On the time bound for convex decomposition of simple polygons

Mark Keil
Department of Computer Science
University of Saskatchewan

Jack Snoeyink*
Department of Computer Science
University of British Columbia

Abstract

We show that a decomposition of a simple polygon having n vertices, r of which are reflex, into a minimum number of convex regions without the addition of Steiner vertices can be computed in $O(n + r^2 \min\{r^2, n\})$ time and space. A Java demo is available at `http://www.cs.ubc.ca/spider/snoeyink/demos/convdecomp`

1 Introduction

Suppose that P is a simple polygon in the plane with n vertices, r of which are *reflex*—their interior angles are greater than π. The *minimum convex decomposition problem* asks for a decomposition of the interior of P into the minimum number of convex regions.

There is an anomaly in the literature on minimum convex decomposition: If segments used in the decomposition can end at arbitrary *Steiner points*, as in Figure 1, then Chazelle and Dobkin [2, 3] have shown that a minimum decomposition can be computed by dynamic programming in $O(n + r^3)$ time. On the other hand, if the segments must end at vertices of the polygon, as in Figure 2, then the best published bound is a dynamic programming algorithm of Keil [9, 10] that computes a minimum decomposition in $O(nr^2 \log r)$ time. As Chazelle and Dobkin noted in 1985 [3], this is asymptotically slower for any non-constant r, even though allowing Steiner points usually makes optimization problems more difficult.

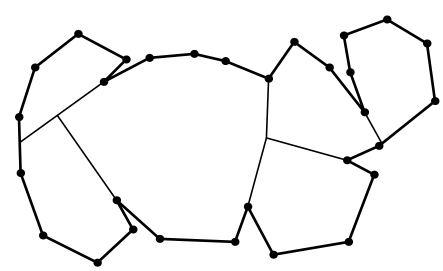

Figure 1: Minimum convex decomposition with Steiner points

After some definitions, we show in Section 3 that Keil's dynamic programming algorithm can be implemented using stacks in place of a search structure, which removes a $\log r$ factor from the running time. In Section 4 we show that the input can be reduced, in $O(n + r^2 \log n)$ time, to a polygon that has the same minimum decomposition but only $\min\{r^2, n\}$ sides. Thus, a minimum decomposition of P can be computed in $O(n + \min\{nr^2, r^4\})$ time, matching Chazelle and Dobkin's running time at least when $r = O(\sqrt[4]{n})$. If we apply a similar reduction when Steiner points are allowed, we obtain a polygon that has only $O(r)$ sides. This may explain why the problem without Steiner points appears harder.

*Supported in part by an NSERC Research Grant, the Killam Foundation, and the CIES. This work was performed while on leave at INRIA Sophia Antipolis.

A related problem is that of finding a convex decomposition with "minimum ink"—that is, having minimum total edge length. We can simplify the search structure from Keil's minimum ink algorithm [9] to achieve $O(n^2r^2)$ time, but cannot reduce the dependence on n. Greene [7] already achieved this time bound with a double dynamic programming algorithm that was a contemporary of Keil's.

2 Notation for simple polygons

Assume that we are given the n vertices of a polygon $P = \{p_0, p_1, \ldots, p_{n-1}\}$ in counter-clockwise (ccw) order. We assume that P is *simple:* that is, the only intersections between the polygon *edges,* the segments $\overline{p_i p_{i+1}}$ that form the boundary of P, are at the shared endpoint of adjacent edges.

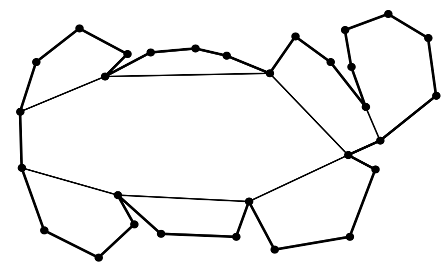

Figure 2: A minimum convex decomposition without Steiner points

A *diagonal* of P is a segment that joins two vertices of P and remains strictly inside P—we use the notation d_{ij} for the diagonal $\overline{p_i p_j}$ with $i < j$. By convention, we will say that $d_{0(n-1)}$ is also a diagonal, and sometimes we will allow d_{ij} to denote a polygon edge when $j = i + 1$. Diagonals for a given vertex p_i can be found by computing the *visibility polygon* for p_i, which can be done in linear time [6]. The following observation is important for visibility algorithms, and will be important for us.

Observation 1 *The order of diagonals ccw around a vertex p_i is the same as the ordering of their other endpoints ccw around P.*

A vertex of P is *reflex* if its interior angle is greater than π. It is not hard to see that a minimum convex decomposition without Steiner points must use diagonals that have at least one reflex vertex. In fact, an easy way to obtain a decomposition into at most four times the maximum number of convex pieces [7, 8] is to start with any triangulation of P, consider each diagonal in turn, and remove a diagonal unless doing so forms a reflex angle at one of its endpoints.

3 Dynamic programming for convex decomposition

To solve the minimum convex decomposition problem we use *dynamic programming*, which is an algorithmic paradigm that finds the best solution to a problem by combining optimal solutions to subproblems [1].

3.1 Defining subproblems

We define a subproblem for each diagonal d_{ik} that uses at least one reflex vertex; let P_{ik} denote the polygonal line $p_i, p_{i+1}, \ldots, p_k$. Note that by adding the diagonal d_{ik} to P_{ik} we obtain a polygon inside P. The size of a subproblem is the number of vertices in P_{ik}; solving the subproblems from smallest to largest gives us a minimum decomposition of $P_{0(n-1)}$, which is also a minimum decomposition of P.

We associate with P_{ik} a *weight* w_{ik}, which is the minimum number of diagonals needed to obtain a convex decomposition. By convention, $w_{i(i+1)} = -1$. Since there may be exponentially-many decompositions that attain weight w_{ik}, Keil [9] suggests storing only certain equivalence classes of decompositions. Each class can be represented by a pair of diagonals or edges incident to p_i and p_k that bound the convex region adjacent to d_{ik}, since this pair forms the interface between subproblems on both sides of d_{ik}. Specifically, we store the *narrowest pairs* whose convex region in a small neighborhood of d_{ik} does not contain the convex region of any other minimum decomposition of P_{ik}. By Observation 1, we can test for narrowest pairs by simply testing indices.

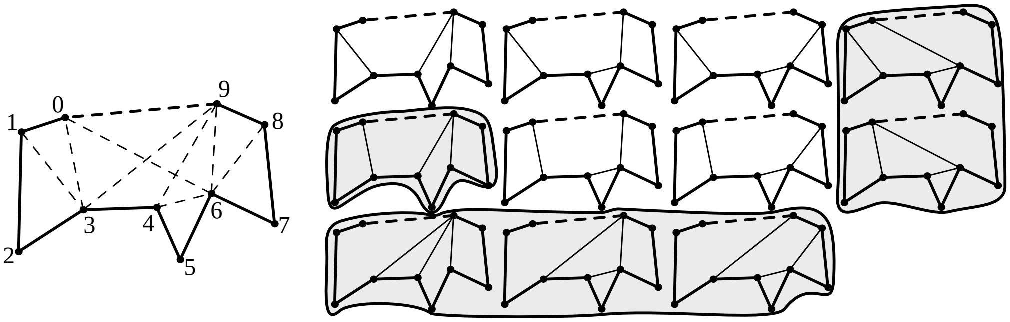

Figure 3: Polygon P_{09} and the diagonals used in its eleven minimum convex decompositions; decompositions with *narrowest pairs* are circled

Figure 3 shows the example of P_{09} with its eleven convex decompositions that use only three diagonals: in the rows we choose one of d_{03}, d_{13}, or d_{39}, and in columns we choose d_{49} with d_{69} or choose d_{46} with d_{69}, d_{68}, or d_{06}. The narrowest decompositions, circled, fall into three equivalence classes. As we compute indices for narrowest pairs, we store them on a stack $\mathcal{S}_{ik}$ so that the segments from bottom to top are in ccw order around p_i and p_k. Stack $\mathcal{S}_{09}$ for Figure 3 would contain $(1,3)$, $(3,4)$, and $(6,8)$, from bottom to top. Thus, we would know that diagonal d_{06} and edge d_{89} formed the narrowest pair that was furthest counter-clockwise.

3.2 Solving subproblems

It is important to observe that any minimum convex decomposition can be found by investigating subproblems defined by diagonals for which at least one endpoint is a reflex vertex. One way to make this apparent is to define a *canonical triangulation* for any minimum convex decomposition.

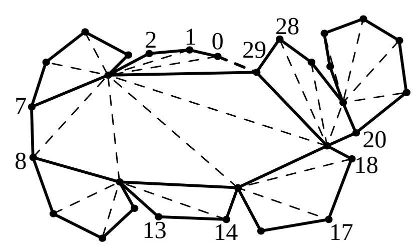

Figure 4: Canonical triangulation from minimum decomposition

We assume that vertex p_0 is reflex either by convention, or by renumbering vertices of P if P is not already convex. Then any minimum convex decomposition can be completed to a *canonical triangulation*: in each convex region, connect the reflex vertex with lowest index to all vertices with higher index. If any vertex remains unconnected in a region, then the vertex with highest index in the region is reflex; connect it to the remaining vertices. In Figure 4, the vertices between p_8 and p_{12} and between p_{20} and p_{26} connect to the higher index; all other regions connect to the lowest.

In a canonical triangulation, each diagonal d_{ik}, with $i < k$, has three properties:

1. If p_i is reflex, then the adjacent triangle $\triangle p_i p_j p_k$, with $i < j < k$, either has $j = k - 1$ or d_{jk} is a diagonal used in the convex decomposition.
2. If p_i is not reflex, then p_k must be. The adjacent triangle $\triangle p_i p_j p_k$, with $i < j < k$, either has $j = i + 1$ or d_{ij} is a diagonal used in the convex decomposition.
3. The diagonals with endpoints in P_{ik} define a canonical triangulation of P_{ik}.

We can use these properties to systematically explore the canonical triangulations of the minimum decompositions of P_{ik} that have narrowest pairs. We assume that we have, for each subproblem P_{xy} that is smaller than P_{ik}, the narrowest pairs for all minimum convex decompositons of P_{xy}. These are in ccw (increasing) order in stack $\mathcal{S}_{xy}$ and in cw (decreasing) order in stack $\mathcal{T}_{xy}$. (The data structure can be implemented as a single list with two independent "stack" pointers that start on either end.) We use stack $\mathcal{T}_{ij}$ or $\mathcal{S}_{jk}$, depending on whether p_i is a reflex vertex or not, to produce the narrowest pairs for minimum decompositions of P_{ik} in ccw order on stack $\mathcal{S}_{ik}$. Then we make $\mathcal{T}_{ik}$ from $\mathcal{S}_{ik}$, so that both are available for subsequent computation.

A. p_i **reflex:** Minimum decompositions use, for some $i < j < k$, the diagonal or edge d_{jk}, a decomposition of P_{jk}, and a decomposition of P_{ij}, perhaps with the diagonal d_{ij}. To find the narrowest decompositions, we consider, in increasing order, the vertices j with $i < j < k$ and d_{ij} and d_{jk} in the visibility graph.

Popping the cw-ordered stack $\mathcal{T}_{ij}$ will go through the pairs in ccw order: find the last pair (s, t) such that d_{tj} and d_{jk} do not form a reflex angle at p_j. If, as in the upper half of Figure 5, there is no such pair (s, t), or if d_{is} and d_{ik} form a reflex angle at p_i, then we must use diagonal d_{ij} to obtain a convex decomposition of P_{ik} with weight $w_{ij} + w_{jk} + 2$ and narrowest pair (j, j). (Recall our convention that the weight of any polygon edge $w_{i(i+1)} = -1$.) Otherwise, as in the lower half of Figure 5, we obtain a convex decomposition of P_{ik} with weight $w_{ij} + w_{jk} + 1$ and narrowest pair (s, j).

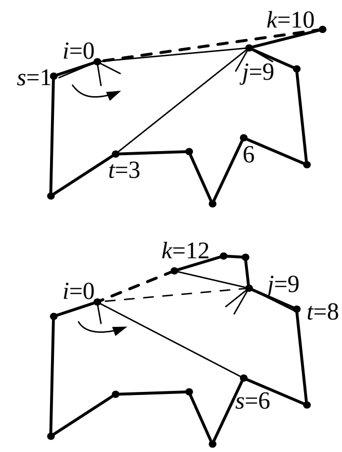

Figure 5: p_i reflex: use d_{jk} and perhaps d_{ij}

To build the ccw-ordered stack $\mathcal{S}_{ik}$ of narrowest pairs for P_{ik} is easy since the second element is always the loop index j. For each pair (x, j) that achieves the minimum weight, push (x, j) if the first element of the pair on top of $\mathcal{S}_{ik}$ is $< x$. Otherwise the pair on top of the stack $\mathcal{S}_{ik}$ is narrower.

Because we use d_{jk} in the decomposition, either $j = k - 1$ so that d_{jk} is a polygon edge or at least one of p_j and p_k is reflex.

B. p_i **not reflex:** This case is symmetric except that we know that p_k is reflex. Since minimum decompositions use d_{ij}, either p_j is reflex or d_{ij} is a polygon edge. We consider, in increasing order, the index $j = i + 1$ and indices of reflex vertices p_j with $i + 1 < j < k$ for which d_{ij} and d_{jk} are in the visibility graph.

Popping the ccw-ordered stack $\mathcal{S}_{jk}$ will go through the pairs in cw order: find the last pair (s, t) such that d_{ij} and d_{js} do not form a reflex angle at p_j. If there is no such pair (s, t), or if d_{tk} and d_{ik} form a reflex angle at p_k, then we must use diagonal d_{jk} to obtain a convex decomposition of P_{ik} with weight $w_{ij} + w_{jk} + 2$ and narrowest pair (j, j). Otherwise, we obtain a convex decomposition of P_{ik} with weight $w_{ij} + w_{jk} + 1$ and narrowest pair (j, t).

To build the ccw-ordered stack $\mathcal{S}_{ik}$ of narrowest pairs for P_{ik} is again easy since now the first element is the loop index j. For each pair (j, x) that achieves the minimum weight, while the second element of the pair on top of $\mathcal{S}_{ik}$ is $\geq x$, pop $\mathcal{S}_{ik}$. Then push (j, x).

Given polygon $P = \{p_0, p_1, \ldots, p_{n-1}\}$, compute a minimum convex decomposition by dynamic programming. This procedure shows the flow of control; "Type" subroutines are described in the text.

```
Procedure MCD(P)
    Initialize weights of edges w_i(i+1) = -1;
    for visible(i, i+2) { w_i(i+2) = 0; push (i+1, i+1) on S_i(i+2); }

    for size = 3 to n do {
        for reflex vertices p_i with i + size ≤ n do
            k = i + size; if visible(i, k) {
                if (p_k reflex) for j = i+1 to k-1 do TypeA(i, j, k);
                else {                    /* p_j must be reflex or j = k-1 */
                    for reflex p_j with i < j < k-1 do TypeA(i, j, k);
                    TypeA(i, k-1, k); }
            }
        for reflex vertices p_k with size ≤ k < n do
            i = k - size; if (p_i not reflex and visible(p_i, p_k)) {
                TypeB(i, i+1, k); /* Do edge, j = i+1, then do reflex */
                for reflex p_j with i+1 < j < k do TypeB(i, j, k);
            }
        Stack S_ik is complete in ccw order; form stack T_ik in cw order;
    }
```

Algorithm 1: MCD algorithm

3.3 Correctness and analysis

Algorithm 1 shows the flow of control for the dynamic programming.

Theorem 2 *Given a simple polygon with n vertices, r of which are reflex, we can solve the minimum convex decomposition problem in $O(nr^2)$ time and space.*

Proof: It is not difficult to bound the total running time of Algorithm 1: the algorithm calls subroutines `TypeA`(i, j, k) or `TypeB`(i, j, k) for triples $i < j < k$ that are indices of at least two reflex vertices, or one reflex vertex and polygon edge. Thus, there are less than nr^2 calls. The work done in each subroutine is constant plus the number of pairs popped from stacks; since each subroutine adds at most one pair to two stacks, at most $O(nr^2)$ elements can be popped. Thus, $O(nr^2)$ bounds the total time. The memory requirements in the worst case are dominated by the $O(nr^2)$ space for the stacks.

To prove the correctness of Algorithm 1 we can argue by induction that we inspect the canonical triangulations and find the narrowest pairs in ccw order for each minimum convex decomposition of P_{ik}, where p_i or p_k is reflex. The key is that by our flow of control—solving subproblems from smallest to largest—pairs popped from a stack will never be needed again. For example, if while solving P_{ik} we pop the ccw-ordered stack $\mathcal{S}_{jk}$, then Observation 1 implies that any $P_{i'k}$ with $i < i'$ that uses subproblem P_{jk} will have $d_{i'j}$ clockwise of d_{ij}, and thus would also require popping $\mathcal{S}_{jk}$. ■

4 Biased convex decompositions

To reduce the dependence on n, the input size, we can look for a decomposition of a special form. We say that a minimum convex decomposition is *biased* if diagonals that ends at a convex vertex can neither be moved to the next vertex ccw, nor deleted and replaced by a reflex-reflex diagonal (RR-diagonal) while maintaining a convex decomposition.

We single out two special types of diagonals that end at convex vertices; these definitions are easier to illustrate (Figure 6) than to write down. A diagonal $\overline{rp}$, with reflex vertex r and convex vertex p, is a *reflex extension*, or *RE-diagonal,* if the extension through r of the edge after r in ccw order first hits vertex p or the edge after p. Similarly, diagonal $\overline{rp}$ is a *diagonal extension,* or *DE-diagonal,* if the extension through r of an RR-diagonal or RE-diagonal incident on r first hits vertex p or the edge after p.

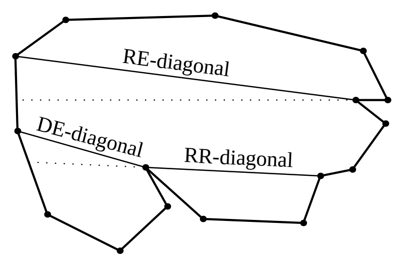

Figure 6: Types of diagonals

Note that an RE-diagonal or DE-diagonal cannot be moved in a convex decomposition; doing so would create a reflex angle with a polygon edge (in the RE-diagonal case) or with an RR-diagonal or RE-diagonal (in the DE-diagonal case) that is incident to its reflex vertex. In fact, these are the only possible obstructions to moving the *lead diagonal* at any convex vertex of P—the diagonal that bounds the same face as the next polygon edge ccw from that vertex.

Lemma 3 *In a biased decomposition of a polygon P, the lead diagonal at any convex vertex $p_i \in P$ must be an RE-diagonal or DE-diagonal.*

Proof: Let $\overline{rp_i}$ be the lead diagonal under consideration in the convex decomposition. By the definition of lead diagonal, the next vertex p_{i+1} is also on the boundary of the convex region on the left of $\overline{rp_i}$. Therefore, $\overline{rp_{i+1}}$ is a diagonal of P that does not intersect any other diagonals of the decomposition.

On the other hand, because the decomposition is biased, we cannot replace diagonal $\overline{rp_i}$ with $\overline{rp_{i+1}}$. Because the replacement cannot cause non-convexity at p_i or at p_{i+1}, it must do so at r.

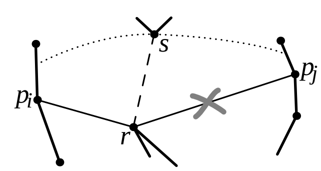

Figure 7: Replacing $\overline{rp_j}$ with $\overline{rs}$

We argue that if the non-convexity at r is not caused by a polygon edge or an RR-diagonal, then it is caused by an RE-diagonal at r. Assume, therefore, that non-convexity is caused by a diagonal $\overline{rp_j}$, with p_j a convex vertex. If the convex region that has p_i, r, and p_j on its boundary also has another reflex vertex s, then adding the RR-diagonal $\overline{rs}$ as in Figure 7, would allow us to delete $\overline{rp_i}$ or $\overline{rp_j}$—one of these deletions will maintain the convex decomposition. Thus, by the definition of biased decompositions, we conclude that the portion of the boundary from p_j ccw to p_i is a convex chain of polygon edges. Since p_j cannot move ccw, it must be an RE-diagonal.

This completes the proof that $\overline{rp_i}$ must be an RE-diagonal or DE-diagonal. ∎

As an easy corollary, any convex decomposition can be converted to a biased decomposition by deleting, moving and replacing diagonals. We are not concerned about the time complexity of this process, just that it is sufficient to look for a minimum convex decomposition among those that are biased.

Corollary 4 *Any convex decomposition can be converted to a biased decomposition by a finite number of steps that delete, move and replace diagonals.*

Proof: Each operation either decreases the number of non-RR-diagonals, or advances a non-RR-diagonal's endpoint. No endpoint is revisited. ∎

Notice that we now know a subset of vertices of P that can be used as endpoints of diagonals in a minimum, biased, convex decomposition—the set of vertices that can be endpoints of RE-diagonals or DE-diagonals. We can explicitly construct this set by ray shooting, which takes $\log n$ time for each ray extended in P [4, 5]. A tempting idea, therefore, is to reduce P to a polygon that uses just a few more vertices, and run the dynamic programming algorithm on the reduced polygon.

This idea works well for the decomposition that allows Steiner points—it becomes clear that the time complexity is of the form $O(n + r \log n + T(r))$. In polygon P, mark the edges incident on reflex vertices, then shoot inside P from every reflex vertex along the extensions of the incident edges, and mark the edges hit. Form polygon P' by omitting from P all vertices not incident to marked edges; P' has at most $7r$ vertices.

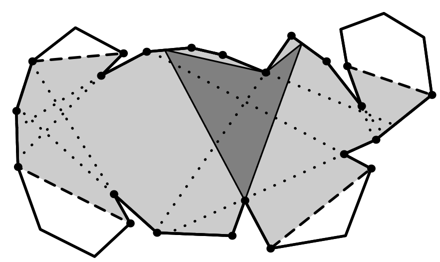

Figure 8: Forming P'

To prove that P' is simple, consider deleting vertices one by one and stop when the first intersection occurs. This must be by an edge passing over a reflex vertex. As illustrated in Figure 8, however, the shortest path that joins the extensions of edges at any reflex vertex is the same in P' as it is in P since it turns only at reflex vertices. The extensions and this path certify that no edge of P' crosses a reflex vertex.

Chazelle and Dobkin's algorithm for convex decomposition with Steiner points uses single "RE-diagonals" and X-configurations—embedded trees that join reflex vertices and have all angles bounded by π. A biased decomposition will move the RE-diagonals so that they end on edges of P that are included in P'. The X-configurations will not be affected, but will remain in P'. Thus, a minimum decomposition of P' gives a minimum decomposition of P.

To perform a similar reduction for the non-Steiner problem, we would have to add, in the worst case, $\min\{n, r^2\}$ endpoints for DE-diagonals, since there are potentially $r(r-1)$ RR-diagonals and r RE-diagonals that must be extended. Thus, after $O(n + \min\{n, r^2\} \log n)$ time for ray shooting and other preprocessing, the dynamic programming algorithm runs in $O(\min\{n, r^2\} r^2)$ time, giving $O(n + \min\{nr^2, r^4\})$ time overall.

Theorem 5 *Given a simple polygon with n vertices, r of which are reflex, we can solve the minimum convex decomposition problem in $O(n + \min\{nr^2, r^4\})$ time and space.*

5 Acknowledgment

We thank Mariette Yvinec and Sylvan Lazard for discussions.

References

[1] R. E. Bellman. *Dynamic Programming.* Princeton University Press, Princeton, NJ, 1987.

[2] B. Chazelle and D. P. Dobkin. Decomposing a polygon into its convex parts. In *Proc. 11th Annu. ACM Sympos. Theory Comput.*, pages 38–48, 1979.

[3] B. Chazelle and D. P. Dobkin. Optimal convex decompositions. In G. T. Toussaint, editor, *Computational Geometry*, pages 63–133. North-Holland, Amsterdam, Netherlands, 1985.

[4] B. Chazelle, H. Edelsbrunner, M. Grigni, L. J. Guibas, J. Hershberger, M. Sharir, and J. Snoeyink. Ray shooting in polygons using geodesic triangulations. *Algorithmica*, 12:54–68, 1994.

[5] B. Chazelle and L. J. Guibas. Visibility and intersection problems in plane geometry. *Discrete Comput. Geom.*, 4:551–581, 1989.

[6] H. ElGindy and D. Avis. A linear algorithm for computing the visibility polygon from a point. *J. Algorithms*, 2:186–197, 1981.

[7] D. H. Greene. The decomposition of polygons into convex parts. In F. P. Preparata, editor, *Computational Geometry*, volume 1 of *Adv. Comput. Res.*, pages 235–259. JAI Press, London, England, 1983.

[8] S. Hertel and K. Mehlhorn. Fast triangulation of the plane with respect to simple polygons. *Inform. Control*, 64:52–76, 1985.

[9] J. M. Keil. Decomposing a polygon into simpler components. *SIAM J. Comput.*, 14:799–817, 1985.

[10] J. M. Keil and J.-R. Sack. Minimum decompositions of polygonal objects. In G. T. Toussaint, editor, *Computational Geometry*, pages 197–216. North-Holland, Amsterdam, Netherlands, 1985.