

On Convex Polygons in Cartesian Products*

Jean-Lou De Carufel¹, Adrian Dumitrescu², Wouter Meulemans³,
Tim Ophelders³, Claire Pennarun⁴, Csaba D. Tóth⁵, and
Sander Verdonschot⁶

- 1 University of Ottawa, Canada
jdecaruf@uottawa.ca
- 2 University of Wisconsin-Milwaukee, USA
dumitres@uwm.edu
- 3 TU Eindhoven, The Netherlands
{w.meulemans|t.a.e.ophelders}@tue.nl
- 4 LIRMM, CNRS & Université de Montpellier, France
claire.pennarun@lirmm.fr
- 5 California State University Northridge, Los Angeles, CA, USA
csaba.toth@csun.edu
- 6 Carleton University, Ottawa, Canada
sander@cg.scs.carleton.ca

Abstract

We study several problems concerning convex polygons whose vertices lie on a grid defined by the Cartesian product of two sets of n real numbers, using each coordinate at most once. First, we prove that all such grids contain a convex polygon with $\Omega(\log n)$ vertices and that this bound is asymptotically tight. Second, we present two polynomial-time algorithms that find the largest convex polygon of a restricted type. These algorithms give an approximation of the unrestricted case. It is unknown whether the unrestricted problem can be solved in polynomial time.

1 Introduction

A fast way to generate a random convex polygon, based on a proof by Pavel Valtr [7], first generates two random sets of n integer coordinates before significantly transforming the y -coordinates to produce a convex n -gon with the original x -coordinates. What happens if we do not transform the y -coordinates, and instead ask for a convex polygon with the original x - and y -coordinates?

Formally, we say that two sets, X and Y , each containing n real numbers, form a *grid* $X \times Y$. A grid *supports* a convex polygon P if for every vertex of P , its x -coordinate is in X and its y -coordinate is in Y , and no two vertices of P share an x - or y -coordinate.

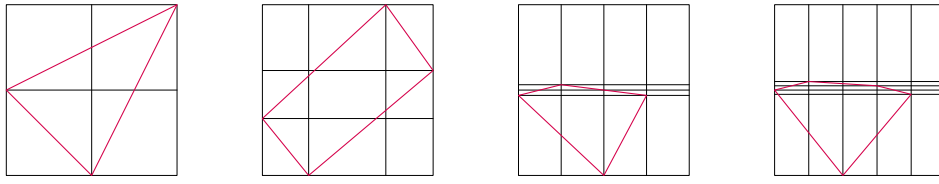
It turns out that not every $n \times n$ grid supports a convex n -gon. In fact, this is true already for $n = 5$ (see Figure 1). This raises several interesting questions. Can we quickly decide whether a grid supports a convex n -gon? Or can we find the largest k such that it supports a convex k -gon? And what is the largest k such that *any* $n \times n$ grid supports a convex k -gon? We initiate the study of these questions.

There is a rich history of problems involving convex subsets, including the famous Happy Ending Problem: that any set of five points in the plane in general position contain four

* This work was initiated at the 2017 Fields Workshop on Discrete and Computational Geometry. Meulemans is partially supported by the Netherlands eScience Center (NLeSC, grant 027.015.G02). Ophelders is supported by the Netherlands Organisation for Scientific Research (NWO) under project no. 639.023.208. Research by Tóth was partially supported by the NSF awards CCF-1422311 and CCF-1423615.

34th European Workshop on Computational Geometry, Berlin, Germany, March 21–23, 2018. This is an extended abstract of a presentation given at EuroCG'18. It has been made public for the benefit of the community and should be considered a preprint rather than a formally reviewed paper. Thus, this work is expected to appear eventually in more final form at a conference with formal proceedings and/or in a journal.

39:2 On Convex Polygons in Cartesian Products



■ **Figure 1** Maximum-size supported convex polygons of respective sizes 3, 4, 4, and 5 in $n \times n$ grids, where n is between 3 and 6.

points in convex position. Generalizing this result, Erdős and Szekeres conjectured that every set of $2^{n-2} + 1$ points in general position contains n points in convex position and that this is tight [2]. While this conjecture has been proven only for $n \leq 6$ [6], and the current best upper bound is $2^{n+o(n)}$ [5], the asymptotics are known to be correct for the lower bound: a set of n points in general position always contains $\Omega(\log n)$ points in convex position [2].

Algorithmically, the problem of finding the largest convex subset of a set of n points in the plane in general position can be solved in $O(n^3)$ time [1]. While this approach likely generalizes to finding the largest convex subset in a grid, it is not clear how to include the restriction that each coordinate is used at most once. On the negative side, it was recently shown that the problem of finding the largest convex subset in a point set in \mathbb{R}^d for dimensions $d \geq 3$ is NP-hard [3].

The remainder of the paper is structured as follows. First, in Section 2, we give an asymptotically tight lower bound on the maximum size convex polygon supported by an $n \times n$ grid. Then, in Section 3, we provide algorithms to find, for a given grid, the largest supported convex polygon of two special types. These algorithms give a constant-factor approximation of the size of the largest supported convex polygon.

2 General bounds

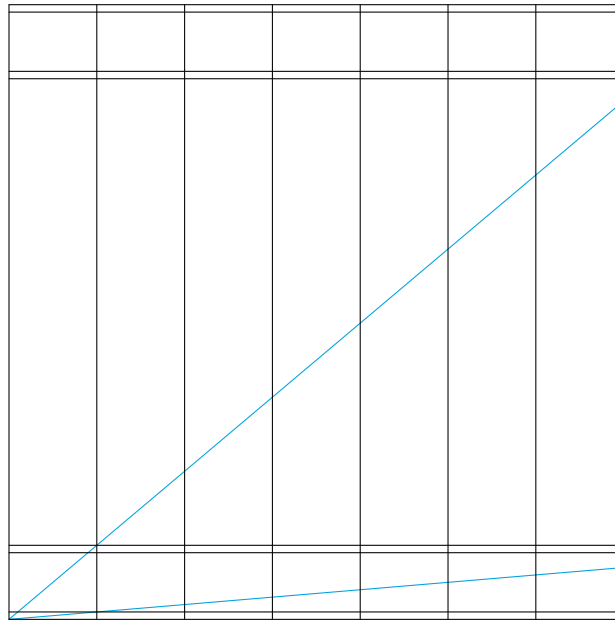
We consider two special types of convex polygons. We classify a convex polygon P with vertices $((x_1, y_1), \dots, (x_k, y_k))$ (in clockwise order), as follows:

- *Convex caps* come in four types $\{\curvearrowright, \curvearrowleft, \smile, \frown\}$. We have
 - $P \in \curvearrowright$ if and only if $(x_i)_{i=1}^k$ is increasing;
 - $P \in \curvearrowleft$ if and only if $(y_i)_{i=1}^k$ is increasing;
 - $P \in \smile$ if and only if $(x_i)_{i=1}^k$ is decreasing;
 - $P \in \frown$ if and only if $(y_i)_{i=1}^k$ is decreasing.
- *Convex chains* come in four types $\{\nearrow, \searrow, \swarrow, \nwarrow\}$. We have
 - $\nearrow = \curvearrowright \cap \smile$, $\searrow = \curvearrowleft \cap \frown$, $\swarrow = \smile \cap \curvearrowleft$, $\nwarrow = \frown \cap \smile$.

Figure 1 illustrates some maximum-size supported convex polygons for various grids. For $n \times n$ grids with $n \leq 4$, the largest supported convex polygon always has size n . For $n > 4$, this size can be less than n (as for $n = 5$ in Figure 1). Interestingly, for $n = 6$, there always exists a supported convex polygon of size at least 5.

► **Lemma 2.1.** *Every 6×6 grid $X \times Y$ supports a convex polygon of size at least 5.*

Proof. Let $X' = X \setminus \{\min(X), \max(X)\}$ and $Y' = Y \setminus \{\min(Y), \max(Y)\}$. The 4×4 grid $X' \times Y'$ supports a convex chain P' of size 3 between two opposite corners of $X' \times Y'$. Then one x -coordinate $x' \in X'$ and one y -coordinate $y' \in Y'$ are not used by P' . Without loss of generality, assume that $P' \in \nearrow$. Then the convex polygon containing the points of P' and $(x', \min(Y))$ and $(\max(X), y')$ is a supported convex polygon of size 5 on $X \times Y$. ◀



■ **Figure 2** An 8×8 grid without convex chains of size greater than $4 = \log_2 8 + 1$. $X = \{1, \dots, 8\}$, $Y = \{0, 7, 63, 70, 511, 518, 574, 581\}$. Two lines through pairs of grid points are drawn in blue.

More generally, by Lemma 2.2, every grid supports a convex chain of size $\Omega(\log n)$. We show in Lemma 2.3 that this bound is asymptotically tight: for each n , there exists a grid for which the maximum convex chain has size $O(\log n)$. Since every convex cap consists of two convex chains (some of which may be empty), and each convex polygon is composed of two convex caps, the same asymptotic bounds hold for maximum convex caps and polygons.

► **Lemma 2.2.** *Every $n \times n$ grid $X \times Y$ supports a convex polygon of size $\Omega(\log n)$.*

Proof. By Payne and Wood [4], every set of k points with at most ℓ collinear contains a set of $\Omega(\sqrt{k/\log \ell})$ points in general position. Here, we have n^2 points with at most n collinear, so there is a set of $\Omega(\sqrt{n^2/\log n}) = \Omega(n/\sqrt{\log n})$ points in general position. By Suk [5], every set of $2^{k+o(k)}$ points in general position contains a set of k points in convex position. Hence, we can find a subset of $\Omega(\log(n/\sqrt{\log n})) = \Omega(\log n)$ points in convex position. Eliminating points with the same x - or same y -coordinate reduces the size by at most 75%, so this asymptotic bound also holds when coordinates in X and Y may be used at most once. ◀

For the upper bound, we construct a family of grids without any large convex chain. For $n = 8$, this grid is depicted in Figure 2.

► **Lemma 2.3.** *For every $n \in \mathbb{N}$, there exists an $n \times n$ grid $X \times Y$ that does not support any convex chain of size greater than $\lceil \log n \rceil + 1$.*

Proof. Let $g(n)$ be the maximum value such that for all X and Y of size n , the grid $X \times Y$ supports a convex polygon of size $g(n)$; clearly $g(n)$ is nondecreasing. Let k be the minimum integer such that $n \leq 2^k$. We show that $g(2^k) \leq k + 1$ to establish that $g(n) \leq g(2^k) \leq k + 1$.

Without loss of generality, assume that $n = 2^k$, and let $X = \{1, \dots, n\}$. For a k -bit integer m , let m_i be the bit at its i -th position, such that $m = \sum_{i=0}^{k-1} m_i 2^i$. Let $Y = \{\sum_{i=0}^{k-1} m_i (n^{i+1} - 1) \mid 0 \leq m \leq n - 1\}$. Both X and Y are symmetric: $X = \{\max(X) + 1 - x \mid$

39:4 On Convex Polygons in Cartesian Products

$x \in X\}$ and $Y = \{\max(Y) - y \mid y \in Y\}$. Thus, it suffices to show that no $P \in \mathcal{C}$ of size greater than $k + 1$ exists.

Consider $p = (x, y)$ and $p' = (x', y') \in X \times Y$ with $y = \sum_{i=0}^{k-1} m_i(n^{i+1} - 1)$ and $y' = \sum_{i=0}^{k-1} m'_i(n^{i+1} - 1)$. The slope of the line between p and p' is $\text{slope}(p, p') = \frac{\sum_{i=0}^{k-1} (m'_i - m_i)(n^{i+1} - 1)}{x' - x}$. Let j be the largest index such that $m_j \neq m'_j$. Assume that $x < x'$ and $y < y'$, then $1 \leq x' - x \leq n - 1$ and we bound the slope as follows:

$$\begin{aligned} \frac{n^{j+1} - 1}{n - 1} &\leq \frac{n^{j+1} - 1 + \sum_{i=0}^{j-1} (m'_i - m_i)(n^{i+1} - 1)}{x' - x} = \text{slope}(p, p') \\ &= \frac{\sum_{i=0}^j (m'_i - m_i)(n^{i+1} - 1)}{x' - x} \leq \frac{\sum_{i=0}^j n^{i+1} - 1}{1} = n \frac{n^{j+1} - 1}{n - 1} - 1 < n \frac{n^{j+1} - 1}{n - 1}. \end{aligned}$$

Hence, $\text{slope}(p, p') \in I_j = [\frac{n^{j+1}-1}{n-1}, n \frac{n^{j+1}-1}{n-1}]$. Consider the family of intervals I_0, I_1, \dots, I_{k-1} defined analogously. For $n > 1$, we have $\max(I_j) < \min(I_{j+1})$. Suppose for a contradiction that some $P \in \mathcal{C}$ is of size greater than $k + 1$. Then, since the slopes of the first $k + 1$ edges of P decrease monotonically, there must be three consecutive vertices $p = (x, y)$, $p' = (x', y')$, and $p'' = (x'', y'')$ of P such that both $\text{slope}(p, p') \in I_j$ and $\text{slope}(p', p'') \in I_j$. Let $y = \sum_{i=0}^{k-1} m_i(n^{i+1} - 1)$ and $y' = \sum_{i=0}^{k-1} m'_i(n^{i+1} - 1)$ and $y'' = \sum_{i=0}^{k-1} m''_i(n^{i+1} - 1)$. Then j is the largest index such that $m_j \neq m'_j$, and also the largest index such that $m'_j \neq m''_j$. Because $m < m' < m''$, we have $m_j < m'_j < m''_j$, which is impossible since each of m_j, m'_j and m''_j is either 0 or 1. Hence, there are no convex chains of size greater than $k + 1$. \blacktriangleleft

3 Algorithms

In this section, we describe polynomial time algorithms for finding convex chains and caps of maximum size, as well as polynomial time approximation algorithms for finding the maximum size of a convex polygon. We make use of the following general observation:

► **Observation 3.1.** *If a supported convex polygon P is in a set of $\mathcal{C}, \searrow, \swarrow, \nearrow, \nwarrow, \curvearrowright, \curvearrowleft, \cup, \cap$, then any subsequence of P also lies in that set.*

Convex chains. Given a grid $X \times Y$, we provide an algorithm to compute a supported convex chain $P \in \searrow$ of maximum size. For this, we use a dynamic program to compute for each edge $(p_1, p_2) \in E = (X \times Y)^2$, the maximum size $R(p_1, p_2)$ of a chain of \searrow with p_1 and p_2 as first two vertices, or (p_1) if $p_1 = p_2$ (in which case $R(p_1, p_2) = 1$). By Observation 3.1, removing the first vertex from a chain of \searrow again yields a chain of \searrow .

► **Observation 3.2.** *If $A = (a_1, \dots, a_k) \in \searrow$ and $B = (b_1, \dots, b_\ell) \in \searrow$ with $k \geq 2, \ell \geq 2$ and $a_{k-1} = b_1$ and $a_k = b_2$, then $(a_1, \dots, a_{k-2}, b_1, \dots, b_\ell)$ also lies in \searrow and has size $k + \ell - 2$.*

Conversely, by Observation 3.2, for a chain P in \searrow , adding a vertex v at the front yields a chain of \searrow if $(v, p_1, p_2) \in \searrow$: the x - (resp., y -) coordinate of v is less (resp., greater) than those of vertices of P , so distinctness is maintained. Therefore we can find the maximum size of a chain starting with p_1 and p_2 based on chains without p_1 as follows:

$$R(p_1, p_2) = \begin{cases} -\infty & \text{if } p_1 \neq p_2 \text{ and } (p_1, p_2) \notin \searrow \\ 1 & \text{if } p_1 = p_2 \\ \max_{(p_1, p_2, v) \in \searrow \text{ or } v=p_2} R(p_2, v) + 1 & \text{otherwise.} \end{cases}$$

Since the x -coordinate of the first vertex of a \searrow chain is less than those of subsequent vertices, this formula is well defined. For sequences of constant size, membership in \searrow

can be checked in constant time. So the running time to compute $R(e)$ for all edges is $O(|E| \cdot |X \times Y|) = O(n^6)$, and the space complexity is $O(|E|) = O(n^4)$. This algorithm can easily be adapted to find the maximum size convex chains in $\{\nearrow, \searrow, \swarrow, \nwarrow\}$ with the same time and space complexity. We thus conclude the following:

► **Lemma 3.3.** *For a given $n \times n$ grid, we can compute a maximum size convex chain in $O(n^6)$ time and $O(n^4)$ space.*

Convex caps. To compute the maximum size of a convex cap in \curvearrowright , we compute the maximum size of two convex chains that use distinct y -coordinates. Specifically, for two edges $l = (l_1, l_2)$ and $r = (r_1, r_2)$, we compute the maximum total size $C(l, r)$ of a pair of chains $A \in \nearrow$ and $B \in \searrow$ such that their vertices use distinct y -coordinates and such that A ends with vertices l_1 and l_2 (or $A = (l_1)$ if $l_1 = l_2$), and B starts with vertices r_1 and r_2 (or $B = (r_1)$ if $r_1 = r_2$). To compute $C(l, r)$, we reuse the algorithm of Lemma 3.3 to compute $L(p_1, p_2)$ (resp., $R(p_1, p_2)$), the size of a largest convex chain P in \nearrow (resp., \searrow), ending (resp., starting) in vertices p_1 and p_2 , or $P = (p_1)$ if $p_1 = p_2$.

The desired quantity $C(l, r)$ can now be computed using a dynamic program. The main idea is that we can always safely eliminate the highest vertex of the two chains, to find a smaller subproblem, as this vertex cannot be (implicitly) part of the optimal solution to a subproblem. In particular, if l is a single vertex and it is highest, we can simply use the value of $R(r_1, r_2)$, incrementing it by one for the one vertex of l . Analogously, we handle the case if r is or both l and r are a single vertex. The interesting case is when both chains end in an edge. Here, we observe that we can easily check whether l and r use unique coordinates. If not, then this subproblem is invalid; otherwise, we may find a smaller subproblem by eliminating the highest vertex and checking all possible subchains that could lead to it.

With the reasoning above, we obtain the recurrence below. The first case eliminates invalid edges and combinations that use a coordinate more than once or that do not give a cap. After the first, it holds that $l \in \nearrow$, $r \in \searrow$ and that l and r use unique coordinates.

$$C(l, r) = \begin{cases} -\infty & \text{if } l_1 \neq l_2 \text{ and } l \notin \nearrow, \text{ or} \\ & r_1 \neq r_2 \text{ and } r \notin \searrow, \text{ or} \\ & \{l_1.y, l_2.y\} \cap \{r_1.y, r_2.y\} \neq \emptyset \\ 2 & \text{otherwise, if } l_1 = l_2 \text{ and } r_1 = r_2 \\ L(l_1, l_2) + 1 & \text{otherwise, if } r_1 = r_2 \text{ and } l_2.y < r_1.y \\ R(r_1, r_2) + 1 & \text{otherwise, if } l_1 = l_2 \text{ and } l_2.y > r_1.y \\ \max_{(v, l_1, l_2) \in \nearrow \text{ or } v=l_1} C((v, l_1), r) + 1 & \text{otherwise, if } l_2.y > r_1.y \\ \max_{(r_1, r_2, v) \in \searrow \text{ or } v=r_1} C(l, (r_2, v)) + 1 & \text{otherwise, } l_2.y < r_1.y. \end{cases}$$

We can compute $C(l, r)$ for all l and r in $O(|E|^2 |X \times Y|) = O(n^{10})$ time and $O(|E|^2) = O(n^8)$ space. With $C(l, r)$, we can easily find the size of a maximum size cap P in \curvearrowright , using the observation below, and analogous observations for the special case $k = 1$ and/or $\ell = 1$.

► **Observation 3.4.** *If $A = (a_1, \dots, a_k) \in \nearrow$ and $B = (b_1, \dots, b_\ell) \in \searrow$ with $k \geq 2$, $\ell \geq 2$ and $(a_{k-1}, a_k, b_1, b_2) \in \curvearrowright$ and A and B use distinct y -coordinates, then $(a_1, \dots, a_{k-2}, b_1, \dots, b_\ell)$ lies in \curvearrowright and has size $k + \ell$.*

► **Lemma 3.5.** *For a given $n \times n$ grid, we can compute a maximum size convex cap in $O(n^{10})$ time and $O(n^8)$ space.*

Convex n -chains and n -caps. If we are solely interested in deciding whether a convex chain or cap exists that has $|X| = |Y| = n$ vertices, we can improve upon the previous algorithms

considerably. Let $X = \{x_1, \dots, x_n\}$ and $Y = \{y_1, \dots, y_n\}$ with $x_i < x_{i+1}$ and $y_i < y_{i+1}$. To test whether \curvearrowright supports a chain of size n , it suffices to test the chain $((x_1, y_1), \dots, (x_n, y_n))$, which can be done in linear time.

To test whether \curvearrowleft supports a convex cap of size n , we adapt the algorithm of Lemma 3.5. Suppose P is a cap of \curvearrowleft of size n . For $k < n$, let $A_k \in \curvearrowright$ and $B_k \in \curvearrowleft$ be the subchains of A_P and B_P obtained after discarding vertices with y -coordinates greater than y_k . Let (l_1, l_2) be the last edge of A_k and let (r_1, r_2) be the first edge of B_k . Then h, i and j exist such that $i < j < k$ and $l_1.x = x_h, l_2.x = x_{h+1}, r_1.x = x_{n-k+h+2}, r_2.x = x_{n-k+h+3}$ and $\{l_1.y, l_2.y, r_1.y, r_2.y\} = \{y_i, y_j, y_{k-1}, y_k\}$. Suppose we adapt the formula for $C(l, r)$ to consider only entries of this form, and adapt the formulas for L and R to consider only entries of the form $((x_{i-1}, y_{i-1}), (x_i, y_i))$ and $((x_{n-i}, y_{i-1}), (x_{n-i+1}, y_i))$, respectively. We then obtain $O(|Y|^3|X|)$ possible values for (l, r) , and the corresponding values can be computed in $O(|Y|^4|X|) = O(n^5)$ time and $O(n^4)$ space. Testing whether \curvearrowleft supports a cap of size n can be done within the same time and space bounds.

Approximation. Although an efficient algorithm for computing the maximum size of a supported convex polygon is left as an open problem, the algorithms above provide constant-factor approximations. A convex cap $P \in \curvearrowleft$ is composed of two convex chains ($A_P \in \curvearrowright$ and $B_P \in \curvearrowleft$ as defined before), which are themselves caps in \curvearrowleft , and one of which has at least half the size of P . Hence, the algorithms to compute the maximum size of an x - and y -monotone chain provide a factor $\frac{1}{2}$ -approximation on the size of the largest cap. Similarly, a convex polygon P is composed of four x - and y -monotone convex chains, one of which contains at least a quarter of the vertices of P . Furthermore, P is composed of a convex cap and a convex cup, one of which contains at least half of the vertices of P . Thus, the algorithms in Lemma 3.3 and Lemma 3.5, respectively, yield factor $\frac{1}{4}$ - and $\frac{1}{2}$ -approximations for the maximum size of a convex polygon supported by $X \times Y$.

References

- 1 V. Chvátal and G. Klincsek. Finding largest convex subsets. In *Proc. 11th SE Conf. on Combin., Graph Theory and Comp*, volume 29, pages 453–460, 1980.
- 2 P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio mathematica*, 2:463–470, 1935.
- 3 Panos Giannopoulos, Christian Knauer, and Daniel Werner. On the computational complexity of Erdős-Szekeres and related problems in \mathbb{R}^3 . In *European Symposium on Algorithms*, pages 541–552. Springer, 2013.
- 4 Michael S. Payne and David R. Wood. On the general position subset selection problem. *SIAM Journal on Discrete Mathematics*, 27(4):1727–1733, 2013.
- 5 Andrew Suk. On the Erdős-Szekeres convex polygon problem. *Journal of the American Mathematical Society*, 30:1047–1053, 2017.
- 6 George Szekeres and Lindsay Peters. Computer solution to the 17-point Erdős-Szekeres problem. *The ANZIAM Journal*, 48(2):151–164, 2006.
- 7 Pavel Valtr. Probability that n random points are in convex position. *Discrete & Computational Geometry*, 13(3-4):637–643, 1995.