ERDŐS–SZEKERES: UPPER BOUNDS FOR CONVEX HEPTAGONS AND OCTAGONS

MICHAEL SCHROEDER

ABSTRACT. The Erdős–Szekeres number $\mathrm{ES}(n)$ is the least N such that every set of N points in the plane, with no three collinear, contains n points in convex position. We prove $\mathrm{ES}(7) \leq 81$ and $\mathrm{ES}(8) \leq 319$. Both bounds follow from a common geometric partition and exact finite certificates for the resulting cup–cap configurations. The heptagon proof combines a five-point transversal with exclusions that preserve shared geometric witnesses. The octagon proof counts missing Ferrers states using hereditary convexity and four-cap compression. We also establish the $(7, 5, 6)$ case of the Erdős–Tuza–Valtr conjecture: the largest planar set in general position with distinct abscissae avoiding a convex heptagon, a five-cap, and a six-cup has 25 points. All three upper bounds are formalized in Lean, including the geometric reductions and the validity of the finite encodings. The accompanying archive provides sources, exact certificates, and independent checkers. Finite Lean checks use native evaluation with an explicitly documented trust boundary.

1. INTRODUCTION

Let $\mathrm{ES}(n)$ denote the least integer N such that every set of at least N points in $\mathbb{R}^2$, with no three collinear, contains n points in convex position. The polygon need not be empty. Erdős and Szekeres [6, 7] proved

$$2^{n-2} + 1 \leq \mathrm{ES}(n) \leq \binom{2n-4}{n-2} + 1$$

and conjectured equality with the lower bound. Szekeres and Peters [17] established $\mathrm{ES}(6) = 17$ by a computer proof. The next two conjectured values are 33 and 65.

Theorem 1.1. *Every set of* 81 *points in the plane in general position contains a convex heptagon, and every such set of* 319 *points contains a convex octagon. Consequently,*

$$33 \leq \mathrm{ES}(7) \leq 81, \qquad 65 \leq \mathrm{ES}(8) \leq 319.$$

Date: 19 September 2026.

2020 *Mathematics Subject Classification.* Primary 52C10; Secondary 05D10, 68V15.

Key words and phrases. Erdős–Szekeres problem, convex polygons, cups and caps.

ORCID: 0009-0004-3249-0195. Version 1, 19 September 2026.

DOI: 10.5281/zenodo.22842891.

Our contribution is a finite strengthening of the cup–cap method, with a complete formal connection to planar geometry. We use the classical projective reduction, a residual partition, and injective encodings by Ferrers down-sets. The heptagon proof rules out a balanced fifty-point equality case while keeping one shared system of actual witnesses. The octagon proof uses a different certificate: weighted omission inequalities on small families whose occupied states force convexity. Both arguments retain every chronological order allowed by their necessary constraints.

Here is the numerical structure. Write $M_n(a,b)$ for the maximum size of a set avoiding a convex n-gon, an a-cap, and a b-cup, in that parameter order. Put $C_n = M_n(n-1,n-1)$, and let D_n be the maximum size of a set containing no $(n-2)$-cap, no convex n-gon, and no point simultaneously ending an $(n-2)$-cup and starting an $(n-3)$-cap. Both witnesses in the last condition must lie in the same set. Sections 2–3 prove

$$\mathrm{ES}(n) \le C_n + D_n + 2 \qquad (n \ge 6). \tag{1.1}$$

The savings in the two avoiding maxima are as follows.

n	C_n: baseline	here	D_n: baseline	here	Total saving
7	70	49	41	30	$21 + 11 = 32$
8	252	193	154	124	$59 + 30 = 89$

All entries are upper bounds. Before imposing polygon-capacity rows, the cup–cap count gives $\binom{2n-6}{n-3}$ balanced states. The two-sided residual construction in Section 7 gives $\binom{2n-6}{n-4} - \binom{2n-8}{n-5}$ states: in the $(n-4)\times(n-2)$ rectangle, the forbidden junction excludes the maximal cell $(1, n-2)$. At $n = 7, 8$ these give respectively $70 + 41 + 2 = 113$ and $252 + 154 + 2 = 408$. The two-point adjustment includes one deleted projective root and the conversion from an avoiding maximum to a forcing threshold.

We also prove the sharp restricted equality $M_7(5,6) = 25$ (Theorem 5.2), a case of the conjecture of Erdős, Tuza, and Valtr [8]. The conjectured value is $\sum_{i=1}^{3} \binom{5}{i}$. Baek–Balko [2, Section 2.1] record the previously known cases as $n \ge a + b - 3$ or $a = 4$ or $b = 4$; $(n, a, b) = (7, 5, 6)$ lies outside these ranges. The lower construction is classical [7, 8]; the upper bound is the new certificate. Reflection gives the mirrored case $(7, 6, 5)$.

1.1. **Previous work.** Tóth and Valtr's projective method [18, 19] underlies the reduction used below. The explicit finite bound of Nassajian Mojarrad and Vlachos [13, Theorem 1.1] is

$$\mathrm{ES}(n) \le \binom{2n-5}{n-2} - \binom{2n-8}{n-3} + 2 \qquad (n \ge 6), \tag{1.2}$$

which gives 113 and 408. Our bounds improve these by 32 and 89, respectively. Norin and Yuditsky [14] obtain the asymptotic factor 7/8 by encoding incoming and outgoing chains. Their three-chain argument is used in our residual partition. Suk [16] proves $\mathrm{ES}(n) = 2^{n+o(n)}$; Holmsen, Nassajian Mojarrad, Pach and Tardos [9] refine the upper bound to $2^{n+O(\sqrt{n \log n})}$. These asymptotic expressions do not by themselves supply explicit improvements at $n = 7, 8$.

The strong- and weak-polygon analogues of the Erdős–Szekeres conjecture fail for arbitrary red–blue colorings of ordered triples already at $k = 7$: Balko–Valtr [3] prove $C_{\text{strong}}(7) > 33$, and Baek–Balko [2] prove $C_{\text{weak}}(7) > 33$. Here C_{strong} and C_{weak} are the corresponding forcing thresholds. A weak k-gon consists of differently colored monotone paths with common endpoints and vertex counts summing to $k + 2$; a strong k-gon additionally requires disjoint interiors. The signotope extension remains open: Baek–Balko [2, Section 3] relate it to the Goodman–Pollack conjecture. Our results concern real planar sets; a signotope interpretation of the complete proof has not been audited. Down-set encodings of ordered paths are established machinery [12]. Baek [1] proves that $\binom{n-1}{2} + 2$ points force a four-cap or a convex n-gon; its $n = 8$ case is a numerical input to our block-compression argument. Baek and Balko [2] obtain sharp split-polygon thresholds and results for decomposable configurations. A split polygon permits unmatched left endpoints, so its threshold cannot be substituted for the unrestricted convex-polygon threshold.

Recent computational work also emphasizes the distinction between an encoding and a complete exclusion. Dumitru [5] reports refutations of anchored thirty-three-point subfamilies, without an exhaustive cover. Damásdi, Dong, Scheucher and Zeng [4] construct small saturated sets, showing why failure to extend one example need not establish a maximum. Koshelev and Koshka [10] develop SAT/ASP models and fixed-abscissa realization searches for related variants. Infeasibility with fixed abscissae does not exclude arbitrary coordinates. None of these restrictions is imposed in our final theorem.

Our finite comparison uses (1.2).[1]

Computer-assisted and formal geometric proofs have substantial precedent. Marić [11] formalized the small convex-polygon cases, and Subercaseaux and coauthors [15] formalized the empty-hexagon number in Lean. The latter concerns a different emptiness condition, but its separation of geometric encoding from propositional refutation is directly relevant. Our final Lean statements concern ordinary convex hulls and arbitrary real coordinates. Their precise computational trust is stated in Section 8; external peer review is not claimed.

[1]Table 1 of [10] lists 92 and 324 under Norin–Yuditsky's asymptotic result. We have not located their finite derivation in the cited proof, which discards a nonzero deletion term. Our bounds improve those entries by 11 and 5 as well. The literature supplement records the table's fitted formula and its exact relation to (1.2).

2. Geometric conventions and the two-point shift

First apply an invertible shear $(x, y) \mapsto (x + ty, y)$ separating all abscissae; only finitely many parameters t fail. Convexity and general position are preserved. Write $p < q$ when $x(p) < x(q)$, and $m(p, q) = (y(q) - y(p))/(x(q) - x(p))$. A k-cup is an increasing sequence of k points whose consecutive slopes strictly increase; a cap has strictly decreasing slopes. Length always counts vertices.

Every subsequence of a cup is a cup, and likewise for caps: a chord slope is a positive weighted average of the slopes of the edges it spans. A cup and cap sharing both endpoints have disjoint interiors, on opposite sides of the common chord. Their union is in convex position and has the sum of their lengths minus two vertices. These statements are used in the planar category.

Proposition 2.1 (Projective reduction)**.** *Let Q_n be the maximum size of a set avoiding an $(n-1)$-cap and a convex n-gon. Then* $\mathrm{ES}(n) \leq Q_n + 2$*.*

Proof. Take an n-gon-free set S and a hull vertex o. After a translation and rotation, $o = (0, 0)$ and every other point has positive second coordinate. For sufficiently small $t > 0$, apply

$$T_t(x, y) = \left(\frac{x - t}{y + t^2}, \frac{1}{y + t^2} \right).$$

This projective map is invertible and has positive denominator on $\mathrm{conv}(S)$, so it preserves convex position. The nonroot images stay bounded, with distinct limiting first coordinates x/y: an equality would make two original points collinear with o. The root becomes $(-1/t, 1/t^2)$. For nonroot images $(X_i, Y_i) < (X_j, Y_j)$,

$$\mathrm{orient}(T_t(o), T_t(p_i), T_t(p_j)) = \frac{X_j - X_i}{t^2} + \frac{Y_j - Y_i}{t} + X_i Y_j - Y_i X_j > 0$$

for small enough t. Thus the root is leftmost and all its triples are cups. Deleting it leaves no $(n-1)$-cup; reflection in a horizontal line gives a set counted by Q_n. Hence $|S| \leq Q_n + 1$ and $\mathrm{ES}(n) \leq Q_n + 2$. □

3. A common residual partition

We use the three-chain argument of Norin and Yuditsky [14, Claim 3.3].

Lemma 3.1 (Three-chain bridge)**.** *Let $n \geq 6$. Suppose an $(n-2)$-cup F ends at s, an $(n-1)$-cup P starts at s and ends at t, and an $(n-2)$-cap J starts with (s, q). If $q \neq t$, their union contains an $(n-1)$-cap or a convex n-gon. Other overlaps between these chains are permitted.*

Proof. Let f be the penultimate point of F, and let v, u, t be the final three points of P. If (f, s, q) is a cap, prepend f to J. Otherwise F, q is an $(n-1)$-cup and $m(f, s) < m(s, q)$.

Suppose first that $q < t$. If t is above the line sq, then F, q, t is an n-cup. If it is below that line, then q is above the chord st, so P and (s, q, t) are opposite chains with common endpoints and yield an n-gon. In the latter case q is not internal to P, whose internal points lie below st.

Now let $q > t$. If q is above ut, append q to P. Otherwise q is below ut. If t is below sq, then (u, t, q) is a cap and $m(t, q) > m(s, q) > m(q, J_3)$; following (u, t) by J without s gives an $(n-1)$-cap.

It remains that t is above sq and q below ut. If u is above sq, then F, u, t is an n-cup, using $m(f, s) < m(s, q) < m(s, u)$ and the cup triple (s, u, t). Otherwise u is below sq. If q is above vu, replace the last point t of P by q; the resulting $(n-1)$-cup and the cap (s, t, q) give an n-gon. Finally, if q is below vu, then (v, u, q) is a cap and $m(u, q) > m(s, q) > m(q, J_3)$, giving an $(n-1)$-cap by following (v, u) with J without s.

General position excludes every line equality in this partition of cases. When $q > t$, the cap suffix starting at q lies after all of P, and the nonterminal points of F lie before s. Thus the chains displayed in the argument have distinct vertices where required. $\square$

The decision map below lists the eight terminal cases of lemma 3.1. Use the first test that holds; $J' = J \setminus \{s\}$ retains its order.

First applicable test	Forced configuration
(f, s, q) is a cap	(f, J): $(n-1)$-cap
$q < t$ and t above sq	(F, q, t): n-cup
$q < t$	$P \cup (s, q, t)$: convex n-gon
q above ut	(P, q): n-cup
t below sq	(u, t, J'): $(n-1)$-cap
u above sq	(F, u, t): n-cup
q above vu	$(P \setminus \{t\}, q) \cup (s, t, q)$: n-gon
Otherwise	(v, u, J'): $(n-1)$-cap

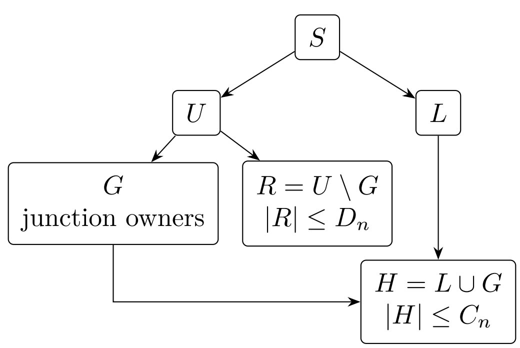

FIGURE 1. The residual partition. First $S = U \sqcup L$, then $U = G \sqcup R$. Transferring G to the balanced part gives $S = H \sqcup R$. Membership in G uses witnesses inside U.

Theorem 3.2 (Residual transfer)**.** *For $n \geq 6$, $Q_n \leq C_n + D_n$ and consequently* $\mathrm{ES}(n) \leq C_n + D_n + 2$.

Proof. Let S avoid an $(n-1)$-cap and a convex n-gon. The case $|S| \leq 1$ is immediate. At each point x, let $p(x)$ be the point minimizing the slope of the incident unoriented line. The minimizer is unique by general position. Partition S into

$$U = \{x : p(x) > x\}, \qquad L = \{x : p(x) < x\}.$$

The map p takes U into L and L into U. For example, if $x < y = p(x) < z = p(y)$, then $m(y,z) < m(x,y)$, and the weighted-average identity gives $m(x,z) < m(x,y)$, a contradiction. The reversed argument handles the other direction. A cap ending in U extends to its chosen right neighbor in L; a cup starting in L extends to its chosen left neighbor in U. Thus U has no $(n-2)$-cap and L has no $(n-1)$-cup.

Let $G \subseteq U$ be the points ending an $(n-2)$-cup and starting an $(n-3)$-cap, both in U. Put $H = L \cup G$ and $R = U \setminus G$. The set R satisfies the definition of D_n, since any forbidden junction inside R would also occur in U.

Suppose that H contains an $(n-1)$-cup with endpoints s, t. If $s \in L$, extend it left to an n-cup. If $t \in G$, extend its $(n-3)$-cap in U once into L, giving an $(n-2)$-cap starting at t. At its junction with the supposed cup, one orientation yields an n-cup and the other an $(n-1)$-cap. Hence $s \in G$ and $t \in L$. Take the actual $(n-2)$-cup in U ending at s and extend its actual outgoing $(n-3)$-cap once into L. The second point q of this cap lies in its original cap in U; extending that cap only appends its last point. Therefore $q \neq t$. Lemma 3.1 applies and gives a contradiction. Hence H has no $(n-1)$-cup, and inherits the other avoidances, so $|H| \leq C_n$. Now $S = H \sqcup R$ gives the result. The threshold statement follows from proposition 2.1. □

4. The heptagon residual has a five-point transversal

Write $B = M_7(5,6)$. Call a point a *right root* if it starts no three-cap in the specified ambient set. Every triple starting at a right root is a cup, so it can be prepended to any cup to its right.

Lemma 4.1 (Endpoint shift)**.** *Every endpoint of a six-cup in a set counted by D_7 is a right root.*

Proof. Let a six-cup end with u, s and suppose (s,q,r) is a three-cap. If (u,s,q) is a cup, append q to obtain a seven-cup. Otherwise u ends the five-point prefix of the six-cup and starts the four-cap (u,s,q,r), a forbidden junction. Both possibilities are impossible. □

Lemma 4.2 (Anchoring)**.** *In a heptagon-free set, suppose $(a, b, t_3, t_4, t_5, t_6)$ is a six-cup and a is a right root. Every six-cup ending at b starts at a.*

Proof. If a start $x < a$ existed, (x,a,b) would have to be a cap, since otherwise x could be prepended to the displayed cup. That cap and the six-cup from x to b would form a heptagon. If $x > a$, the right root a could be prepended to the six-cup from x to b. Only $x = a$ remains. □

Theorem 4.3 (Five-point deletion)**.** *Every set counted by D_7 has a subset of at most five points meeting every six-cup. In particular $D_7 \leq B + 5$.*

Proof. Let E be all actual six-cup endpoints. By lemma 4.1, every triple in E is a cup; thus $|E| \leq 6$. If $|E| \leq 5$, delete E. Otherwise write $E = (e_1, \dots, e_6)$. This is a six-cup, and lemma 4.2 says that every six-cup ending at e_2 starts at e_1. Delete $E \setminus \{e_2\}$. A surviving six-cup would have an original endpoint in E, hence endpoint e_2, but would have to start at the deleted point e_1. This is impossible. The remaining set has no five-cap, six-cup, or heptagon, and has size at most B. □

5. Finite states and the balanced heptagon exclusion

Down-set encodings of monotone paths are classical [12, 2].

The certificate conventions in this section use *cup first, cap second.* This differs from the order of the arguments of $M_n(a, b)$. For reference, the two stored conventions are

Certificate	Coordinates of a cell
Section 5: heptagon	(cup, cap)
Sections 6–7: octagon	(cap, cup)

For $u < v$, let $d(u, v) = (r(u, v), b(u, v))$, where $r+1$ and $b+1$ are the longest cup and cap lengths ending at the exact pair (u, v). Define

$$F(v) = {\downarrow}\{d(u, v) : u < v\}. \tag{5.1}$$

All these ranks are actual maxima in the stated ambient set.

Lemma 5.1 (Difference and chronology)**.** *For $u < v$, $d(u, v) \in F(v) \setminus F(u)$. Consequently states are injective, and strict inclusion of two occupied states forces the same order on their owners.*

Proof. If $d(u, v) \in F(u)$, some actual predecessor $t < u$ has $d(t, u) \geq d(u, v)$ coordinatewise. According to the orientation of (t, u, v), extending an attaining cup or cap contradicts one of the two maxima at (u, v). Thus the difference is nonempty. Equal states, or an earlier state containing a later one, would contradict this conclusion. □

With no a-cap or b-cup, the states lie in $[b-2] \times [a-2]$, whose number of down-sets is $\binom{a+b-4}{a-2}$, including the empty state. A Ferrers tuple denotes column heights, with trailing zeros where required by the data format. The finite models permit every chronological order compatible with inclusion; they do not sort incomparable states into a preferred order.

Theorem 5.2 (A sharp ETV case; computer-assisted)**.** *Every set of 26 points in $\mathbb{R}^2$, in general position and with distinct x-coordinates, contains a* *five-cap, a six-cup, or seven points in convex position. There is a 25-point set avoiding all three. Thus $M_7(5,6) = 25$; reflection also gives $M_7(6,5) = 25$.*

Certificate reduction. For the lower bound, the Erdős–Szekeres construction [7], in the restricted form of Erdős–Tuza–Valtr [8], gives $\sum_{i=1}^{3} \binom{5}{i} = 25$ points with none of the three configurations; see also Baek–Balko [2, Section 4]. The archive also contains an exact rational 25-point witness and a separate integer checker confirming maximum cup length five, maximum cap length four, and no convex heptagon among all $\binom{25}{7} = 480\,700$ seven-point subsets. This lower-bound check is separate from the Lean upper-bound theorem.

A hypothetical 26-point counterexample has 26 distinct actual states in $[4] \times [3]$, a universe of 35 states, and hence nine missing states. The empty state and the one-cell state occur at the first two points; the third has one of the two two-cell states. In the certificate's area-sorted dictionary, an eligible hole set therefore avoids indices 0 and 1 and does not contain both 2 and 3. There are exactly

$$\binom{33}{9} - \binom{31}{7} = 35\,937\,525$$

eligible hole sets.

The certificate derives necessary hole clauses from pair-label differences, exact corner attainment, shared predecessor witnesses, actual owner order, four-point sign constraints, and common-endpoint paths. A five-cup and four-cap with common endpoints give a heptagon, so excluding just this type is a necessary restriction. The Lean coverage stage refutes a CNF consisting of 91,720 derived hole clauses, a cardinality encoding, and one padding tautology: 461 variables and 93,169 clauses in total. Its LRAT proof has 381,647 additions. Separately, two enumerations, recursive and fixed-popcount, verify that none of the eligible hole sets survives. Appendix B describes the primitive semantics and proof format. $\square$

Consider now a balanced avoiding set S of size 50. Let $S_\cup$ and $S_\cap$ be its actual five-cup and five-cap starters, and $E_\cup$ and $E_\cap$ the corresponding endpoint sets.

Lemma 5.3 (Equality partition)**.** *There is an integer t such that*

$$S = S_\cup \sqcup E_\cap = S_\cap \sqcup E_\cup, \qquad |S_\cup| = |S_\cap| = |E_\cup| = |E_\cap| = 25,$$

and, with

$$L = S_\cup \cap S_\cap, \quad X = S_\cup \cap E_\cup, \quad Y = S_\cap \cap E_\cap, \quad R = E_\cup \cap E_\cap,$$

one has $|X| = |Y| = t$ and $|L| = |R| = 25 - t$.

Proof. The set $S_\cup$ contains no five-cap: joining the end of such a cap to an actual outgoing five-cup extends one of them to a six-chain. Its complement contains no five-cup. By theorem 5.2 and reflection, both parts have size at most 25, and hence exactly 25. Color exchange and order reversal give the same endpoint counts. A point ending a five-cap cannot start a five-cup, by the same junction argument, so the disjoint pairs each exhaust S. Taking their four intersections gives the stated cardinalities. □

Lemma 5.4 (Population certificates)**.** *In a planar set with no six-cup, six-cap, or heptagon:*

(i) at most thirteen points both end and start five-cups;
(ii) any subset avoiding five-cups and five-caps, whose last point starts a five-cup in the ambient set, has at most nineteen points;
(iii) if N_0 consists of points ending neither five-chain and N_2 of points ending both, then

$$\min(|N_0|, |N_2|) \leq 16.$$

Necessary reductions and finite certificates. For (i), select fourteen supposed through-cup points, a five-cup ending at the first, and a five-cup starting at the last. Exactly 22 points are selected: fourteen core points and four on each side. Any core four-cap extends at both ends to a six-cap, so the core has cap length at most three. Its cap paths extend by one at each end, whereas its cup paths remain valid ambient paths. The source records exact core recurrences and these valid implications for ambient ranks, without requiring ambient attainment in the selection. Its refutation has 3,009 additions.

For (ii), select a supposed 20-point core and the four other vertices of an actual five-cup from its last point. Incoming ranks ending in the core are recomputed there; neither color reaches five. The resulting 24-point source is refuted in 79,173 additions.

For (iii), the incoming states of N_0 are the twenty ideals of $[3]^2$. Those of N_2 are the twenty ideals $(4, 1 + h_1, 1 + h_2, 1 + h_3)$ with h an ideal of $[3]^2$. Select seventeen owners from each family, retaining the ambient first two points. A named-owner source permits all such selections and all their inclusion-compatible orders. Pair ranks are ambient and paths are selected. Its refutation has 102,563 additions.

All three formulas include the necessary four-point packet constraints and common-endpoint polygon exclusions. The compact formulas and refutations are embedded in the Lean sources. Their geometric interpretations and logical soundness are proved there; section 8 and appendix A describes the interface. Thus each supposed configuration supplies an assignment to a refuted formula. □

Corollary 5.5. *In lemma 5.3,* $9 \leq t \leq 13$*. Every twenty ambient five-cup starters contain an internal five-cup, and a 25-point starter class has at least six actual internal five-cup sources.*

Proof. The through-cup bound gives $t \leq 13$. In the equality partition, $N_0 = L$ and $N_2 = R$, so $25-t \leq 16$ and $t \geq 9$. Any set of twenty cup starters has no internal five-cap by the junction argument. If it had no five-cup either, part (ii) of lemma 5.4, applied to its last point, would be contradicted. Finally, deleting at most five internal sources from a 25-point starter class would leave at least twenty ambient starters with no internal five-cup. Deletion cannot create a new path, so at least six sources must exist. □

Proposition 5.6 (Balanced heptagon bound)**.** $C_7 \leq 49$.

Proof. Assume $|S| = 50$ and let $K = S_\cup$. Retain the 25 points of K and the four other points of an actual five-cup from $\max K$. For a point of K, no ambient five-cap can end there, so its actual incoming state lies in the 35-state universe $[4] \times [3]$. Fifteen of these states have width four; their active population is precisely t. The low-width owners are also cap starters by lemma 5.3. All 25 states are actual ambient states, not states recomputed after selecting these 29 points.

Reflect the whole configuration if necessary to make its first triple a cup, and select the corresponding starter class anew. The four possible initial packets, in order $012, 013, 023, 123$, are $1000, 1100, 1110, 1111$. A cup and cap share at most two vertices. The initial three-cup therefore forces the first five ambient points to have no ending five-cap, and hence to belong to K. In the 1111 sector the first six points similarly belong to K. Exhaustive literal enumeration retains all seven five-point extensions of 1000 and all 118 six-point extensions of 1111 without a six-cup. The 1100 and 1110 sectors remain in the primary formula until their exclusions are derived.

The primary formula has 151,801 variables and 4,017,399 clauses. It includes corollary 5.5, the six-source condition, free order for incomparable owners, shared ambient pair labels, selected paths, exact core outgoing paths, and the prefix alternatives. Exact incoming attainment is imposed only at retained ambient prefix points, where every predecessor is present. The variable meanings and necessary clause families are listed in Appendix A.

A single proof over this unaugmented primary formula derives the empty clause. It contains 1,989,465 additions and 84,721,022 ordered hints. Its derivation includes the exclusions of $t = 13, 9, 12$, the two mixed prefixes, and both final cases $t = 10, 11$. These case facts are derived within the proof; the final contradiction has no case assumption. The compact source uses 190,932 indexed clauses of the primary formula; the same number of proof additions suffices. Checked replay of this compact proof establishes unsatisfiability. The formal proof additionally checks each retained clause against a proved necessary schema under one common assignment.

Every supposed planar S supplies a truthful assignment to the retained source, so unsatisfiability is a contradiction. Any larger avoiding set would contain an avoiding fifty-point subset. □

Combining theorems 3.2, 4.3 and 5.2 and proposition 5.6 gives

$$\mathrm{ES}(7) \leq 49 + (25 + 5) + 2 = 81.$$

The classical construction [7] supplies the lower endpoint $\mathrm{ES}(7) \geq 33$; it is separate from the formalized upper bound.

6. A common interface for octagon capacities

For the octagon arguments, coordinates are *cap first, cup second*. This is the transpose of the heptagon certificate convention, not a change in the parameter order of $M_n(a,b)$.

Definition 6.1 (Admissible labels and states)**.** Let P be a planar point set in general position. Assign a label $d(u,v) = (p(u,v), q(u,v)) \in \mathbb{Z}^2_{>0}$ to every ordered pair $u < v$, and a finite down-set $I_v \subset \mathbb{Z}^2_{>0}$ to every point. The assignment is admissible if

$$p(u,v) < p(v,w) \qquad \text{for each cap triple } u < v < w, \tag{6.1}$$

$$q(u,v) < q(v,w) \qquad \text{for each cup triple } u < v < w, \tag{6.2}$$

$$d(u,v) \in I_v \setminus I_u \qquad \text{for each } u < v. \tag{6.3}$$

A family of states has *admissible capacity* b if every octagon-free set with any such assignment occupies at most b of its states.

The states are distinct, and strict inclusion forces the same owner order: the opposite order would make (6.3) an empty directed difference.

Ordinary actual incoming ranks, with their incoming down-sets, provide an admissible assignment by lemma 5.1. More generally, labels satisfying (6.1)–(6.2) provide one by taking $I_v = \downarrow\{d(u,v) : u < v\}$. The same domination proof applies. The states in an admissible assignment need not equal these generated down-sets. Restricting to a subset of points and *retaining* its labels and assigned states preserves admissibility. This is the interface needed for transport; a bound using actual attainment cannot automatically be transported in this way.

Lemma 6.2 (Forced signs)**.** *Suppose the chronological owners of states A, B, C exist, and put $\Delta_1 = B \setminus A$, $\Delta_2 = C \setminus B$. If*

$$\max_{z \in \Delta_2} z_2 \le \min_{z \in \Delta_1} z_2,$$

their triple is a cap. Interchanging coordinates forces a cup. Both directed differences must be nonempty.

Proof. The two consecutive pair labels belong to Δ_1 and Δ_2. A cup would strictly increase the second coordinate, contradicting the displayed inequality. The other assertion is dual. □

6.1. **Hereditary convex families.**

A cyclic state list is certified by checking every triple in its cyclic order under all six chronological permutations. Empty directed pair differences exclude an order. Simultaneous forcing of both signs also excludes an order. Every other order must have the sign required by the cyclic orientation, with permutation parity accounted for.

Lemma 6.3 (Hereditary convexity)**.** *Every occupied subset of a certified cyclic family is in convex position. Consequently such a family K has admissible capacity seven.*

Proof. For at least three occupied states, every triple of their owners is counterclockwise in the inherited cyclic order. For each directed edge between consecutive selected vertices, every other selected vertex is strictly on its left. Each edge is supporting, so all selected vertices are hull vertices. Eight occupied states would give an octagon. □

This test is hereditary because a triple's test uses only those three states. A proof for a whole path with fixed endpoints would not, by itself, justify a capacity after some endpoints are omitted.

7. The two octagon omission certificates

Write $h_A = 1$ when state A is absent. A valid occupied capacity b_K on a family K gives the omission inequality

$$\sum_{A \in K} h_A \geq |K| - b_K.$$

We first give the arithmetic rule used in both finite proofs.

Lemma 7.1 (Exact weighted contradiction)**.** *Suppose omission rows are $\sum_{v \in K} h_v \geq a_K$. At a branch node, let O be the states fixed missing, Z the states fixed occupied, V the unassigned states, and k the remaining hole budget. For nonnegative integer weights w_K and an integer $\Lambda > 0$, put*

$$r_K = a_K - |K \cap O|, \qquad W = \sum_K w_K r_K,$$
$$\ell(v) = \sum_{K \ni v} w_K, \qquad P = \sum_{v \in V} \max(0, \ell(v) - \Lambda).$$

Every completion using at most k further holes satisfies $W \leq \Lambda k + P$. Thus $W > \Lambda k + P$ certifies the node as impossible.

Proof. Adding the weighted residual inequalities and using binary values gives

$$W \leq \sum_{v \in V} \ell(v) h_v \leq \Lambda \sum_{v \in V} h_v + \sum_{v \in V} \max(0, \ell(v) - \Lambda) \leq \Lambda k + P.$$

All operations are exact integer arithmetic; P accounts for excess load. □

7.1. **The balanced bound.** Ordinary cap/cup incoming states inject a balanced avoiding set into the 252 down-sets of $[5]^2$. The supplied catalogue contains 2,685 families: 1,584 of size eight, 780 of size nine, and 321 of size ten. Every family has a certified cyclic order and hence capacity seven by lemma 6.3.

For example, the following ten-state cyclic list requires at least three omissions:

$$(3,2,1,0,0),\ (3,2,1,1,1),\ (3,2,2,1,1),\ (3,3,2,1,1),$$
$$(4,3,2,1,1),\ (5,3,2,1,1),\ (5,3,2,1,0),\ (5,3,2,0,0),$$
$$(5,3,1,0,0),\ (5,2,1,0,0).$$

The check covers all 120 triples and their chronological permutations, as illustrated in Figure 2. The independent checker rejects 472 of the 720 orders and forces the required sign in all 248 others.

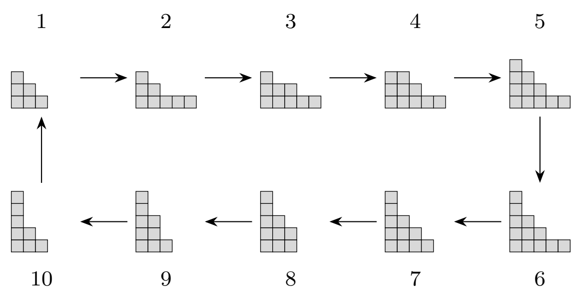

FIGURE 2. The ten Ferrers states in their certified cyclic order. The arrows specify the order of the certificate, not the chronological order of points. At most seven states can be occupied.

Proposition 7.2 (Balanced octagon bound)**.** $C_8 = M_8(7,7) \leq 193$.

Proof. The complete binary certificate assumes at most 58 missing states. It has 1,590 branches and 1,591 weighted leaves, with maximum depth 14. Each split assigns one previously unassigned Boolean variable and includes both children. Every node is reached exactly once, and all leaves satisfy lemma 7.1 in their actual branch context. There are 306,922 positive weight terms and 119 leaves with $P > 0$. The minimum strict margin is 21.

For the smallest-margin leaf, the remaining budget is $k = 52$, $\Lambda = 10^6$, $W = 52{,}000{,}021$, and $P = 0$. Thus a completion would require $52{,}000{,}021 \leq 52{,}000{,}000$. The other 1,590 leaves are checked in the same exact fashion. Every hypothetical hole set with size at most 58 follows a path to a contradiction. At least 59 states are absent, leaving at most $252 - 59 = 193$ points. □

7.2. **A two-sided residual state space.** For a set R counted by D_8, let $C^+(u,v)+1$ be the longest cap starting at the exact pair (u,v) and $Q^-(u,v)+1$ the longest cup ending there, both in the entire fixed R. Define

$$d(u,v) = (5 - C^+(u,v), Q^-(u,v)), \qquad \Psi(v) = {\downarrow}\{d(u,v) : u < v\}. \tag{7.1}$$

These coordinates lie in $[4] \times [6]$. A cap triple $u < v < w$ gives $C^+(u,v) \geq C^+(v,w) + 1$, so the first coordinate strictly increases; a cup triple strictly increases the second. Thus these labels and states are admissible and injective. Future-cap statistics have precedent in Baek's work [1, Theorem 3.2 and Definition 3.4]; the useful specialization here is to the residual forbidden junction.

Proposition 7.3 (Residual state restriction)**.** *Every residual state belongs to the 154-element universe*

$$\mathcal{U} = \left\{ h \in \mathbb{Z}^4 : \begin{array}{l} 6 \geq h_1 \geq h_2 \geq h_3 \geq h_4 \geq 0, \\ h_1 < 6 \text{ or } h_1 = h_2 = 6 \end{array} \right\}.$$

Proof. No actual pair label is $(1, 6)$. Such a label would exhibit a five-cap starting at (u, v) and a seven-cup ending at (u, v). Removing v from the latter gives a six-cup ending at u, where the five-cap starts. This is the forbidden residual junction.

Every maximal cell of an actual incoming down-set is attained by some incoming pair. Hence $(1, 6)$ cannot be maximal. Among the $\binom{10}{4} = 210$ down-sets of $[4] \times [6]$, exactly those with $h_1 = 6$ and $h_2 \leq 5$ have that maximal cell. There are $\binom{8}{3} = 56$ such shapes, leaving $210 - 56 = 154$. □

A state containing $(1, 6)$ remains allowed when a larger cell dominates it. An exact 154-point control realizes all the allowed states while satisfying the chain and junction restrictions; it contains a checked octagon. Thus further omissions genuinely require polygon avoidance.

Proposition 7.4 (General residual baseline)**.** *For $n \geq 6$,*

$$D_n \leq \binom{2n-6}{n-4} - \binom{2n-8}{n-5}.$$

Consequently, the cup–cap estimate for C_n and theorem 3.2 recover the finite bound (1.2)*.*

Proof. Use the labels $(n-3-C^+(u,v), Q^-(u,v))$ in $[n-4]\times[n-2]$, with both ranks computed in the residual set. The growth and injection argument is unchanged. A label $(1, n-2)$ would give an $(n-3)$-cap starting at u and an $(n-1)$-cup ending at the pair (u, v), hence the forbidden junction at u. So $(1, n-2)$ cannot be a maximal cell. The excluded down-sets have first column $n-2$ and remaining $n-5$ columns of height at most $n-3$, giving $\binom{2n-8}{n-5}$ exclusions. Finally, Pascal's identity and binomial symmetry give

$$\begin{aligned} &\binom{2n-6}{n-3} + \binom{2n-6}{n-4} - \binom{2n-8}{n-5} + 2 \\ &\qquad = \binom{2n-5}{n-2} - \binom{2n-8}{n-3} + 2. \end{aligned}$$

This is a derivation of the known baseline, not a new bound for general n. □

7.3. **Transporting four-cap capacities.** The following lemma supplies numerical information from Baek's theorem without assuming that transported labels are actual longest ranks.

Lemma 7.5 (Horizontal block compression)**.** *For $\rho = 1, \ldots, s$, let nonnegative integer thresholds $f_\rho(0) < \cdots < f_\rho(r)$ lie in a larger rank rectangle and satisfy*

$$f_{\rho+1}(i) \le f_\rho(i), \qquad f_{\rho+1}(i) \le f_\rho(i-1) + 1 \quad (1 \le i \le r),$$

where the first inequality also includes $i = 0$. Map a source ideal with row widths w_ρ to the ideal with row widths $f_\rho(w_\rho)$. In any admissible assignment, the owners of states in this image admit labels in $[r] \times [s]$ satisfying strict growth. Consequently they contain neither an $(r+2)$-cap nor an $(s+2)$-cup.

Proof. Threshold monotonicity makes every image a down-set, and strictness makes the state map injective. Directed differences are precisely the unions of intervals belonging to the corresponding source difference. Compress a pair label (R, ρ) to (i, ρ) when $f_\rho(i-1) < R \le f_\rho(i)$. Such an interval exists uniquely because the label is in a directed image difference.

Cup growth is unchanged. At a cap triple, suppose the outgoing compressed cell (j, σ) had $j \le i$, where (i, ρ) is incoming. The middle source ideal contains the incoming cell and excludes the outgoing one, so its down-set property forces $\sigma > \rho$. Therefore

$$R_{\text{out}} \le f_\sigma(j) \le f_{\rho+1}(i) \le f_\rho(i-1) + 1 \le R_{\text{in}},$$

contrary to cap growth. Thus the compressed first coordinate increases on every cap triple. Bounded strictly increasing ranks exclude the stated chains. □

With $r = 2$ and $s = 6$, the source has 28 states. The image owners have no four-cap and inherit octagon avoidance. Baek's theorem [1, Theorem 1.6] says that 23 points force a four-cap or an octagon. Each image therefore has occupied capacity 22, and requires six omissions. The residual catalogue contains 23 explicit maps into $\mathcal{U}$. The finite validation checks the threshold conditions, images, and directed differences for all 23 maps.

7.4. **The residual omission theorem.** The final formal catalogue has 1,187 omission inequalities: 1,164 hereditary convex families, each with demand $|K| - 7$, and 23 four-cap block images, each with demand six. The original discovery archive split the first group into 973 transported families and 191 cycles. In the formalization, every one of the 973 transported supports also receives an explicit cyclic certificate. Thus all 1,164 rows use lemma 6.3; their historical transport provenance is no longer an additional proof dependency. The remaining 23 rows use lemma 7.5 and the port of Baek's theorem.

Proposition 7.6 (Residual octagon bound)**.** *$D_8 \le 124$.*

Proof. Assume at most 29 of the 154 states are missing. The complete tree has 7,314 binary branches and 7,315 weighted leaves, maximum depth 21. Every split includes both values of one previously unassigned state, and every node is reached exactly once. Each leaf satisfies lemma 7.1 with its recorded context. There are 662,561 positive weight terms; the minimum strict integer margin is one. Thus no Boolean vector with at most 29 holes satisfies all the valid inequalities. At least 30 states are absent and $|R| \le 154-30 = 124$. $\square$

Finally, theorem 3.2 and propositions 7.2 and 7.6 give

$$Q_8 \le 193 + 124 = 317, \qquad \boxed{\mathrm{ES}(8) \le 193 + 124 + 2 = 319}.$$

The classical lower construction gives $\mathrm{ES}(8) \ge 65$ [7].

8. Formalization and reproducibility

The two final declarations are

`ES7FinalVerified.es7_le_81`, `ES8Final.es8_le_319`.

Appendix C prints their types and the definitions of general and convex position, together with the exact axiom output. The companion module `StatementAudit.lean` proves that `ConvexPosition` is equivalent to mathlib's `ConvexIndependent` for the inclusion of the point set, and restates both bounds with that standard predicate. It adds no native-evaluation dependency.

Each theorem quantifies over a finite subset of $\mathbb{R}^2$. General position requires nonzero oriented area for every distinct triple; the conclusion is a subset of cardinality seven or eight, each of whose points lies outside the convex hull of the others. The statements have no additional numerical or geometric hypotheses. The geometric reductions are formalized for $n = 7$ and $n = 8$. The classical lower constructions in theorem 1.1 are cited separately, not claimed as formalized here.

8.1. **The proof boundary.** There are three distinct obligations. First, geometric necessity sends every hypothetical planar counterexample to one assignment satisfying the finite source. Second, a general checker-soundness theorem turns acceptance of a certificate into a contradiction for that source. Third, the concrete checker accepts the supplied finite data. The first two are ordinary Lean proofs. The third uses `native_decide` for the large certificates and finite catalogues. The five final heptagon refutations use Lean's standard-library `Std.Tactic.BVDecide.LRAT` checker and its `check_sound` theorem. The earlier restricted primitive-replay stage has its own proved RUP layer. No SAT solver or numerical optimizer is trusted to assert unsatisfiability.

The standard axioms in both final dependency lists are

`propext`, `Classical.choice`, `Quot.sound`.

The heptagon theorem additionally uses 19 native-evaluation checks, and the octagon theorem uses eight. There is no `sorryAx` and no additional mathematical axiom. Native evaluation additionally trusts Lean's compilation and execution infrastructure; this is not a kernel-only evaluation of every finite computation. Exact axiom names, source hashes, and the toolchain are recorded in the archive. The 208 bundled project modules contain no occurrence of `implemented_by`, `extern`, `csimp`, or `unsafe`. This is a scoped source check, not an absence claim for the transitive runtime: Lean's standard array operations, for example, have external implementations. The compiler, its rewrites, and runtime remain within the stated native-evaluation trust boundary.

For the heptagon closure, the source decoder checks all 190,932 retained clauses against proved schemas. One normalized configuration supplies the state owners, the common chronological order, all actual ambient ranks, the induced-core ranks, and one shared terminal five-cup. A source classifier proposes schema identifications but is untrusted: Lean checks literal decoding, clause equality, schema validity, and complete coverage before applying the refutation.

For octagons, both omission trees and all their capacity rows are connected to the same geometric occupancy assignment. In particular, the 973 historically transported residual supports now have direct cyclic proofs. The block rows use an audited port of Baek's own Lean formalization, with its Apache-2.0 attribution retained. The port's combinatorial boundary predicate is connected to the real-plane convex-hull predicate by an explicit theorem.

8.2. **Lessons from formalization.** The completed proofs make three distinctions especially useful for further computations.

Actual and inherited ranks. A longest-path value has an attaining witness in a specified set. Restriction can destroy that witness. The heptagon proof therefore recomputes core recurrences, while keeping ambient ranks separately. The octagon proof deliberately uses the weaker admissible-label interface, which is hereditary and requires no new attainment after transport.

Local validity and shared existence. Individually feasible witnesses need not coexist. The heptagon source uses one interpretation of every owner, sign, order, and path flag. Prefix coverage is proved before exclusions are applied; no preferred order of incomparable states is imposed. In the octagon proof, triplewise cyclic certification survives missing endpoints and therefore justifies multi-omission rows.

Projective geometry and cardinality. The formal projective map is applied only on a positive-denominator chart. Convex combinations are transported with renormalized positive weights, not unchanged affine coefficients. A finite-exception shear first separates abscissae, sorting is bijective, and the deleted root is treated separately. The three-chain exception $q \neq t$ is discharged by actual membership in the two parts of the partition. These steps establish both the ordinary planar meaning and the two-point adjustment.

8.3. **Finite data and reconstruction.** The heptagon refutations use reverse unit propagation: assuming the negation of a proposed clause, the listed earlier clauses must force a conflict. Soundness and the final empty clause are checked in Lean. The main proof sizes are shown below; the restricted-25 certificate has its own primitive, replay, and coverage stages.

Obligation	Retained source clauses	Proof additions
Restricted hole coverage	93,169	381,647
Through-cup population ≤ 13	5,077	3,009
Terminal core ≤ 19	25,074	79,173
Joint endpoint population	59,479	102,563
Balanced fifty-point exclusion	190,932	1,989,465

The coverage formula has 461 variables. The octagon certificates have the following sizes. Each branch covers both Boolean values, and every weighted leaf has a strict integer contradiction in its own context.

Obligation	States	Rows	Branches	Leaves
Balanced bound 193	252	2,685	1,590	1,591
Residual bound 124	154	1,187	7,314	7,315

The archive contains 208 project modules, including embedded certificates and the Baek port. It pins Lean 4.32.0-rc1 and a specific mathlib commit. The complete library was rebuilt from source without precompiled project modules in 1,290 seconds on an ARM64 macOS host; peak memory was not recorded. This release preserves those proof sources byte for byte and includes a separately compiled statement-equivalence and parser-guard module. The latter proves that a nonempty numeric source row cannot parse as an empty clause; all five literal sources pass the independent check of that premise. The accompanying audit distinguishes the full build from later source-identical extraction checks. From the archive root, with the pinned mathlib checkout and its dependencies already built, run

```
python3 tools/verify_release.py --mathlib /path/to/mathlib
```

This verifies file identities, rebuilds the project in a fresh directory, and checks both final declarations and their exact allowed axiom lists. The archive README gives the dependency setup and PDF build commands. Certificate discovery is not needed for replay. Independent checks are

```
python3 tools/check_octagon_finite.py
python3 tools/check_heptagon_replays.py
python3 tools/check_examples_and_counts.py
python3 tools/check_residual_lower.py
```

The first command reconstructs both octagon state universes and all directed differences, certifies all 3,849 cyclic families and 23 block maps, compares the JSON and Lean data, and replays all 8,906 leaves with complete branch coverage. This finite audit runs without Lean or native evaluation; it uses the geometric lemmas proved above. It is an independent executable check, not a change to the Lean theorem's dependency list.

The second command compiles a separate C++ RUP checker. All four population/closure proofs pass unchanged. The coverage proof contains one satisfied, redundant hint (473786 in step 473998); the command reproduces its strict rejection and then replays with exactly that hint omitted. All clauses and other hints remain unchanged, as do the sealed Lean files. The final coverage replay has 381,647 additions; the closure replay has 1,989,465. This checker is not itself formally verified. A replay through an external verified LRAT implementation and a second-platform full build remain further audits. The third command checks the ten-state example, exact 25-point control, and comparison arithmetic. The fourth checks the full residual predicate and every seven-point subset of the 26-point lower example in Section 9. The capacity, strict-replay, and residual-control checkers include rejection controls.

8.4. **Archive identification.** The companion bundle is

`es7_es8_bounds_81_319_v1.zip`.

The SHA-256 below identifies the unchanged formal manifest. Its path in the bundle is

`formal/MANIFEST.json`.

`81d3fb7e1d49ce780cd0ddb69506f4833c9a507c550dc5d57168ac2132e036c1`

The bundle includes the complete paper sources, proof sources, certificates, checker commands, and hash manifest. The paper and companion archive share the DOI 10.5281/zenodo.22842891.

9. Further directions

Either $D_7 \le 29$ or $C_7 \le 48$ would give $\mathrm{ES}(7) \le 80$. There are, however, unconditional limits to improving the two capacities separately. The classical restricted constructions [7, 8] give $C_7 \ge 30$, $C_8 \ge 62$, and $M_8(6,6) \ge 50$. A set counted by $M_8(6,6)$ has no six-cup, hence no residual junction, so $D_8 \ge 50$. The archive also supplies an exact 26-point residual set. Its five-cup endpoints and four-cap starting points are disjoint; the checker verifies this condition and excludes a heptagon in all $\binom{26}{7} = 657\,800$ seven-point subsets. Thus $D_7 \ge 26$. Consequently,

$$C_7 + D_7 + 2 \ge 58, \qquad C_8 + D_8 + 2 \ge 114.$$

These are floors of the independently optimized additive bound (1.1), not lower bounds on ES(7) or ES(8). Of the 48-point gap from 81 to the conjectural 33, at most $81 - 58 = 23$ points can be recovered by improving the separate capacities; at least $58 - 33 = 25$ require their joint compatibility or another argument.

The conjecture of Erdős, Tuza, and Valtr would give:

Quantity	Known lower bound	Proved upper bound	Under ETV
C_7	≥ 30	≤ 49	$= 30$
D_7	≥ 26	≤ 30	$= 26$
C_8	≥ 62	≤ 193	$= 62$
D_8	≥ 50	≤ 124	≤ 57

Its proposed value for $M_n(a,b)$ is $\sum_{i=n-b}^{a-2} \binom{n-2}{i}$ [8, 1]. For residual upper bounds we use $D_n \leq M_n(n-2,n)$; the equality for D_7 additionally uses the checked lower example. Thus ETV would make $C_7 + D_7 + 2 = 58$ and place $C_8 + D_8 + 2$ in $[114, 121]$. The latter interval is conditional; only its lower endpoint is established unconditionally here.

The additive bound discards compatibility between the two parts. Retaining their shared witnesses, or bounding the asymmetric class Q_n directly, may overcome these limits. For octagons, admissible state differences also offer a way to certify further hereditary capacities. These are open tasks; no incomplete search enters the results above.

Acknowledgments. This work benefited from research assistance by OpenAI's AI systems, including proof exploration, exact computational checks, and manuscript preparation. All responsibility for the content and conclusions of this work rests with the author.

Appendix A. Necessary CNF semantics

This appendix specifies the source contract for the three population lemmas and the balanced exclusion. The formal source decoders match each stored clause to a proved necessary schema. Catalogue order indexes variables; actual owner order is represented separately.

For a fixed chronological selection, C_{ijk} means that (i,j,k) is a cup. For named state owners, the catalogue-sorted key identifies one variable, but its truth value means the orientation in the *actual chronological order.* Order guards specify that order whenever it matters. Literal negation gives the cap case.

Family	Truthful interpretation and reason for necessity
Owner activation	An owner exists exactly when its state is occupied. Prefix counters implement exact population or threshold counts; their catalogue order does not constrain geometric chronology.
Order	Comparable states have forced order by lemma 5.1. Free pair variables and transitivity retain all orders of incomparable states. Inactive potential owners can be inserted into a linear extension.
Pair labels	If both owners exist in the indicated order, exactly one label in the destination-minus-source state difference is selected. Its value is the actual ambient incoming pair rank. Otherwise all those pair variables are false.
Rank growth	A cup triple increases the cup coordinate; a cap triple increases the cap coordinate. Disallowed label pairs therefore force the opposite color.
Selected paths	A flag records an actual path with the stated first vertex and final pair. Base and extension implications cover every actual selected path. Maximum clauses exclude six-chains, and endpoint summaries exclude opposite paths of total length nine.
Ambient outgoing ranks	Threshold flags represent actual outgoing pair maxima. Nesting, growth, and incoming/outgoing concatenation budgets follow from their actual paths. Selected attainment is not imposed on ranks whose witnesses may be outside the selection.
Exact core paths	A core path threshold is the disjunction of shared successor or predecessor witnesses, each conjoining a path threshold and its joining color. Both directions are encoded only when the whole core is retained.
Endpoint extensions	An actual ending five-cup extends every outgoing cap one step to the left, or a six-cup would occur. The dual rule holds for starters and for the other color.
Packets	For each chronological quadruple, the signs in lexicographic triple order have at most one change. Clauses forbid alternating three-term subsequences of that four-sign packet. All selected owner activations and necessary order comparisons guard these clauses.

Family	Truthful interpretation and reason for necessity
First-pair restrictions	Pairs starting at the ambient first vertex have coordinates $(1,1)$. Pairs starting at the second have both coordinates at most two. Those vertices are retained explicitly.
Prefix attainment	For a retained initial ambient point i, $d_c(i,j) = \max(\{1\} \cup \{d_c(h,i)+1 : h < i,\ C(h,i,j) = c\})$. Every such h is present, so reverse attainment is valid there. All clauses are guarded by the chosen prefix alternative.

A.1. **Outgoing and endpoint schemas.** Use the guarded pair-label variables G_{ij} and $L_{ij,z}$ defined in Appendix B. In particular,

$$G_{ij} \Rightarrow \bigvee_{z \in F_j \setminus F_i} L_{ij,z}, \qquad L_{ij,z} \wedge L_{ij,z'} \Rightarrow z = z'.$$

For active chronological owners $i < j < t$, let $U_k^c(i,j)$ assert an actual ambient c-chain of at least k vertices starting with (i,j). With $c = 1$ for cup and $c = 0$ for cap, write z_c for the corresponding coordinate. The schemas include

$$U_{k+1}^c(i,j) \Rightarrow U_k^c(i,j),$$
$$\chi_{ijt}^c \wedge U_k^c(j,t) \Rightarrow U_{k+1}^c(i,j),$$
$$L_{ij,z} \wedge \chi_{ijt}^c \wedge U_k^c(j,t) \Rightarrow z_c + k \leq 5.$$

The last implication is the no-six-chain budget. Every encoding includes its activation and order guards. No converse recurrence is imposed when the next witness may lie outside the selected owners.

If E_i^1 means that an actual five-cup ends at i, and $\mathrm{Path}^0(i,j,\ldots,t)$ is any ambient cap, then

$$E_i^1 \wedge \mathrm{Path}^0(i,j,\ldots,t) \Rightarrow \exists h < i : \mathrm{Path}^0(h,i,j,\ldots,t).$$

Choose h to be the penultimate point of the same attaining five-cup. The triple (h,i,j) must be a cap, or appending j to that cup would give six points. This extends the cap while preserving its last pair. Reversal gives the starter extension preserving the first pair; color exchange gives both dual rules. Witnesses remain in their specified ambient set, not necessarily in the selected core.

For completeness, the two-color vertex budget follows from a shared incoming pair attaining a cell $(x,y) \in F(v)$. If an outgoing pair (v,w) had cup and cap lengths at least $6-x$ and $6-y$, respectively, the joining triple would concatenate same-color paths into a six-chain. Thus those two outgoing thresholds cannot both hold. A single incoming pair dominates both coordinates of the cell; this step does not choose two unrelated predecessors.

In the balanced equality model, low-width members of K also start five-caps. Every cup ending at such a point therefore extends rightward by one while preserving its first pair, which supplies the low-endpoint extension clauses. Internal five-cup sources are defined from exact core outgoing paths, and the source counter requires at least six by corollary 5.5. The terminal witness consists of an actual five-cup from the last core point, so its four later points have a fixed relative position without restricting any other ambient witness.

The initial four packet alternatives cover all geometric possibilities after reflection. The later five- and six-point lists are generated exhaustively, retaining all packet-valid alternatives with the stated prefix and chain limits. The mixed sectors are still represented in the primary formula. Their removal occurs only in the final checked derivation. Thus neither an assumed preferred order nor an unproved prefix exclusion enters the primary source.

Appendix B. Restricted-25 certificate interface

The restricted source uses hole variables, actual owner-order variables, and pair-label variables. A label simultaneously records the existence and order of its owners and its two actual ranks. Each maximal corner of an incoming state has one incoming pair attaining it. Backward witnesses bind that predecessor pair and the joining triple together; their consequences do not select unrelated witnesses.

The primitive layer proves label choice, uniqueness, rank growth, path extension, endpoint restrictions, four-point sign constraints, and common-endpoint polygon exclusions. All clauses are guarded by the existence and required order of their owners. A verified reverse unit-propagation layer derives the hole clauses. Finally, the LRAT refutation of the 93,169-clause coverage CNF rejects the nine-hole case. Its 91,720 hole clauses are selected from the proved primitive replay; the remaining clauses encode cardinality and padding. The two exhaustive enumerations of the eligible hole sets in theorem 5.2 are separate computational evidence. The geometric assignment, primitive soundness, logical replay, and coverage are separate formal obligations. The catalogue's literal and color conventions are decoded explicitly.

B.1. **Witness equations.** Here are the shared-witness relations underlying the compressed terms above. Write A_i for occupation of state i, O_{ij} for the order of its owner before that of j. Let $L_{ij,z}$ assert that the pair's actual incoming label is $z \in F_j \setminus F_i$. Put

$$G_{ij} = A_i \wedge A_j \wedge O_{ij}.$$

Then

$$L_{ij,z} \Rightarrow G_{ij}, \qquad G_{ij} \Rightarrow \bigvee_{z \in F_j \setminus F_i} L_{ij,z}, \qquad \neg L_{ij,z} \vee \neg L_{ij,z'} \quad (z \neq z').$$

For a maximal corner z of F_j, attainment says

$$\neg A_j \vee \bigvee_{i:\, z \notin F_i} L_{ij,z}.$$

Let χ^c_{ijk} mean that the chronologically ordered triple has color c, with $c = 1$ for cup. If $z_c \geq z'_c$, growth gives

$$\neg L_{ij,z} \vee \neg L_{jk,z'} \vee \neg \chi^c_{ijk}.$$

To encode reverse attainment, let $W^c_{tij,r}$ mean that the *same* three active owners satisfy $t < i < j$, χ^c_{tij}, and $d_c(t,i) + 1 = r$. For $z_c \geq 2$ the clauses are

$$\begin{aligned} L_{ij,z} &\Rightarrow \bigvee_t W^c_{tij,z_c}, \\ W^c_{tij,r} &\Rightarrow \chi^c_{tij}, \\ W^c_{tij,r} &\Rightarrow \bigvee_{y \in F_i \setminus F_t:\, y_c + 1 = r} L_{ti,y}. \end{aligned}$$

Candidate sums range only over distinct owners and permitted labels. The corner and backward-witness schemas are checked in

`ES7ExactClauses.expected`.

For a retained induced core K, $T^c_k(i,j)$ means that a c-chain of at least k vertices starts with the exact pair (i,j) in K. Its base is $T^c_2(i,j) = G_{ij}$ for core owners, and for $k \geq 3$,

$$T^c_k(i,j) \;\longleftrightarrow\; \bigvee_{t \in K} \big(A_i \wedge O_{ij} \wedge \chi^c_{ijt} \wedge T^c_{k-1}(j,t)\big).$$

Each conjunction has its own witness variable. Encoding both directions of this equivalence makes these *exact core outgoing paths*, while ambient path thresholds may only have one-way implications after selection. The internal-source flag is $S_i \leftrightarrow \bigvee_{j \in K} T^1_5(i,j)$ and the balanced source imposes $\sum_{i \in K} S_i \geq 6$. The recurrence is proved in `ES7ClosureCoreSemantics`.

Finally, for the four chronological packet entries e_1, e_2, e_3, e_4, every $a < b < c$ imposes

$$(\neg e_a \vee e_b \vee \neg e_c) \;\wedge\; (e_a \vee \neg e_b \vee e_c).$$

It forbids 101 and 010 subsequences. Named-owner versions prepend the negations of the required activation and order guards. A disjunction $v \leftrightarrow \bigvee_i w_i$ is encoded by $(\neg v \vee \bigvee_i w_i)$ and $(\neg w_i \vee v)$; a conjunction is encoded dually. Empty disjunctions are false. These formulas specify the witness coupling. The archive remains the authoritative specification of literal numbering, omitted constant variables, and the exact compact clause list; Lean checks that list against these semantics before replay.

Appendix C. Final statements and their trust boundary

The final types below are the output of Lean's `#print`, with line wrapping only and proof terms omitted. The type of `Point` is $\mathbb{R} \times \mathbb{R}$.

```
theorem ES7FinalVerified.es7_le_81 :
  ∀ (S : Finset ES8Geometry.Point),
  ES7Final.GeneralPosition S → 81 ≤ S.card →
  ES7Geometry.HasConvexHeptagon ↑S

theorem ES8Final.es8_le_319 :
  ∀ (S : Finset ES8Geometry.Point),
  ES8Final.GeneralPosition S → 319 ≤ S.card →
  ES8Geometry.HasConvexOctagon ↑S
```

Both general-position definitions unfold to the following expression:

```
fun S => ∀ a ∈ S, ∀ b ∈ S, ∀ c ∈ S,
  a ≠ b → b ≠ c → a ≠ c → ES8Geometry.orient a b c ≠ 0
```

Here `orient p q r` is $(q_x - p_x)(r_y - p_y) - (q_y - p_y)(r_x - p_x)$. The two polygon predicates unfold respectively to

```
fun S => ∃ T, T.card = 7 ∧ ↑T ⊆ S ∧ ES8Geometry.ConvexPosition ↑T
fun S => ∃ T, T.card = 8 ∧ ↑T ⊆ S ∧ ES8Geometry.ConvexPosition ↑T
```

The underlying definition and the new bridge are

```
def ConvexPosition (S : Set Point) : Prop :=
  ∀ p ∈ S, p ∉ convexHull ℝ (S \ {p})

theorem convexPosition_iff_convexIndependent (S : Set Point) :
  ConvexPosition S ↔ ConvexIndependent ℝ ((↑) : S → Point) := by
  exact convexIndependent_set_iff_notMem_convexHull_sdiff.symm
```

The bridge uses only the three standard axioms listed below. The companion module applies it to both endpoints. Thus their convexity conclusions can be stated using mathlib's standard predicate without strengthening a hypothesis or trusting another finite evaluation.

C.1. The restricted theorem and its chain predicates.

The restricted upper bound uses the following declaration and definitions. The upper bound is formalized; the lower bound is supplied by the cited construction and the independently checked 25-point example.

```
theorem ES7CapacityStage.restricted25 :
  ES7Residual.RestrictedBound 25
```

```
def ES7Residual.RestrictedBound (K : ℕ) : Prop :=
  ∀ {n} (c : ES8BalancedGeometry.Configuration n),
    (¬(colorConfig c.cup true).HasNCap 5 (Finset.range n)) →
    (¬(colorConfig c.cup true).HasNCup 6 (Finset.range n)) →
    (¬HasConvexHeptagon (Set.range c.point)) → n ≤ K
```

```
structure ES8BalancedGeometry.Configuration (n : ℕ) where
  point : Fin n → Point
  xOrder : StrictMono (fun i => (point i).1)
  generalPosition : ∀ u v w : Fin n,
    u < v → v < w → orient (point u) (point v) (point w) ≠ 0
```

For valid indices, `c.cup u v w` is true exactly when the orientation determinant is positive; out-of-range indices return false. The notation `colorConfig` refers to `ES8ActualRanks.colorConfig`:

```
def colorConfig (cup : ℕ → ℕ → ℕ → Bool) (color : Bool) : Config
     ℕ :=
  { Cup3 := fun a b c => cup a b c = color,
    DecidableCup3 := inferInstance }
```

In the Baek port's `Config` namespace, with a linearly ordered vertex type α and `C : Config` α, the predicates are

```
def Cap3 (a b c : α) : Prop := ¬C.Cup3 a b c
def Cap (l : List α) : Prop :=
  l.IsChain (· < ·) ∧ l.Chain3' C.Cap3
def Cup (l : List α) : Prop :=
  l.IsChain (· < ·) ∧ l.Chain3' C.Cup3
def NCap (n : ℕ) (l : List α) : Prop :=
  C.Cap l ∧ l.length = n
def NCup (n : ℕ) (l : List α) : Prop :=
  C.Cup l ∧ l.length = n
def HasNCap (n : ℕ) (S : Finset α) : Prop :=
  ∃ l : List α, C.NCap n l ∧ l.In S
def HasNCup (n : ℕ) (S : Finset α) : Prop :=
  ∃ l : List α, C.NCup n l ∧ l.In S
```

Here `IsChain` tests consecutive order comparisons, `Chain3'` tests consecutive triples, and `l.In S` means that every listed vertex lies in S. Strict order ensures distinct vertices. For actual points these are precisely the slope definitions of cups and caps in Section 2. The archived statement audit prints the fully qualified definitions directly from Lean.

C.2. **Exact final axiom reports.** The following is the complete `#print axioms` output for the two original endpoints. The 19 and eight names ending in `native_decide.ax_1_1` are the disclosed native checks.

```
'ES7FinalVerified.es7_le_81' depends on axioms: [propext,
 Classical.choice,
 Quot.sound,
 ES7ClosureData.accepted._native.native_decide.ax_1_1,
 ES7ClosureLiteral.closure_accepted._native.native_decide.ax_1_1,
 ES7CoverageData.accepted._native.native_decide.ax_1_1,
 ES7CoverageLiteral.accepted._native.native_decide.ax_1_1,
 ES7ExactData.accepted._native.native_decide.ax_1_1,
 ES7Joint17Data.accepted._native.native_decide.ax_1_1,
 ES7Joint17Literal.joint17_accepted._native.native_decide.ax_1_1,
 ES7JointCatalogProof.catalog_checked._native.native_decide.ax_1_1,
 ES7OrderedData.accepted._native.native_decide.ax_1_1,
 ES7PrefixCoverage.extended_coverage._native.native_decide.ax_1_1,
 ES7PrefixData.checked_patterns._native.native_decide.ax_1_1,
 ES7PrefixData.pattern_records._native.native_decide.ax_1_1,
 ES7ReplayData.replay_accepted._native.native_decide.ax_1_1,
 ES7SeedData.accepted._native.native_decide.ax_1_1,
 ES7SourceCatalog.hole_map_checked._native.native_decide.ax_1_1,
 ES7Terminal20Data.accepted._native.native_decide.ax_1_1,
 ES7Terminal20Literal.terminal20_accepted._native.native_decide.
   ax_1_1,
 ES7Through14Data.accepted._native.native_decide.ax_1_1,
 ES7Through14Literal.accepted._native.native_decide.ax_1_1]
'ES8Final.es8_le_319' depends on axioms: [propext,
 Classical.choice,
 Quot.sound,
 ES8BlockValidation.blocks_checked._native.native_decide.ax_1_1,
 ES8BlockValidation.coverage_checked._native.native_decide.ax_1_1,
 ES8CycleValidation.cycles_checked._native.native_decide.ax_1_1,
 ES8CycleValidation.table_checked._native.native_decide.ax_1_1,
 ES8ResidualCycleValidation.cycles_checked._native.native_decide.
   ax_1_1,
 ES8ResidualCycleValidation.table_checked._native.native_decide.
   ax_1_1,
 ES8.Certificates.Native.balanced_accepted._native.native_decide.
   ax_1_1,
 ES8.Certificates.Native.residual_accepted._native.native_decide.
   ax_1_1]
```

References

[1] J. Baek, On the Erdős–Tuza–Valtr conjecture, *European Journal of Combinatorics* **124** (2025), 104085. doi:10.1016/j.ejc.2024.104085.

[2] J. Baek and M. Balko, The Erdős–Szekeres conjecture revisited, *Journal of Combinatorial Theory, Series A* **222** (2026), 106195. doi:10.1016/j.jcta.2026.106195. Conference version: SoCG 2025, LIPIcs 332, 13:1–13:15.

[3] M. Balko and P. Valtr, A SAT attack on the Erdős–Szekeres conjecture, *European Journal of Combinatorics* **66** (2017), 13–23. doi:10.1016/j.ejc.2017.06.010.

[4] G. Damásdi, Z. Dong, M. Scheucher, and J. Zeng, Saturation results around the Erdős–Szekeres problem, *European Journal of Combinatorics* **131** (2026), 104236. doi:10.1016/j.ejc.2025.104236.

[5] B. Dumitru, Notes on the 33-point Erdős–Szekeres problem, arXiv:2512.24061v1 (2025).

[6] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Mathematica* **2** (1935), 463–470.

[7] P. Erdős and G. Szekeres, On some extremum problems in elementary geometry, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **3–4** (1960–1961), 53–62.

[8] P. Erdős, Z. Tuza, and P. Valtr, Ramsey-remainder, *European Journal of Combinatorics* **17** (1996), 519–532. doi:10.1006/eujc.1996.0045.

[9] A. F. Holmsen, H. Nassajian Mojarrad, J. Pach, and G. Tardos, Two extensions of the Erdős–Szekeres problem, *Journal of the European Mathematical Society* **22** (2020), 3981–3995. arXiv:1710.11415.

[10] V. Koshelev and A. Koshka, Combinatorial geometry of Erdős–Szekeres type problems: SAT/ASP modeling and linear subreduction, arXiv:2604.20120v1 (2026).

[11] F. Marić, Fast formal proof of the Erdős–Szekeres conjecture for convex polygons with at most 6 points, *Journal of Automated Reasoning* **62** (2019), 301–329. doi:10.1007/s10817-017-9423-7.

[12] G. Moshkovitz and A. Shapira, Ramsey theory, integer partitions and a new proof of the Erdős–Szekeres theorem, *Advances in Mathematics* **262** (2014), 1107–1129.

[13] H. Nassajian Mojarrad and G. Vlachos, An improved upper bound for the Erdős–Szekeres conjecture, *Discrete & Computational Geometry* **56** (2016), 165–180. doi:10.1007/s00454-016-9791-5.

[14] S. Norin and Y. Yuditsky, Erdős–Szekeres without induction, *Discrete & Computational Geometry* **55** (2016), 963–971. arXiv:1509.03332.

[15] B. Subercaseaux, W. Nawrocki, J. Gallicchio, C. Codel, M. Carneiro, and M. J. H. Heule, Formal verification of the empty hexagon number, in *ITP 2024*, LIPIcs **309** (2024), 35:1–35:19. doi:10.4230/LIPIcs.ITP.2024.35.

[16] A. Suk, On the Erdős–Szekeres convex polygon problem, *Journal of the American Mathematical Society* **30** (2017), 1047–1053. arXiv:1604.08657.

[17] G. Szekeres and L. Peters, Computer solution to the 17-point Erdős–Szekeres problem, *ANZIAM Journal* **48** (2006), 151–164. doi:10.1017/S144618110000300X.

[18] G. Tóth and P. Valtr, Note on the Erdős–Szekeres theorem, *Discrete & Computational Geometry* **19** (1998), 457–459. doi:10.1007/PL00009363.

[19] G. Tóth and P. Valtr, The Erdős–Szekeres theorem: upper bounds and related results, in *Combinatorial and Computational Geometry*, MSRI Publications **52**, Cambridge University Press, 2005, 557–568.

Independent Researcher