Area Optimization of Simple Polygons *

SÁNDOR P. FEKETE † WILLIAM R. PULLEYBLANK ‡

Abstract

We discuss problems of optimizing the area of a simple polygon for a given set of vertices P and show that these problems are very closely related to problems of optimizing the number of points from a set Q in a simple polygon with vertex set P. We prove that it is NP-complete to find a minimum weight polygon or a maximum weight polygon for a given vertex set, resulting in a proof of NP-completeness for the corresponding area optimization problems. We show that we can find a polygon of more than half the area $AR(conv(P))$ of the convex hull $conv(P)$ of P, and demonstrate that it is NP-complete to decide whether there is a simple polygon of at least $(\frac{2}{3}+\varepsilon)AR(conv(P))$. Finally, we prove that for $1 \leq k \leq d$, $2 \leq d$, it is NP-hard to minimize the volume of the k-dimensional faces of a d-dimensional simple nondegenerate polyhedron with a given vertex set, answering a generalization of a question stated by O'Rourke in 1980.

*Portions of this research were supported by the Natural Sciences and Engineering Research Council of Canada, and C.P. Rail.

†Department of Applied Mathematics and Statistics, SUNY Stony Brook, NY 11794–3600

‡IBM Research, T.J. Watson Research Center, Yorktown Heights, NY 10598

9th Annual Comptutational Geometry,5/93/CA,USA

1 Introduction

From a geometrical point of view, the Euclidean Travelling Salesman Problem is to find a polygon with a given set of vertices that has shortest perimeter. Since it is not hard to see that an optimal polygon cannot be self-intersecting (an easy consequence of the triangle inequality), we may restrict our attention to simple polygons that have a given set of vertices.

It seems very natural to look for a simple polygon with a given set of vertices which minimizes another basic geometric measure: the enclosed area. We call this the problem "Minimum Area Polygon" (MAP) and write MAXP for the related problem of *maximizing* the enclosed area.

The problem MAP has gained some importance for questions related to pattern recognition, which are often concerned with simple polygons. In 1989, Suri asked for the complexity of MAP. While there has been some research on extremal polygons on a given vertex set (see Boyce, Dobkin, Drysdale and Guibas [2] and Eppstein, Overmars, Rote and Woeginger [7]), the main attention has focussed on finding subpolygons with certain special properties, such as convexity.

The area of a polygon with rational vertices can be calculated in linear time by adding the area of the triangles in a triangulation. This contrasts sharply to determining the perimeter of a given polygon (see Garey, Graham and Johnson [10]). Since there is no efficient way known comparing the sum of a finite number of square roots with a given rational, it is still an open problem whether the Euclidean Travelling Salesman Problem belongs to the class NP.

On the other hand, there are geometric aspects that make it seem harder to optimize the area enclosed by a polygon than to minimize its perimeter: We can think of a tour as a set of line segments and of a polygonal region as a set of triangles. But while the vertices of a short edge are necessarily at a close distance, the same need not be true for a triangle with a small area. The points in a triangle can be very far apart, while the enclosed area is arbitrarily small. These difficulties were mentioned by Boyce, Dobkin, Drysdale and Guibas in [2], where it was noted that "small perimeter implies that the vertices are well localized, but small area does not." Another difficulty described in [7] by Eppstein, Overmars, Rote and Woeginger was the problem of possible self-intersection: While it automatically does not occur in a polygon with minimal perimeter, it can be algorithmically problematic to maintain simplicity when building up a polygon from triangles.

It may be surprising that until now, it has not been shown that it is NP-hard to find a minimum area (simple) polygon with a given set of vertices. This proof is a major subject of Section 4. In fact, we prove that it is NP-complete to determine whether there exists a simple polygon with a given vertex set whose area equals a very elegant combinatorial lower bound provided by Pick's theorem.

From the described problems of dealing with the area of polygons, one has to expect that the technical details of such a proof require a lot of careful analysis, even when a construction with a simple underlying idea is found. Therefore it should not be too surprising that our NP-completeness proof for Minimum Area Polygons is long and technically involved.

The main results derived in this paper are

- It is NP-complete to determine a minimum area simple polygon for a given set of vertices.
- It is NP-complete to determine a maximum area simple polygon for a given set of vertices.
- For a set P of n vertices, a simple polygon of more than half the area of the convex hull $conv(P)$ can be found in time $O(n \log n)$. The factor $\frac{1}{2}$ and the time bound are tight.
- For $\varepsilon < \frac{1}{3}$ and positive, it is NP-complete to decide whether a set P of vertices allows a simple polygon of area more than $(\frac{2}{3} + \varepsilon)conv(P)$.
- All the results remain true if we replace "area" by the following objective function: For a set Q of m points, optimize the number of points of Q that are contained in a polygon with vertex set P.
- For $1 \leq k \leq d$ and $2 \leq d$, it is NP-hard to minimize the volume of the k-dimensional faces of a simple nondegenerate d-dimensional polyhedron with a given set of vertices in d-dimensional space.

The rest of the paper is organized as follows. Section 2 shows that questions of area optimization and lattice point separation are closely related by Pick's theorem. Section 3 discusses some special cases. Section 4 gives an outline of the proof that it is NP-complete to find a minimum area polygon for a given vertex set. Section 5 contains results on area maximization. Section 6 sketches the proof that it is NP-hard to find a simple nondegenerate d-dimensional polyhedron for a given set of vertices with minimum volume of its k-dimensional faces.

2 Pick's Theorem

Let $\mathcal{P}$ be a polygon, given by a set of vertices and a set of edges. We call $\mathcal{P}$ *simple* if any vertex is only contained in two edges and nonadjacent edges do not intersect. Now consider a simple polygon $\mathcal{P}$ with grid points as vertices. What is its enclosed area $AR(\mathcal{P})$?

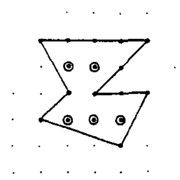


Figure 1: Pick's theorem

Perhaps one of the most surprising and elegant answers is provided by Pick's theorem:

Theorem 2.1 (Pick [14]) *Let $\mathcal{P}$ be a simple polygon with integer vertices; let $i(\mathcal{P})$ be the number of grid points contained in the interior of $\mathcal{P}$ and let $b(\mathcal{P})$ be the number of grid points on the boundary of $\mathcal{P}$. Then*

$$AR(\mathcal{P}) = \frac{1}{2}b(\mathcal{P}) + i(\mathcal{P}) - 1.$$

An elegant proof can be found in Coxeter [4].

Pick's theorem yields a combinatorial interpretation for minimizing or maximizing the area of a polygon. Any grid point that is contained in the boundary contributes $\frac{1}{2}$ to the area of the polygon, any grid point in the interior contributes 1. The best we can do when minimizing the area is to avoid including any grid points other than the given n vertices, thus getting a polygon of area $\frac{n}{2} - 1$.

If we want to *maximize* the area, we have to include as many additional grid points as possible into the polygon, in a way that each of them contributes as much as possible. Since no grid point on the boundary of the convex hull of the given vertex set can be contained in the interior, they can at most contribute $\frac{1}{2}$. Any other grid point that is not given as a vertex contributes 1 when contained in the interior of the polygon.

We summarize these upper and lower bounds:

Theorem 2.2 *Let P be a set of n points in the plane that all have integer coordinates. Let $h_i(P)$ denote the number of points of the integer grid that are not contained in P and strictly inside the convex hull, and $h_b(P)$ the number of grid points not in P that are on the boundary of the convex hull. Then for any simple polygon $\mathcal{P}$ on the vertex set P, we have*

$$\frac{n}{2} - 1 \leq AR(\mathcal{P}) \leq \frac{n}{2} + \frac{h_b(P)}{2} + h_i(P) - 1.$$

These bounds suggest the following stronger versions of MAP and MAXP:

Grid Avoiding Polygon (GAP) *Given n grid points in the plane. Is there a simple polygon on this vertex set that does not contain any other grid points on its boundary or in its interior, and therefore has area $\frac{n}{2} - 1$?*

Grid Approximating Polygon (GAXP) *Given n grid points in the plane. Is there a simple polygon on this vertex set that contains all grid points in the convex hull as well as possible: The grid points on the boundary of the convex hull lie on the boundary and all other grid points belonging to the convex hull lie in the interior? Such a polygon must have area $\frac{n}{2} + \frac{h_b(P)}{2} + h_i(P) - 1$.*

Even though these problems may seem combinatorially easier (all we have to do is avoid including integer points or making sure that they are all included in the polygon), we will see Section 4 that they are NP-complete, and hence as difficult as the original problem.

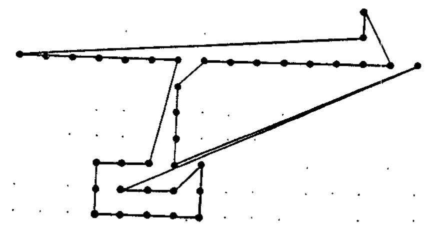
3 Special Cases

In this section, we discuss a few easy special cases for optimizing the area of a simple polygon and point out a source of difficulties in more general situations.

Corollary 3.1 *Given a* convex *set P of grid vertices, i.e. a set for which each grid point contained in the convex hull* conv(P) *belongs to P. Then any simple polygon with vertex set P has area $\frac{n}{2} - 1$.*

In this situation, we do not have any choice about including additional grid points, so Pick's theorem guarantees an area of $\frac{n}{2} - 1$. A set P of grid vertices has $h_b(P) = h_i(P) = 0$ if and only if it forms a convex subset of the integer grid. Thus, Corollary 3.1 characterizes exactly the situation where lower and upper bound from Theorem 2.2 coincide.

Consider some "horizontal" direction, and the orthogonal "vertical" direction. (We do *not* require the "horizontal" direction to be parallel to an axis of the integer grid.) A set $M \subseteq R^2$ is said to be *horizontally convex*, if any horizontal line intersects M in a connected set. We say that a set of grid points is horizontally convex, if no horizontal line segment between two points in the set contains a grid point that is not in the set. A horizontal line l *splits* a set of grid points P, if each of the open halfplanes bounded by l contains at least one point of P. We say that a horizontally convex set P of grid points has *thickness* t, if any line l that splits P and contains a point of the integer grid contains at least t points of P.



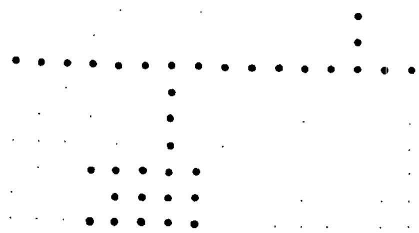


Figure 2: Two horizontally convex sets, with and without a Grid Avoiding Polygon

Theorem 3.2 *Given a set P of grid vertices. If P is horizontally convex and has thickness at least 2 for some horizontal orientation, then P has a Grid Avoiding Polygon.*

Proof: Consider the subsets of grid vertices that are contained in the same horizontal line. We can think of these subsets $S_1, \dots, S_m$ as ordered by their vertical coordinates. The points in subset S_i can be ordered by their horizontal coordinates. Let p_{ij} denote the jth point in S_i and k_i the number of points in S_i. Then connect the points by the polygon

$$\langle p_{11}, \dots, p_{m1}, \dots, p_{mk_m}, p_{(m-1)2}, \dots, p_{(m-1)k_{(m-1)}}, \\ \dots, p_{22}, \dots, p_{2k_2}, p_{1m_1} \dots, p_{11} \rangle.$$

By construction, this polygon does not contain any other grid points. ∎

The constructed polygon is monotone with respect to the vertical coordinates. Conversely, the existence of a monotone Grid Avoiding Polygon requires the vertex set to be horizontally convex.

If we drop the thickness condition, matters can get difficult, as shown in Figure 2. Both point sets are horizontally convex. Careful analysis shows that P_2 does not allow a Grid Avoiding Polygon. The additional point in P_1 allows a narrow triangle of area $\frac{1}{2}$ that picks up a point far away, thereby "deblocking" the point set P_2 and allowing the polygon shown. These "probes" make it very difficult to localize arrangements of points. We show in Section 4 how they can be controlled and used to show the NP-completeness of GAP.

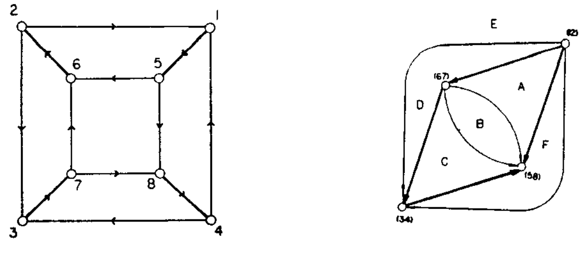


Figure 3: A cubic digraph D and its modification $\overline{D}$

4 Minimal Area

In this section we demonstrate that deciding the existence of a grid-avoiding polygon is NP-complete:

Theorem 4.1 GAP *is NP-complete.*

The problem is contained in NP. To show that it is NP-hard, we give a reduction of HAMILTONIAN CYCLE IN PLANAR CUBIC DIRECTED GRAPHS. That is, for a given planar cubic digraph D, we construct a point set P_D in polynomial time, such that P_D admits a GRID AVOIDING POLYGON if and only if D has a Hamiltonian Cycle.

The idea for this construction is as follows:

After some minor rearrangements of the planar cubic directed graph, it is suitably embedded in the plane, such that all edges are rectilinear sets of line segments. Then the embedding is suitably scaled up and perturbed, to guarantee that there are no collinearities between non-adjacent end points of line segments. Finally, these end points are replaced by suitable sets of grid points ("boxes"), such that a Hamiltonian path in the graph corresponds to a very narrow polygon that does not encounter any other grid points. Each set of points corresponding to an edge that is used in a Hamiltonian path is collected in one connected "branch" of the polygon, while the point sets corresponding to an edge that is not used by the path are split into two sets that are contained in two separate branches of the polygon. The layout of the points is chosen in a way that these branches can only be put together in a certain way without including any extra grid points.

As pointed out before, dealing with areas instead of distances makes it very hard to localize neighbours in a set of points. We achieve the desired localization by the perturbation mentioned above.

4.1 Planar Cubic Digraphs

Consider any planar cubic digraph D. It is straightforward to see that we can make the following assumptions when considering Hamiltonicty of D: All vertices must have either indegree 2 or outdegree 2. Let the first type of vertices be called *in-vertices*, the second *out-vertices.* An edge is *mandatory* for a vertex, if and only if it is the only incoming or outgoing edge; otherwise it is *optional* for the vertex. Furthermore, we may assume that any edge is either mandatory or optional for both its vertices. This implies that D is bipartite, since any mandatory edge goes from an in-vertex to an out-vertex and any optional edge goes from an out-vertex to an in-vertex. The optional edges of D form a set of vertex-disjoint cycles in the undirected graph G, obtained by replacing all arcs of D with edges. We denote by m the number of vertices of D. Then D has $\frac{3m}{2}$ edges.

4.2 Embedding the Digraph

In a next step, D is embedded in the plane, such that all edges are rectilinear paths, and the optional edges have precisely one bend, while the mandatory edges have three or five bends. Furthermore, all end points of line segments ("nodes") will have coordinates which are multiples of $\frac{1}{5}$ in the range between $\frac{3}{5}$ and m.

We start by identifying the cycles formed by the optional edges. As noted above, they are vertex-disjoint and even, so we can easily choose one (of two possible) perfect matchings in each cycle and contract all its edges, resulting in a 4-regular planar digraph $\overline{D}$ with $\frac{m}{2}$ vertices, m edges and $\frac{m}{2}+2$ faces. (See Figure 3.)

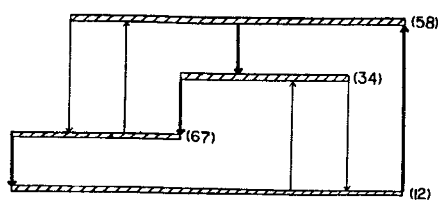


Figure 4: A rectilinear planar layout for $\overline{D}$

Using the approach of Rosenstiehl and Tarjan [16], this digraph $\overline{D}$ can be embedded in the plane as a planar rectilinear layout with no two collinear line segments. We modify the planar rectilinear layout to get an embedding of D where every optional edge has precisely one bend, every mandatory edge has three or five bends and all coordinates are multiples of $\frac{1}{5}$ in the range between $\frac{3}{5}$ and m. One mandatory edge adjacent to the exterior face of the embedding is split by introducing a source s and a sink t; there is a Hamiltonian cycle in D, if and only if in the resulting digraph D' there is a Hamiltonian path. An example of the resulting embedding is shown in Figure 5.

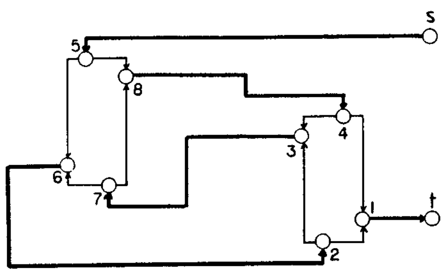


Figure 5: An embedding of D

4.3 Perturbing the Embedding

In the following, a *node* is a grid point in the embedding that separates two straight line segments. So a node either represents a vertex of D', or is a grid point where two line segments meet that belong to the representation of the same edge of D'. We let $\overline{m}$ denote the number of nodes. It is not hard to see that $\overline{m} \leq \frac{9m}{2}$.

Next, we perturb the embedding of D'. We start by multiplying all the distances in the embedding by a factor of $5N^3$, where $N := \overline{m}^{\overline{m}}$.

We think of x-coordinates as coordinates in the horizontal direction and y-coordinates as coordinates in the vertical direction. In the next step, partition the nodes into *vertical* classes. Two nodes belong to the same vertical class if they have identical y-coordinates and if they are connected by a path of at most two line segments of this y-coordinate. So each class consists of two or three nodes, where three nodes occur in the case of an out-vertex, with the two adjacent nodes representing the bends of the two adjacent optional edges. This implies that there are less than $\frac{\overline{m}}{2}$ vertical classes.

Shift all points in the ith vertical class by the vector $(0, \overline{m}^i)$. Similarly, define horizontal classes and shift the points in the ith horizontal class by the vector $(N\overline{m}^i, 0)$.

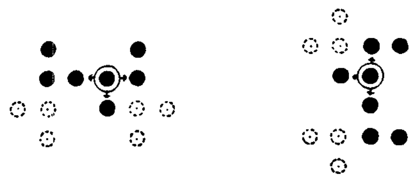


Figure 6: Horizontal and vertical switch

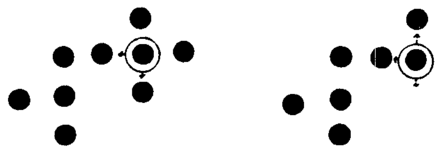


Figure 7: The two types of link boxes

4.4 Replacing Nodes by Point Sets

Finally, we replace each node by an appropriate set of points that form a connected subset of the integer grid — called a *box*.

Every degree 3 node is replaced by a *switch* box — see Figure 6. The circled point denotes the location of the node, the arrows indicate the edges adjacent to the node. Dotted locations correspond to points that may or may not be in the box, depending on the adjacent optional edges — we discuss this below. Out-vertices (represented by nodes with both optional edges leaving horizontally) are replaced by a horizontal switch; in-vertices (with optional edges entering vertically) by a vertical switch.

Every node representing a bend in an optional edge is replaced by one of the *link* boxes shown in Figure 7. The choice of link box depends on the configuration of the optional edge and the two adjacent mandatory edges. There are four different cases; we show two of them in Figures 8 and 9. Depending on the case, we also have to add three points to one of the adjacent boxes; again see Figures 8 and 9.

We use *bend* boxes to represent the nodes at bends of

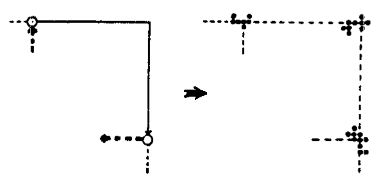


Figure 8: How to choose the boxes for an optional edge: Case 1

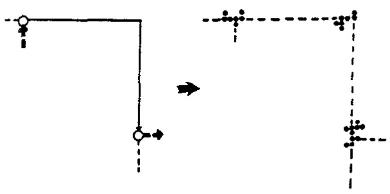


Figure 9: How to choose the boxes for an optional edge: Case 2

Figure 10: Odd 3-bends and 4-bends

Figure 11: Even 3-bends and 4-bends

Figure 12: The two terminal boxes

the mandatory edges. Again, they are positioned such that the circled point is placed on the node and the two line segments for the mandatory edge run as indicated.

Running through any mandatory edge e from its start to its end vertex, we encounter a sequence of 3 or 5 nodes. The first is odd, the second is even, etc. The parity of the bend boxes is chosen accordingly. The type is chosen in the following way: If at its first bend, e makes a right-hand turn, we place a 3-bend; if it is a left-hand turn, we place a 4-bend. Again, the circled position p is placed on the node and the points p_1 and p_2 indicate the direction of the mandatory edge. On any following bend we use the opposite bend type if the edge makes the same type of turn and the same bend type if the edge makes the opposite type of turn. This implies that the points of an odd bend box are always to the right of the mandatory edge. (See Figure 13 for the overall situation.)

The nodes corresponding to s and t are replaced by *terminal boxes* as shown in Figure 12. (With respect to later usage in the following section on area maximization, we make sure that the points t_1 and t_2 of the box for t are extremal in horizontal direction. t can be moved horizontally as far to the right as necessary.) The grid points in all these boxes form a point set $\overline{P}$. A straightforward estimate yields $n := \mid \overline{P} \mid \leq 13\overline{m}$.

We claim: *The resulting point set $\overline{P}$ is the vertex set of a simple polygon of area $\frac{n}{2} - 1$ if and only if D has a Hamiltonian cycle.*

It is not hard to see that any Hamiltonian path H in D' induces a polygon of area $\frac{n}{2} - 1$; see Figures 5 and 14 for an illustration.

Conversely, we can show that any simple polygon on $\overline{P}$ of area $\frac{n}{2} - 1$ induces a Hamiltonian path H in D'. Careful analysis of possible connections by triangles in the polygon ultimately shows that they form a particular structure that can be used to construct a Hamiltonian Path. See Figures 15 and 16 for the two basic ways to connect points from different boxes. For full technical details, see [9]. ∎

Corollary 4.2 *MAP is NP-complete.*

Corollary 4.3 *It is NP-complete to decide whether a set P of vertices allows a simple polygon that contains no point of a given set Q.*

Proof: We can use the same construction as in our NP-completeness proof of GAP to obtain a point set P of size n. Choose a set Q of size $O(n^3)$ in the following way: For any triangle Δ with vertices from P that contains a grid point not in P, add a point q_Δ to Q. Then a Q-avoiding polygon with vertex set P is the same as a grid-avoiding polygon. ∎

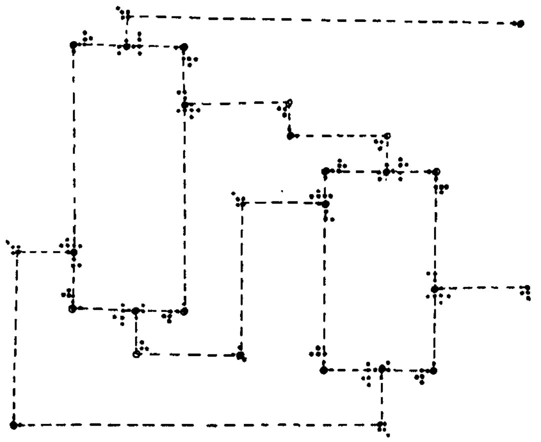


Figure 13: The point set $\overline{P}$

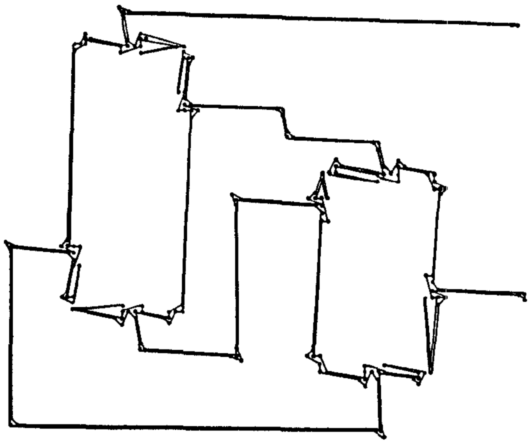


Figure 14: A small polygon for $\overline{P}$

5 Maximal Area

5.1 GAXP is NP-complete

Using our results on GAP, we show that GRID APPROXIMATING POLYGON is NP-complete.

Theorem 5.1 *GAXP is NP-hard.*

Proof: See Figure 17. Consider the point set P in the NP-hardness proof of GAP. In any simple polygon with vertex set P of area $\frac{n}{2}-1$, the points $t_1 := (t_x, t_y - 1)$ and $t_2 := t_1 - (0,1)$ in the terminal box for t are connected to each other. By construction, all other grid points lie to the left of the vertical line through t_1 and t_2. Then add the points

$$\begin{array}{lcll} p_1 & := & t_2 & -(0, N^4) \\ p_2 & := & t_1 & +(0, N^4) \\ p_3 & := & p_1 & -(2N^4+1, 0) \\ p_4 & := & p_2 & -(2N^4+1, 0) \end{array}$$

to P to get the set $\overline{P}$. It is straightforward to see that there is a simple polygon $\mathcal{P}$ on the vertices P that satisfies GAP if and only if there is a simple polygon $\overline{\mathcal{P}}$ on the vertices $\overline{P}$ that satisfies GAXP. (The polygon $\overline{\mathcal{P}}$ is simply the complement of $\mathcal{P}$ in the square with vertices p_1, p_2, p_3, p_4.) ∎

Similarly as for GAP and MAP, we get the following corollaries from the NP-completeness of GAXP.

Corollary 5.2 *MAXP is NP-complete.*

Corollary 5.3 *It is NP-complete to decide whether a vertex set P allows a simple polygon that contains all points of a set Q.*

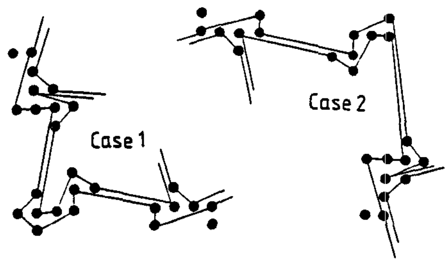


Figure 15: An optional edge contained in H

5.2 A $\frac{1}{2}$-Approximation for MAXP

A lower time bound for any approximation algorithm for MAP is given from the difficulty of just finding a simple polygon on a given set of n vertices. It is well known that finding a simple polygon with a given set of vertices has a lower bound of lower bound of $\Omega(n \log n)$ in the model of algebraic computation trees. Drysdale and Jaromczyk [6] have established a lower bound of $\Omega(n \log n)$ for finding a largest polygon with k vertices chosen from a set of n points in the plane, where k is a given parameter.

In the following, we describe an $O(n \log n)$ method to obtain a simple polygon on a given set of points P whose area is bigger than half the area of the convex hull, *conv*(P), of P. Since the area of the convex hull is an upper bound for any solution of MAXP, this yields a fast approximation method for MAXP.

Theorem 5.4 *Let P be a set of $n+1$ points in the plane. In time $O(n \log n)$ we can determine a simple polygon $\mathcal{P}$ on P that has area larger than $\frac{1}{2}AR(\mathrm{conv}(P))$.*

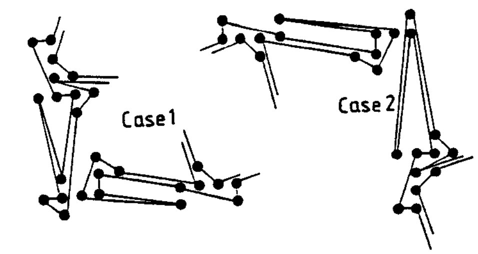


Figure 16: An optional edge not contained in H

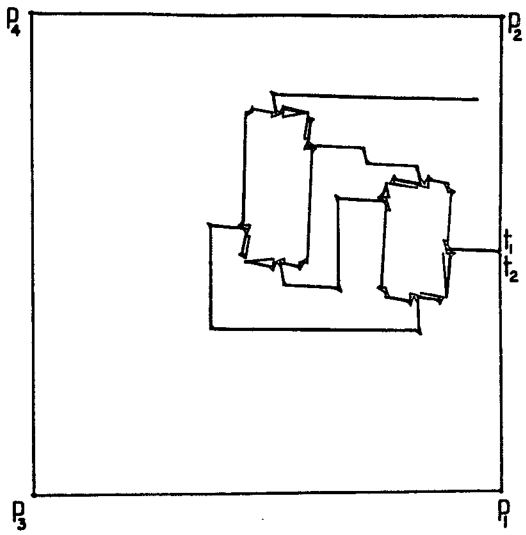


Figure 17: GAXP solves GAP: Turning a small polygon inside out

Proof: Let p_0 be a point on the convex hull of P. In time $O(n \log n)$, sort the points p_i of P by the slope of the lines $l(p_0, p_i)$, such that the neighbours of p_0 on the convex hull are the first and the last point, respectively. If there is a set of points for which the slope is the same, break the tie by ordering them in increasing distance from p_0, except when those points have the smallest of all slopes, in which case we take them in order of decreasing distance from p_0. (The latter corresponds to the line through p_0 and p_n.) Connecting the points p_i in this order yields a simple polygon $\mathcal{P}_1$ on P — see Figure 18. If $AR(\mathcal{P}_1) > \frac{1}{2}AR(conv(P))$, we are done. Suppose this is not the case. Then the set $Q := conv(P) \setminus \mathcal{P}_1$ has area at least $\frac{1}{2}AR(conv(P))$. Let $\overline{P}$ be the set of points that lie on the boundary of $conv(P)$. In the following, we write $\overline{q}_i$ for a point $q_i \in P$ if and only if $q_i \in \overline{P}$. Q consists of $h \geq 1$ polygons $Q_j := \langle \overline{q}_1^{(j)}, q_2^{(j)}, \ldots, q_{z_j-1}^{(j)}, \overline{q}_{z_j}^{(j)}, \overline{q}_1^{(j)} \rangle$, where $\overline{q}_1^{(j)}$ and $\overline{q}_{z_j}^{(j)}$ are the only vertices of Q_j that belong to $\overline{P}$. Now by construction of $\mathcal{P}_1$, for any j, the points $q_2^{(j+1)}$ and $q_{z_j-1}^{(j)}$ can see each other in the polygon $\mathcal{P}_1$

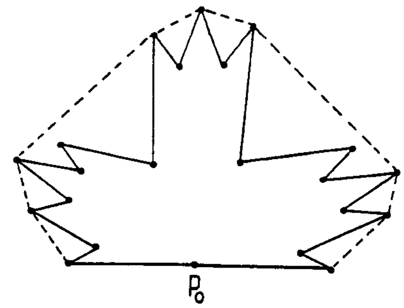


Figure 18: The simple polygon $\mathcal{P}_1$

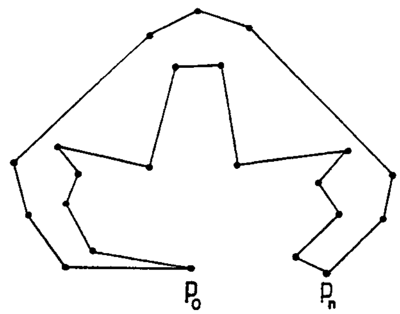


Figure 19: The simple polygon $\mathcal{P}_2$

and the edge from $\overline{p}_n$ to $q_{z_h-1}^{(h)}$ lies in $\mathcal{P}_1$. This implies that the polygon

$$\mathcal{P}_2 = \langle \overline{p}_0, q_2^{(1)}, \ldots, q_{z_1-1}^{(1)}, q_2^{(2)}, \ldots, q_{z_j-1}^{(j)}, q_2^{(j+1)}, \ldots, q_2^{(h)}, \ldots, q_{z_h-1}^{(h)}, \overline{p}_n, \ldots, \overline{p}_1, \overline{p}_0 \rangle$$

is simple: First travel all the points inside the hull, then go back along the hull — see Figure 19.

Since $\mathcal{P}_2$ contains Q as a proper subset, we get $AR(\mathcal{P}_2) > \frac{1}{2}AR(conv(P))$, concluding the proof. ∎

All the estimates involved are tight, even if we try *all* possible starting points p_0 on the convex hull — see the example shown in Figure 20. The points q_1, q_2, q_3 form a suitable approximation of an equilateral triangle spanning the convex hull. The sets P_1, P_2, P_3 all consist of points that are close to the midpoints of the sides of the triangle.

Corollary 5.5 *The same approach can be applied to get a $\frac{1}{2}$-approximation method for maximizing the number of points of a given set Q that are contained in a polygon with vertex set P. As before, the bound is tight.*

5.3 A Bound on Approximation Factors

We can use a similar idea as in Theorem 5.1 to show that there is little hope of finding a simple polygon that encloses more than $\frac{2}{3}$ of the area of the convex hull:

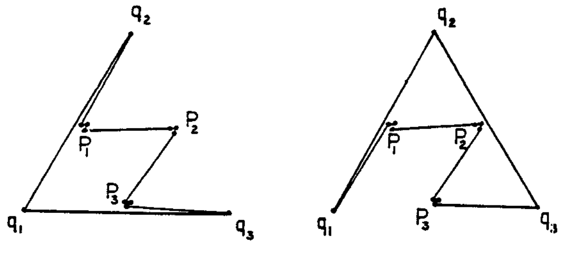


Figure 20: An example for which the bound $\frac{1}{2}$ is tight

Theorem 5.6 *Let $0 < \varepsilon < \frac{1}{3}$. It is NP-complete to decide whether a vertex set P admits a simple polygon of area at least $(\frac{2}{3} + \varepsilon)AR(\text{conv}(P))$.*

Sketch: The idea is similar to the proof of Theorem 5.1. For an instance of GAP with n points as constructed in the NP-completeness proof, add a suitable set of points outside the convex hull, such that deciding the existence of a polygon with at least $\frac{2}{3} + \varepsilon$ of the area of the convex hull solves the GAP instances constructed in the NP-completeness proof of GAP. See Figure 21. We can assume that the n points lie inside the box B with vertices $(\frac{N^4}{2}, \frac{N^4}{2})$, $(\frac{N^4}{2}, -\frac{N^4}{2})$, $(\frac{3N^4}{2}, \frac{N^4}{2})$, $(\frac{3N^4}{2}, -\frac{N^4}{2})$, where $N = n^n$. The terminal box for t is moved to the right so that the points t_1 and t_2 have the coordinates $(2N^4, 1)$ and $(2N^4, 0)$. Finally, add the points $p_1 = (M, -M)$, $p_2 = (M, M)$ and $p_3 = (-2M, 0)$, where

$$M = \frac{4N^4 - 1}{6\varepsilon} + \frac{2N^4 - n + 1}{4N^4 - 1}.$$

With these parameters it is straightforward to check that there is a polygon with at least $\frac{2}{3} + \varepsilon$ of the area of the convex hull if and only if the GAP instance constructed in the NP-completeness proof of GAP allows a grid avoiding polygon. ∎

6 Higher Dimensions

In this section we show that in any fixed dimensions k and d, finding a simple d-dimensional polyhedron with a given set of vertices that has minimal volume of its k-dimensional faces is NP-hard. This answers and generalizes a question stated by O'Rourke [13], [12].

Definition 6.1 *A d-dimensional polyhedron $\mathcal{P}$ is called* simple, *if it is homotopic (topologically equivalent) to a d-dimensional sphere.*

It is feasible *for a given vertex set P, if every vertex of $\mathcal{P}$ belongs to P and every point in P is contained in at least $d-1$ different faces of $\mathcal{P}$.*

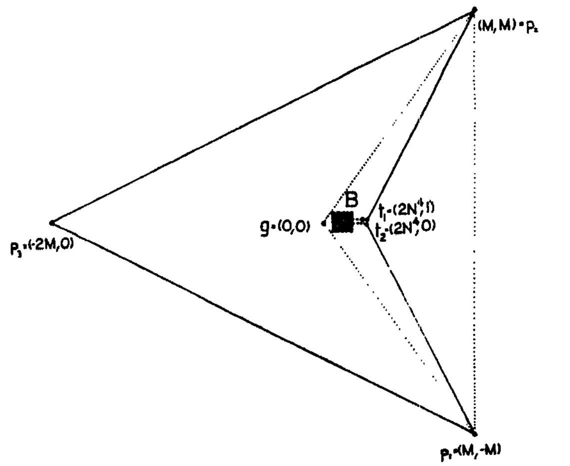


Figure 21: Finding a polygon with area at least $(\frac{2}{3} + \varepsilon)AR(conv(P))$ solves GAP

The generalization of simplicity from the two-dimensional situation is clear. The reason for considering feasibility in the stated form is the following: We want all points in the given set P to carry some significance for the polyhedron. In the two-dimensional case, however, a "vertex" of a simple polygon is usually not required to be locally extreme; it can very well be the common end point of two adjacent collinear edges. We account for this situation with the above definition. Any point in P is at least contained in an edge of $\mathcal{P}$.

O'Rourke [13], [12] has proposed the following problem:

MINIMAL SURFACE POLYHEDRON (SURF):
Given a finite set P of points in 3-dimensional Euclidean space. Among all simple polyhedra that are feasible for a vertex set P, find one with the smallest surface area.

This is a special case of the following problem:

MINIMAL FACE POLYHEDRON (FACE):
Let $2 \le d$ and $1 \le k \le d$. Given a finite set P of points in d-dimensional Euclidean space. Among all simple polyhedra that are feasible for vertex set P, find one with the smallest volume of its k-dimensional faces.

It turns out that FACE is NP-hard for any choice of d and k. We describe the basic idea of the proof for the special case of SURF:

Theorem 6.2 SURF *is NP-hard.*

Proof: We describe a reduction of HAMILTONIAN CYCLE IN GRID GRAPHS. Take any instance of HCGG, i.e. a grid graph G with n vertices. This grid graph

G can be canonically represented by a set P_G of n points in the plane $E_2 = \{(x, y, 0) \mid x, y \in R\}$. Let $g = (x_g, y_g, 0)$ be the center of mass of the points in P_G and $p^* = (x_g, y_g, H)$, where $H = 2n^3$.

Now for every simple polyhedron $\mathcal{P}$ on P, the domain $\mathcal{P}_2 = \mathcal{P} \cap E_2$ must be a simple two-dimensional polygon, i.e. the edges of $\mathcal{P}_2$ form a tour. A straightforward ananlysis of the given parameters shows that for the set $P := P_G \cup \{p^*\}$, there is a polyhedron with the required properties of surface area $n^4 + n^2 + \frac{1}{8}$ or less *if and only if* there is a (Euclidean) tour of length n or less in P, i.e. a Hamiltonian cycle in the grid graph. ∎

In a similar way, we can prove the NP-hardness of FACE:

Theorem 6.3 FACE *is NP-hard.*

Sketch: We consider 3 cases:
(A) $k = d$
(B) $k = 1$
(C) $1 < k < d$.

Consider a a set P_G of n points in the plane $E_2 = \{(x_1, x_2, 0, \ldots, 0) \mid x_1, x_2 \in R\}$, representing an instance of MAP in case (A), or a grid graph (cases (B) and (C).) Let g be the center of mass of the points in P_G and $p_i = g + H\dot{e}_i$ for $i = 3, \ldots, d$, where $H = 9d^2n^3$ and e_i denotes the ith unit vector. Then consider the point set $P_d = P_G \cup \bigcup_{i=3}^{d} \{p_i\}$.

In case (A), it is easy to show by induction that there is a simple polyhedron feasible for the vertices P_d of volume at most

$$VOL_d = \frac{2}{d!} H^{d-2} \left(\frac{n}{2} - 1\right),$$

if and only if there is a simple polygon on the vertices of P_G with area at most $\frac{n}{2} - 1$.

For the other two cases, the proof is similar to the hardness proof of FACE, but the calculations get more involved. ∎

Acknowledgements: The authors would like to thank Jit Bose, David Shallcross, Joe O'Rourke and Gerhard Woeginger for helpful discussions.

References

[1] A.Aggarwal, H.Imai, N.Katoh, S.Suri, Finding k Points with Minimum Diameter and Related Problems, *Journal of Algorithms*, 12 (1991), 38–56.

[2] J.E.Boyce, D.P.Dobkin, R.L.Drysdale, L.J. Guibas, Finding Extremal Polygons, *SIAM Journal of Comput.*, 14 (1985), 134–147.

[3] B.Chazelle, Triangulating a Simple Polygon in Linear Time, *Discrete and Computational Geometry*, 6 (1991), 485–524.

[4] H.S.M.Coxeter, *Introduction to Geometry*, Wiley, 1969.

[5] D.P.Dobkin, R.L.Drysdale, L.J.Guibas, Finding Extremal Polygons, in: *Advances in Computing Research*, Vol.1 (1983), 181–214.

[6] R.L.Drysdale, J.W.Jaromczyk, A Note on Lower Bounds for the Maximum Area and Maximum Perimeter k-gon Problems *Information Processing Letters*, 32 (1989), 301–303.

[7] D.Eppstein, M.Overmars, G.Rote, G.Woeginger, Finding minimum area k-gons, *Discrete and Computational Geometry*, 7 (1992), 45–58.

[8] P.Erdős, P.M.Gruber, J.Hammer, *Lattice Points* , Longman Scientific & Technical (copublished with Wiley, New York), 1989.

[9] S.P.Fekete, *Geometry and the Travelling Salesman Problem*, Ph.D. thesis, University of Waterloo, Waterloo, Ontario, Canada, 1992.

[10] M.R.Garey, R.L.Graham and D.S.Johnson, Some NP-complete geometric problems, *Proceedings of the 8th Annual Symposium on Theory of Computing*, Association for Computing Machinery, 1976, 10–22.

[11] M.R.Garey and D.S.Johnson, *Computers and Intractability, A guide to the theory of NP-completeness*, Freeman, San Francisco, 1979.

[12] J.O'Rourke, *Polyhedral Object Models from 3D Points*, Universität Hamburg Technical Report IFI-HH-M-77/80, July 1980.

[13] J.O'Rourke, Personal Communication, 1992.

[14] G.Pick, Geometrisches zur Zahlenlehre, *Naturwissenschaftliche Zeitschrift Lotus* (Prag) (1899), 311–319.

[15] J.Plesńik, The NP-Completeness of the Hamiltonian Cycle Problem in Planar Digraphs with Degree Bound Two, *Information Processing Letters*, 8 (1979), 199–201.

[16] P.Rosenstiehl, R.E.Tarjan, Rectilinear Planar Layouts and Bipolar Orientations, *Discrete and Computational Geometry*, 1 (1986), 342–351.