Preliminary Observations on the Happy Ending Problem

Warren Jin

July 2023

1 Introduction and Literature Review

Given N points in the plane, we wish to check whether there must exist a convex polygon with n vertices. The lowest number of points required is denoted $k(n)$, such that $\exists$ convex n-gon $\forall N \geq k(n)$. For simplicity, we can assume that no 3 points are collinear. This problem was posed by Erdős and Szekeres in this 1935 paper, where they conjectured that $k(n) = 2^{n-2} + 1$ (which I will refer to as the ES conjecture).[1] They managed to prove that $k(n) \geq 2^{n-2} + 1$ in a later paper.[2] A related problem is to find the infimum of the number of convex n-gons in a set of N points, which I will denote by $f(N, n)$. The case $f(N, 4)$ was considered by Hu, Chen, and Zhu, who found exact values for $N \leq 9$ and a lower bound for $N > 9$.[3] I present an alternative method to find the lower bound of $f(N, 4)$ that coheres with Hu et al's finding on the growth rate of the function, and generalize it to any choice of n.

2 Some observations

For this article I will introduce some notation. Noting that if the power set of N points is taken to be a poset $(P, \subseteq)$, then every subset of the same number of points is pairwise incomparable. Thus, the number of elements on height H equals $\binom{N}{H}$. Evidently, the height of this poset is N. We can also see that P is a lattice, since the joins and meets for any two arbitrary elements are their union and intersection respectively, which are both elements of the power set.

Let x_i be an element of P with i elements. Define a function $A : P \to \mathbb{R}$, where $A(x)$ is the supremum of the possible areas enclosed by a cycle over the elements of x. We let $A(x_1) = A(x_2) = 0$. A cycle which gives this supremum does not have any edge crossings, because if the edges v_1v_2 and v_3v_4 crossed, then either v_1v_3 and v_2v_4 form a non-crossing cycle of larger area, or v_1v_4 and

[1] `http://www.numdam.org/item/CM_1935__2__463_0.pdf`

[2] `https://www.renyi.hu/~p_erdos/1960-09.pdf`

[3] `http://archive.ymsc.tsinghua.edu.cn/pacm_download/21/86-2010On_the_Minimum_Number_of_Convex_Quadrilaterals_in_Point_Sets_of_Given_Numbers_of_Points.pdf`

v_2v_3 will achieve this. This process can be iterated as many times as necessary to obtain various cycles over x, from which it is possible to select one with the largest area.

Notice that for any polygon X, any other polygon with vertices chosen from X must also be convex and has a smaller area than X if and only if X is convex. We can then define a poset $(Q, \geq)$ such that $A(x) \in Q$ if $x \in P$. Clearly Q is a subset of $\mathbb{R}$ and is thus a total order, but right now I am only interested in the relations in Q of corresponding comparable elements in P, which is to say that we only need to consider $(A(x), A(y))$ for $x \subseteq y$ or vice versa. By the observation above, consider a subposet of P formed by taking some element at height i and all the chains up to height i with this element as their join. The i-sided polygon corresponding to this element is convex if and only if the aforementioned subposet is isomorphic to the subposet of Q formed by applying A to every element. For an element x_i exhibiting these properties, we call it a convex element of P.

From this we have a restatement of the happy ending problem: given the posets P and Q, as well as the function A, which itself has many properties that are still unknown to me, we want to find an integer $k(n)$ such that a convex element x_n must exist in P if the height of $P \geq k(n)$.

3 Lower bound for $f(N, n)$

Consider the poset P for a given power set of N elements. In general, we only consider cases where $N \geq k(n)$. Take the elements on the heights $H = N, k(n)$, and n. For each height, the number of elements would be $\binom{1,N}{k(n),}$ and $\binom{N}{n}$ respectively. The number of elements on $H = k(n)$ that some element on $H = n$ is a subset of is equal to the number of ways that one can choose another $k(n) - n$ points out of $N - n$ available ones.

Now consider the number of elements on $H = n$ which must be convex. Since every $k(n)$-element subset must be comparable to a convex element on $H = n$, and every such element is comparable to exactly $\binom{N-n}{k(n)-n}$ unique $k(n)$-element subsets, then minimally we would need

$$\left\lceil \binom{N}{k(n)} \div \binom{N-n}{k(n)-n} \right\rceil = \left\lceil \frac{N!(k(n)-n)!}{k(n)!(N-n)!} \right\rceil$$

convex elements on $H = n$ to satisfy the requirements.

By the result in the original paper where $k(4) = 5$, we have

$$f(N, 4) = \left\lceil \frac{N!}{120(N-4)!} \right\rceil$$

The growth rate of $f(N, 4)$ by this calculation is given as such:

$$\frac{f(N, 4)}{f(N-1, 4)} \approx \frac{N}{N-4}$$

for large enough N. This is the same result found in Hu et al. However, this formula gives $f(7,4) = 7$, which is worse than the actual bound $f(7,4) = 9$. The disparity is due to the fact that the minimal number of convex elements required does not match the actual graph in the plane, as the additional point added from N to $N+1$ creates more convex elements than expected by our method. Additionally, the possible graphs of 7 points containing only 9 convex quadrilaterals must contain a convex pentagon.

Regardless of the actual value of $f(N,4)$, the best bounds currently are $O(N^4)$. At $H = 5$, there are $\binom{N}{5}$ elements, which is $O(N^5)$. Given that the number of (x_4, x_5) pairs in P is $(N-4)f(N,4) = O(N^5)$, it is not possible to use the pigeonhole principle alone to prove the happy ending problem for case $n = 5$. A similar argument applies for every n.

4 Some sketchy guesses

Assuming that $k(n)$ and $f(N,n)$ are accurately formulated:

1. For some set of $k(n) - 1$ points which does not contain a convex polygon with n sides, the convex hull of the entire set is an $n-1$ sided polygon. This is a **very** sketchy guess.
2. For $N \geq k(n)$, a set of N points which does not contain a convex polygon with n sides contains more than $f(N,n)$ convex $n-1$ sided polygons.
3. $f(N,n)$ is $O(N^n)$.