Discrete Comput Geom 23:73–110 (2000)
DOI: 10.1007/s004549910005

Discrete & Computational
Geometry

On Simple Polygonalizations with Optimal Area*

S. P. Fekete

Department of Mathematics, TU Berlin,
Str. des 17. Juni 136,
D-10623 Berlin, Germany

Abstract. We discuss the problem of finding a simple polygonalization for a given set of vertices P that has optimal area. We 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 and 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. This answers a generalization of a question stated by Suri in 1989. Finally, we turn to higher dimensions, where we prove that, for $1 \leq k \leq d$, $2 \leq d$, it is NP-hard to determine the smallest possible total 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.

1. Introduction

While the classical geometric Travelling Salesman Problem is to find a (simple) polygon with a given set of vertices that has shortest perimeter, it is natural to look for a simple polygon with a given set of vertices that minimizes another basic geometric measure: the enclosed area.

Minimum Area Polygonalization (MIN-AREA). Given a finite set P of points in the Euclidean plane. Among all the simple polygons with vertex set P, find one with minimal enclosed area.

The related problem of *maximizing* the enclosed area of a polygonalization is called MAX-AREA.

* A preliminary version of this paper, titled "Minimum Area Polygons," was presented at the 1993 ACM Symposium on Computational Geometry.

Problems of the type MIN-AREA and MAX-AREA arise in the context of like pattern recognition, image reconstruction, and clustering; higher-dimensional variants play a role in the modeling of technical objects, as well as optimal surface design.

The complexity of these problems has been open for a while. In 1980 O'Rourke [16] considered the complexity of a three-dimensional variant, where a given set of vertices has to be covered with a simple polyhedralization of small surface area. At the first Canadian Conference on Computational Geometry in 1989, Suri posed the complexity of MIN-AREA as an open problem. In addition, there has been some research on optimal polygonalizations of a given vertex set (see [2], [4], [5], and [25]) that focussed on finding subpolygons with certain special properties, e.g., convexity. Typically, the results are fairly general and "area" is only one special case of a measure of simple polygons for which an algorithm works—"perimeter" usually being among the others.

One particularly nice aspect of the area of a simple polygon is provided by *Pick's theorem*. It states that the area $AR(\mathcal{P})$ of any simple polygon $\mathcal{P}$ with integral vertices can be expressed as a simple linear function of the number of grid points that $\mathcal{P}$ encounters. This yields strong connections to problems dealing with the separation of point sets, and it yields an easy lower bound for the area, which is met iff no new grid points are encountered. As a consequence, we get a close connection to problems arising from the separation of point sets by means of polygonal curves which have been considered by Mitchell and Suri [14], Mitchell [13], Mata and Mitchell [12], and Aggarwal and Suri [1].

When dealing with structures of small area, one encounters specific difficulties. Most notably, edges in a polygon with small area need not be short. This makes it difficult to restrict potential neighbors of a point in a good polygonalization, inhibiting local search methods for efficient algorithms on one hand, but also straightforward component design for a proof of NP-hardness on the other. (See Fig. 1 for an illustrative example.)

The main result of this paper is to resolve the open questions by O'Rourke and Suri by giving proofs of NP-completeness for their respective problems. There are consequences for related problems.

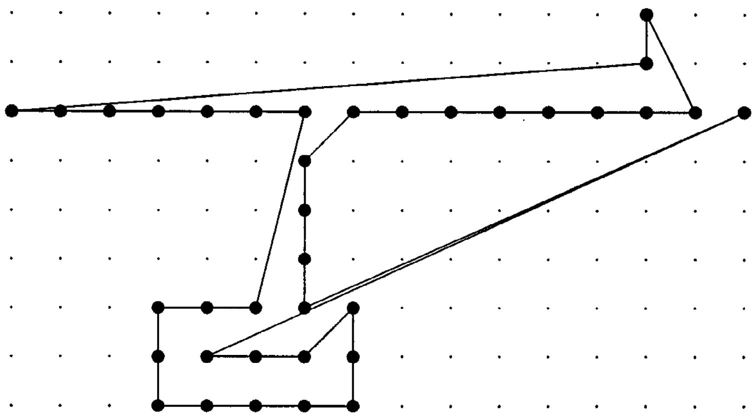
The rest of the paper is organized as follows:

In Section 2 we give a description of Pick's theorem. We note that Pick's theorem yields easy upper and lower bounds for the area of a simple polygon on a given vertex set.

Section 3 contains an NP-completeness proof for the problem GRID-EMPTY: Is there a simple polygon that connects a given set of grid points and does not contain any other grid points on its boundary or in its interior? Since this question is a strong version of MIN-AREA, the result implies NP-completeness of MIN-AREA.

In Section 4 we shift our attention to the problem MAX-AREA of finding a simple polygonalization of a given point set with maximal area. We show how a similar NP-completeness result of MAX-AREA follows directly from the result for GRID-EMPTY.

In Section 5 we consider related problems in higher dimensions. We show that for 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 [16].



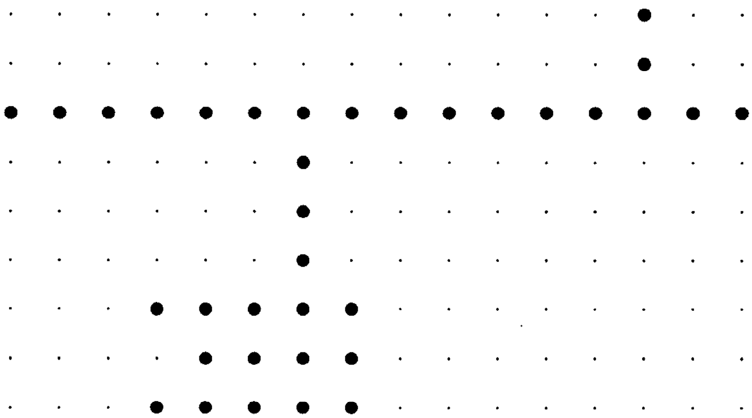


Fig. 1. Two horizontally convex sets, with and without a grid-empty polygonalization.

The concluding Section 6 gives a brief discussion of connections to other problems.

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 $AR(P)$?

Algorithmically, this problem can be solved quite efficiently, e.g., see [18]. A surprising and elegant answer of a different type is provided by Pick's theorem (see

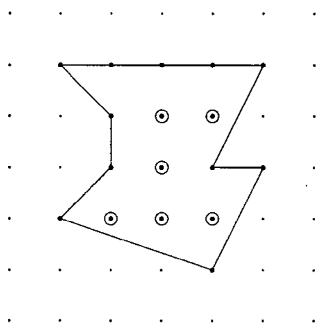


Fig. 2. Pick's theorem.

Fig. 2):

Theorem 2.1 [19]. *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*

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

An elegant proof can be found in [3]. For a discussion of alternative approaches see the article by Niven and Zuckermann [15]. There are numerous generalizations to other than the orthogonal grid, e.g., by Ren and Reay [23]; see [22] for a generalization to higher dimensions.

Pick's theorem yields a combinatorial interpretation for finding a polygon with minimal or maximal possible area. (See Fig. 3.) Any grid point that is contained in

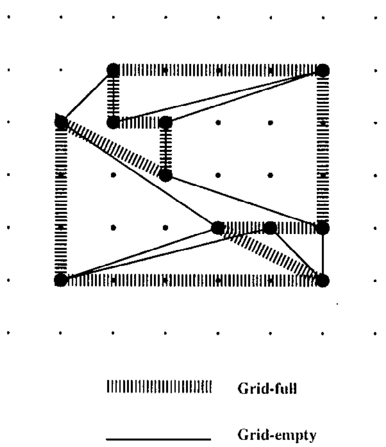


Fig. 3. A set of grid vertices with a grid-empty polygonaliztion and a grid-full polygonalization.

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 $n/2 - 1$.

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 will contribute 1 when contained in the interior of the polygon.

We summarize this upper and lower bound:

Theorem 2.2. *Let P be a set of n points in the plane that all have integer coordinates. Let $h_{\mathrm{i}}(P)$ denote the number of points of the integer grid that are not contained in P and strictly inside the convex hull, and let $h_{\mathrm{b}}(P)$ be 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*

$$\frac{n}{2} - 1 \le AR(\mathcal{P}) \le \frac{n}{2} + \frac{h_{\mathrm{b}}(P)}{2} + h_{\mathrm{i}}(P) - 1.$$

These bounds suggest the following questions that are closely related to MIN-AREA and MAX-AREA:

Grid-Empty Polygonalization (GRID-EMPTY). 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, which is equivalent to having area $n/2 - 1$?

Grid-Full Polygonalization (GRID-FULL). Given n grid points in the plane. Is there a simple polygon on this vertex set that contains as many additional grid points as well as possible, which is equivalent to having area $n/2 + h_{\mathrm{b}}(P)/2 + h_{\mathrm{i}}(P) - 1$?

We show in the following section that these problems are NP-complete. This implies that MIN-AREA or MAX-AREA cannot be solved in polynomial time, unless P = NP.

3. The Complexity of Minimum Area Polygonization

In this section we give a proof of our main result:

Theorem 3.1. GRID-EMPTY *is NP-complete.*

It is clear that the problem is in NP. To show that it is NP-hard, we give a reduction of HAMILTONIAN CYCLE IN CUBIC PLANAR DIRECTED GRAPHS, which was shown to be NP-complete by Plesńik [21]. That is, for a given cubic planar digraph D, we construct a point set P_D in polynomial time, such that P_D can be described in polynomial space, and P_D admits a GRID-EMPTY POLYGONALIZATION if and only if D has a Hamiltonian cycle.

The idea for this construction is as follows:

After some minor rearrangements of the cubic planar 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. The endpoints of line segments are replaced by suitable sets of grid points ("clusters"). The resulting point set is perturbed, in order to guarantee that there are no collinearities between nonadjacent endpoints of line segments. It turns out that a Hamiltonian path in the graph corresponds to a very narrow polygonization of P_D 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 clusters 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. (This is somewhat similar to the idea contained in the NP-hardness proof for HAMILTONIAN CYCLE IN GRID GRAPHS described in [10], [11], and [20]. See [8] and [9] for related techniques.)

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 in the Introduction, dealing with areas instead of distances makes it hard to localize neighbors in a set of points. We achieve the desired localization by the perturbation mentioned above.

3.1. *Basic Observations on Cubic Planar Digraphs*

Throughout, we use the notation $\langle u, v\rangle$ to denote a directed arc from u to v in a digraph; a directed path from v_1 to v_n, consisting of edges $\langle v_1, v_2\rangle$, $\langle v_2, v_3\rangle, \ldots, \langle v_{n-1}, v_n\rangle$, is denoted by $\langle v_1, \ldots, v_n\rangle$.

Now consider any cubic planar digraph D. There are a few easy assumptions that we can make about the digraph D when we want to test it for Hamiltonicity: if D has a vertex with in-degree 3 or out-degree 3, there can be no Hamiltonian Circuit, so all vertices must have either in-degree 2 or out-degree 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. So optional edges for a vertex come in incoming or outgoing pairs.

Assume there was an edge that was mandatory for one of its endpoints v_1, but optional for its other endpoint v_2. We could delete the other optional edge of v_2 without changing the Hamiltonicity of the graph. The resulting vertices of degree 2 and their two adjacent edges can be replaced by a single edge. We can continue this process until all edges are either mandatory or optional for both their vertices. (See Fig. 4.)

So 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. Let V be the set of vertices of D, and $|V| = m$; since D is cubic, it has $3m/2$ edges; as a planar graph, it has $m/2+2$ faces. Consider the undirected graph G with vertex set V, obtained by replacing all arcs of D with edges. Then the optional edges of D induce a set of vertex-disjoint cycles in G.

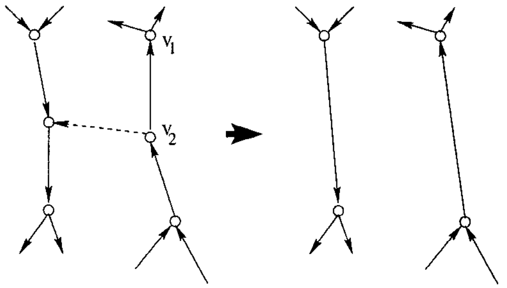


Fig. 4. Making all edges mandatory or optional for both their vertices.

3.2. *Embedding the Cubic Planar Digraph*

In this section we show the following lemma:

Lemma 3.2. *Let D be a cubic planar digraph with in-degree 1 or 2 for each vertex, and each edge optional or mandatory for both its end vertices. Then D can be drawn in the plane, such that all edges are represented by rectilinear paths, with precisely two line segments for the optional edges, and at most eleven line segments for the mandatory edges. Furthermore, all endpoints of line segments ("joints") have coordinates which are multiples of $\frac{1}{8}$ in the range between 0 and m, and no two line segments are collinear.*

Proof. We use three intermediate steps to produce the required embedding of D:

(a) Contract a perfect matching of D to get a 4-regular planar digraph $\overline{D}$.
(b) Represent $\overline{D}$ by a planar drawing.
(c) Use the drawing of $\overline{D}$ to get a drawing of D with the required properties.

(a) 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. For an edge $e = \langle v_1, v_2 \rangle$ in a chosen matching, let $o_+ = \langle v_3, v_2 \rangle$ be the optional edge adjacent to v_2, let $e_- = \langle v_2, v_4 \rangle$ be the mandatory edge adjacent to v_2, let $o_- = \langle v_1, v_5 \rangle$ be the optional edge adjacent to v_1, and let $e_+ = \langle v_6, v_1 \rangle$ be the mandatory edge adjacent to v_1. Now all selected edges e are contracted: replace v_1 and v_2 by a single vertex $v_{1,2}$, and replace the edges o_+, e_-, o_-, e_+ by $\langle v_3, v_{1,2} \rangle$, $\langle v_{1,2}, v_4 \rangle$, $\langle v_{1,2}, v_5 \rangle$, $\langle v_6, v_{1,2} \rangle$. For easier notation, we still write o_+, e_-, o_-, e_+ for the corresponding new edges. The resulting 4-regular planar digraph $\overline{D}$ has $m/2$ vertices, m edges, and $m/2 + 2$ faces, which can be identified with the faces of D. (See Fig. 5; a graph $\overline{D}$ to D is shown in Fig. 6(a).)

(b) For this step, we make use of a particular way of representing a planar graph in the Euclidean plane: in a *rectilinear planar layout*, every vertex is represented by a horizontal line segment, every edge is represented by a vertical line segment. Two vertices are connected by an edge if and only if the corresponding horizontal line segments have nonempty intersection with the vertical line segment representing the edge. (See Fig. 7(b) for a planar rectilinear layout for the graph shown in Fig. 5.)

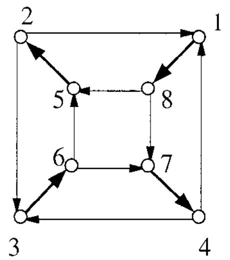


Fig. 5. A cubic planar digraph D.

Rosenstiehl and Tarjan have described in [24] how every planar graph can be represented by such a layout. Their method uses a so-called *bipolar labeling* of a bridgeless planar graph: this is a vertex labeling (from 1 to $m/2$ in the case of $\overline{D}$), such that orienting each edge from lower to higher label produces a digraph with a unique source 1 and a unique sink $m/2$, and there is an edge e^* from source to sink. (Without loss of generality we may assume that e^* is mandatory.) This labeling induces an edge labeling from 1 to m by labeling e^* with m and using the lexicographic order of vertex labels for all other edges. Furthermore, the bipolar vertex labeling also induces a bipolar order of the dual graph. See Fig. 6(b) for a bipolar orientation D^* of the graph $\overline{D}$. (The bipolar labeling of the dual is indicated by the alphabetic order of the letter labels; note that the oriented edges in the dual cross the edges in the primal from left to right, with the exception of e^*.) Once a labeling is obtained (which is possible in linear time), we get a planar rectilinear layout: let v be a vertex with vertex label $y(v)$, and let $\min(v)$ and $\max(v)$ be the lowest and highest edge labels of edges incident to v. Then v is represented by the horizontal line segment from $(\min(v), y(v))$ to $(\max(v), y(v))$. Any edge $e = \langle v_i, v_j \rangle$ with edge label $x(e)$ is represented by a vertical line segment from $(x(e), y(v_i))$ to $(x(e), y(v_j))$. (A proof that this is indeed a feasible rectilinear layout uses the dual bipolar orientation and can be found in [24].)

In this representation of D^*, we can orient all edges appropriately to get a planar rectilinear layout of $\overline{D}$; clearly, no two segments are collinear. Note that the edge e^* is represented by the segment $[(m, 1), (m, m/2)]$, marking the rightmost edge of the bounding box of the layout.

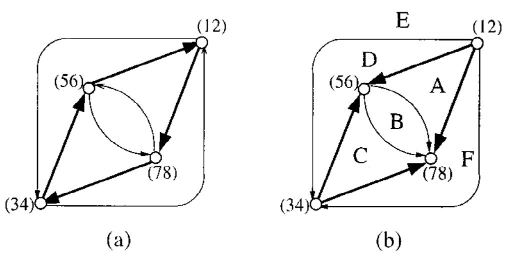


Fig. 6. (a) A 4-regular planar digraph $\overline{D}$; (b) a bipolar orientation D^* for $\overline{D}$.

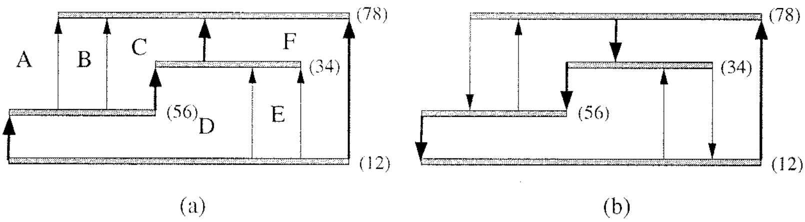


Fig. 7. A planar rectilinear layout for (a) D^* and (b) $\overline{D}$.

(c) Now we use the planar rectilinear layout to get an embedding of D. For easier notation, we identify vertices and edges with the line segments or points representing them, and we continue to identify the edges o_+, o_-, e_+, e_- in $\overline{D}$ with the corresponding edges in D.

For an intuitive idea of this step, see Fig. 8.

Every vertex segment v has two edge segments adjacent to it that correspond to optional edges, an edge o_+ coming in and an edge o_- going out. Similarly, every vertex segment is adjacent to one incoming mandatory edge segment e_+ and one outgoing mandatory edge segment e_-.

Let $\langle v_1, v_2 \rangle$ be the directed optional edge in D that was contracted to v. In D, the two edges o_+ and e_- bound a face f that corresponds to a face f in $\overline{D}$, which must still be bounded by o_+ and e_-. Similarly, o_- and e_+ must bound the same face g in $\overline{D}$ and D.

For all vertex segments v, v_1 is placed on the vertex segment v between $(x(o_-), y(v))$ and $(x(o_+), y(v))$, at a distance of $\frac{1}{8}$ from $(x(o_-), y(v))$; let $x(v_1)$ be the corresponding x-coordinate. Place v_2 on the edge segment $o_+ = \langle u, v \rangle$ between $(x(o_+), y(v))$ and

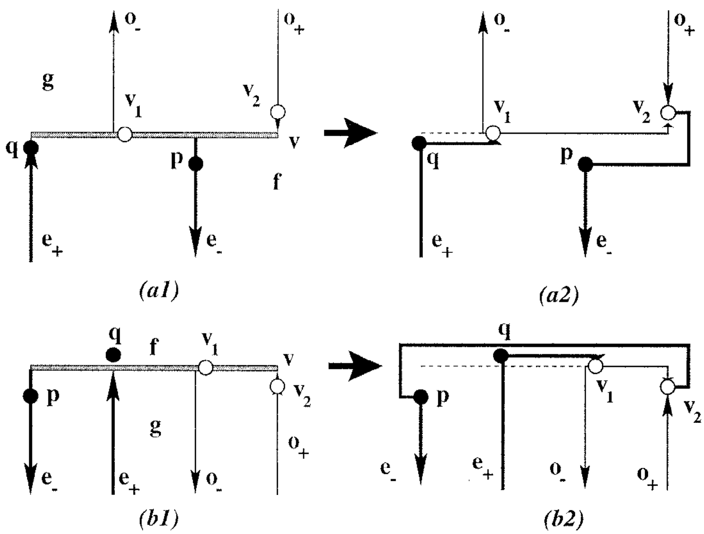


Fig. 8. Replacing vertex segments by pairs of edges and rerouting the edges.

$(x(o_+), y(u))$, at a distance of $\frac{1}{4}$ from $(x(o_-), y(v))$; let $y(v_2)$ be the corresponding x-coordinate. Now we reroute the edges o_+, o_-, e_+, e_- around any vertex segment v to get a feasible drawing.

The optional edge from v_1 to v_2 is represented by the path $\langle v_1, (x(o_+), y(v)), v_2\rangle$. Like v_2, the vertex v_5 in the edge $o_- = \langle v_1, v_5\rangle$ has been placed at a vertical distance of $\frac{1}{4}$ from the end of o_-; then the edge o_- is represented by the path $\langle v_1, (x(o_-), y(v)), v_5\rangle$. Like v_1, the vertex v_3 in the edge $o_+ = \langle v_3, v_2\rangle$ has been placed at a horizontal distance of $\frac{1}{8}$ from the end of o_+, at y-coordinate $y(v_3)$; then the edge o_+ is represented by the path $\langle v_3, (x(o_+), y(v_3)), v_2\rangle$. Because these paths use pieces of the edges of the planar drawing, and none of them more than once, this is a feasible routing.

Now consider the edge $e_- = \langle v_2, v_4\rangle$ in D (see Fig. 8(a1), (b1)). Without loss of generality assume that e_- reaches the vertex segment v from below. Since o_+ and e_- bound the same face f of $\overline{D}$ and f has width at least 1 in each direction, there is an axis-parallel path from v_2 to $p = (x(e_-), y(v) - \frac{3}{8})$ that does not leave f; more precisely, we can move from v_2 in the x-parallel direction until we have reached a horizontal distance of $\frac{1}{4}$ from the boundary of f, then trace the boundary of f at a distance of $\frac{1}{4}$ by moving parallel to o_+, v, e_-, until reaching a horizontal distance of $\frac{1}{4}$ from p, from where the path connects horizontally to v_2. This path consists of at most five line segments (with the extreme case shown in Fig. 8(b2)), and since it stays within f, it does not intersect any line segments of other paths. Moreover, it keeps a distance of $\frac{1}{4}$ to the vertex segment v.

Finally, consider the edge $e_+ = \langle v_6, v_1\rangle$ in D (again, see Fig. 8). If v_1 lies on the boundary of the face g incident to e_+ and o_- (this case is not shown in the figure), we can proceed as in the previous paragraph and trace the boundary of g at a distance of $\frac{1}{8}$. This path consists of at most five line segments, and since it stays within g, it does not intersect any line segments of other paths. Therefore, it remains to deal with the case where v_1 does not lie on the boundary of g, implying that the edge segment o_- is closer to the edge segment e_+ than the edge segment o_+. Edge segment o_- reaches the vertex segment v from above (shown in Fig. 8(a1)), or from below (Fig. 8(b1)). In case (a), let $q = (x(e_+), y(v) - \frac{1}{8})$, in case (b), consider $q = (x(e_+), y(v) + \frac{1}{8})$. Then $\langle q, (x(v_1), y(v) \mp \frac{1}{8}), (x(v_1), y(v)), v_1\rangle$ is a rectilinear path from q to v_1. By construction, q is the only point on the boundary of g encountered by this path, so o_- cannot be intersected. Since the path stays within distance $\frac{1}{8}$ from the edge segment v, it cannot intersect the path for e_- that was described in the preceding paragraph. Since it always remains closer to the edge segment o_- than to the edge segment o_+, it cannot intersect the path for o_+.

As a result, we get a drawing of D where every optional edge is represented by a rectilinear path of two line segments, every mandatory edge is represented by a rectilinear path of at most eleven line segments, and all coordinates determining line segments are multiples of $\frac{1}{8}$, in the range between 0 and m. It is easy to check that step (c) does not introduce any new collinearities.

This proves the lemma. □

We noted above that the edge segment for the mandatory edge $e^* = \langle v_a, v_z\rangle$ is given by $[(m, 1), (m, m/2)]$. If we replace the edge $\langle v_a, v_z\rangle$ by the two edges $\langle s, v_z\rangle$ and $\langle v_a, t\rangle$, we get a digraph D' that has a Hamiltonian path from s to t if and only if D

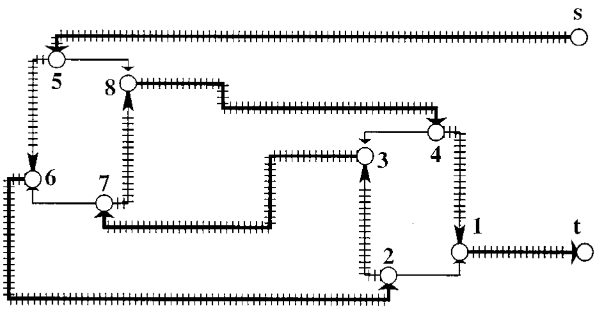


Fig. 9. A rectilinear drawing of D', and a Hamiltonian path.

has a Hamiltonian cycle. (For reasons that become clear in Section 4 on maximal area, we make the position of t extremal by moving t to $(m+1, t_y)$, thereby extending the adjacent line segment by 1.) (See Fig. 9 for a drawing of D', with a Hamiltonian path indicated by marked arcs.) Clearly, the described embedding for D implies a similar embedding for D'.

3.3. *Replacing Nodes by Point Sets*

We proceed to construct a point set P_D that will serve as an input for GRID-EMPTY. In the following, a *joint* is a point in the embedding that separates two straight line segments. So a joint either represents a vertex of D', or it is a point where two line segments meet that belong to the representation of the same edge of D'. We let M denote the set of joints, and $\overline{m}$ the number of joints. It is not hard to see that $\overline{m} \leq 12m$.

In order to have integer coordinates and sufficient space for the following construction, we start by multiplying all the coordinates in the embedding by a factor of $8N^8$, where $N := (13\overline{m})^{39\overline{m}}$. The number of bits to describe N is polynomial in m, so the resulting input for GRID-EMPTY is still polynomial in m.

At each joint, we place an appropriate set of points that form a connected subset of the integer grid—called a *cluster*. We use four different types of clusters:

- **terminal clusters** for the end vertices s and t,
- **bend clusters** for the (at most ten) bends in a path representing a mandatory edge,
- **link clusters** for the (unique) link in a path representing an optional edge,
- **switch clusters** for the degree 3 vertices of D'.

The joints corresponding to s and t are replaced by *terminal clusters* as shown in Fig. 10. The circled positions correspond to the location of the joints, and the underlying grid is spaced at distance 1. (With respect to later usage in the following section on maximum area, we make sure that the points t_1 and t_2 of the clusters for t are extremal in the horizontal direction. Note that t can be moved horizontally as far to the right as necessary.)

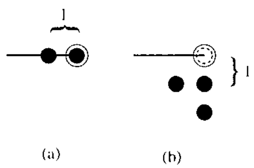


Fig. 10. The terminal cluster for (a) s and (b) t.

The joints at bends in the paths representing mandatory edges are replaced by one of four types of *bend clusters* as shown in Fig. 11. Again, they are positioned such that the circled point is placed on the joint and the two line segments for the mandatory edge run as indicated. The particular choice of type is done as follows. Running through any mandatory edge e from its start to its end vertex, we encounter a sequence of up to ten joints. The first has an odd parity bit, the second even, etc. The turn bit is 1 at a left-hand turn, and 0 at a right-hand turn. Note that this implies that the points of an odd bend cluster are always to the right of the mandatory edge. The only function of the combination of a parity bit and a turn bit is to allow a feasible connection of successive bend clusters along a mandatory edge. (See Fig. 15 for the overall situation.)

Every joint corresponding to the unique link in the path representing an optional edge is represented by one of two types of *link clusters*—see Fig. 12. The circled point denotes the location of the joint, the underlying lines indicate the path representing the optional edge. (In (b), the vertical line runs either above or below the node, depending on the case shown in Fig. 14.) The indicated triple of points (called a *tab*) inhibits the choice of an additional tab triple in the switch cluster that is adjacent in the shown direction—see below. The choice of link cluster depends on the configuration of the optional edge and the two adjacent mandatory edges. There are four different cases (shown in Fig. 14) that will be discussed further below.

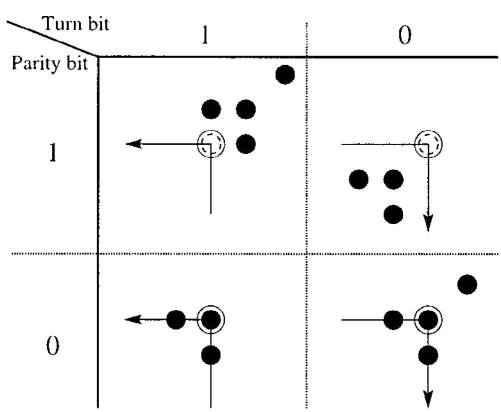


Fig. 11. Bend clusters.

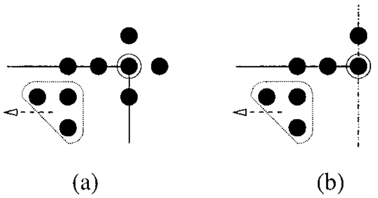


Fig. 12. The two types of link clusters.

Every joint corresponding to a vertex of degree 3 in D' is replaced by a *switch cluster*, consisting of seven, ten, or thirteen points as shown in Figure 13. Out-vertices (represented by joints with both optional edges leaving horizontally) are replaced by a horizontal switch cluster, in-vertices (with optional edges entering vertically) by a vertical switch. In addition to the seven basic points (shown solid in the figure), we add a tab triple at the dotted location, iff the adjacent link cluster in the indicated direction does not have an inhibiting tab triple. Each switch cluster has two special points called *pivot points*; as it turns out, they prevent any grid-empty polygon from containing a path from the points of a switch cluster to the points of more than two of the three adjacent switch clusters. (This idea is illustrated in Figs. 18 and 19 and will be formalized and proven by a series of lemmas.)

Finally, we specify the specific choice of link cluster for a link ℓ in the drawing of D'. Any optional edge $o = \langle u, v\rangle$ in D' forms a path of length 3 with an incoming mandatory edge e_+ and an outgoing mandatory edge. It has orthogonal turns at the locations for u, ℓ, and v. We only describe the situation of a right-hand turn at u—the situation for a left-hand turn is symmetric. There are four different cases—see Fig. 14(a)–(d):

(a) The turn at ℓ is right-hand, the turn at v is left-hand.
(b) The turn at ℓ is right-hand, the turn at v is right-hand.
(c) The turn at ℓ is left-hand, the turn at v is right-hand.
(d) The turn at ℓ is left-hand, the turn at v is left-hand.

The grid points in all these clusters form a point set P. A straightforward estimate yields $n := |P| \leq 13\overline{m}$.

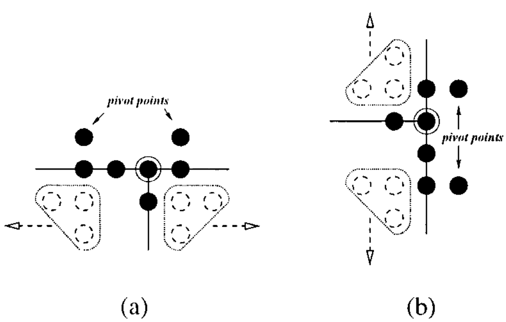


Fig. 13. Horizontal and vertical switch.

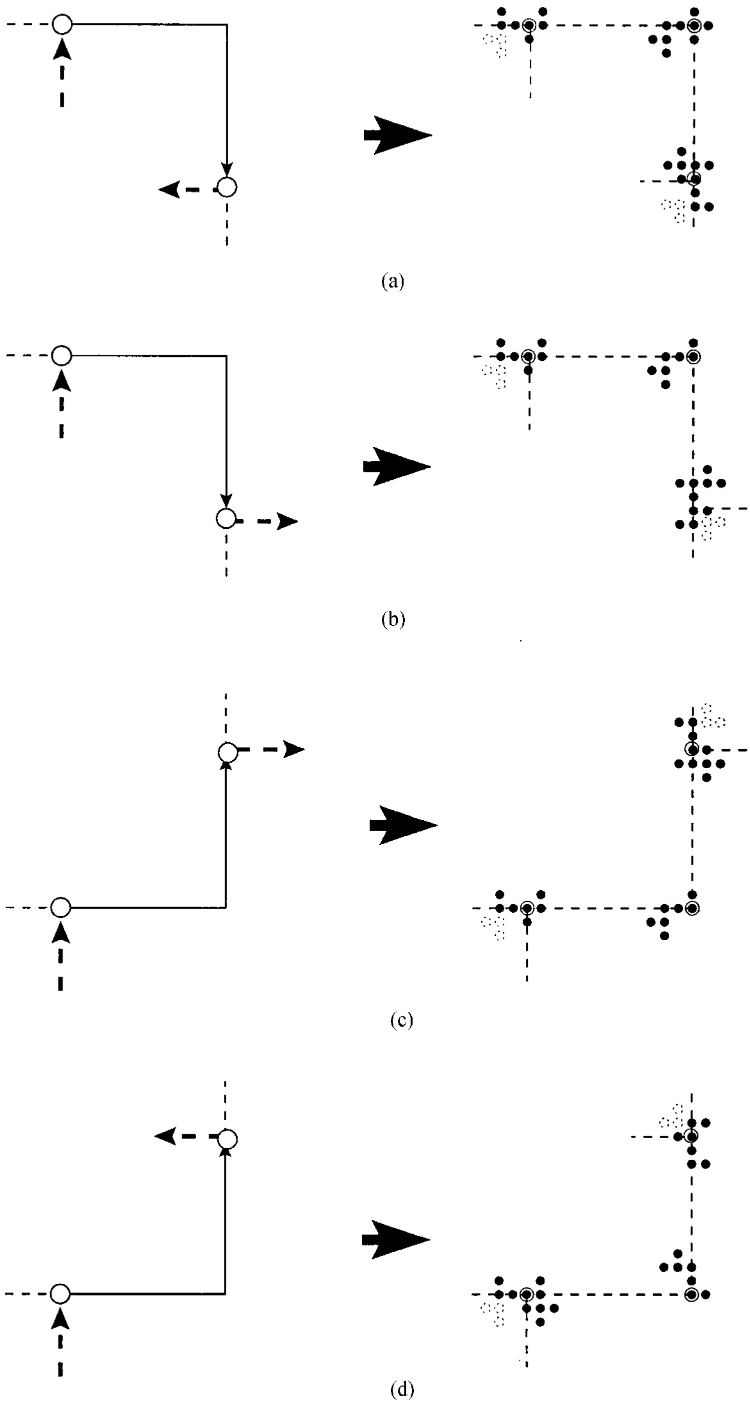


Fig. 14. How to choose the clusters for an optional edge.

3.4. *Perturbing the Point Set*

As noted in the Introduction, excluding stray connections is an important issue. We preempt these unwanted connections in P by a suitable perturbation of the point set.

We think of x-coordinates as coordinates in the horizontal direction and y-coordinates as coordinates in the vertical direction. Now partition the clusters into *horizontal classes*. Two clusters belong to the same horizontal class if their corresponding joint locations (circled in the figures) have identical y-coordinates. Since the original drawing given by Lemma 3.2 does not have any collinear line segments, each class consists of two or three joints, where three joints occur in the case of an out-vertex with the two adjacent joints representing the bends of the two adjacent optional edges. This implies that there are less than $\overline{m}/2$ vertical classes. We shift all points in the ith horizontal class by the vector $(0, N^2 n^{2i})$.

Similarly, define vertical classes and shift the points in the ith horizontal class by the vector $(N^4 n^{2i}, 0)$. We denote by P_D the point set resulting from both sets of shiftings. (See Fig. 15. Note that shiftings have been simplified in the drawing to save space.)

These perturbations have a very useful consequence:

Lemma 3.3. *Let $\Delta = (p, q, r)$ be a triangle with p, q, $r \in P_D$ and $AR(\Delta) = \frac{1}{2}$. Then p, q, r are from the same cluster, or Δ has an edge of length* 1 *whose altitude has length* 1.

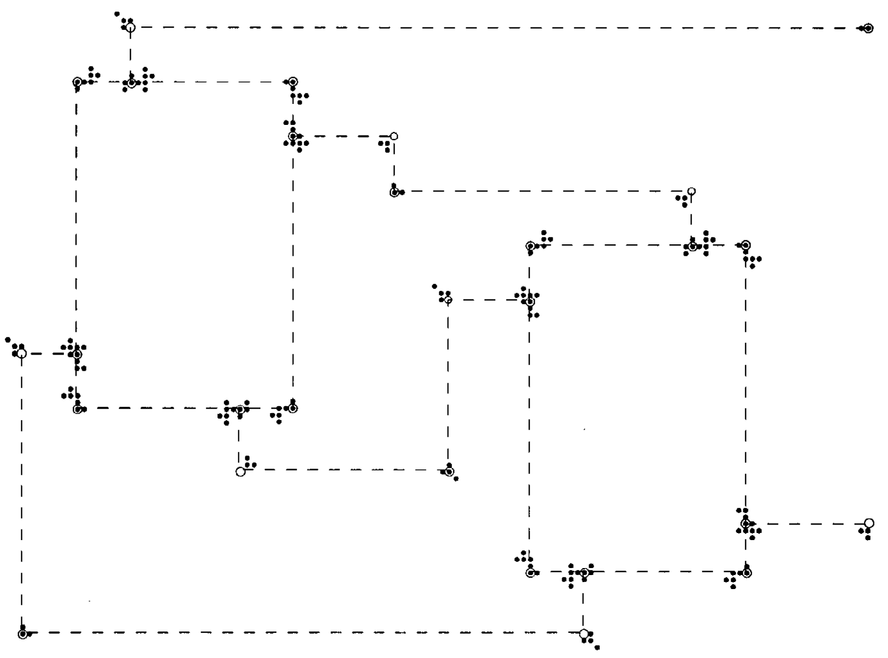


Fig. 15. The point set P_D.

Proof. Let $p=(p_1,p_2)$, $q=(q_1,q_2)$, and $r=(r_1,r_2)$. By our choice of coordinates and shiftings we know that the position of the joint of p's cluster can be written as $c(p) = (x_{c(p)}, y_{c(p)}) = N^8(c_1,c_2)+(N^4n^{2i_1}, N^2n^{2i_2})$; similarly, we get $c(q)=(x_{c(q)}, y_{c(q)}) = N^8(c_3,c_4)+(N^4n^{2i_3}, N^2n^{2i_4})$, and $c(r)=(x_{c(r)}, y_{c(r)}) = N^8(c_5,c_6)+(N^4n^{2i_1}, N^2n^{2i_2})$. For the vector $d(p)=(d_1,d_2)=p-c(p)$ from p to $c(p)$, we know by construction of the clusters that $|d_i|\leq 3$, and the same holds for $d(q)=(d_3,d_4)=q-c(q)$ and $d(r)=(d_5,d_6)=r-c(r)$.

Thus,

$$p = N^8(c_1,c_2)+(N^4n^{2i_1}, N^2n^{2i_2})+(d_1,d_2),$$

$$q = N^8(c_3,c_4)+(N^4n^{2i_3}, N^2n^{2i_4})+(d_3,d_4),$$

$$r = N^8(c_5,c_6)+(N^4n^{2i_5}, N^2n^{2i_6})+(d_5,d_6).$$

Now

$$\begin{aligned}
\tfrac{1}{2} &= AR(\Delta) = \tfrac{1}{2}\left|(q_1-p_1)(r_2-p_2)-(r_1-p_1)(q_2-p_2)\right| \\
&= \tfrac{1}{2}\Big|\big(N^8(c_3-c_1)+N^4(n^{2i_3}-n^{2i_1})+(d_3-d_1)\big) \\
&\qquad \times \big(N^8(c_6-c_2)+N^2(n^{2i_6}-n^{2i_2})+(d_6-d_2)\big) \\
&\qquad - \big(N^8(c_5-c_1)+N^4(n^{2i_5}-n^{2i_1})+(d_5-d_1)\big) \\
&\qquad \times \big(N^8(c_4-c_2)+N^2(n^{2i_4}-n^{2i_2})+(d_4-d_2)\big)\Big| \\
&= \tfrac{1}{2}\Big|N^{16}((c_3-c_1)(c_6-c_2)-(c_5-c_1)(c_4-c_2)) \\
&\qquad + N^{12}((c_6-c_2)(n^{2i_3}-n^{2i_1})-(c_4-c_2)(n^{2i_5}-n^{2i_1})) \\
&\qquad + N^{10}((c_3-c_1)(n^{2i_6}-n^{2i_2})-(c_5-c_1)(n^{2i_4}-n^{2i_2})) \\
&\qquad + N^{8}((c_3-c_1)(d_6-d_2)-(c_5-c_1)(d_4-d_2)) \\
&\qquad + N^{6}((n^{2i_3}-n^{2i_1})(n^{2i_6}-n^{2i_2})-(n^{2i_5}-n^{2i_1})(n^{2i_4}-n^{2i_2})) \\
&\qquad + N^{4}((n^{2i_3}-n^{2i_1})(d_6-d_2)-(n^{2i_5}-n^{2i_1})(d_4-d_2)) \\
&\qquad + N^{2}((n^{2i_6}-n^{2i_2})(d_3-d_1)-(n^{2i_4}-n^{2i_2})(d_5-d_1)) \\
&\qquad + ((d_3-d_1)(d_6-d_2)-(d_5-d_1)(d_4-d_2))\Big|.
\end{aligned}$$

Since $N\geq n^{3n}$, this implies the following equations:

$$\begin{aligned}
(c_3-c_1)(c_6-c_2) &= (c_5-c_1)(c_4-c_2), && (1)\\
(c_6-c_2)(n^{2i_3}-n^{2i_1}) &= (c_4-c_2)(n^{2i_5}-n^{2i_1}), && (2)\\
(c_3-c_1)(n^{2i_6}-n^{2i_2}) &= (c_5-c_1)(n^{2i_4}-n^{2i_2}), && (3)\\
(c_3-c_1)(d_6-d_2) &= (c_5-c_1)(d_4-d_2), && (4)\\
(n^{2i_3}-n^{2i_1})(n^{2i_6}-n^{2i_2}) &= (n^{2i_5}-n^{2i_1})(n^{2i_4}-n^{2i_2}), && (5)\\
(n^{2i_3}-n^{2i_1})(d_6-d_2) &= (n^{2i_5}-n^{2i_1})(d_4-d_2), && (6)\\
(n^{2i_6}-n^{2i_2})(d_3-d_1) &= (n^{2i_4}-n^{2i_2})(d_5-d_1), && (7)\\
(d_3-d_1)(d_6-d_2) &= (d_5-d_1)(d_4-d_2)+1. && (8)
\end{aligned}$$

Note $i_j=i_\ell \Leftrightarrow c_j=c_\ell$ and consider the ways to satisfy (2): if both sides disappear, we get the cases

(A) $c_6 = c_2 \wedge c_4 = c_2 \Rightarrow p, q, r$ are from the same horizontal class.
(B) $i_3 = i_1 \wedge c_4 = c_2 \Rightarrow p, q$ are from the same cluster.
(C) $c_6 = c_2 \wedge i_5 = i_1 \Rightarrow p, r$ are from the same cluster.
(D) $i_3 = i_1 \wedge i_5 = i_1 \Rightarrow p, q, r$ are from the same vertical class.

If neither side of (2) disappears, then $i_5 \neq i_3$ and $c_i \leq n$ imply a contradiction, hence it follows that

(E) $i_3 = i_5 \wedge c_4 = c_6 \Rightarrow q, r$ are from the same cluster.

Without loss of generality consider cases (A) and (B). In both cases, consider (6). Because of (8), $d_6 = d_2$ and $d_4 = d_2$ cannot both be true. In case (A), we get the subcases

(A1) $i_3 = i_1 \Rightarrow p, q$ are from the same cluster,
(A2) $i_5 = i_1 \Rightarrow p, r$ are from the same cluster,

if both sides of (6) disappear, otherwise the subcase

(A3) $i_5 = i_3 \vee d_6 = d_4 \Rightarrow q, r$ are from the same cluster.

In case (B), (6) implies the two subcases

(B1) $i_5 = i_1 \Rightarrow p, q, r$ are from the same vertical class,
(B2) $d_4 = d_2$. Note that the left side of (7) must disappear. Since $d_3 = d_1$ implies the contradiction $p = q$, we conclude $i_6 = i_2 \Rightarrow p, q, r$ are from the same horizontal class.

Thus, we may consider without loss of generality that

(X) p, q are from the same cluster, p, q, r are from the same horizontal class.

Since the right side of (6) must disappear, we get the cases

(X1) $i_5 = i_1 \Rightarrow p, q, r$ are from the same cluster.
(X2) $d_4 = d_2$. From (8), it follows that $|d_3 - d_1| = 1$ (meaning that p and q form a horizontal edge of length 1) and $|d_6 - d_2| = 1$ (meaning that the vertical distance of r from this edge is 1).

This concludes the proof. □

As an immediate implication, we get

Corollary 3.4. *Let $e = (p, q)$ with $p = (x_1, y_1)$ and $q = (x_2, y_2)$ be an edge of a grid-empty polygonization $\mathcal{P}$ of P_D, with p, q from different clusters. Then $|x_1 - x_2| = 1$ or $|y_1 - y_2| = 1$.*

Proof. In a triangulation of $\mathcal{P}$, e must be an edge of some triangle. Since all these triangles are empty, they must have area $\frac{1}{2}$, so the claim follows from Lemma 3.3. □

For later use when analyzing connections between different clusters, we note another easy consequence:

Corollary 3.5. *Let $e = (p_1, p_2)$ be an edge between different clusters C_1 and C_2. Then there must be a point p_3 at distance 1 from p_1 or p_2, such that $AR(p_1, p_2, p_3) = \frac{1}{2}$.*

In light of this, we call a *bridgehead* from a cluster C_1 to a different cluster C_2 a pair of points p_1, p_3 in C_1 that has distance 1, if there is at least one point p_2 in C_2, such that p_2 has distance 1 from the line through p_1 and p_3.

3.5. *A Hamiltonian Path Induces an Empty Polygon*

Now we show the following:

Lemma 3.6. *Suppose that D' has a Hamiltonian path H from s to t. Then P_D has a simple polygonalization $\mathcal{P}$ that does not encounter any grid points not in P_D, so $AR(\mathcal{P}) = n/2 - 1$.*

Proof. Let the Hamiltonian path H be given as a sequence of vertices $s = v_{i_0}, v_{i_1}, \ldots, v_{i_n}, v_{i_{n+1}} = t$. s and t are represented by the terminal clusters T_0 and T_{n+1} in P_D, any other vertex v_i corresponds to a switch cluster W_i. We construct $\mathcal{P}$ as a pair P_1, P_2 of disjoint polygonal paths from T_0 via $W_1, \ldots, W_n$ to T_{n+1}—see Fig. 16 for the overall picture in our running example. P_1 has the interior of $\mathcal{P}$ to its right, P_2 to its left.

As shown in Fig. 17, both P_1 and P_2 follow the bend clusters along a mandatory edge $\langle v_i, v_j\rangle$ from switch cluster W_i to switch cluster W_j, collecting all points of these bends along the way, while avoiding any grid points not in P_D. (The indicated edge e_i is the unit edge between the last points of P_1 and P_2 in W_i, and e_j is the unit edge between the first points of P_1 and P_2 in W_j; the shaded polygonal region bounded by e_i, P_1, P_2, and e_j does not contain any grid points not in P_D.)

An optional edge $\langle v_i, v_k\rangle$ in H is followed by both paths P_1 and P_2 as shown in Fig. 18. (Shown are cases (a) and (b) from Fig. 14—cases (c) and (d) are symmetric to case (b). The curved parts of the layout are symbolic for the detour pieces that are shown in Fig. 19 for better analogy—see the description below.) We start at the unit edge e_i (formed by the first points of P_1 and P_2 in W_i) and proceed to the unit edge e_k (formed by the last points of P_1 and P_2 in W_k). On the way, all points of W_i, W_k, and the link cluster L_{ik} for the edge $\langle v_i, v_k\rangle$ are visited, as are some subsets S_i and S_k of the other two link clusters adjacent to W_i and W_k. The shaded polygonal region bounded by e_i, P_1, P_2, and e_k does not contain any grid points not in P_D. An optional edge $\langle v_i, v_j\rangle$ not in H corresponds to a link cluster L_{ij}. L_{ij} is partitioned into two subsets S_i and S_j; as shown in Fig. 19, S_i and the leftover pivot point p_i of the adjacent switch cluster W_i are collected by a detour of one of the paths P_1, P_2 when traversing W_i.

When merging the paths P_1 and P_2 for the mandatory and optional edges in H, P_1 and P_2 meet after collecting all points of P_D; furthermore, merging the polygonal subregions for all mandatory and optional edges in H at the separating unit edges e_i yields a region

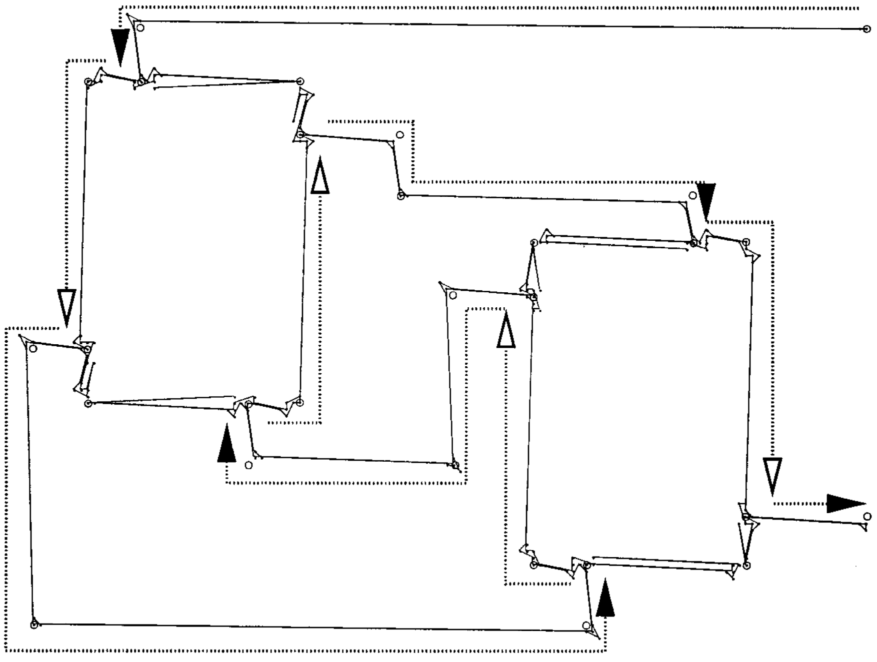


Fig. 16. A Hamiltonian path in D' and the corresponding grid-empty simple polygonalization for P_D; mandatory edges are indicated by solid arrowheads, optional edges by hollow arrowheads.

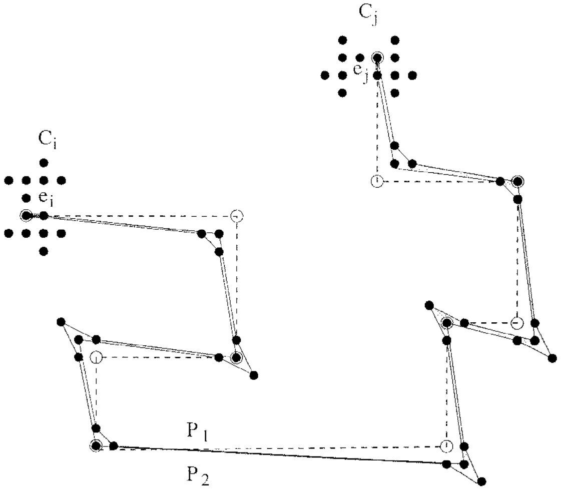


Fig. 17. How to follow an optional edge contained in H.

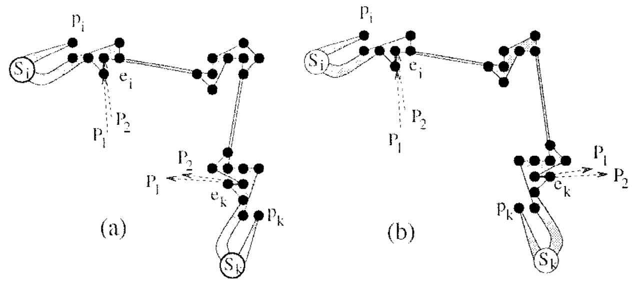


Fig. 18. How to follow an optional edge contained in H.

that does not contain any grid points not in P_D. Thus, the two paths form a simple polygon $\mathcal{P}$ of the claimed properties. □

3.6. *An Empty Polygon Induces a Hamiltonian Path*

In this section we prove the converse of Lemma 3.6:

Lemma 3.7. *Suppose that P_D has a simple polygonalization $\mathcal{P}$ that does not encounter any grid points not in P_D, so $AR(\mathcal{P}) = n/2 - 1$. Then D' has a Hamiltonian path H from s to t.*

In the following we assume that $\mathcal{P}$ is a grid-empty simple polygonalization of P_D. In Section 3.6.1 we analyze the geometry of possible edges. We show with the help of Corollary 3.4 and further analysis of the cluster structure that edges of a grid-empty polygon may only occur between adjacent clusters. This allows us to concentrate on connections between switch clusters in Section 3.6.2. From this we derive that in a certain well-defined sense, the edges of a grid-empty polygon can connect the points of a switch cluster to at most two other, adjacent switch clusters. These connections can thus be used to construct a Hamiltonian path.

Each of the subsections progresses by a sequence of lemmas.

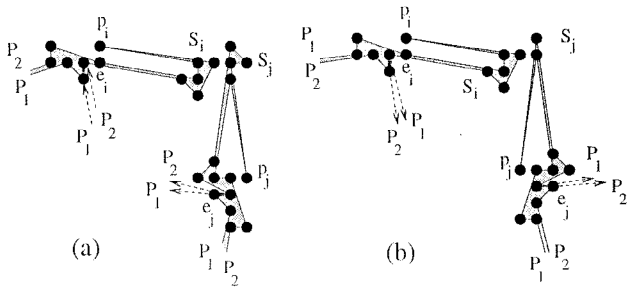


Fig. 19. The detours for an optional edge not contained in H.

3.6.1. *Local Connections*. We say that a quadrangle $\Pi = (p_0, p_1, p_2, p_3)$ is a *bridge* between two different clusters W_1 and W_2, iff $p_0, p_1 \in W_1$, $p_2, p_3 \in W_2$, Π is fully contained in $\mathcal{P}$, and $AR(\Pi) = 1$. A *probe* from cluster W_1 to cluster W_2 is a triangle $\Delta = (p_1, p_2, p_3)$ fully contained in $\mathcal{P}$, such that $p_1, p_3 \in W_1$ and $p_2 \in W_2$.

Lemma 3.8. *Any bridge* $\Pi = (p_0, p_1, p_2, p_3)$ *is a parallelogram, and the edges* $e_1 = (p_0, p_1)$ *and* $e_2 = (p_0, p_1)$ *have unit length; moreover, the lines through* e_1 *and* e_2 *are distance* 1 *apart.*

Proof. Both triangles $\Delta_1 = (p_0, p_1, p_3)$ and $\Delta_2 = (p_1, p_2, p_3)$ lie fully in $\mathcal{P}$, so they must both be grid-empty. By construction, any axis-parallel connection between points from different clusters contains grid points not in $\mathcal{P}$, so both triangles are nondegenerate. Thus, we can apply Pick's theorem and conclude that both triangles have area $\frac{1}{2}$. From Lemma 3.3, we conclude that both Δ_1 and Δ_2 have base and altitude of size 1; since they share the edge (p_1, p_3), they must be congruent, and the claim follows. □

Lemma 3.9. *Let* $e_1 = (p_1, p_2)$ *be an edge in* $\mathcal{P}$, *and suppose* p_1 *and* p_2 *are from two different clusters* C_1 *and* C_2. *Then there is either a probe or a bridge that contains* e_1 *as an edge.*

Proof. See Fig. 20. Let $e_0 = (p_0, p_1)$ and $e_2 = (p_2, p_3)$ be the edges adjacent to e_1 in $\mathcal{P}$. Assume that e_1 is not an edge of a probe, so $p_0 \notin C_2$, or the triangle (p_0, p_1, p_2) is not contained in $\mathcal{P}$, or $p_3 \notin C_1$, or the triangle (p_1, p_2, p_3) is not contained in $\mathcal{P}$. We will show that there are two points $p_4 \in C_1$ and $p_5 \in C_2$, such that the quadrangle $\Pi = (p_1, p_2, p_4, p_5)$ is a bridge.

We write $p_i = (x_i, y_i)$. e_1 is an edge of some triangle $\Delta = (p_1, p_2, p_4)$ in a triangulation of $\mathcal{P}$. By Lemma 3.3 we may assume without loss of generality that p_1 and p_2 are from the same horizontal class, and that p_4 is from the same cluster as p_1. Furthermore, Lemma 3.3 implies that $|y_1 - y_2| = 1$ and $|x_4 - x_1| = 1$; without loss of generality, consider $y_2 = y_1 + 1$ and $x_4 = x_1 - 1$.

By assumption, (p_1, p_2, p_4) is not contained in a probe, so (p_2, p_4) must be a diagonal of $\mathcal{P}$. Consider the two grid points $q_1 = (x_1 + n, y_1)$ and $q_2 = (x_1 + n, y_1 + 1)$. By construction of P_G, the distance between different clusters is larger than n, and no two points of the same cluster can have distance n, so $q_1, q_2 \notin P_G$. Consider the intersection point s_1 of the line $f = \overline{q_1 q_2}$ with the edge e_1. By our assumptions on Δ, there is a

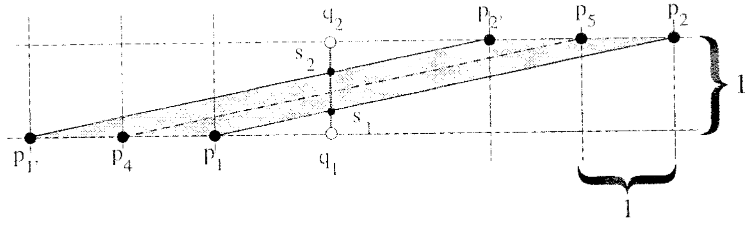


Fig. 20. A bridge between different clusters.

segment $\overline{s_1 s_2}$ of f above s_1 that belongs to $\mathcal{P}$, so there must be another edge of $\mathcal{P}$ that separates s_1 from q_2; let $e'_1 = (p'_1, p'_2)$ be the lowest of these edges, i.e., the edge that intersects f in s_2. Since s_2 lies strictly between q_1 and q_2, it follows from Corollary 3.4 that $y'_1 = y_1$ and $y'_2 = y_1 + 1$. Since Δ is not contained in a probe, we know that e'_1 is disjoint from e_1, and we conclude $x'_1 < x_1$ and $x'_2 < x_2$.

Now consider the quadrangle $\Pi' = (s_1, p_2, p'_2, s_2)$ and the point $p_5 = (x_2 - 1, y_1 + 1)$. By our minimality assumption on s_2, no edge of $\mathcal{P}$ crosses $\overline{s_1 s_2}$. Furthermore, no edge of $\mathcal{P}$ can cross the polygon edges $\overline{s_1 p_2}$ or $\overline{p'_2 s_2}$, and no points of P_G are contained strictly inside of Π'. Therefore, no edge of $\mathcal{P}$ can intersect the interior of Π', meaning that all points strictly inside of Π' belong to $\mathcal{P}$. Since the boundary of any open subset of $\mathcal{P}$ belongs to $\mathcal{P}$, we conclude that all points of Π' belong to $\mathcal{P}$. This includes p_5. Similarly, all points of the quadrangle $\Pi'' = (p_1, s_1, s_2, p'_s)$ belong to $\mathcal{P}$. The claim follows. $\square$

From the above proof, it is straightforward to deduce the following:

Corollary 3.10. *Let $\Pi = (p_1, p_2, p_3, p_4)$ with $p_i = (x_i, y_i)$ be a bridge with $p_1, p_4 \in C_1$ and $p_2, p_3 \in C_2$. Then there are two edges $e'_1 = (p'_1, p'_2)$ and $e''_1 = (p''_1, p''_2)$ in $\mathcal{P}$, such that $y'_1 = y''_1 = y_1$ and $y'_2 = y''_2 = y_2 = y_1 + 1$, the quadrangle $A = (p'_1, p'_2, p''_2, p''_1)$ is fully contained in $\mathcal{P}$, and $\Pi \subseteq A$.*

Proof. Consider two points s_1 and s_2 like in the proof of Lemma 3.9; then e'_1 is the lowest edge intersecting $\overline{s_1 s_2}$ above the interior of Π, and e''_1 is the highest edge intersecting $\overline{s_1 s_2}$ below the interior of Π. The rest is shown like in the proof of Lemma 3.9. $\square$

For a bridge Π, the edges e'_1 and e''_1 enclosing Π are called the *rails* of the bridge. As we noted in Corollary 3.5, connections between different clusters can only occur at bridgeheads, and bridges require a pair of bridgeheads, one in each cluster involved.

Lemma 3.11. *There cannot be a bridge between two different link clusters.*

Proof. By Corollaries 3.10 and 3.4, L_1 and L_2 must be from the same class, say, horizontal. Let W be the switch cluster in the same class as L_1 and L_2, with the joint positioned at $y = 2$. Supppose that the bend cluster B adjacent to W is positioned below $y = 0$.

There are three cases:

(I) *W has no tab triples*. Then both L_1 and L_2 must be chosen as in one of cases (a)–(c) of Fig. 14. Without loss of generality, consider case (a), as the link clusters arising from cases (b) and (c) are merely subsets of the ones from case (a). We see from Fig. 21(I) that any bridge Π between L_1 and L_2 must connect vertices at $y = 1$ and $y = 2$. Let p_0 be the first among the (circled) points of W at $y = 2$ that is reached along P_1 by an edge $e = (p_i, p_0)$. By Corollary 3.4, p_i must be from B, L_1, or L_2; p_i cannot be from B, as e would intersect the rails of the bridge Π. By Lemma 3.9, this means that there must be a probe or a bridge from L_1 or L_2 that contains e as an edge. As noted above, this requires a bridgehead of L_1 or L_2 to a circled point. The only such bridgeheads are at

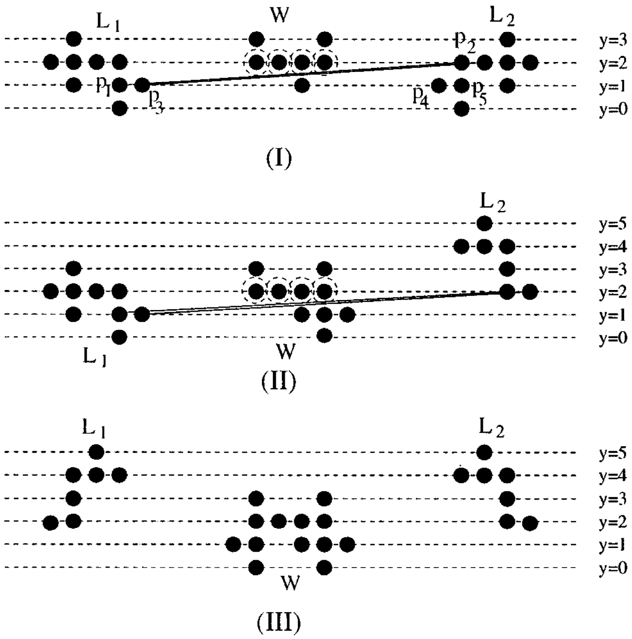


Fig. 21. There cannot be a bridge between two different link clusters.

$y = 1$, as indicated in the figure by the pairs (p_1, p_3) and (p_4, p_5). Now it is easy to see that a bounding edge of a probe or the rail of a bridge originating from either of these bridgeheads would intersect the interior of Π, which is a contradiction to $\Pi \subset \mathcal{P}$.

(II) *W has one tab triple*. Then consider L_1 chosen as in case (a) of Fig. 14, and L_2 as in case (d). (See Fig. 21(II).) Similar to the line of argument in (I), we get a contradiction.

(III) *W has two tab triples*. Then both L_1 and L_2 must be chosen as in case (d) of Fig. 14. As Fig. 21(III) shows, there are no corresponding bridgeheads in L_1 and L_2, so there cannot be a bridge between link clusters of this type. □

We know from Corollary 3.4 that a bridge can only exist between clusters from the same horizontal or vertical class. Since Lemma 3.11 excludes bridges between different link clusters, only the following possibilities remain:

Corollary 3.12. *Any bridge from a nonswitch cluster connects a link cluster to an adjacent switch cluster, or a bend cluster to an adjacent terminal, switch, or bend cluster.*

3.6.2. *Connections between Switch Clusters*. With the help of the results from the previous section, we can proceed to establish properties of connections between switch clusters. These are used to construct a Hamiltonian path.

The boundary of $\mathcal{P}$ can be considered a (closed) directed path P_1 starting at the terminal cluster T_0 that has the interior of $\mathcal{P}$ to its right. Let W_1 be the switch cluster next to T_0, and let W_n be the switch cluster next to T_{n+1}.

Definition 3.13. With $V_\psi = \{T_0, W_1, \dots, W_n, T_{n+1}\}$, we define a digraph $\Psi = (V_\psi, A_\psi)$ on the set of terminal and switch clusters by its arc set. For $i, j \in \{1, \dots, n\}$, we have $\langle W_i, W_j\rangle \in A_\psi$ iff W_i is the last switch cluster encountered by P_1 before hitting W_j for the first time. Furthermore, the only arc incident to T_0 is $\langle T_0, W_1\rangle$, and the only arc incident to T_{n+1} is $\langle W_n, T_{n+1}\rangle$.

For any W_i, let $P_1^{(i)}$ be the subpath of P_1 that starts at T_0 and ends at the first point of W_i.

With $V_\gamma = M$ and $i \in \{0, \dots, n+1\}$ we describe a digraph $\Gamma^i = (V_\gamma, A_\gamma^i)$ on the set M of clusters by its set of arcs: $\langle C_i, C_j\rangle \in A_\gamma^i$ iff there is an edge of $e \in P_1^{(i)}$ from C_i to C_j that does not form a probe with another edge $f \in P_1^{(i)}$.

Clearly, Ψ is a spanning arborescence rooted at T_0, and T_{n+1} is a leaf. We will show that Ψ is a directed path from W_0 to W_{n+1} running through all vertices. For this purpose, it suffices that there is no vertex of out-degree 2. The graph Γ will help us to analyze the connections between switch clusters.

Lemma 3.14. *For $i \in \{1, \dots, n\}$, let $\langle W_i, W_j\rangle \in A_\psi$. Then there must be a path $a_0, \dots, a_k$ from W_i to W_j in Γ^j, with $a_0, \dots, a_{k-1}$ corresponding to rails of bridges in $\mathcal{P}$.*

Proof. By definition, any consecutive pair of edges bounding a probe returns to the same cluster. Recall that, by Corollary 3.4, there cannot be an edge directly from W_i to W_j. This means that after removing all consecutive pairs of edges in $P_1^{(j)}$ that form a probe, there must still be a set of edges e_ℓ, $\ell = 1, \dots, k$, in $P_1^{(j)}$ connecting W_i to a nonswitch cluster $C_{ij}^{(1)}$, connecting $C_{ij}^{(1)}$ to a nonswitch cluster $C_{ij}^{(2)}$, etc., and connecting nonswitch cluster $C_{ij}^{(k-1)}$, to W_j. By definition, this means that there is a path from W_i to W_j in Γ^j. Since none of the edges e_ℓ, $\ell = 1, \dots, k-1$, is contained in a probe, we know from Lemma 3.9 that they must be rails of bridges. $\square$

Furthermore, we get

Corollary 3.15. *Let $\langle W_i, W_j\rangle \in A_\psi$ with $i, j \in \{1, \dots, n\}$. Then W_i and W_j correspond to vertices of D that are adjacent.*

Proof. By the previous lemma, there is a path from W_i to W_j in Γ^j, with all but possibly the last arc corresponding to bridges between nonswitch clusters. Then the claim follows from Corollary 3.12. $\square$

This allows us to identify arcs in Ψ with edges in D. Clearly, this implies the following:

Corollary 3.16. *Let $W_i \in V_\psi$ with $i \in \{1, \ldots, n\}$, and let L_1, L_2, B be the clusters adjacent to the switch cluster W_i. If W_i has out-degree 2 in Ψ, then there must be bridges from W_i to L_1, L_2, B.*

Lemma 3.17. *Let $W_i \in V_\psi$ with $i \in \{1, \ldots, n\}$, and let L_1, L_2, B be the clusters adjacent to the switch cluster W_i. Let W_i have out-degree 2 in Ψ, and let e_1, e_1' be the rails of the bridge between W_i and L_1, let e_2, e_2' be the rails of the bridge between W_i and L_2, and let e_3, e_3' be the rails of the bridge between W_i and B. Then for $k = 1, 2, 3$, e_k and e_k' cannot be connected by edges of P_1 that stay within W_i.*

Proof. Let W_j be the switch cluster with $\langle W_j, W_i\rangle \in A_\psi$, and let W_i', W_i'' be the switch clusters with $\langle W_i, W_i'\rangle \in A_\psi$, $\langle W_i, W_i''\rangle \in A_\psi$. It is easy to check in Fig. 22 that no switch cluster can have more than three bridges. If there were a connection between e_k and e_k' within W_i, P_1 would enter and leave W_1 through the same bridge, and return to the same switch cluster it came from. By definition, this cannot be one of the clusters W_i' or W_i''; on the other hand, it cannot be W_j, since then P_1 cannot return to W_i to visit W_i' or W_i'' without visiting either one of them. $\square$

Lemma 3.18. *Let W have out-degree 2 in Ψ. Then there must be a probe from an adjacent link cluster to W.*

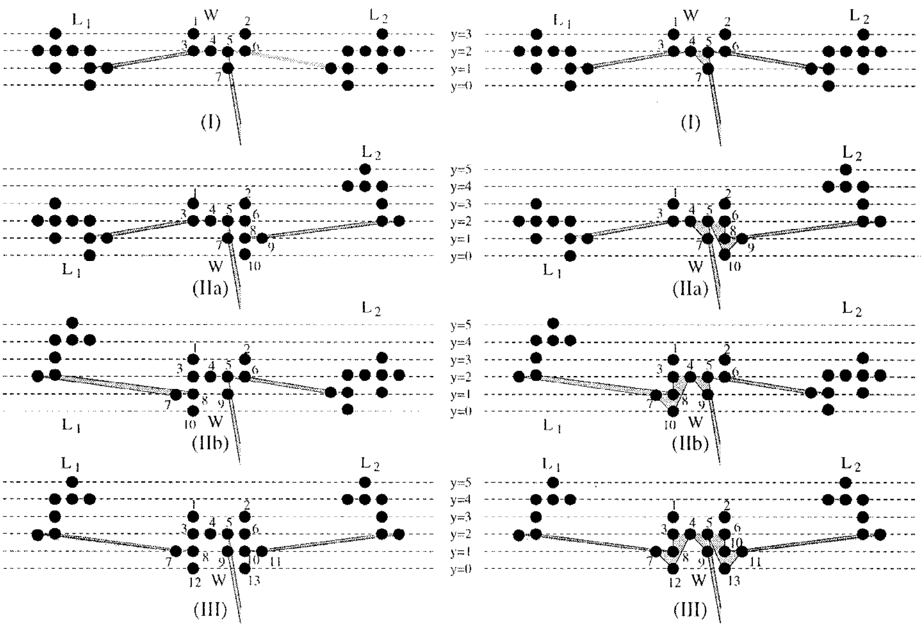


Fig. 22. Three bridges to a switch cluster: (left) basic setup; (right) resulting connections

Proof. See Fig. 22. As in the proof of Lemma 3.11, we distinguish cases according to the number of tab triples of W, which depends on the choice of nearby link clusters; since the bridgehead to the next bend cluster is assymetric, we distinguish cases (IIa) and (IIb). As before, we may consider without loss of generality link clusters of type (a) when we have to deal with type (a), (b), or (c).

We analyze the possible other edges $\mathcal{P}$ that are incident to W; at each step, we use that

(i) we must not intersect any edges that have already been deduced,
(ii) no point in P_G can be incident to more than two edges,
(iii) by Corollary 3.4, edges to points outside of the cluster can only occur at bridgeheads,
(iv) by Lemma 3.17, we must not create a connection between rails of the same bridge, and
(v) no edge must contain grid points outside of P_D.

(I) *W does not have any tab triples*. Refer to Fig. 22(I). Since the only bridgehead to a bend cluster B is formed by vertices 5 and 7, the rails of the bridge from B must end at these vertices. Since no rail from L_2 may intersect the rail to 5, the rails from L_2 must end at 5 and 6. This leaves 3 and 4 as the only endpoints of rails from L_1, and we get the situation as shown in the left figure.

As the only bridgehead involving 7 is already used, it follows from (iii) that 7 must be connected to a point in W. By (i), it must be 4 or 5, by (ii), it cannot be 5, leaving 4. Because of (v) and (ii), the only connections of 1 and 2 in W is to 3 or 4. If one of them is connected to 3 and 4, the other must be connected to two points outside of W; since it is not part of a bridgehead, it must be contained in a probe from the outside, i.e., from an adjacent link cluster.

(IIa) *W has one tab triple close to the bridgehead to the bend cluster*. Refer to Fig. 22(IIa). As in case (I), the rails of the bridge from B must end at 5 and 7, and the rails from L_1 must end at 3 and 4. This leaves only 8 and 9 for the bridge from L_2, and we have the situation as shown on the left of the figure.

As in (I), 7 must be connected to 4. Similarly, 9 must be connected to 11. Point 11 is not part of a bridgehead, so it must be connected to a point in W. By (i), it can only be connected to 5 or 8, by (iv), 8 is excluded, leaving 5. As the only bridgehead involving 8 has already been used, then by (i) and (ii), 8 must be connected to 6. Again, we conclude that one of the points 1 or 2 must be contained in a probe from an adjacent link cluster.

(IIb) *W has one tab triple away from the bridgehead to the bend cluster*. Refer to Fig. 22(IIb). As in case (I), the rails of the bridge to B must end at 5 and 7, and the rails from L_2 must end at 5 and 6. This leaves only 7 and 8 for the bridge from L_1, and we have the situation as shown on the left of the figure.

As 9 to 10 in (IIa), 7 must be connected to 10. By (iii), 10 must be connected inside of W, and, by (i), it must be to one of 8, 4, 5, 9. Point 8 is ruled out by (iv), 5 by (ii), and a connection to 9 would imply that the grid point between 8 and 9 that is not in P_D would be enclosed by $\mathcal{P}$; so 10 is connected to 4. With the known arguments, we conclude that 8 is connected to 3, and 9 to 4. Again, we conclude that one of the points 1 or 2 must be contained in a probe from an adjacent link cluster.

(III) *W has two tab triples*. Refer to Fig. 22(III). As in case (I), the rails of the bridge to B must end at 5 and 9, and the rails from L_2 must end at 10 and 11. Points 7 and 8 form the only bridgehead to L_1, and we have the situation as shown on the left of the figure.

Like in case (IIa), we conclude that 11 is connected to 13, 13 to 5, and 10 to 6. Like in case (IIb), we conclude that 7 is connected to 12, 12 to 4, 9 to 4, and 8 to 3. Again, we conclude that one of the points 1 or 2 must be contained in a probe an adjacent link cluster.

This concludes the proof. □

Lemma 3.19. *There cannot be a switch W with out-degree* 2 *in* Ψ.

Proof. See Fig. 23. We assume the same y-coordinates as in the previous lemma. By the preceding discussion, we know that both adjacent link clusters L_1 and L_2 must have a bridge to switch cluster W, and a probe or an edge to the other adjacent switch clusters W_1 and W_2. This uses a bridgehead from either link cluster to both their adjacent switch clusters. Furthermore, we know by Lemma 3.18 that, from one of them, there must be

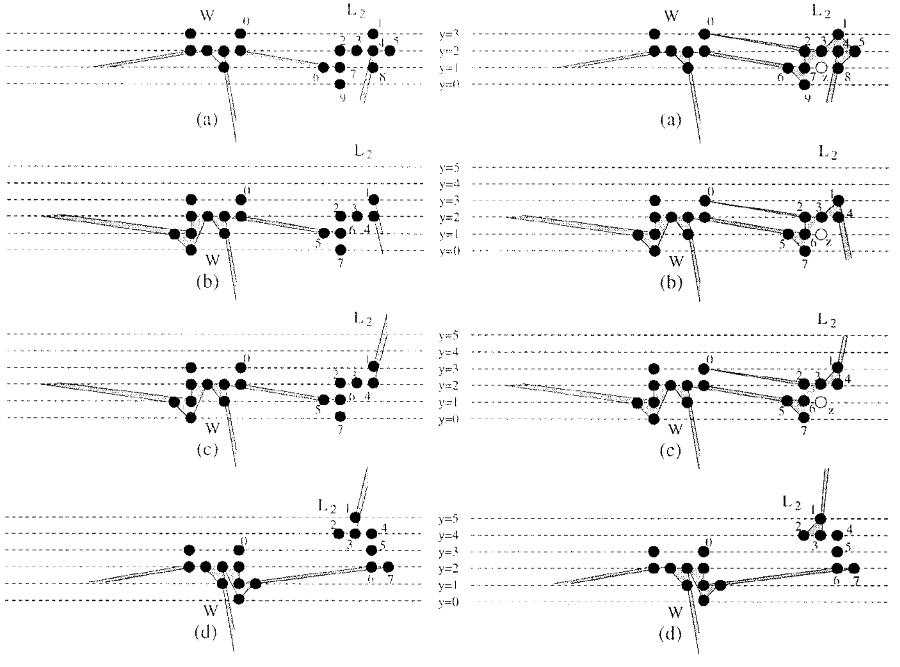


Fig. 23. Two connections for a link cluster.

a probe to W to one of the two points of W at $y = 3$. Without loss of generality, we assume that this point is 0 and that it is contained in a probe from L_2.

As before, we analyze the possible other edges $\mathcal{P}$ that are incident to W. At each step, we use that

(i) we must not intersect any edges that have already been deduced,
(ii) no point in P_G can be incident to more than two edges,
(iii) by Corollary 3.4, edges to points outside of the cluster can only occur at bridgeheads,
(iv) we must not create a connection within L_2 between rails of the same bridge in order to guarantee a path from W to W_2 via L_2, and we must not create a second path between the endpoints of a probe, and
(v) no edge must contain grid points outside of P_D.

Now distinguish cases according to the type of link cluster:

(a) Refer to Fig. 23(a). The only bridge from W to L_2 uses points 6 and 7. Since an edge between 1 and W_2 (located below L_2 in the figure) would separate 4, 5, and 8 from W, the bridgehead from L_2 to W_2 must be 4 and 8, and we have the situation as in the left part of the figure.

Since the only possible connection of 6 outside of L_2 is already used, it follows from (i) and (iv) that 6 must be connected to 9. By (i) and (iv), 8 can only be connected to 5. Again by (i) and (iv), 5 can only be connected to 0 or 1; since an edge to 0 would separate 1 from all other points, it must be to 1. Now an edge between 4 and 0 separates 1 from all feasible neighbors for a second connection. This leaves only the bridgehead 2 and 3 to collect 0 by a probe from L_2. Then, by (i), 2 must be connected to 7. By (iv), 3 cannot be connected to 9, and connecting it to 4 separates 1 from any feasible neighbors, 3 must be connected to 1. This leaves 4 and 9 as the only potential neighbors. Their connection runs over the grid point $z \notin P_D$, and we have a contradiction.

(b) Refer to Fig. 23(b). The only bridge from W to L_2 uses points 5 and 6, the only bridgehead from L_2 to W_2 (located below L_2 in the figure) is 1 and 4, and we have the situation as in the left part of the figure.

Similar to the previous case, it follows that 5 must be connected to 7. An edge between 1 and 0 separates 1 from all feasible neighbors, so the probe from L_2 to 0 must use the bridgehead 2 and 3. By (i), 2 must be connected to 6. By (iv), 3 cannot be connected to 7. An edge between 3 and 4 separates 1 from all feasible neighbors, so 3 must be connected to 1. This leaves 4 and 9 as the only potential neighbors. Their connection runs over the grid point $z \notin P_D$, and we have again a contradiction.

(c) Refer to Fig. 23(c). The argument is the same as in case (b).

(d) Refer to Fig. 23(d). The only bridge from W to L_2 uses points 6 and 7, the only bridgehead from L_2 to W_2 (located above L_2 in the figure) is 1 and 3, and we have the situation as in the left part of the figure.

Since the only possible connection of 1 outside of L_2 is already used, it follows from (i) and (iv) that 1 must be connected to 2. As a probe to 0 from 6 and 7 violates (iv), the probe must be from 2, 3, 4. A probe from 2 and 3 violates (iv), a probe from 3 and 4 separates 2 from any potential neighbors, and a probe from 2 and 4 forces the connection between 3 and 4, contradicting (iv).

This concludes the proof. □

This implies that Ψ is a path, and each switch cluster W has two neighbors in Ψ.

Lemma 3.20. *Ψ cannot contain two consecutive arcs that correspond to optional edges in D.*

Proof. Consider a switch cluster W that is connected to W_1 and W_2 in Ψ optional edges. If W had a bridge or a probe to its adjacent bend cluster, the proofs of Lemmas 3.18 and 3.19 still yield a contradiction, so suppose this was not the case. Then the switch cluster W' that shares the mandatory edge through B with W has to be connected to B in Γ^n; since W' and W cannot be adjacent in Ψ, W' must also have two neighbors via optional edges, so we get the same type of contradiction for W' instead of W. □

Summarizing, we see that no vertex in the spanning arborescence Ψ (induced by connections from terminal clusters and from switch clusters) has out-degree more than 2 (Lemma 3.19), so Ψ is a Hamiltonian path. Since the corresponding edges in D' alternate between optional and mandatory (Lemma 3.20), it follows that from each pair of optional edges to a vertex, precisely one is chosen, such that we get a Hamiltonian path in D'. This concludes the proof of Lemma 3.7, and thus Theorem 3.1.

4. Maximal Area

For some practical purposes, one might be more interested in simple polygons with a large enclosed area than in those with a small area. As we will see, the relation between the problems MAX-AREA and MIN-AREA is very close, which allows for a simple proof of NP-hardness. More precisely, we can use our results on GRID-EMPTY to show that GRID-FULL is NP-hard.

Theorem 4.1. GRID-FULL POLYGON *is NP-hard.*

Proof. See Fig. 24. Consider the point set P in the NP-hardness proof of GRID-EMPTY. Note that in any simple polygon with vertex set P of area $n/2 - 1$, 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. (Note that we made sure of this in the previous section in order to guarantee this property.) Then add the points

$$\begin{aligned} p_1 &:= t_2 - (0, N^{10}),\\ p_2 &:= t_1 + (0, N^{10}),\\ p_3 &:= p_1 - (2N^{10} + 1, 0),\\ p_4 &:= p_2 - (2N^{10} + 1, 0) \end{aligned}$$

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 GRID-EMPTY if and only if there is a simple polygon $\overline{\mathcal{P}}$ on the vertices $\overline{P}$ that satisfies Pick's bound for GRID-FULL. (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.) □

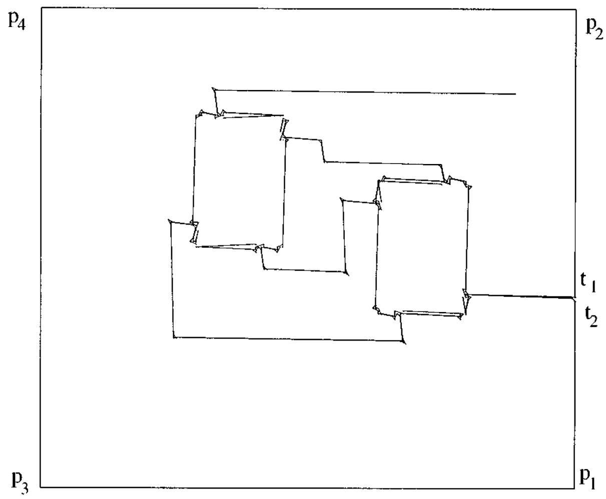


Fig. 24. GRID-FULL solves GRID-EMPTY: turning a small polygon inside out.

5. Higher Dimensions

In this section we study several higher-dimensional generalizations. 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 [17], [16].

We start by defining the polyhedral objects that will be considered:

Definition 5.1. 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 $\mathcal{P}$.

The generalization from the two-dimensional situation is clear: any point in P is required to be contained in an edge of $\mathcal{P}$.

5.1. *Minimizing Surface Area*

The following problem arises in optimal surface design and has been proposed by O'Rourke [16]:

Minimum Surface Polyhedron (SURF). Given a finite set P of points in three-dimensional Euclidean space. Among all simple polyhedra that are feasible for a vertex set P, find the one with the smallest surface area.

This is a special case of the following problem:

Minimum 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 the 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 first show that the special case SURF is NP-hard by giving a reduction of the problem HAMILTONIAN CYCLE IN GRID GRAPHS (HCCG), which was shown to be NP-complete by Itai et al. [10]. A *grid graph* G is given by a set P_G of n grid points in the plane, with two vertices adjacent iff the corresponding points have distance 1. In this geometric representation, any Hamiltonian path in G corresponds to a TSP tour of length n for P_G; clearly, such a tour induces an orthogonal simple polygon. If on the other hand, there is no tour of P_G that uses only unit length edges, it has to use an edge of length at least $\sqrt{2}$, so a shortest TSP tour must have length at least $n - 1 + \sqrt{2}$.

Theorem 5.2. SURF *is NP-hard.*

Proof. Take any instance of HCGG, i.e., a grid graph G with n vertices. Any point $(x_i, y_i) \in P_G$ is represented by the point $(x_i, y_i, 0)$ in $\mathbb{R}^3$, yielding the set P'_G. 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 $H = 2n^3$. (See Fig. 25.)

We show 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.

If there is a tour of length n, it induces a polyhedron with a surface area of

$$A + \sum_{i=1}^{n} \frac{\sqrt{H^2 + d_i^2}}{2},$$

where A is the area enclosed by the tour and d_i is the distance of $(x_g, y_g, 0)$ from the ith tour edge. Clearly, we have $A \le n^2$ and $d_i < n$. Thus we get the surface area to be at

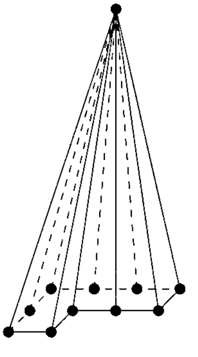


Fig. 25. Extending a planar tour into a cone in 3-space.

most

$$n^2 + \sum_{i=1}^{n} \frac{\sqrt{H^2 + n^2}}{2}.$$

Now since $H = 2n^3$ and

$$H^2 + n^2 < \left(H + \frac{n^2}{2H}\right)^2,$$

we have

$$\sqrt{H^2 + n^2} < H + \frac{1}{4n},$$

so the surface area is at most

$$n^2 + n\left(\frac{H}{2} + \frac{1}{8n}\right) = n^4 + n^2 + \tfrac{1}{8}.$$

To see the converse, assume that there is no tour of length n. First note that 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. (Each vertex in P'_G must be close to interior points of $\mathcal{P}$; furthermore, the set of inner points of $\mathcal{P}$ close to p^* has to be simply connected and nonempty.) This implies that the edges of $\mathcal{P}_2$ must form a tour. Let A be the enclosed area of this tour. Discounting this face of $\mathcal{P}$, we can partition the surface of $\mathcal{P}$ into a set of triangles by connecting p^* to all other vertices in P.

If there is no tour of length n, the length of the shortest tour must be at least $n+\sqrt{2}-1 > n + \frac{1}{3}$ and the surface area

$$A + \sum_{i=1}^{n} \frac{s_i\sqrt{H^2 + d_i^2}}{2}$$

(where s_i is the length of the ith edge) is at least

$$0 + \frac{H(n + \frac{1}{3})}{2} = n^4 + \frac{n^3}{3}.$$

This is bigger than $n^4 + n^2 + \frac{1}{8}$ as soon as n exceeds 3. $\square$

5.2. *Minimizing Face Volume*

We conclude this section by proving the NP-hardness of FACE: the idea is to add a set of $(d-2)$ points at a large distance to an instance of a two-dimensional problem to transform it into a d-dimensional one. We use the following easy lemma:

Lemma 5.3. *Let $1 \leq k$ and $P = \{e_1, \ldots, e_{k+1}\}$, where e_i denotes the ith unit vector. Then the k-dimensional regular simplex S_k spanned by P has k-dimensional volume*

$$VOL_k(S_k) = \frac{\sqrt{k+1}}{k!}.$$

Theorem 5.4. FACE *is NP-hard*.

Proof. We distinguish the following three cases:

(A) $k = d$.
(B) $k = 1$.
(C) $1 \leq k < d$.

For all these cases we use the following:

Let P_G be a set of n points in the plane $E_2 = \{(x_1, x_2, 0, \dots, 0) \mid x_1, x_2 \in \mathbb{R}\}$ that either represents an instance of MIN-AREA (case (A)) or a grid graph G (cases (B) and (C)). Let g be the center of mass of the points in P_G and $p_i = g + H \cdot e_i$ for $i = 3, \dots, d$, where $H = 9d^2n^3$ and e_i denotes the ith unit vector. We write $P_j := \bigcup_{i=3}^{j}\{p_i\} \cup P_G$ and $E_j := \{(x_1, \dots, x_d) \in \mathbb{R}^d \mid x_{j+1}, \dots, x_d = 0\}$. For any d-dimensional simple polyhedron $\mathcal{P}_d$ feasible for the vertex set P_d, the corresponding j-dimensional subpolyhedron induced on P_j is denoted by $\mathcal{P}_j := \mathcal{P}_d \cap E_j$.

As in the previous proof, we may assume that $\mathcal{P}_2 = \mathcal{P} \cap E_2$ is a simple polygon. Furthermore note that $\mathcal{P}_j$ is simple if and only if $\mathcal{P}_{j-1}$ is simple. Finally, each of the points p_i is connected to any other vertex by an edge of $\mathcal{P}_d$.

Now consider case (A).

By Theorem 3.1, it is NP-complete to decide whether a given set of n grid points allows a simple polygon of area $n/2-1$. By Pick's theorem, we may assume that if there is no such polygon, there is none of area less than $(n-1)/2$.

We claim 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 P_G with area at most $n/2 - 1$.

This claim follows by induction over j: consider the (j-dimensional) volume $VOL_j(\mathcal{P}_j)$ and note that

$$VOL_j(\mathcal{P}_j) := VOL_{j-1}(\mathcal{P}_{j-1})\frac{H}{j}, \quad \text{so}$$

$$VOL_j(\mathcal{P}_j) = \frac{2}{j!}H^{j-2}\left(\frac{n}{2} - 1\right).$$

Next consider case (B).

We show that there is a simple polyhedron $\mathcal{P}_d$ that is feasible for the vertex set P_d with sum of edge lengths $LEN(\mathcal{P}_d)$ not exceeding

$$n + \frac{(d-2)(d-3)}{2}H\sqrt{2} + (d-2)nH + \frac{n^3(d-2)}{2H}$$

if and only if there is a (Euclidean) tour of length n or less in P_G, i.e., a Hamiltonian cycle in the grid graph.

If there is a tour of length n, it induces a polyhedron $\mathcal{P}$, such that the sum of edge lengths in $\mathcal{P}_{d-k}$ satisfies

$$LEN(\mathcal{P}_j) = LEN(\mathcal{P}_{j-1}) + \sum_{i=3}^{j-1} dist(p_j, p_i) + \sum_{q \in P_G} dist(p_j, q).$$

Now

$$dist(p_j, p_i) = H\sqrt{2}$$

and, for $q \in P_G$,

$$dist(p_j, q) = \sqrt{H^2 + s_q^2},$$

where $s_q = dist(q, g) \le n$, and

$$\left(H + \frac{n^2}{2H}\right)^2 > H^2 + n^2,$$

thus

$$\sum_{q \in P_G} dist(p_j, q) < nH + \frac{n^3}{2H}.$$

Using these relations, we get

$$LEN(\mathcal{P}_d) < n + \frac{(d-2)(d-3)}{2} H\sqrt{2} + (d-2)nH + \frac{n^3(d-2)}{2H}.$$

Now assume that there is no tour of length n in P_G. Then it follows from the previous estimates that

$$LEN(\mathcal{P}_d) \ge n + \sqrt{2} - 1 + \sum_{i=3}^{d} ((i-2)H + nH), \quad \text{so}$$

$$LEN(\mathcal{P}_d) \ge n + \sqrt{2} - 1 + \frac{(d-2)(d-3)}{2} H\sqrt{2} + (d-2)nH.$$

Since $\sqrt{2} - 1 > (d-2)n^3/2H$, the claim holds.

Finally, consider case (C).

For a simple polygon $\mathcal{P}_d$ on P_d, let $FACE_k(\mathcal{P}_d)$ denote the sum of volumes of its k-dimensional faces. For easier notation, we write

$$\begin{aligned}
V_{k+1} &= \binom{d-2}{k+1} \frac{\sqrt{k+1}}{k!} H^k, \\
V_k &= \binom{d-2}{k} \left(nH + \frac{n^3}{2H}\right) \frac{\sqrt{k}}{(k-1)!} H^{k-1}, \\
V_{k-1} &= \binom{d-2}{k-1} n \left(1 + \frac{n^2}{2H^2}\right) \frac{\sqrt{k-1}}{k!} H^{k-1}, \\
V_{k-2} &= \binom{d-2}{k-2} \frac{2}{(k-2)!} H^{k-2} n^2.
\end{aligned}$$

We show that there is a simple polyhedron $\mathcal{P}_d$ feasible for the vertex set P_d with

$$FACE_k(\mathcal{P}_d) < V_{k+1} + V_k + V_{k-1} + V_{k-2}$$

if and only if there is a (Euclidean) tour of length n or less in P_G, i.e., a Hamiltonian cycle in the grid graph.

Consider the k-dimensional faces f of $\mathcal{P}_d$. There are four cases:

Case 1: *f is determined by a set of $k+1$ of the $d-2$ points p_i*. In this case f is a regular k-dimensional simplex and we get

$$VOL_k(f) = \frac{\sqrt{k+1}}{k!} H^k.$$

There are $\binom{d-2}{k+1}$ faces of this form, for a total volume of V_{k+1}.

Case 2: *f is determined by a set of k of the $d-2$ points p_i and one of the points in P_G*. In this case f consists of a regular $(k-1)$-dimensional simplex and a single point at distance at most

$$\sqrt{n^2 + H^2} < H + \frac{n^2}{2H}.$$

We get

$$\frac{\sqrt{k}}{(k-1)!} H^k \leq VOL_k(f) < \left(1 + \frac{n^2}{2H^2}\right) \frac{\sqrt{k}}{(k-1)!} H^k.$$

There are $n\binom{d-2}{k}$ faces of this form, so the total volume of the upper bound is V_k.

Case 3: *f is determined by a set of $k-1$ of the $d-2$ points p_i and an edge in $\mathcal{P}_2$*. Let (q_1, q_2) be the edge of length s_i and $l_1 = dist(q_1, g)$. The volume of the $(k-1)$ dimensional simplex $S_{k-2}^{(q_1)}$ formed by q_1 and the $k-1$ points p_i is

$$VOL_{k-1}(S_{k-2}^{(q_1)}) = \frac{\sqrt{l_1^2 + H^2}}{k-1} VOL_{k-2}(S_{k-2}).$$

The distance of q_2 from $S_{k-2}^{(q_1)}$ lies between $s_i(1 - k/H)$ and s_i. Since

$$H \leq \sqrt{l_1^2 + H^2} < H + \frac{n^2}{2H},$$

we get

$$s_i\left(1 - \frac{k}{H}\right)\frac{\sqrt{k-1}}{k!} H^{k-1} \leq VOL_k(f) < s_i\left(1 + \frac{n^2}{2H^2}\right)\frac{\sqrt{k-1}}{k!} H^{k-1}.$$

There are $\binom{d-2}{k-1}$ faces of this form for each edge of $\mathcal{P}_2$.

If we sum this over all (tour) edges for a fixed simplex S_{k-2}, we get a lower bound of

$$s(T)\left(1 - \frac{k}{H}\right)\frac{\sqrt{k-1}}{k!} H^{k-1}$$

and an upper bound of

$$s(T)\left(1+\frac{n^2}{2H^2}\right)\frac{\sqrt{k-1}}{k!}H^{k-1},$$

where $s(T)$ is the length of the tour. For all the simplices, this yields a total upper bound of V_{k-1}.

Case 4: *f is determined by a set of $k-2$ of the $d-2$ points p_i and the simple polygon* $\mathcal{P}_2$. In this case we get

$$VOL_k(f) = VOL_k(\mathcal{P}_k)$$

and therefore (see Case (A))

$$VOL_j(f) = \frac{2}{(k-2)!}H^{k-2}A,$$

where A is the area enclosed by $\mathcal{P}_2$; since $0 < A < n^2$, we get

$$0 < VOL_k(f) < \frac{2}{(k-2)!}H^{k-2}n^2.$$

There are $\binom{d-2}{k-2}$ faces of this form, for a total upper bound of V_{k-2}. □

Now assume there is a tour of length n. Using the above estimates, we see that this tour induces a polyhedron $\mathcal{P}_d$, such that

$$FACE_k(\mathcal{P}_d) < V_{k+1} + V_k + V_{k-1} + V_{k-2}.$$

For the converse assume that there is no tour of length n. As noted before, this implies $s(T) \geq n + \sqrt{2} - 1$. From our above lower bounds, we get

$$\begin{aligned}
FACE_k(\mathcal{P}_d) \;>\; & V_{k+1} + \binom{d-2}{k} nH \frac{\sqrt{k}}{(k-1)!}H^{k-1} \\
& + \binom{d-2}{k-1}(n+\sqrt{2}-1)\left(1-\frac{k}{H}\right)\frac{\sqrt{k-1}}{k!}H^{k-1} \\
= \;& V_{k+1} + V_k + V_{k-1} + V_{k-2} + \left(\sqrt{2}-1\right)\binom{d-2}{k-1}\frac{\sqrt{k-1}}{k!}H^{k-1} \\
& - \binom{d-2}{k}\left(\frac{n^3}{2}\right)\frac{\sqrt{k}}{(k-1)!}H^{k-2} \\
& + \binom{d-2}{k-1}\left(\frac{n^3}{2}\right)\frac{\sqrt{k-1}}{k!}H^{k-3} \\
& + \binom{d-2}{k-1}(n+\sqrt{2}-1)k\frac{\sqrt{k-1}}{k!}H^{k-2} \\
& + \binom{d-2}{k-2}\frac{2}{(k-2)!}H^{k-2}n^2.
\end{aligned}$$

By multiplying both sides of the easy inequality

$$(d-2)\frac{\sqrt{k}}{\sqrt{k-1}}\frac{n^3}{2}+\frac{n^3}{2H}+(n+\sqrt{2}-1)k+\frac{2(k-1)^2kn^2}{\sqrt{k-1}} < 9(\sqrt{2}-1)d^2n^3 = (\sqrt{2}-1)H$$

with $(((d-2)!\sqrt{k-1})/((k-1)!\,k!))H^{k-2}$, we conclude that the inequality

$$\begin{aligned}&\binom{d-2}{k}\binom{n^3}{2}\frac{\sqrt{k}}{(k-1)!}H^{k-2}+\binom{d-2}{k-1}\binom{n^3}{2}\frac{\sqrt{k-1}}{k!}H^{k-3}\\ &\quad+\binom{d-2}{k-1}(n+\sqrt{2}-1)k\frac{\sqrt{k-1}}{k!}H^{k-2}+\binom{d-2}{k-2}\frac{2}{(k-2)!}H^{k-2}n^2\\ &\quad<\left(\sqrt{2}-1\right)\binom{d-2}{k-1}\frac{\sqrt{k-1}}{k!}H^{k-1}\end{aligned}$$

holds. Thus, it follows that $FACE_k(\mathcal{P}_d) > V_{k+1}+V_k+V_{k-1}+V_{k-2}$, concluding the proof. □

6. Conclusion

In this paper we show that various problems of optimal volume are NP-hard. By means of Pick's theorem, these problems are closely related to problems about inclusion of grid points. Therefore, our proof techniques have consequences for other problems, where the objective is to find a simple closed polygonal curve with few edges that separates two finite sets of points. Complexity of this problem was stated as an open problem by Mitchell [13], and Mitchell and Suri [14]. Using a construction similar to the one in Section 3, it is possible to show hardness of this problem. Details will be described in a future paper.

Since the above problems are NP-hard, it is natural to consider approximation algorithms. Some upper and lower bounds on possible factors for the problem MAX-AREA can be found in [7].

Acknowledgments

This paper is based on parts of my 1992 Ph. D. thesis [6]. I am very grateful to Bill Pulleyblank for his support, supervision, inspiration, and many great conversations that helped me to complete it. Many thanks to David Avis, Charlie Colbourn, Bill Cunningham, and Anna Lubiw for their feedback and comments. Joe O'Rourke asked me about higher-dimensional problems; as a referee, he continued to have an important impact on this work by giving constructive, extensive, and detailed feedback, which turned this into a presentable paper. A second anonymous referee also provided helpful comments. Joe Mitchell told me about the problems of point separation, kept asking me to complete this paper, and gave some helpful final comments. Mike Houle helped to improve the presentation of the paper.

References

1. P. K. Aggarwal and S. Suri. Surface approximation and geometric partitions. *Proc. Fifth Annual ACM–SIAM Symp. on Discrete Algorithms*, 1994, pp. 34–43.
2. J. E. Boyce, D. P. Dobkin, R. L. Drysdale, and L. J. Guibas. Finding extremal polygons. *SIAM Journal of Computing*, **14** (1985), 134–147.
3. H. S. M. Coxeter. *Introduction to Geometry*. Wiley, New York, 1969.
4. D. P. Dobkin, R. L. Drysdale, and L. J. Guibas. Finding extremal polygons. *Advances in Computing Research*, **1** (1983), 181–214.
5. D. Eppstein, M. Overmars, G. Rote, and G. Woeginger. Finding minimum area k-gons. *Discrete & Computational Geometry*, **7** (1992), 45–58.
6. S. P. Fekete. Geometry and the Travelling Salesman Problem. Ph.D. thesis, University of Waterloo, Waterloo, Ontario, 1992.
7. S. P. Fekete and W. R. Pulleyblank. Minimum area polygons. *Proc. Ninth Annual ACM Symp. on Computational Geometry*. 1993, pp. 173–182.
8. M. R. Garey, R. L. Graham, and D. S. Johnson. Some NP-complete geometric problems. *Proc. 8th Annual ACM Symp. on Theory of Computing*, 1976, pp. 10–22.
9. M. R. Garey and D. S. Johnson. *Computers and Intractability, a Guide to the Theory of NP-completeness*. Freeman, San Francisco, CA, 1979.
10. A. Itai, C. H. Papadimitriou, and J. L. Szwarcfiter. Hamilton paths in grid graphs. *SIAM Journal of Computing*, **11** (1982), 676–686.
11. D. S. Johnson and C. H. Papadimitriou. Computational complexity. In: E. Lawler, J. Lenstra, A. Rinnooy Kan, and D. Shmoys, eds., *The Traveling Salesman Problem*. Wiley, New York, 1985, pp. 37–85.
12. C. Mata and J. S. B. Mitchell. Approximation algorithms for geometric tour and network design problems. *Proc. Eleventh Annual ACM Symp. on Computational Geometry*, 1995, pp. 360–369.
13. J. S. B. Mitchell. Approximation algorithms for geometric separation problems. Technical report, AMS Dept., SUNY, Stony Brook, NY, July 1993.
14. J. S. B. Mitchell and S. Suri. Separation and approximation of polyhedral objects. *Computational Geometry Theory and Applications*, **5** (1995), 95–114.
15. I. Niven and H. S. Zuckermann. Lattice points and polygonal area. In: A. K. Stehney, T. K. Milnor, J. E. D'Atri, and T. F. Banchoff, eds., *Selected Papers on Geometry*. Mathematical Association of America, Washington, DC, 1979, pp. 218–233.
16. J. O'Rourke. Polyhedral object models from 3D points. Technical Report IFI-HH-M-77/80, Universität Hamburg, July 1980.
17. J. O'Rourke. Private communication, 1992.
18. J. O'Rourke. *Computational Geometry in C*. Cambridge University Press, Cambridge, 1998.
19. G. Pick. Geometrisches zur Zahlenlehre. *Zeitschrift des Vereins "Lotos*," Prague, 1899, pp. 311–319.
20. C. H. Papadimitriou and U. V. Vazirani. On two geometric problems related to the travelling salesman problem. *Journal of Algorithms*, **5** (1984), 231–246.
21. J. Plesńik. The NP-completeness of the Hamiltonian Cycle Problem in planar digraphs with degree bound two. *Information Processing Letters*, **8** (1979), 199–201.
22. J. E. Reeve. On the volume of lattice polyhedra. *Proceedings of the London Mathematical Society*, **3** (1957), 378–395.
23. D. Ren and J. R. Reay. The boundary characteristic and Pick's theorem in the Archimedean planar tilings. *Journal of Combinatorial Theory, Series A*, **44** (1987), 110–119.
24. P. Rosenstiehl and R. E. Tarjan. Rectilinear planar layouts and bipolar orientations. *Discrete & Computational Geometry*, **1** (1986), 342–351.
25. J. S. Salowe. Selecting the kth largest-area convex polygon. In: F. Dehne, J.-R. Sack, and N. Santoro, eds., *Algorithms and Data Structures* (WADS '89 Proceedings). Lecture Notes in Computer Science, vol. 382. Springer-Verlag, Berlin, 1989, pp. 243–250.

Received June 26, 1997, *and in revised form February* 13, 1999, *and May* 19, 1999.