Around and around: Computing the shortest loop *

John Hershberger
DEC Systems Research Center

Jack Snoeyink[†]
Department of Computer Science
Utrecht University

Abstract

We show that the *funnel algorithm*, devised to compute shortest paths in simple polygons, can be used to compute shortest loops in triangulated 2-manifolds. The time and space required is linear in the number of triangles that the path intersects. This work can be seen as generalizing the problems of computing relative convex hulls and minimum perimeter in-polygons.

1 Introduction

The relative convex hull of two simple polygons is the shape of a rubber band that includes one polygon and excludes the other. The minimum perimeter in-polygon of a convex polygon is the polygon with minimum perimeter that touches every edge of the convex polygon. The computation of either of these polygons can be transformed into the more general question of computing the shortest loop of a given homotopy class in a triangulated 2-manifold. In this paper, we show that the funnel algorithm of Lee and Preparata [6], which was devised to compute the shortest path between two points in a simple polygon, can be used to solve this problem. The time and space required is proportional to the number of triangles intersected by a representative of the homotopy class.

Before we go further, let us define our problem more precisely. We make similar definitions in a companion paper on minimum length paths between two points [5]. A *boundary-triangulated 2-manifold*, or BTM, is a 2-manifold composed of triangles in which all vertices are boundary vertices. Triangulated polygons are the most important examples of BTMs; in general, however, BTMs need not have planar embeddings.

The length of a path or loop in a BTM is defined to be the sum of the Euclidean (L_2) lengths in each triangle. Notice that a minimum length path will consist of line segments and that a minimum path that crosses several triangles without touching vertices will be a straight line if the triangles are unfolded to lie flat in a plane.

Two loops are said to be *homotopic* in a manifold M if one can be deformed to the other without leaving M. Two paths (or open curves) from p to q are homotopic in M if one can be deformed to the other without leaving M and while keeping points p and q fixed. We can now state our problem:

Problem: Given a loop α in a BTM M, compute a curve homotopic to α of minimum length.

What are probably the two most important special cases of this problem have previously been solved in linear time. Toussaint has studied the problem of computing the relative convex hull of two simple polygons in connection with the separability of polygons under translation [1, 7, 8]. He begins by triangulating the region between the polygons to obtain a BTM that has a particularly nice planar embedding. Czyzowicz et al. [3] have recently solved the "Aquarium Keeper's Problem," a generalization of the problem of computing the minimum perimeter polygon that touches each edge of a given convex polygon. Essentially,

*This research was supported by Digital Equipment Corporation and the ESPRIT Basic Research Action No. 3075 (project ALCOM).

†On leave from the Department of Computer Science of the University of British Columbia.

they use the reflection principle to convert this problem to one of computing the shortest loop around a triangulated annulus or Möbius strip. They solve this in linear time using shortest path maps [4]. We gain some advantage from looking at the more general versions of these problems: our algorithms are trivial extensions of Lee and Preparata's funnel algorithm for computing shortest paths in simple polygons and our proofs are slightly easier.

In the next section, we examine some properties of shortest paths in BTMs and reduce the problem of computing the shortest loop of a given homotopy type to the problem of finding the shortest loop around a *band*—a restricted type of BTM. We also review the funnel algorithm. Sections 3 and 4 deal with the two cases of this problem—orientable and non-orientable bands, respectively.

2 Preliminaries: bands, funnels, turns, and cuts

The *dual graph* of a BTM has a vertex for each triangle; edges connect vertices whose corresponding triangles share a common edge. We define a *band* to be a BTM whose dual graph is a single cycle. In this section, we investigate some properties of shortest loops and bands. We reduce the general shortest loop problem to the case of finding the shortest loop that is homotopic to the cycle of a band.

A loop or path α intersects triangulation edges in a sequence. We can define the *canonical loop* (or path) for a sequence of triangulation edges to be the curve that visits the midpoints of the edges in order. By looking at deformations to the canonical loop, one can prove:

Lemma 2.1 *Any two loops (or paths with same starting and ending points) with the same sequences are homotopic.*

In the full paper, we also prove:

Lemma 2.2 *The sequence of triangulation edges intersected by the shortest loop of a homotopy class can be obtained by repeatedly removing pairs of repeated edges.*

Proof: If an edge appears twice in succession in a sequence of a loop, then the loop can be shortened by running along the edge rather than crossing it. The proof that removing edges gives the sequence of the shortest loop is left to the reader. ∎

Given these results, it is easy to construct a covering space for the triangles intersected by the shortest loop in time proportional to the length of the sequence. For more on covering spaces see [5].

Theorem 2.1 *In a BTM M with a loop α, we can compute a band whose shortest loop is the lift of the shortest loop homomorphic to α in M. Computation time and space is proportional to the number of times α intersects a triangle of M.*

Proof: We can construct the sequence of triangulation edges intersected by the shortest loop in M by casting out duplicates according to lemma 2.2. We then make a copy of a triangle for each time it is visited by the sequence and glue these copies together along the edges of the sequence so the resulting manifold is a band whose cycle hits the edges in sequence. Since cutting any edge of the sequence makes the manifold simply connected, the shortest loop around this band visits each edge of the sequence. ∎

Now, to compute the shortest loop of a given homotopy class, it is enough to find the shortest loop around a band. We solve two cases of that problem in sections 3 and 4, but first we review the *funnel algorithm* [2, 4, 6] for shortest paths and define the concepts of *turn angles* and *cut manifolds.*

If a path from a point p gives a triangulation edge sequence with no duplications, the funnel algorithm can walk through the sequence and maintain the shortest paths from p to the endpoints of the current edge. The shortest paths from p to an edge $\overline{uv}$ may travel together for a while. At some point a they diverge and are concave until they reach u and v, as shown in figure 1. The region bounded by $\overline{uv}$ and the concave chains to

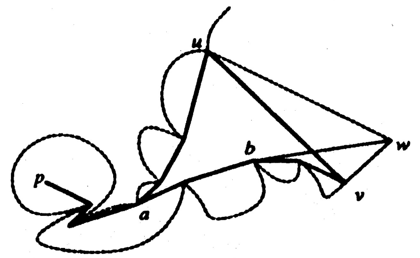

Figure 1: Splitting a funnel

a is called the *funnel;* a is the *apex* of the funnel. We store the vertices of a funnel in a double-ended queue, a *deque*. To update the funnel when the next edge $\overline{uw}$ arrives, we pop points from the deque until we reach b, the tangent to w, then we push w. If the apex of the funnel is popped during the process, then b becomes the new funnel apex. Since each vertex is pushed once and popped at most once, the total time to handle a sequence of n edges is proportional to n. Notice also that once the sequence and the starting and ending points are given, the shortest path is unique.

The *turn angle* (figure 2) of an oriented piecewise-linear path with given starting and ending points in a BTM M is measured by following the orientation of the path and summing the angles of its turns. Each turn has an angle $-\pi < \theta < \pi$; (locally) right turns are negative and left are positive. The *turn angle of a loop* is the turn angle of the path around the loop starting and ending at the orientation of some edge—which edge is chosen does not affect the angle.

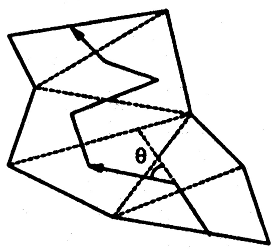

Figure 2: Turn angle

Finally, if we cut a band M along any non-boundary triangulation edge e, we obtain a simply connected manifold M_{cut} whose boundary has two copies of e. The shortest loop around M becomes a shortest path in M_{cut} between two copies of a point $p \in e$. Czyzowicz et al. [3] show how to use shortest path maps to compute the shortest path between two copies in linear time—we use somewhat lighter artillery in the following sections.

If one follows the canonical loop around a band from an edge e back to e, then one finds that the orientation of e has either remained the same or reversed. We will handle these cases separately in the following two sections.

3 Orientable bands

In this section, we show how to find the shortest loop around an orientable band. After defining the *inner boundary* of the band, we state a procedure using the funnel algorithm [6] to compute the shortest loop by walking around the inner boundary twice. We prove its correctness in the rest of the section.

The boundary of an orientable band M consists of two closed curves, σ_r to the right and σ_l to the left of M's cycle. According to the next lemma, the turn angle of the shortest loop in an orientable band equals the turn angle of the canonical loop or of either boundary curve.

Lemma 3.1 *In an orientable BTM, two homotopic loops with the same orientation that have no self intersections have the same turn angle.*

If the turn angle of M is negative, then we say that σ_r is the *inner boundary*, otherwise σ_l is the inner boundary.

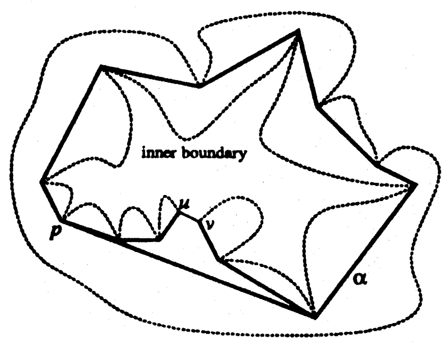

Figure 3: Around the inner boundary

The following procedure computes the shortest loop:

1. Let $\overline{uv}$ be a line segment of the inner boundary.

2. Use the funnel algorithm to compute the shortest path α from u to v that winds around the band twice. (See figure 3.)
3. Let p be a vertex that appears twice on the path; the path from p to p is the shortest loop.

This algorithm is based on the fact that once we identify a point p on a shortest loop, we can compute the loop by computing the shortest path from p back around to p. Lemma 3.2 says that there is a shortest loop touching a vertex of the inner boundary.

Lemma 3.2 *There is a shortest loop that touches a vertex of the inner boundary.*

Proof: If the turn angle of a band M is positive, then the shortest loop must make a left turn. It can only do so by turning at a vertex of the left or inner boundary. The case of a negative turn angle is symmetric.

If the turn angle of the band M is zero then any shortest loop turns as much to the right as to the left. Thus, if it turns at all, it turns at vertices of both the inner and outer boundaries. If the shortest loop does not turn, then cut the band M along a triangulation edge e—the two copies of e are parallel and the shortest loop becomes a straight line segment ℓ between corresponding points of the copies of e. Without changing the length of the segment ℓ, one can translate ℓ to the left until it touches a vertex of the inner boundary. ∎

With this lemma, we can prove correctness.

Theorem 3.1 *Given an orientable band M composed of n triangles, the above procedure correctly computes the shortest loop around M in linear time.*

Proof: Let p be the vertex on the inner boundary of some shortest loop whose existence is proved by lemma 3.2. The shortest path λ from u starts on or inside this shortest loop and reaches p before going completely around the band. Similarly, the shortest path from v reaches p before going around the band in the other direction. Thus, p is reached twice.

The path α can thus be decomposed into three pieces: the shortest path from u to p, denoted α_u; the shortest loop around the band, denoted λ; and the shortest path from p to v, denoted α_v. The vertices of λ are obviously the vertices of the shortest loop. Together α_u and α_v compose the shortest path from u around to v—a vertex appears on this path only once. Thus, any vertex that appears twice on α is on the shortest loop and can be used in place of p. ∎

4 Non-orientable bands

One might think that computing the shortest loop in a non-orientable band would be more difficult. In this section, however, we show how to find the shortest closed curve that winds twice around the band by a reduction to an orientable band. We then show how to obtain the shortest loop from this curve. The result is theorem 4.1.

Theorem 4.1 *Given a non-orientable band M composed of n triangles, one can compute the shortest loop around M in linear time.*

We can conceptually take two copies of M_{cut}, one reversed left to right, and paste them into a single band M_{double}. The band M_{double} is orientable and has turn angle zero: starting from triangulation edge e, you travel through one copy of M_{cut} until you encounter the reversed copy, denoted e_R. Then you travel through the reversed copy of M_{cut} until you reach e again. The turn angles in each copy of M_{cut} have opposite sign. We can use the procedure of the previous section to find the shortest loop in M_{double} that touches the left boundary—call it λ. Notice that λ is the shortest closed curve that winds around M twice, so its length is at most double the length of the shortest loop in M. We shall see that the length is exactly double.

Suppose λ intersects e at a point p. Then the shortest loop touching the right boundary is the shortest path starting and ending at the corresponding point $p_R \in e_R$. In other words, the shortest loop in M_{double} touching the right boundary is λ_R—the loop λ viewed from the perspective of edge e_R. This should not be surprising as M has only one boundary.

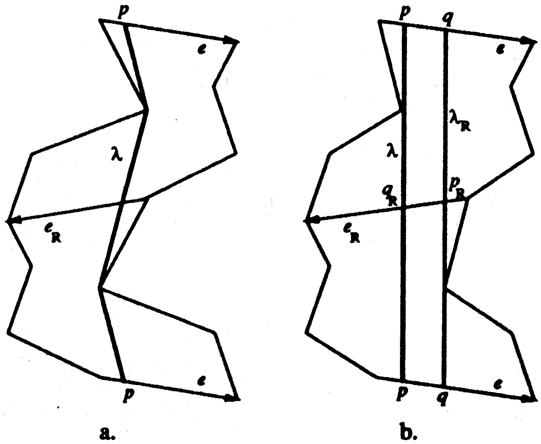

Figure 4: Cases for the shortest loop in M_{double}

We now consider two cases depicted in figure 4. First, if the shortest loop λ in M_{double} makes any turns, then λ makes turns on vertices of both the right and left boundaries. Since the shortest loop touching a given boundary is unique, both loops λ and λ_R are identical. Therefore, λ passes through the point $p_R \in e_R$—that is, λ winds around the shortest loop in M twice.

Second, if the shortest loop λ makes no turns, then by cutting the manifold M_{double} along e, we see that the loops touching the left and right boundaries, λ and λ_R, form two parallel lines. If the intersections with e are points p on the left and q on the right, as shown in figure 4b, then the intersections with e_R are the corresponding points q_R on the right and p_R on the left. The line λ' parallel to λ and λ_R and passing through the midpoint of the segment $\overline{pq}$ is also a shortest loop in M_{double}. Moreover, λ' also passes through the midpoint of $\overline{q_R p_R}$. But these two midpoints are just the corresponding points on two copies of e. As a result, λ' winds around the shortest loop in M twice.

5 Conclusions

We have shown that the problem of finding the Euclidean shortest loop in a BTM can be reduced to the problem of finding the shortest loop around an orientable or non-orientable band. We have shown that these two problems can be solved by extensions of Lee and Preparata's funnel algorithm for shortest paths. Both the reduction and solution use time and space proportional to the size of the description of the homotopy class. This gives alternative linear time algorithms for relative convex hulls and the aquarium keeper's problem.

References

[1] B. Bhattacharya and G. T. Toussaint. A linear algorithm for determining translation separability of two simple polygons. Technical Report SOCS-86.1, School of Computer Science, McGill University, Montreal, 1986.

[2] B. Chazelle. A theorem on polygon cutting with applications. In *Proceedings of the 23rd IEEE Symposium on Foundations of Computer Science*, pages 339–349, 1982.

[3] J. Czyzowicz, P. Egyed, H. Everett, D. Rappaport, T. Shermer, D. Souvaine, G. Toussaint, and J. Urrutia. The aquarium keeper's problem. In *Proceedings of the Second Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 459–464, Jan. 1991.

[4] L. Guibas, J. Hershberger, D. Leven, M. Sharir, and R. Tarjan. Linear time algorithms for visibility and shortest path problems inside triangulated simple polygons. *Algorithmica*, 2:209–233, 1987.

[5] J. Hershberger and J. Snoeyink. Computing minimum length paths of a given homotopy class. Technical Report RUU-CS-90-37, Utrecht Unversity, Department of Computer Science, Dec. 1990.

[6] D. T. Lee and F. P. Preparata. Euclidean shortest paths in the presence of rectilinear barriers. *Networks*, 14(3):393–410, 1984.

[7] G. Toussaint. On separating two simple polygons by a single translation. *Discrete and Computational Geometry*, 4:265–278, 1989.

[8] G. T. Toussaint. Movable separability of sets. In G. T. Toussaint, editor, *Computational Geometry*, volume 2 of *Machine Intelligence and Pattern Recognition*, pages 335–376. North Holland, Amsterdam, 1985.