv}(P) = \left\{ \sum_{i=1}^{m} \lambda_i p_i : \lambda_i \geq 0, \ \sum_{i=1}^{m} \lambda_i = 1 \right\}.$$

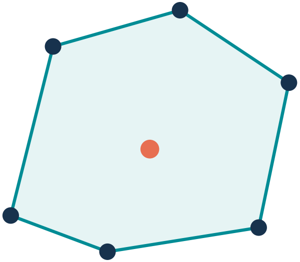
It is the smallest convex set containing P or the **convex hull** of P.



Convex Position



convex position



not in convex position

Equivalent test

P is in convex position iff no point of P belongs to $\mathrm{conv}(P \setminus \{p\})$.

A Question & A General Assumption

Question

How many points do you need to create a k-gon?

Assumption

In general, the answer could be infinitely many if all points are collinear. So we assume that no 3 given points are collinear. From here on, whenever we say non-collinear we mean that no 3 points are collinear.

Answer

Now under our assumption, we can safely say that k vertices are enough to form a k-gon.

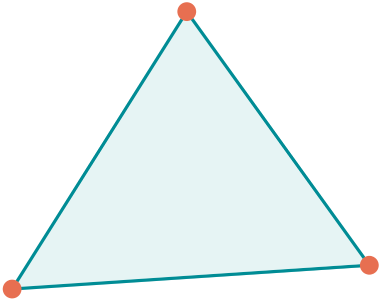
Triangles

Claim

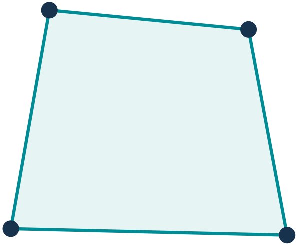
Three non-collinear points form a triangle or 3-gon, and in fact a convex 3-gon.

Why?

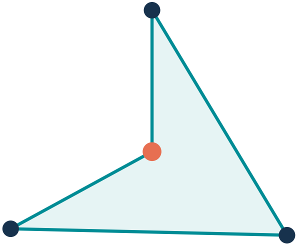
- Three non-collinear points are exactly the vertices of their triangular convex hull.
- Fewer than three points cannot form a triangle.



Four points have two order types



convex quadrilateral



So 4 points form a 4-gon or quadrilateral. But it does not force a convex 4-gon. So we move from the uninteresting question to the interesting one.

The Happy Ending Problem

The Interesting Question

How many non-collinear points do you need to create a convex k-gon?

Extremal function

For $k \geq 3$, let $\mathrm{ES}(k)$ be the least integer M such that every set of M non-collinear points in the plane contains k points in convex position.

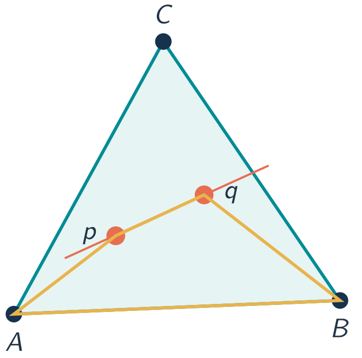
Why five points force a convex quadrilateral?

Let P contain five points.

1. If $\mathrm{conv}(P)$ has at least four vertices in convex position, we are done.
2. Otherwise the hull is a triangle ABC, with two interior points p, q.
3. Two of A, B, C, say A, B, lie on the same side of the line pq.
4. Both AB and pq are boundary edges of $\mathrm{conv}\{A, B, p, q\}$.

Thus A, B, p, q form a convex quadrilateral, so

$$\boxed{\mathrm{ES}(4) = 5}.$$



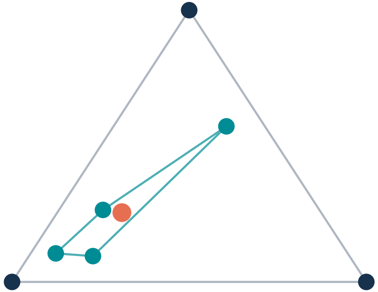
A Pentagon

Classical exact result

$$\boxed{\mathrm{ES}(5) = 9}$$

Every set of 9 non-collinear points contains five points in convex position.

But 8 points can avoid every convex pentagon. The picture shows one coordinate realization; every five-point subset has at most four hull vertices.



The Erdős–Szekeres conjecture

k	$\mathrm{ES}(k)$	$2^{k-2}+1$
3	3	3
4	5	5
5	9	9
6	17	17

Conjecture

The first four values suggest $\mathrm{ES}(k) = 2^{k-2} + 1$. This is the Erdős–Szekeres conjecture.

The Erdős–Szekeres theorem (1935)

Theorem

For every positive integer N, every sufficiently large finite set of points in the plane in general position contains a subset of N points that form the vertices of a convex polygon.

Equivalently,

$$\forall N \geq 3, \qquad \mathrm{ES}(N) < \infty.$$

Detour: Erdős–Szekeres monotone subsequence theorem

Theorem

Every sequence of $rs + 1$ distinct real numbers contains either

- an increasing subsequence of length $r + 1$, or
- a decreasing subsequence of length $s + 1$.

Setting $r = s = n$ gives the memorable form:

$$\boxed{n^2 + 1 \text{ distinct numbers force a monotone subsequence of length } n + 1.}$$

Proof idea: a label for each entry

For each position i, define

$$\begin{aligned} I_i &= \text{length of the longest increasing subsequence ending at } a_i, \\ D_i &= \text{length of the longest decreasing subsequence ending at } a_i. \end{aligned}$$

$$a_i \longrightarrow \text{label } (I_i, D_i)$$

Assume there is no increasing $r+1$-subsequence and no decreasing $s+1$-subsequence, then

$$1 \leq I_i \leq r, \qquad 1 \leq D_i \leq s.$$

Only rs different labels are available.

Proof: two positions cannot share a label

Take $1 \leq i < j \leq rs + 1$.

If $a_i < a_j$

Append a_j to a longest increasing subsequence ending at a_i:

$$I_j \geq I_i + 1.$$

If $a_i > a_j$

Append a_j to a longest decreasing subsequence ending at a_i:

$$D_j \geq D_i + 1.$$

Thus $(I_i, D_i) \neq (I_j, D_j)$ whenever $i \neq j$. We have $rs + 1$ entries but only rs possible labels. So by pigeonhole principle at least 2 of them share a label, which gives a contradiction. $\square$

Can we apply the monotone subsequence theorem directly to points?

Order points from left to right and write the consecutive slopes

$$m_i = \text{slope}(p_i, p_{i+1}).$$

It is tempting to find a monotone subsequence of the m_i.

Where it fails?

We might have a monotone increasing or decreasing subsequence of the slopes m_i but that does not necessarily imply that the slopes we get in the subsequence belong strictly to consecutive edges. But this is a viable direction that works with slight modification, and we exploit this but use a geometric version that removes the bottleneck.

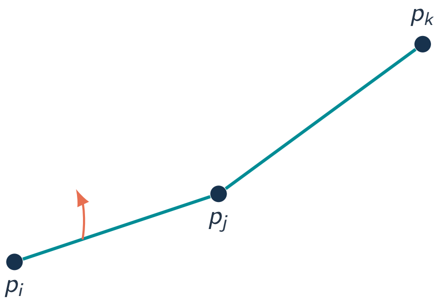
The right geometric encoding: compare adjacent slope

Take

$$x(p_1) < x(p_2) < \cdots < x(p_m).$$

For three consecutive selected points,

$$\text{turn}(p_i, p_j, p_k) > 0 \iff \text{slope}(p_i, p_j) < \text{slope}(p_j, p_k) \iff \text{the slope increases}$$



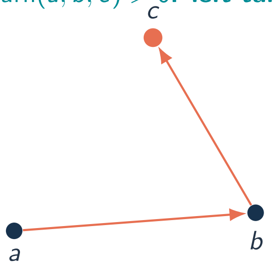
This produces the geometric analogues of increasing and decreasing subsequences: **cups and caps**.

Encoding a triple by its orientation

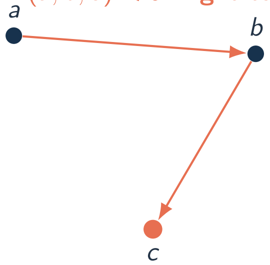
For $a = (a_x, a_y),\ b = (b_x, b_y),\ c = (c_x, c_y)$ define

$$\operatorname{turn}(a, b, c) = \det \begin{pmatrix} b_x - a_x & b_y - a_y \\ c_x - a_x & c_y - a_y \end{pmatrix}.$$

$\operatorname{turn}(a, b, c) > 0$: **left turn**



$\operatorname{turn}(a, b, c) < 0$: **right turn**



Our non-collinearity condition guarantees that $\operatorname{turn}(a, b, c) \neq 0$.

Geometric lemma

Order points by x-coordinate: $p_1, p_2, \ldots, p_N$.

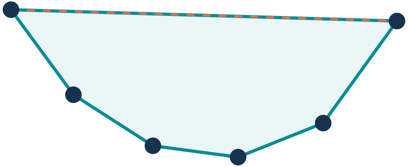
Lemma

If $\text{turn}(p_i, p_j, p_k) > 0$ for every $i < j < k$, then all p_i are in convex position. The same holds with all signs negative.

Indeed,

$$\text{turn}(p_i, p_j, p_k) > 0 \iff \text{slope}(p_i, p_j) < \text{slope}(p_j, p_k)$$

for consecutive points. Thus the chain makes strictly increasing turns. Also, each intermediate point lies below the chord $p_1 p_N$.



Cups and Caps

Let $p_1, \ldots, p_k$ be ordered by increasing x-coordinate.

k-cup

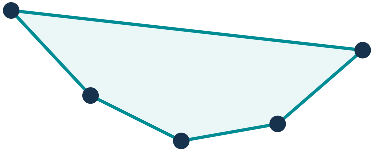
$$\text{slope}(p_1, p_2) < \text{slope}(p_2, p_3) < \cdots < \text{slope}(p_{k-1}, p_k).$$

k-cap

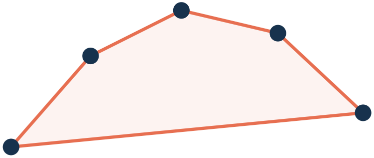
$$\text{slope}(p_1, p_2) > \text{slope}(p_2, p_3) > \cdots > \text{slope}(p_{k-1}, p_k).$$

General position makes all comparisons strict. A cup bends upward; a cap bends downward.

Cups and Caps II



5-**cup: slopes increase**



5-**cap: slopes decrease**

Why every cup or cap is a convex polygon

Consider a k-cup $p_1, \ldots, p_k$.

- Increasing consecutive slopes mean every internal vertex makes a strict left turn.
- Every point $p_2, \ldots, p_{k-1}$ lies below the closing chord $p_1 p_k$.
- Therefore the polygonal chain plus the chord $p_k p_1$ is the boundary of a convex k-gon.

The cap case is the reflected picture.

The cup–cap extremal function

Definition

Let $f(r, s)$ be the least integer M such that every M-point set in general position contains an r-cup or an s-cap.

The goal is to prove

$$\boxed{f(r, s) \leq \binom{r+s-4}{r-2} + 1.}$$

Since an k-cup or k-cap is a convex k-gon,

$$\mathrm{ES}(k) \leq f(k, k).$$

The binomial coefficient will arise from Pascal's identity—but first we need a geometric recurrence.

Boundary cases: $f(r,3)=r$

Take r points from left to right.

- If some three selected points form a 3-cap, we are done.
- Otherwise no consecutive triple turns downward.
- Hence all consecutive slopes strictly increase, so all r points form an r-cup.
- Also note that, an $r-1$ cup contains neither a r cup nor a 3 cap.

Therefore

$$f(r,3)=r, \qquad f(3,s)=s.$$

These agree with the proposed formula:

$$\binom{r-1}{r-2}+1=r.$$

The key recurrence: set up the partition

We prove

$$\boxed{f(r,s) \leq f(r-1,s) + f(r,s-1) - 1.}$$

Let P have $f(r-1,s) + f(r,s-1) - 1$ points. Define

$$L = \{p \in P : \ p \text{ is the leftmost point of an } (r-1)\text{-cup}\}.$$

If $|P \setminus L| \geq f(r-1,s)$, then $P \setminus L$ contains an $(r-1)$-cup or an s-cap. The first is impossible by the definition of L.

Only one case remains.

Otherwise

$$|L| \geq f(r,s-1),$$

so L contains an r-cup or an $(s-1)$-cap.

The key recurrence: the joining step

Suppose L contains an $(s-1)$-cap

$$q_1, q_2, \ldots, q_{s-1}.$$

Because its rightmost point q_{s-1} lies in L, it starts an $(r-1)$-cup

$$q_{s-1} = u_1, u_2, \ldots, u_{r-1}.$$

Compare the two slopes meeting at q_{s-1}.

If

$$\mathrm{slope}(q_{s-2}, q_{s-1}) < \mathrm{slope}(q_{s-1}, u_2),$$

then $q_{s-2}, q_{s-1}, u_2, \ldots, u_{r-1}$ is an r-cup.

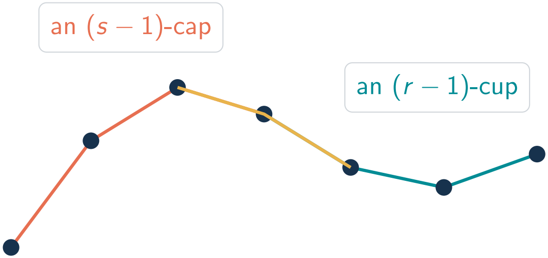
If the inequality is reversed, then

$$q_1, \ldots, q_{s-1}, u_2$$

is an s-cap.

Equality is excluded by general position. Thus one desired pattern always appears.

The recurrence in one picture



$$f(r,s) \leq f(r-1,s) + f(r,s-1) - 1.$$

Pascal's identity solves the recurrence

Assume the bound for smaller values of $r + s$. Then

$$\begin{aligned} f(r,s) &\le f(r-1,s) + f(r,s-1) - 1 \\ &\le \left[\binom{r+s-5}{r-3} + 1\right] + \left[\binom{r+s-5}{r-2} + 1\right] - 1 \\ &= \binom{r+s-4}{r-2} + 1, \end{aligned}$$

where the last equality is Pascal's identity

$$\binom{n}{j-1} + \binom{n}{j} = \binom{n+1}{j}.$$

Together with the boundary cases, this completes the induction. $\square$

The classical Erdős–Szekeres upper bound

Set $r = s = k$. Any $f(k,k)$ points contain a k-cup or a k-cap, hence a convex k-gon. Therefore

$$\boxed{\mathrm{ES}(k) \leq \binom{2k-4}{k-2} + 1.}$$

Using the central-binomial estimate,

$$\binom{2k-4}{k-2} \sim \frac{4^{k-2}}{\sqrt{\pi(k-2)}}.$$

So the proof gives a bound of order $O(\frac{4^k}{\sqrt{k}})$.

How large is the cup–cap bound?

k	exact/known $\mathrm{ES}(k)$	cup–cap upper bound	conjectured
4	5	$\binom{4}{2}+1=7$	5
5	9	$\binom{6}{3}+1=21$	9
6	17	$\binom{8}{4}+1=71$	17
7	unknown	$\binom{10}{5}+1=253$	33
10	unknown	$\binom{16}{8}+1=12871$	257

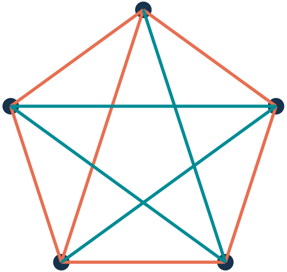
Two-color Ramsey number

The number $R(r, s)$ is the least n such that every red–blue coloring of the edges of the complete graph K_n contains either

- a red complete graph K_r, or
- a blue complete graph K_s.

In the symmetric case, $R(k, k)$ asks:

How many vertices force a monochromatic K_k?



$R(3,3) \leq 6$

Take any red–blue coloring of all edges of K_6 and choose a vertex v.

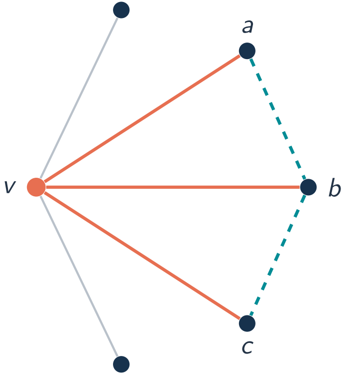
The five edges from v have only two colors. By the pigeonhole principle, at least three have one color; suppose

$$va, \ vb, \ vc \quad \text{are red.}$$

Now inspect the three edges ab, bc, ca:

- if one is red, it forms a red triangle with v;
- if none is red, all three are blue, so abc is a blue triangle.

Thus every coloring of K_6 has a monochromatic triangle.

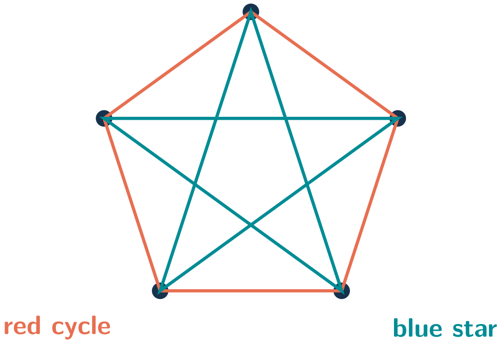


To show that five vertices are not enough, we exhibit a coloring of K_5 with no monochromatic triangle.

- Color the five edges of a pentagon red.
- Color its five diagonals blue.
- The red graph is a 5-cycle, so it has no triangle.
- The blue graph is also a 5-cycle, drawn as a star, so it has no triangle.

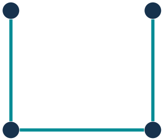
Therefore $R(3,3) > 5$. Together with the previous slide,

$$\boxed{R(3,3) = 6.}$$

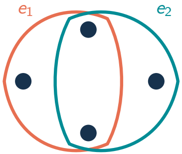


Definition

A **hypergraph** consists of a vertex set V and a collection of subsets of V, called hyperedges. It is k-uniform if every hyperedge contains exactly k vertices.



2-**uniform**
an ordinary graph



3-**uniform**
a 3-uniform hypergraph

For the Happy Ending Problem:

- vertices are planar points;
- every triple is a 3-hyperedge;
- color it by left or right orientation.

The corresponding Ramsey number is written $R_3(N, N)$; the subscript records that the colored hyperedges have size 3.

Ramsey for triples

Finite Ramsey theorem, special case

For every N, there exists an integer $R_3(N, N)$ such that every red–blue coloring of the triples of an $R_3(N, N)$-element set contains N elements whose every triple has the same color.

The number $R_3(N, N)$ is enormous, but **finite**.

What does Ramsey's theorem for triples mean?

For a vertex set V, write $\binom{V}{3} = \{T \subseteq V : |T| = 3\}$ for the collection of all triples. A coloring is a function $\chi : \binom{V}{3} \to \{\text{red}, \text{blue}\}$.

What is $R_3(N, N)$ if put in a simpler but rigorous manner?

$R_3(N, N)$ is the least M with the property that for every set V with $|V| = M$ and every coloring χ, there are a subset $S \subseteq V$ and a color c such that

$$|S| = N \qquad \text{and} \qquad \chi(T) = c \quad \text{for every } T \in \binom{S}{3}.$$

Thus all $\binom{N}{3}$ triples determined by the same N vertices are red, or all are blue.

Finiteness of $R_3(N, N)$

Let $t = R(N, N) < \infty$, using the ordinary graph Ramsey theorem. Begin with a sufficiently large finite reservoir W.

1. Choose $v_1, v_2, \dots, v_t$ successively. After choosing v_j, for every $i < j$ split the remaining W according to the color of $\{v_i, v_j, w\}$ and retain the larger half.
2. Consequently, every pair v_i, v_j has a fixed color c_{ij} satisfying $\chi(\{v_i, v_j, v_k\}) = c_{ij}$ whenever $i < j < k$.
3. Color the edge $v_i v_j$ of K_t by c_{ij}. Since $t = R(N, N)$, there are N vertices whose pair-colors are all the same, say c.
4. Any three of these vertices, indexed $i < j < k$, form a triple of color $c_{ij} = c$. Hence all their triples are monochromatic.

Why the construction terminates

There are only $\binom{t}{2}$ halvings. Starting with, for example, 2^{t^2} vertices leaves enough points to choose all $v_1, \dots, v_t$. Therefore $R_3(N, N) < \infty$.

Proof of finiteness of $ES(N)$

Fix $N \geq 3$ and take at least $M = R_3(N, N)$ non-collinear points.

1. Label the points $p_1, \ldots, p_M$ from left to right.
2. Color each triple $i < j < k$ by the sign of $\text{turn}(p_i, p_j, p_k)$.
3. Ramsey's theorem gives indices $i_1 < \cdots < i_N$ such that every triple among them has the same color.
4. By the geometric lemma, $p_{i_1}, \ldots, p_{i_N}$ are in convex position.

Conclusion

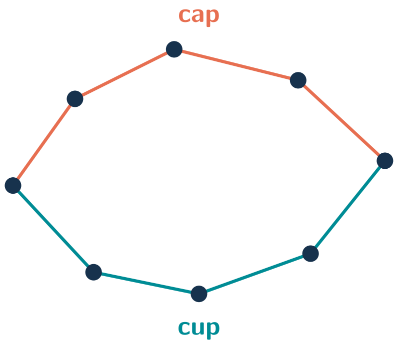
$$\text{ES}(N) \leq R_3(N, N) < \infty.$$

This also proves the Erdős–Szekeres theorem.

Why a convex polygon need not be one cup or one cap

A general convex polygon has two x-monotone boundary chains:

- a lower chain, which is a cup;
- an upper chain, which is a cap.



The lower bound and the conjecture

Erdős and Szekeres constructed, for every k, a set of 2^{k-2} non-collinear points with no k points in convex position. Thus

$$\mathrm{ES}(k) \geq 2^{k-2} + 1.$$

Erdős–Szekeres conjecture

For every $k \geq 3$,

$$\boxed{\mathrm{ES}(k) = 2^{k-2} + 1.}$$

Known general upper bounds

Erdős–Szekeres
$\mathrm{ES}(k) \leq \binom{2k-4}{k-2} + 1$

1935

Suk
$2^{k+O(k^{2/3}\log k)}$

2017

Holmsen–Mojarrad–Pach–Tardos
$2^{k+O(\sqrt{k}\log k)}$

2020

Recent progress

Split k-gon

A split k-gon is the union of an a-cap and a u-cup that share their **rightmost point**, where

$$a + u = k + 2.$$

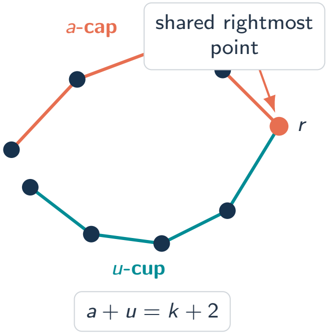
It has $k + 1$ points if only the right endpoint is shared. If the left endpoint is also shared, it has k points in convex position and is an ordinary convex k-gon.

Baek–Balko theorem (2025)

Let $\mathrm{ES}_{\mathrm{split}}(k)$ be the least number of points that forces a split k-gon. For every $k \geq 2$,

$$\boxed{\mathrm{ES}_{\mathrm{split}}(k) = 2^{k-2} + 1.}$$

Thus the conjectured threshold is exact for this relaxed version of convex position.



Further research questions

1. Is $\mathrm{ES}(7) = 33$?
2. Can the upper bound $2^{k+O(\sqrt{k \log k})}$ be pushed all the way to $2^{k-2} + 1$?
3. Which restrictions on triple-orientation colorings capture exactly those arising from planar point sets?
4. What changes for **empty** convex polygons, where the chosen polygon may contain no other points of the original set?

References

P. Erdős and G. Szekeres,
A combinatorial problem in geometry,
Compositio Math. 2 (1935), 463–470.

P. Erdős and G. Szekeres,
On some extremum problems in elementary geometry,
Ann. Univ. Sci. Budapest 3 (1960), 53–62.

W. Morris and V. Soltan,
The Erdős–Szekeres problem on points in convex position—a survey,
Bull. Amer. Math. Soc. 37 (2000), 437–458.

G. Szekeres and L. Peters,
Computer solution to the 17-point Erdős–Szekeres problem,
ANZIAM J. 48 (2006), 151–164.

A. Suk,
On the Erdős–Szekeres convex polygon problem,
J. Amer. Math. Soc. 30 (2017), 1047–1053.

A. F. Holmsen, H. N. Mojarrad, J. Pach, and G. Tardos,
Two extensions of the Erdős–Szekeres problem,
J. Combin. Theory Ser. A 170 (2020), 105132.

J. Baek and M. Balko,
The Erdős–Szekeres conjecture revisited,
SoCG 2025, LIPIcs 332, Article 13.

F. P. Ramsey,
On a problem of formal logic,
Proceedings of the London Mathematical Society, Series 2, 30 (1930), 264–286.