Peeling the Grid

Sariel Har-Peled* Bernard Lidický†

September 24, 2012

Abstract

Consider the set of points formed by the integer $n \times n$ grid, and the process that in each iteration removes from the point set the vertices of its convex-hull. Here, we prove that the number of iterations of this process is $O(n^{4/3})$; that is, the number of convex layers of the $n \times n$ grid is $\Theta(n^{4/3})$.

*Department of Computer Science; University of Illinois; 201 N. Goodwin Avenue; Urbana, IL, 61801, USA; sariel@uiuc.edu; http://www.uiuc.edu/~sariel/. Work on this paper was partially supported by a NSF AF award CCF-0915984.

†Department of Mathematical Sciences; University of Illinois; 1409 W. Green St.; Urbana, IL, 61801, USA; lidicky@illinois.edu.

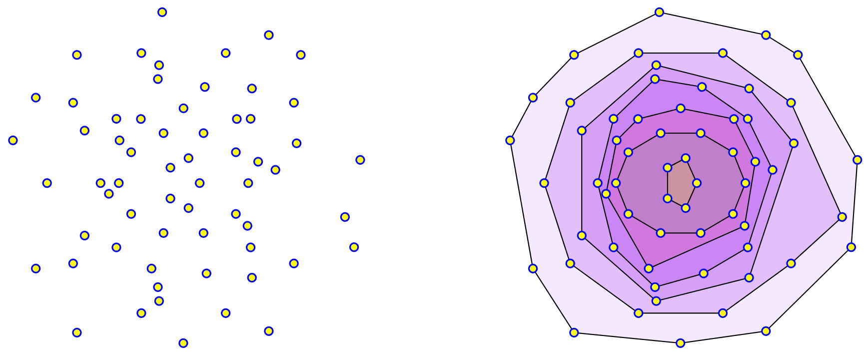

Figure 1: A point set and its decomposition into convex layers.

1 Introduction

For many algorithms, the worst case behavior is rarely encountered in practice. This is because the worst case behavior might require a degenerate and convoluted input. To address this gap between the worst case analysis and a real world behavior, a considerable amount of research was spent on analyzing algorithms and discrete geometric structures under certain assumptions on the input, including (i) realistic input models [dBKSV02], (ii) fatness [AdBES11], (iii) randomness, etc.

Random points. There is a significant amount of work on the geometric behavior of random point sets [RS63, Ray70, WW93, Bár08, OBSC00, JN04]. The question of how the Voronoi diagram or the convex-hull of a point set randomly generated inside a convex domain behaves had received considerable attention. In particular, it is known that for a set of n points chosen uniformly in the unit square, the expected complexity of the convex-hull is $O(\log n)$, and $O(n^{1/3})$ if the domain is a disk (this bound holds for any convex shape).

Grid points. Surprisingly, the known results on uniformly sampled points match the results known for the grid point set. For example, the number of vertices of the convex hull of any subset of the $\sqrt{n} \times \sqrt{n}$ grid is $O\left(n^{1/3}\right)$, which matches the bound for the random points. This phenomena holds for many similar scenarios, see the survey by Bárány [Bár08].

Convex layers. The decomposition of a point set into convex layers is one possible way to measure the depth of a point inside the point set. Formally, the ***convex depth*** of a point p in a point set P is $\mathsf{d}_{\mathsf{p}}(P) = 1$ if p is a vertex of the convex-hull of P, and it is $\mathsf{d}_{\mathsf{p}}(P) = 1 + \mathsf{d}_{\mathsf{p}}(P \setminus V(\mathcal{CH}(P)))$ otherwise, where $\mathcal{CH}(P)$ denotes the convex-hull of P and $V(\mathcal{CH}(P))$ denotes the set of its vertices[1]. This partitions the point set into convex-layers, as depicted in Figure 1. In particular, if the points rise out of physical measurements (that might contain noise), a point with large convex depth is unlikely to be an outlier. This is one

[1]A point of P is a vertex of the convex-hull only if it is a corner of the convex-hull. Formally, p is a vertex of the convex-hull of P is $\mathcal{CH}(P) \neq \mathcal{CH}(P \setminus \{\mathsf{p}\})$.

possible definition of robust statistics for points, although this definition has its limitations, see [RS04] for details. In particular, Chazelle [Cha85] provided an $O(n \log n)$ time algorithm for computing all the convex layers for a set of points in the plane.

For a set of n points picked uniformly inside a bounded convex domain in $\mathbb{R}^d$, it is known that the expected number of convex layers is $\Theta\left(n^{2/(d+1)}\right)$ [Dal04].

Our results. In this paper, we are prove that the number of convex layers of the $n \times n$ grid is $\Theta\left(n^{4/3}\right)$. This bound is quite surprising – indeed, as demonstrated by Figure 2, the peeling process starts out quite slowly, the first three layers having $4, 8, 8$ vertices (independent of the value of n), respectively. A priori, it is not clear why this process accelerates and contains more vertices. Furthermore, the maximum number of vertices in convex position in an $n \times n$ grid is $O\left(n^{2/3}\right)$ (this is well known, see Lemma 2.1). Namely, somewhat surprisingly, a constant fraction of the layers have asymptotically maximum size. Our result matches the known result for random points. Note, that although the bounds are similar, the proof for the random point set does not carry over to the grid case.

We also observe that the number of convex layers is $\Omega(n^2)$ if the grid of $n \times n$ points is allowed to be non-uniform (instead of the integer grid used above). Naturally, in this construction, where every point is on two lines where each has n points.

2 Peeling the grid

Let $P_0 = G_n = \{1, \ldots, n\}^2$, be the $n \times n$ integer grid. In the ith iteration, consider the convex-hull $C_i = \mathcal{CH}(P_{i-1})$, for $i = 1, \ldots$. Let V_i be the set of vertices of C_i. Naturally, we consider a grid point to be a ***vertex*** only if it is a corner of the convex-hull, and as such grid points falling in the middle of edges of C_i, are not in V_i. Now, let $P_i = P_{i-1} \setminus V_i$. In words, we start with the $n \times n$ grid, and peel away the vertices of the convex-hull, and we repeat this process till all the grid points of G_n are removed. Let $\tau(n)$ be the number of iterations, till P_i is an empty set. Here we are interested in the behavior of $\tau(n)$. See Figure 2 for an example of how the generated polygons look like.

2.1 A lower bound on $\tau(n)$

The following is well known, and we include a proof for the sake of completeness.

Lemma 2.1. *Given any convex set C in the plane, it can have at most $O\left(n^{2/3}\right)$ vertices of G_n.*

Proof: Consider a convex set C such that all its vertices are points of G_n. The perimeter of C is at most $4n$. The number of edges of the convex hull of C of length at least (or equal to) μ is at most $4n/\mu$. The number of edges having length smaller than μ is bounded by the number of integer points of distance at most μ from the origin, and this number is bounded by $(2\mu+1)^2 = O(\mu^2)$. As such, the number of vertices of C is at most $O(n/\mu + \mu^2)$. Setting $\mu = \left\lfloor n^{1/3} \right\rfloor$ then implies the claim. ∎

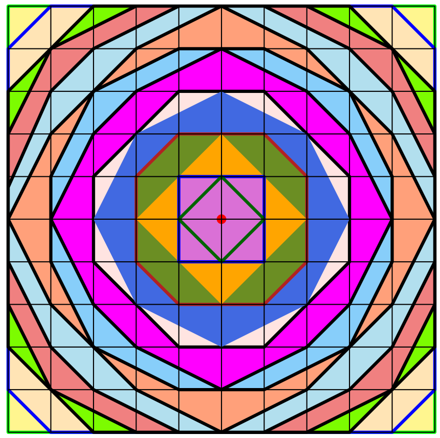

Figure 2: The polygons generated while peeling the 11×11 integer grid.

As such, $|V_i| = O(n^{2/3})$, which implies immediately that $\tau(n) \geq n^2 / \max_i |V_i| = \Omega(n^{4/3})$.

2.2 An upper bound on $\tau(n)$

An integer vector (x, y) is ***primitive*** if $\gcd(x, y) = 1$. For an integer μ, let $\mathcal{V}_\mu$ be the set of all primitive non-zero integer vectors (x, y), where $0 \leq y < x \leq \mu$. The following is well known, and we sketch a proof for the sake of completeness.

Lemma 2.2. *We have $|\mathcal{V}_\mu| \geq c\mu^2$, for some constant $c > 0$.*

Proof: For a fixed x, consider the vectors (x, y) in $\mathcal{V}_\mu$, such that $y < x$, and $\gcd(x, y) = 1$. The number of such vectors is the number of integer values of y that are relative prime to x, and this number is the Euler's totient function $\phi(x)$. As such, $|\mathcal{V}_\mu| \geq \sum_{i=1}^{\mu} \phi(i) \geq c\mu^2$. the last step follows from known bounds, see [HW65]. ■

In the following, we pick μ to be smaller than $n/4$, and n is sufficiently large.

For every vector $v \in \mathcal{V}_\mu$, consider the set L_v of all lines having direction v that intersect the grid points G_n. Every line in L_v contains at most $1 + \lfloor (n-1)/v_x \rfloor$ points of the grid (and most lines in this family contain at least $\lfloor (n-1)/v_x \rfloor$ points of the grid (the only problematic lines are the ones that have short intersection with the square $[1, n]^2$ because of the corners)).

Claim 2.3. *For $n > 10$, $\mu < n/4$ and $v \in \mathcal{V}_\mu$, we have that $|L_v| \leq 4n\mu$.*

Proof: A line $\ell \in L_v$ that intersects G_n has an intersection of length at least n with the enlarged square $[1, 2n]^2$. Specifically, the projection of the intersection on the x axis has length at least n. Since ℓ has direction v and it contains a grid point, it follows that it has grid points on it, that are of distance $||v||$ from each other. On the projection, the distance between these points is v_x. As such, this intersection contains at least $1 + \lfloor n/v_x \rfloor \geq n/\mu$ points of the grid G_{2n} on it. In particular, the number of such lines can be at most $4n^2/(n/\mu) = 4n\mu$. ■

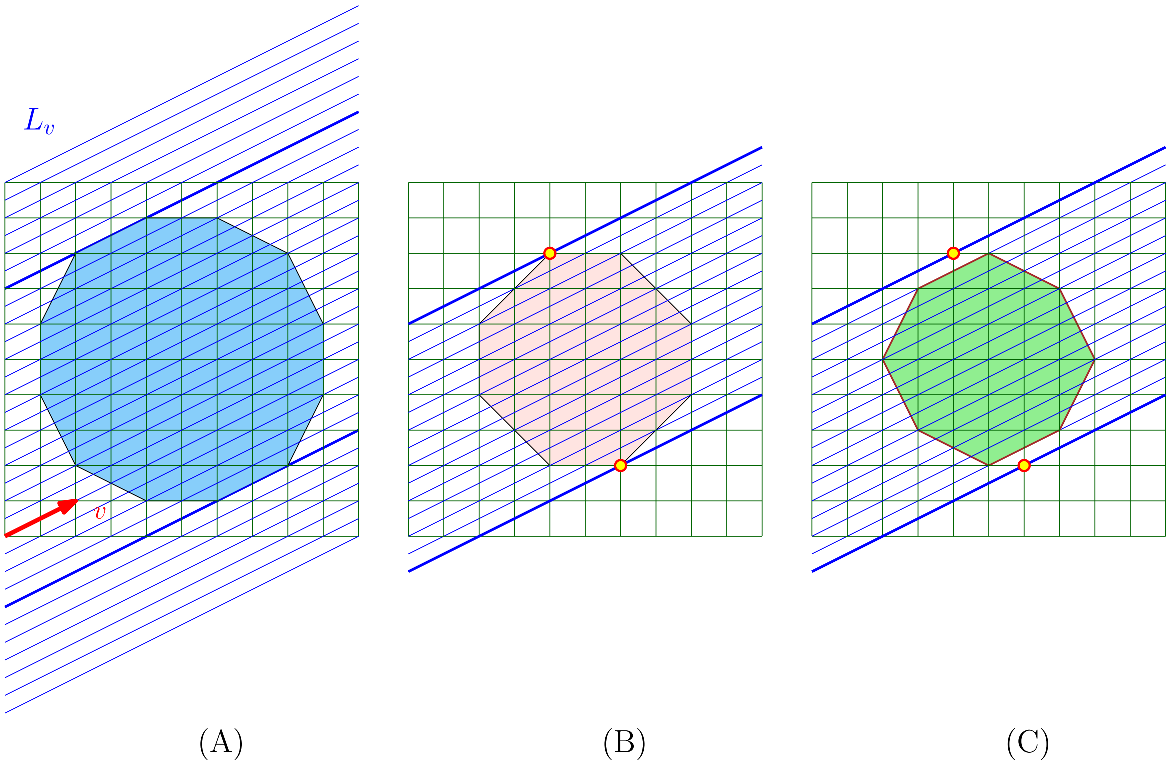

Figure 3: (A) An active direction v, and the set of lines L_v. (B) An inactive iteration for v. (C) The next iteration – the two "old" tangent lines no longer intersect the current convex layer.

Since the lines of L_v cover all the grid points of G_n, and the vertices of C_i are grid points, it follows that L_v always contains two lines that are tangent to C_i. If these two tangent lines intersect ∂C_i along an non-empty edge, then v is ***active*** at iteration i (i.e., v is not active if the two tangents touch C_i at a vertex).

In the following, we slightly abuse notations and use $L_v \cap C_i$ to denote the set of all lines of L_v that have non-empty intersection with C_i.

Claim 2.4. *If v is not active at iteration i, then $|L_v \cap C_{i+1}| \leq |L_v \cap C_i| - 2$.*

Proof: If v is not active at iteration i then a tangent ℓ to C_i from L_v intersects C_i only at a vertex. But this vertex is being removed from the point set when computing P_{i+1}. In particular, the line ℓ no longer intersects C_{i+1}. The same argument also applies to the other tangent. This is demonstrated in Figure 3. ■

Claim 2.5. *Throughout the process, for a vector $v \in \mathcal{V}_\mu$, it can be inactive in at most $2n\mu$ iterations.*

Proof: Every time v is not active, the number of lines of L_v that intersect the active convex-hull decreases by two, by Claim 2.4. By Lemma 2.3 there are at most $4n\mu$ lines in the set L_v, and as such this can happen at most $4n\mu/2$ times. ■

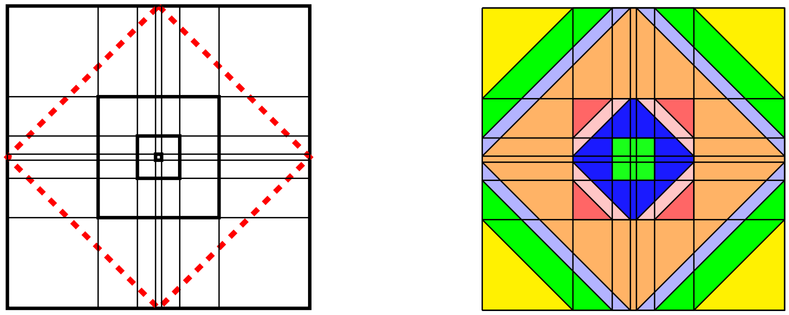

Figure 4: A point set where the peeling process requires $\Omega(n^2)$ steps.

If the process continues more than $M = 4n\mu$ iterations then every vector in $\mathcal{V}_\mu$ is active in at least half of the iterations. In particular, if n_i is the number of active directions at iteration i, then we have that

$$\alpha = \sum_{i=1}^{M} n_i \geq 2n\mu \left|\mathcal{V}_\mu\right| \geq 2cn\mu^3,$$

by Lemma 2.2.

Observe, that if n_i vectors are active at the ith iteration, then the convex hull of C_i has at least $2n_i$ edges (and thus vertices) at iteration i. As such, if we set $\mu = \lceil n^{1/3}/c^{1/3} \rceil = \Theta\big(n^{1/3}\big)$, we have that the total number of vertices of the convex hulls in the first M iterations is at least

$$2\alpha \geq 4cn\mu^3 \geq 4n^2,$$

which is a contradiction, as the initial grid set has at most n^2 points. We conclude that the algorithm must terminate after $M = 4n\mu = O\big(n^{4/3}\big)$ iterations. We thus proved the following.

Theorem 2.6. *Starting with the grid G_n, consider the process that repeatedly removes the convex-hull vertices of the current set of vertices. This process takes $\Theta\big(n^{4/3}\big)$ steps.*

3 Lower bound of $\Omega(n^2)$ for a non-uniform grid

This section is devoted to describing a set M of n^2 points in the plane where the peeling process takes $\Omega(n^2)$ steps. For simplicity assume that $n = 2k$ for some integer k.

Take a collection of k squares $S_1, \ldots, S_k$ where S_i has length of its side 3^i and the squares are positioned such that their centers coincide with the origin. Let L be the set of $4k$ lines that are obtained by extending the segments of the squares into lines. Finally, let M be the set of all intersections of lines in L. Notice that each line contains $2k$ points and that $|L| = 4k^2 = n^2$. See Figure 4.

Let the peeling process partition M into convex sets $C_1, C_2, \ldots$.

Claim 3.1. *For every* C_i *exists* S_j *such that* $C_i \subseteq S_j$*.*

Proof: Let j be the largest index such that $C_i \cap S_j \neq \emptyset$. Notice that C_i is centrally symmetric as M is centrally symmetric and this property is preserved by the peeling process. If $|C_i \cap S_j| = 4$ then $C_i \cap S_j$ are the four corners of S_j and thus $|C_i| = 4$ as C_i is strictly convex. Hence $|C_i \cap S_j| = 8$ and C_i contains points on both vertical and horizontal lines of S_j in every quadrant. Let D be the square with corners being intersections the axis and $\mathcal{CH}(S_j)$. See Figure 4 on the left. Notice that $S_l \subset D \subset \mathcal{CH}(C_i \cap S_j)$ for every $l < j$. Therefore $C_i = C_i \cap S_j \subseteq S_j$. ∎

The previous claim implies that $|C_i| \leq 8$ for every i. Hence the peeling process needs at least $n^2/8 = \Omega(n^2)$ steps.

4 Conclusions

The most natural question left by our work, is the prove similar bounds in higher dimensions. This seems quite challenging, and we leave it as an open problem for further research.

Let us also note for an interested reader that, according to experiments, the layers in the peeling process are getting close to circles as the process is advancing.

Acknowledgments

We thank Robert Jamison for many useful discussions on this and related problems. The authors also would like to thank Imre Bárány for providing relevant information.

References

[AdBES11] B. Aronov, M. de Berg, E. Ezra, and M. Sharir. Improved bound for the union of fat objects in the plane. manuscript, 2011.

[Bár08] I. Bárány. Random points and lattice points in convex bodies. *Bulletin Amer. Math. Soc.*, 45(3):339–365, 2008.

[Cha85] B. Chazelle. On the convex layers of a planar set. *IEEE Trans. Inform. Theory*, IT-31(4):509–517, July 1985.

[Dal04] K. Dalal. Counting the onion. *Random Struct. Alg.*, 24(2):155–165, 2004.

[dBKSV02] M. de Berg, M. J. Katz, A. F. van der Stappen, and J. Vleugels. Realistic input models for geometric algorithms. *Algorithmica*, 34:81–97, 2002.

[HW65] G. Hardy and E. Wright. *The Theory of Numbers.* Oxford University Press, London, England, 4th edition, 1965.

[JN04] F. Jarai-Szabo and Z. Neda. On the size-distribution of poisson voronoi cells. *eprint arXiv:cond-mat/0406116*, June 2004.

[OBSC00] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu. *Spatial tessellations: Concepts and applications of Voronoi diagrams.* Probability and Statistics. Wiley, 2nd edition edition, 2000.

[Ray70] H. Raynaud. Sur l'enveloppe convex des nuages de points aleatoires dans R^n. *J. Appl. Probab.*, 7:35–48, 1970.

[RS63] A. Rényi and R. Sulanke. Über die konvexe Hülle von n zufällig gerwähten Punkten I. *Z. Wahrsch. Verw. Gebiete*, 2:75–84, 1963.

[RS04] P.J. Rousseeuw and A. Struyf. Computation of robust statistics: Depth, median, and related measures. In J. E. Goodman and J. O'Rourke, editors, *Handbook of Discrete and Computational Geometry*, chapter 57, pages 1279–1292. CRC Press LLC, 2nd edition, 2004.

[WW93] W. Weil and J. A. Wieacker. Stochastic geometry. In P. M. Gruber and J. M. Wills, editors, *Handbook of Convex Geometry*, volume B, chapter 5.2, pages 1393–1438. North-Holland, 1993.