Homotopic Fréchet Distance Between Curves

or, Walking Your Dog in the Woods in Polynomial Time

Erin Wolf Chambers[a], Éric Colin de Verdière[b], Jeff Erickson[c], Sylvain Lazard[d], Francis Lazarus[e], Shripad Thite[f]

[a]*Department of Mathematics and Computer Science, Saint Louis University, USA*
[b]*Département d'Informatique, École normale supérieure and CNRS, Paris, France*
[c]*Department of Computer Science, University of Illinois at Urbana-Champaign, USA*
[d]*INRIA Nancy - Grand Est, LORIA, Nancy, France*
[e]*GIPSA-Lab and CNRS, Grenoble, France*
[f]*California Institute of Technology, Center for the Mathematics of Information, USA*

Abstract

The Fréchet distance between two curves in the plane is the minimum length of a leash that allows a dog and its owner to walk along their respective curves, from one end to the other, without backtracking. We propose a natural extension of Fréchet distance to more general metric spaces, which requires the leash itself to move continuously over time. For example, for curves in the punctured plane, the leash cannot pass through or jump over the obstacles ("trees"). We describe a polynomial-time algorithm to compute the *homotopic* Fréchet distance between two given polygonal curves in the plane minus a given set of polygonal obstacles.

Key words: Homotopy, Similarity of curves, Metric space, Homotopic Fréchet distance, Geodesic leash map, Punctured plane

1. Introduction

Given two input curves, it is natural to ask how similar they are to each other. One common measure of curve similarity is the Hausdorff distance, which is the maximum distance between a point on one curve and its nearest neighbor on the other curve. While the Hausdorff metric does measure closeness in space, it does not take into account the flow of the curves, which is important for many applications, such as morphing in computer graphics.

The *Fréchet distance* between two curves, sometimes also called the *dog-leash distance*, is defined as the minimum length of a leash required to connect a dog and its owner as they walk without backtracking along their respective curves from one endpoint to the other. The Fréchet metric takes the flow of the two curves into account; the pairs of points whose distance contributes to the Fréchet distance sweep continuously along their respective curves. This property makes the Fréchet distance a better measure of similarity for curves than

Email addresses: echambe5@slu.edu (Erin Wolf Chambers), Eric.Colin.de.Verdiere@ens.fr (Éric Colin de Verdière), jeffe@cs.uiuc.edu (Jeff Erickson), lazard@loria.fr (Sylvain Lazard), francis.lazarus@gipsa-lab.inpg.fr (Francis Lazarus), shripad@caltech.edu (Shripad Thite)

alternatives for arbitrary point sets such as Hausdorff distance. It is possible for two curves to have small Hausdorff distance but large Fréchet distance. Fréchet distance is used in many different applications; see [2, 1, 3, 30] and the references therein.

When the two curves are embedded in a more complex metric space, such as a polyhedral terrain or some Euclidean space with obstacles, the distance between two points on the curves is most naturally defined as the length of the shortest path between them. Variations on the resulting *geodesic Fréchet distance* have been studied by Efrat *et al.* [15], Maheshwari and Yi [24], and more recently Cook and Wenk [12, 13, 14]. The definition of geodesic Fréchet distance allows the leash to switch discontinuously, without penalty, from one side of an obstacle or a mountain to another.

In this paper, we introduce a continuity requirement on the motion of the leash. We require that the leash cannot switch discontinuously from one position to another; in particular, the leash cannot jump over obstacles, and can sweep over a mountain only if it is long enough. We define the *homotopic Fréchet distance* between two curves as the Fréchet distance with this additional continuity requirement. Our continuity requirement is satisfied automatically for curves inside a simple polygon [12, 13, 15], but not in more general environments like convex polyhedra [24] or the plane with obstacles [14].

The motion of the leash defines a correspondence between the two curves that can be used to morph between the two curves—two points joined by a leash morph into each other [15]. Thus, the homotopic Fréchet distance can be thought of as the minimal amount of deformation needed to transform one curve into the other.

Efficiently computing the homotopic Fréchet distance in general metric spaces is a new open problem. We present a polynomial-time algorithm for a special case of this problem, which is to compute the homotopic Fréchet distance between two polygonal curves in the plane minus a set of polygonal obstacles.

The current paper is structured as follows. In Section 2, we give formal definitions of leash maps, homotopic Fréchet distance, relative homotopy classes and related notions, and then describe some relevant preliminary results in Section 3. In Section 4, we present an algorithm that enumerates a finite set of relative homotopy classes of leashes, such that the homotopic Fréchet distance is realized by a leash within one of these classes. We describe an algorithm to compute the homotopic Fréchet distance between two curves in Section 5. In Section 6, we describe extensions of our algorithm to closed curves and to generalizations of homotopic Fréchet distance. Finally, we conclude by suggesting several open problems.

2. Definitions

Let S be a fixed Hausdorff metric space. A ***curve*** or ***path*** in S is a continuous function from the unit interval $[0,1]$ to S. We will sometimes abuse notation by using the same symbol to denote a curve $A\colon [0,1] \to S$ and its image in S. A ***reparameterization of*** $\boldsymbol{[0,1]}$ is a continuous, non-decreasing, surjection $\alpha\colon [0,1] \to [0,1]$. A reparameterization of a curve $A\colon [0,1] \to S$ is any curve $A \circ \alpha$, where α is a reparameterization of $[0,1]$. The ***length*** of any curve A, denoted $\operatorname{len}(A)$, is defined by the metric of S; in particular, two reparameterizations of the same curve are considered to have the same length.

Given two parameters s and t, an $\boldsymbol{(s,t)}$***-leash*** between two curves A and B is another curve $\lambda\colon [0,1] \to S$ such that $\lambda(0) = A(s)$ and $\lambda(1) = B(t)$. A ***leash*** is an (s,t)-leash for some parameters s and t. If either A or B intersects itself, two distinct leashes may be equal as curves while corresponding to different parameters s and t.

A ***leash map*** is a continuous function $\ell\colon [0,1]^2 \to S$ such that $\ell(\cdot,0)$ is a reparameterization of A, and $\ell(\cdot,1)$ is a reparameterization of B. A leash map describes the continuous motion of a leash between a dog walking along A and its owner walking along B; the curve $\ell(t,\cdot)$ is the leash at time t. The ***length*** of a leash map ℓ, denoted $\operatorname{len}(\ell)$, is the maximum length of any leash $\ell(t,\cdot)$. Finally, the ***homotopic Fréchet distance*** between two curves A and B, denoted $\overline{\mathcal{F}}(A,B)$, is the infimum, over all leash maps ℓ between A and B, of the length of ℓ:

$$\overline{\mathcal{F}}(A,B) := \inf_{\text{leash map } \ell\colon [0,1]^2\to S} \left(\max_{0\leqslant t\leqslant 1} \; \operatorname{len}(\ell(t,\cdot)) \right).$$

In contrast, the classical [2] and geodesic [14] Fréchet distances (for which the leash is not required to move continuously) are defined only in terms of reparameterizations and distances:

$$\mathcal{F}(A,B) := \inf_{\substack{\text{reparameterizations} \\ \alpha,\beta\colon [0,1]\to[0,1]}} \left(\max_{0\leqslant t\leqslant 1} \; \operatorname{dist}(A(\alpha(t)), B(\beta(t))) \right),$$

where $\operatorname{dist}(a,b)$ denotes the distance between points a and b in the ambient metric space.

In spaces where shortest paths vary continuously as their endpoints move, such as the Euclidean plane or the interior of a simple polygon, the two notions $\overline{\mathcal{F}}$ and $\mathcal{F}$ are equivalent. In general, however, the homotopic Fréchet distance $\overline{\mathcal{F}}(A,B)$ between two curves A and B could be larger (but never smaller) than the classical Fréchet distance $\mathcal{F}(A,B)$ between them.

Leash maps are closely related to the standard topological notion of *homotopy*. Two curves λ and λ' with the same endpoints are ***homotopic*** if λ can be continuously deformed into λ' without moving the endpoints. More formally, λ and λ' are homotopic if there is a continuous function $h\colon [0,1]^2 \to S$ such that $h(u,0) = \lambda(u)$ and $h(u,1) = \lambda'(u)$ for all $u \in [0,1]$, and $h(0,v) = \lambda(0) = \lambda'(0)$ and $h(1,v) = \lambda'(1) = \lambda'(1)$ for all $v \in [0,1]$. It is easy to prove that being homotopic is an equivalence relation over the set of curves with any fixed pair of endpoints, and thus determines ***homotopy classes***.

An (s,t)-leash λ and an (s',t')-leash λ' are ***homotopic relative to A and B***, or simply ***relatively homotopic***, if λ can be continuously deformed into λ' while keeping each endpoint of the leash on its respective curve. More formally, λ is relatively homotopic to λ' if there are three continuous functions $\alpha,\beta\colon [0,1] \to [0,1]$ and $h\colon [0,1]^2 \to S$, such that $\alpha(0) = s$, $\alpha(1) = s'$, $\beta(0) = t$, $\beta(1) = t'$, and such that $h(u,0) = \lambda(u)$ and $h(u,1) = \lambda'(u)$ for all $u \in [0,1]$, and $h(0,v) = A(\alpha(v))$ and $h(1,v) = B(\beta(v))$ for all $v \in [0,1]$. Again, for any fixed curves A and B, being relatively homotopic is an equivalence relation, which defines ***relative homotopy classes*** of leashes. Clearly, all leashes $\ell(t,\cdot)$ determined by a single leash map ℓ are relatively homotopic.

3. Preliminaries

In this paper, we develop a polynomial-time algorithm to compute the homotopic Fréchet distance between two polygonal curves A and B in the Euclidean plane $\mathbb{E}^2$ minus a set P of polygonal obstacles. In most of the paper, in order to avoid some technicalities, we assume that the obstacles in P are open sets; however, we also consider the special case of point

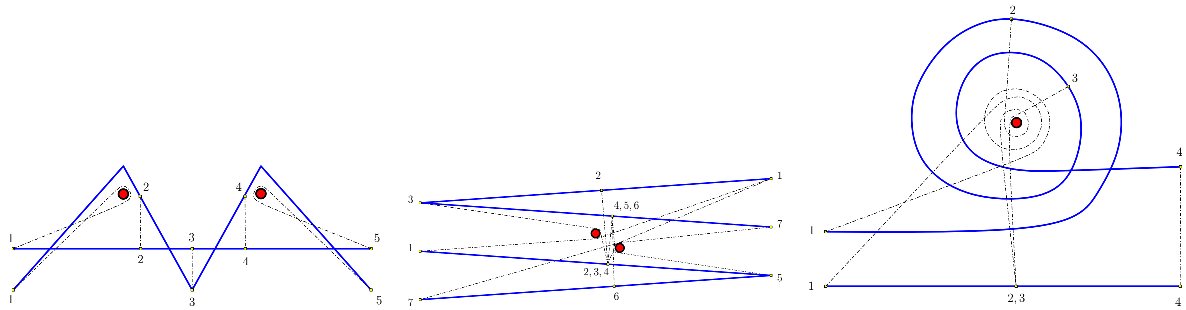

Figure 1: Leash maps for three inputs. Dashed curves between matching numbers represent intermediate leashes.

obstacles in Section 4.4. To simplify our exposition, we assume that no three vertices of the input (vertices of polygons in P or vertices of A and B) are collinear; this assumption can be enforced algorithmically using standard perturbation techniques [29]. Figure 1 illustrates leash maps for a few sample inputs where P is a set of very small polygonal obstacles.

Let $\mathcal{E}$ denote the space $\mathbb{E}^2 \setminus P$ with the metric defined by shortest path distances. In any leash map between curves A and B in $\mathcal{E}$, the moving leash can neither intersect nor jump over any obstacle in P. Curves A and B may self-intersect and intersect each other. For ease of exposition, we will assume that the closures of the obstacle polygons in P are disjoint from each other and from the curves A and B; however, our algorithms can be easily adapted to avoid this restriction.

Let $a_0, a_1, \ldots, a_m$ denote the ordered sequence of vertices of A; these points define a unique parameterization $A\colon [0,1] \to \mathcal{E}$ whose restriction to any range of the form $[(i-1)/m, i/m]$ is an affine map onto the corresponding edge $a_{i-1}a_i$. Similarly, the vertices $b_0, b_1, \ldots, b_n$ of B define a unique piecewise-affine parameterization $B\colon [0,1] \to \mathcal{E}$. Finally, let k denote the total number of vertices in all obstacle polygons.

3.1. Universal Cover

Given a topological space S, its *universal cover* $\widetilde{S}$ is a simply connected topological space that locally resembles S, but is (usually) infinitely larger. Each point x in S corresponds (in general) to an infinite number of points in the universal cover $\widetilde{S}$, one for each homotopy class of curves with both endpoints equal to x. Universal covers are a fundamental concept in topology and using this standard tool considerably simplifies many of our proofs, especially in Section 4.3.

More formally, a continuous function $p\colon \widetilde{S} \to S$ is a ***covering map*** if every point $x \in S$ has an open neighborhood U such that $p^{-1}(U)$ is the union of disjoint open sets $\bigcup_i V_i$, and the restriction of p to each open set V_i is a homeomorphism from V_i to U [26]. If there is a covering map from $\widetilde{S}$ to S, then $\widetilde{S}$ is called a ***covering space*** of S. A point $\widetilde{x}$ in $\widetilde{S}$ is called a ***lift*** of its image $p(\widetilde{x})$ in S; similarly, a path $\widetilde{\alpha}$ in $\widetilde{S}$ is a ***lift*** of its image $p(\widetilde{\alpha})$ in S. Unless the covering map p is a homeomorphism, each point and path in S has several lifts in $\widetilde{S}$. The ***universal cover*** $\widetilde{S}$ is the unique simply connected covering space of S. Two paths α and β in S are homotopic (with fixed endpoints) if and only if they have lifts $\widetilde{\alpha}$ and $\widetilde{\beta}$ with the same endpoints in the universal cover $\widetilde{S}$. The universal cover $\widetilde{S}$ naturally inherits the metric properties of S; for any path $\widetilde{\pi}$ in $\widetilde{S}$, its length is defined as the length of its projection $p \circ \widetilde{\pi}$ in S, where $p\colon \widetilde{S} \to S$ is the covering map. In particular, a path in S is as short as possible in its homotopy class (with fixed endpoints) if and only if it lifts to a globally shortest path in $\widetilde{S}$. For further details, see Munkres [26].

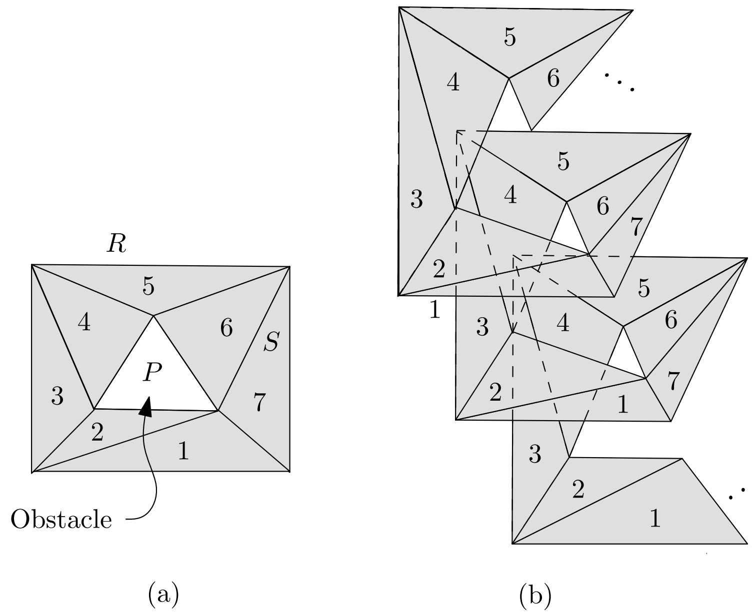

Figure 2: (a) Triangulation Δ of a region S, which is a rectangle minus a triangle. (b) Triangulation $\tilde{\Delta}$ of the universal cover $\tilde{S}$ of S.

We sketch an equivalent constructive definition of the universal cover of a large bounded subset of $\mathcal{E}$, due to Hershberger and Snoeyink [21]; see Figure 2. Let R be a large bounding rectangle strictly containing the obstacles in P and let $S = R \setminus P$. (The homotopic Fréchet distance within S is equal to the homotopic Fréchet distance in $\mathcal{E}$.) Let Δ be any triangulation of S whose vertices are exactly those of P and R and whose set of edges contains the edges of P and R. For example, we can take Δ to be the constrained Delaunay triangulation [9] of the edges of S, minus the triangles inside the obstacles. A triangulation $\widetilde{\Delta}$ of the universal cover $\widetilde{S}$ of S can be obtained by the following incremental construction. Initialize $\widetilde{\Delta}$ with a copy $\tilde{t}$ of some triangle t in Δ. Mark the edges of $\tilde{t}$ that do not correspond to edges of t that are boundary edges of S. While there is a triangle $\tilde{\tau}$ (a copy of some $\tau \in \Delta$) in $\widetilde{\Delta}$ with a marked edge $\tilde{e}$ (a copy of an edge e of τ), glue along $\tilde{e}$ in $\widetilde{\Delta}$ a copy $\tilde{\sigma}$ of the triangle of Δ sharing the edge e with τ. Unmark $\tilde{e}$ and mark the other edges of $\tilde{\sigma}$ that do not correspond to boundary edges of S. Unless Δ is a triangulated disk (i.e., P is empty), the while loop proceeds ad infinitum, and the dual triangulation of $\widetilde{\Delta}$ extends to an infinite tree. For example, in Figure 2, the dual graph of the lifted triangulation $\widetilde{\Delta}$ is an infinite path.

Although universal covers are a convenient tool for proving our results, we emphasize that our algorithm never explicitly constructs the universal cover.

3.2. Geodesics

A ***geodesic*** in a metric space S is a path that is locally as short as possible. More formally, a geodesic is a path $\pi\colon [0,1] \to S$ such that for every parameter $t \in [0,1]$, the restriction of π to a sufficiently small neighborhood of t is a globally shortest path. A path in $\mathcal{E}$ is a geodesic if and only if every lift in $\widetilde{\mathcal{E}}$ is also a geodesic.

The fact that the dual graph of the triangulation of $\widetilde{\mathcal{E}}$ is a tree has important consequences for geodesics and shortest paths in $\widetilde{\mathcal{E}}$, as noted by Hershberger and Snoeyink [21] and other authors [18, 6, 16, 4, 5]. Specifically, because $\widetilde{S}$ is simply connected and locally Euclidean, shortest paths between points in $\widetilde{\mathcal{E}}$ are unique; indeed, every geodesic in $\widetilde{\mathcal{E}}$ is a globally shortest path. Shortest paths in $\widetilde{\mathcal{E}}$ are piecewise linear curves whose internal vertices are lifts of obstacle vertices. Furthermore, shortest paths in $\widetilde{\mathcal{E}}$ vary continuously as the endpoints move continuously.

Hershberger and Snoeyink [21] describe how to algorithmically maintain a shortest path σ in $\widetilde{\mathcal{E}}$ as the endpoints move continuously, by storing the sequence of lifted obstacle vertices that lie on σ in a *double-ended queue* or *deque*. These obstacle vertices partition σ into a sequence of straight line segments. If the first or last segment of σ collides with a lifted obstacle vertex, that vertex is pushed onto the appropriate end of the deque. Conversely, if the first two or last two segments of σ become collinear, the obstacle vertex joining those two segments is removed from the appropriate end of the deque. Except at these critical events, only the first and last segments of the geodesic change as the endpoints of the geodesic move.

3.3. Geodesic Leash Maps

A ***geodesic leash map*** is a leash map $\ell\colon [0,1]\times[0,1]\to\mathcal{E}$ in which every leash $\ell(t,\cdot)$ is a geodesic. We next prove that for any leash map ℓ, there is a geodesic leash map ℓ' in the same relative homotopy class that is no longer than ℓ.

Lemma 3.1. *Suppose ℓ is a leash map between two curves A and B. There is a geodesic leash map ℓ' between A and B such that, for all $t\in[0,1]$, the leash $\ell'(t,\cdot)$ is the shortest path homotopic to $\ell(t,\cdot)$ with fixed endpoints. Additionally, the length of ℓ' is at most the length of ℓ.*

Proof. We lift ℓ to the universal cover $\widetilde{\mathcal{E}}$ of $\mathcal{E}$, obtaining a leash map $\widetilde{\ell}$ between the lifts $\widetilde{A}$ and $\widetilde{B}$ of A and B respectively. For each $t\in[0,1]$, let $\widetilde{\ell'}(t,\cdot)$ be the globally shortest path between the endpoints of $\widetilde{\ell}(t,\cdot)$. Because shortest paths in $\widetilde{\mathcal{E}}$ vary continuously as their endpoints move continuously, $\widetilde{\ell'}$ is a continuous function in both arguments, and therefore a (geodesic) leash map in $\widetilde{\mathcal{E}}$. The projection ℓ' of $\widetilde{\ell'}$ back to $\mathcal{E}$ is a (geodesic) leash map between A and B. For each t, the leash $\ell'(t,\cdot)$ is the shortest path in $\mathcal{E}$ that is homotopic with fixed endpoints to $\ell(t,\cdot)$, so $\operatorname{len}(\ell'(t,\cdot))\leqslant\operatorname{len}(\ell(t,\cdot))$. It follows that $\operatorname{len}(\ell')\leqslant\operatorname{len}(\ell)$. □

This lemma implies that the homotopic Fréchet distance $\overline{\mathcal{F}}(A,B)$ is the infimum, over all relative homotopy classes h, of the *classical* Fréchet distance, where distances are defined by shortest paths in h:

$$\mathcal{F}_h(A,B) := \inf_{\substack{\text{reparameterizations}\\ \alpha,\beta\colon[0,1]\to[0,1]}} \left(\max_{t\in[0,1]} \operatorname{dist}_h(A(\alpha(t)),\, B(\beta(t)))\right)$$

$$\overline{\mathcal{F}}(A,B) := \inf_{\text{relative homotopy class } h} \mathcal{F}_h(A,B).$$

Here, $\operatorname{dist}_h(u,v)$ denotes the length of the shortest path from u to v in relative homotopy class h.

For the rest of the paper, we restrict our attention to geodesic leashes and geodesic leash maps. We call a relative homotopy class h ***optimal*** if $\overline{\mathcal{F}}(A, B) = \mathcal{F}_h(A, B)$. In Section 4, we show that there is at least one optimal relative homotopy class. We also prove a structural result about optimal relative homotopy classes, which leads to a polynomial-time algorithm to enumerate a subset of relative homotopy classes, at least one of which is optimal. Section 5 describes our polynomial-time algorithm to compute the Fréchet distance within a single homotopy class. Combining these two subroutines gives us a polynomial-time algorithm to compute the homotopic Fréchet distance.

4. Structural Properties of Optimal Relative Homotopy Classes

4.1. Minimality

For any relative homotopy class h and any parameters $s, t \in [0, 1]$, let $\sigma_h(s, t)$ denote the shortest path in h between points $A(s)$ and $B(t)$. We define a partial order $\preceq$ on relative homotopy classes as follows: For any two relative homotopy classes h and h', we write $h \preceq h'$ if and only if $\text{len}(\sigma_h(s, t)) \leq \text{len}(\sigma_{h'}(s, t))$ for all parameters s and t. We write $h \prec h'$ whenever $h \preceq h'$ but $h' \npreceq h$.

Lemma 4.1. *For any relative homotopy classes h and h', if $h \preceq h'$, then $\mathcal{F}_h(A, B) \leqslant \mathcal{F}_{h'}(A, B)$.*

Proof. Let ℓ' be any leash map in relative homotopy class h': for some reparameterizations α and β of $[0, 1]$, we have $\ell'(\cdot, 0) = A(\alpha(\cdot))$ and $\ell'(\cdot, 1) = B(\beta(\cdot))$.

Let ℓ be the geodesic leash map in relative homotopy class h defined by the same reparameterizations: $\ell(t, \cdot) = \sigma_h(\alpha(t), \beta(t))$ for all t (Lemma 3.1). The definition of $\preceq$ implies that $\text{len}(\ell(t, \cdot)) \leqslant \text{len}(\ell'(t, \cdot))$ for all t; hence $\mathcal{F}_h(A, B) \leqslant \text{len}(\ell) \leqslant \text{len}(\ell')$. But $\mathcal{F}_{h'}(A, B)$ is the infimum of all such $\text{len}(\ell')$; this concludes the proof. □

A relative homotopy class h is *minimal* if $h' \preceq h$ implies $h \preceq h'$. In other words, h is *not* minimal if there is another relative homotopy class h' such that $h' \prec h$.

4.2. Existence of Minimal Relative Homotopy Classes

Lemma 4.2. *For any relative homotopy class h, there is a minimal relative homotopy class h' such that $h' \preceq h$.*

Proof. Assume, for the sake of contradiction, that there is no minimal relative homotopy class h' such that $h' \preceq h$. In other words, for any $h' \preceq h$ (including $h' = h$), h' is not minimal, so there is another relative homotopy class h'' such that $h'' \prec h' \preceq h$. Then, by induction, we can define an infinite descending chain of relative homotopy classes $h = h_0 \succ h_1 \succ h_2 \succ \cdots$. To simplify notation, let $\sigma_n = \sigma_{h_n}(0, 0)$.

Consider the ordered list of obstacle vertices on each path σ_n. There are finitely many such ordered lists, because $\text{len}(\sigma_n) \leqslant \text{len}(\sigma_0)$ for each n. Thus, for some pair of indices $i < j$, the paths σ_i and σ_j have the same endpoints ($A(0)$ and $B(0)$) and visit the same ordered list of obstacle vertices. Thus, the paths σ_i and σ_j are identical, which implies that their relative homotopy classes h_i and h_j are equal. This is a contradiction. □

Lemmas 4.1 and 4.2 together imply that the homotopic Fréchet distance $\overline{\mathcal{F}}(A, B)$ is the infimum of $\mathcal{F}_h(A, B)$ over all *minimal* relative homotopy classes h:

$$\overline{\mathcal{F}}(A, B) = \inf_{\text{minimal relative homotopy class } h} \mathcal{F}_h(A, B).$$

In the remainder of this section, we prove that all minimal relative homotopy classes have a special form, which implies that the number of minimal relative homotopy classes is finite. (Thus, we can finally replace the infimum in the expression above with a minimum.) We also describe how to enumerate, in polynomial time, a finite set of relative homotopy classes that contains an optimal one. Our overall strategy is to compute $\mathcal{F}_h(A, B)$ for each such candidate homotopy class h, and to return the smallest value obtained.

4.3. Structure of Minimal Homotopy Classes

We define a ***direct geodesic*** to be a geodesic in $\mathcal{E}$ that is either (1) a line segment from A to B, or (2) a geodesic that consists of a line segment from A to some obstacle vertex p, a globally shortest path from p to an obstacle vertex q, and a line segment from q to B. We will prove:

Proposition 4.3. *Every minimal relative homotopy class contains a direct geodesic.*

Let h be an arbitrary minimal relative homotopy class. Let $\widetilde{A}$ and $\widetilde{B}$ be lifts of A and B in the universal cover $\widetilde{\mathcal{E}}$, such that for all s and t, the shortest path $\widetilde{\sigma}_h(s, t)$ between $\widetilde{A}(s)$ and $\widetilde{B}(t)$ is a lift of $\sigma_h(s, t)$. Let $\widetilde{P}$ denote the set of all lifts of the *vertices* of obstacles in P; every point in $\widetilde{P}$ lies on the boundary of $\widetilde{\mathcal{E}}$. Let $\widetilde{\pi}_h$ denote the intersection of all shortest paths $\widetilde{\sigma}_h(s, t)$. Proposition 4.3 follows directly from the following pair of lemmas.

Lemma 4.4. *If $\widetilde{\pi}_h = \varnothing$, then h contains a direct geodesic of type (1): a line segment.*

Proof. If the shortest path $\widetilde{\sigma}_h(0, 0)$ is a line segment, then the geodesic $\sigma_h(0, 0)$ is also a line segment, and the proof is complete. Thus, we assume that $\widetilde{\sigma}_h(0, 0)$ passes through at least one vertex in $\widetilde{P}$.

Let $\widetilde{p}_1, \ldots, \widetilde{p}_\kappa$ be the sequence of lifted obstacle vertices on the shortest path $\widetilde{\sigma}_h(0, 0)$. (The vertices $\widetilde{p}_i$ are distinct, although their projections back into the plane might not be.) Because $\widetilde{\pi}_h = \varnothing$, there is, for each i, a pair of parameters (s_i, t_i) such that $\widetilde{\sigma}_h(s_i, t_i)$ does not pass through $\widetilde{p}_i$.

We consider a continuous motion of the parameter point (s, t), starting at $(s, t) = (0, 0)$ and then moving successively to each point (s_i, t_i). Specifically, we define two continuous functions $s\colon [0, \kappa] \to [0, 1]$ and $t\colon [0, \kappa] \to [0, 1]$ such that $s(0) = t(0) = 0$, and for any integer i, we have $s(i) = s_i$ and $t(i) = t_i$. To simplify notation, let $\widetilde{\sigma}(\tau)$ denote the shortest path $\widetilde{\sigma}_h(s(\tau), t(\tau))$.

As the parameter τ ('time') increases, vertices in $\widetilde{P}$ are inserted into and deleted from the deque of obstacle vertices on $\widetilde{\sigma}(\tau)$. If the deque is empty at any time τ, then the shortest path $\widetilde{\sigma}(\tau)$ is a line segment, which implies that the projected path $\sigma(\tau)$ is a line segment in $\mathcal{E}$, concluding the proof. Thus, we assume to the contrary that the deque is never empty. Each vertex $\widetilde{p}_1, \ldots, \widetilde{p}_\kappa$ must be deleted from the deque at least once during the motion (but may be reinserted later).

Suppose $\widetilde{p}$ is the *last* vertex among $\widetilde{p}_1, \ldots, \widetilde{p}_\kappa$ to be removed from the deque for the *first* time. Without loss of generality, we assume $\widetilde{p}$ is first removed from the front of the deque at time τ_1. Let $\widetilde{q}$ denote the second vertex in the deque just before $\widetilde{p}$ is removed; this vertex must exist, because the deque is never empty. The vertex $\widetilde{p}$ lies on the first segment $\widetilde{a}\widetilde{q}$ of $\widetilde{\sigma}(\tau_1)$, where $\widetilde{a} = \widetilde{A}(s(\tau_1))$.

By definition of $\widetilde{p}$, vertex $\widetilde{q}$ must have been pushed onto the *back* of in the deque at some earlier time $\tau_2 < \tau_1$. Just before $\widetilde{q}$ is inserted, the last vertex in the deque must be $\widetilde{p}$. Moreover, $\widetilde{q}$ lies on the last segment $\widetilde{p}\widetilde{b}$ of $\widetilde{\sigma}(\tau_2)$, where $\widetilde{b} = \widetilde{B}(t(\tau_2))$. Thus, there is a line segment $\widetilde{a}\widetilde{b}$ between a point in $\widetilde{A}$ and a point in $\widetilde{B}$. The projection ab of this segment into $\mathcal{E}$ is a line segment in homotopy class h. □

Lemma 4.5. *If $\widetilde{\pi}_h \neq \varnothing$, then h contains a direct geodesic of type (2): the concatenation of a line segment from A to some obstacle vertex p, a globally shortest path from p to an obstacle vertex q, and a line segment from q to B.*

Proof. The path $\widetilde{\pi}_h$ is a shortest path between some pair of lifted obstacle vertices $\widetilde{p}$ and $\widetilde{q}$. (In the special case where $\widetilde{\pi}_h$ is a single point, we have $\widetilde{p} = \widetilde{q} = \widetilde{\pi}_h$.) Now $\widetilde{p}$ and $\widetilde{q}$ are lifts of obstacle vertices p and q (which may be the same point, even if $\widetilde{p}$ and $\widetilde{q}$ are not), and $\widetilde{\pi}_h$ is similarly a lift of some path π_h with endpoints p and q.

Let $\sigma(p,q)$ denote a globally shortest path from p to q, and suppose that it is strictly shorter than π_h. For any parameters s and t, let $\tau(s,t)$ denote the curve obtained from $\sigma_h(s,t)$ by replacing the subpath π_h with $\sigma(p,q)$. All paths $\tau(s,t)$ belong to the same relative homotopy class, which we denote h'. We now easily confirm that $h' \prec h$, contradicting our assumption that h is minimal. We conclude that π_h is the shortest path from p to q.

It remains to show that there is a line segment from some point on A to p. (A similar argument implies that there is a line segment from q to some point on B.) For all s and t, the geodesic $\sigma_h(s,t)$ is the concatenation of a geodesic $\alpha(s)$ from A to p, the shortest path π_h, and a geodesic $\beta(t)$ from q to B. If $\alpha(0)$ is a line segment, our claim is proved. Thus, we assume that $\alpha(0)$ is not a line segment, which implies that the lifted path $\widetilde{\alpha}(0)$ passes through at least one lifted obstacle vertex other than its endpoint $\widetilde{p}$. Let $\widetilde{p}^-$ be the last lifted obstacle vertex on $\widetilde{\alpha}(0)$ before $\widetilde{p}$. Let s_0 be the largest value such that $\widetilde{\alpha}(s)$ contains $\widetilde{p}^-$ for all $0 \leqslant s \leqslant s_0$. Because $\widetilde{p}^-$ is not on the common subpath $\widetilde{\pi}_h$, it is not on every geodesic $\widetilde{\alpha}(s)$, which implies that $s_0 < 1$. The geodesic $\alpha(s_0)$ is a line segment. □

Corollary 4.6. *We can enumerate a set of $O(mnk^4)$ relative homotopy classes that contains at least one optimal relative homotopy class, in $O(mnk^4)$ time.*

Proof. For any points $a \in A$ and $b \in B$, we call the line segment ab ***extremal*** if it satisfies one of the following conditions:

(i) The endpoints are vertices of A and B.

(ii) One endpoint is a vertex of A or B and the segment contains one vertex of P.

(iii) The segment contains two vertices of P.

Every line segment in $\mathcal{E}$ is relatively homotopic to at least one extremal line segment in $\mathcal{E}$. Thus, to enumerate the relative homotopy classes that contain a line segment, it suffices to enumerate the extremal line segments in $\mathcal{E}$.

There are $O(mn)$ extremal segments of type (i), which we can easily enumerate in $O(mn)$ time by brute force. Each vertex $a \in A$ and vertex $p \in P$ determine at most n extremal segments of type (ii), one for each intersection between the ray from a through p and B. Similarly, each vertex $b \in B$ and vertex $p \in P$ determine at most m extremal segments of type (ii). Thus, there are $O(mnk)$ extremal segments of type (ii); again, we can easily enumerate these in $O(mnk)$ time. Finally, any two vertices $p, q \in P$ determine $O(mn)$ extremal segments of type (iii), distinguished by the intersection points of the line through p and q with A and B, so there are $O(mnk^2)$ type-(iii) extremal segments in total. For any obstacle vertices p and q, we can compute the intersection points between the line through p and q and A or B in $O(m+n)$ time, and then enumerate the extremal segments that contain both p and q in $O(mn)$ time, again by brute force.

Altogether, we enumerate $O(mnk^2)$ extremal line segments in $O(mnk^2)$ time. To build all extremal line segments in $\mathcal{E}$, we discard any line segment that intersects any obstacle polygon; this takes $O(mnk^3)$ time in total.

To enumerate all other direct geodesics (of type (2)), we begin by computing shortest paths between every pair of obstacle vertices [22]. (If there is more than one shortest path between any pair of obstacles vertices, we can break ties arbitrarily.) Next, for every obstacle vertex p, we want to find all (relative homotopy classes of) line segments starting at p and ending at a point on A or on B. We compute them as follows: for every obstacle vertex $q \neq p$, we shoot a ray from p in the direction of q until it reaches the interior of an obstacle (or infinity), and then compute all $O(m+n)$ intersections between the resulting line segment (or ray) and the curves A and B. This gives us endpoints of line segments starting at p. To this list of line segments, we also add every segment in $\mathcal{E}$ from p to a vertex of A or B. We now have the complete list of potential initial and final segments of direct geodesics. Finally, we concatenate all $O(mk)$ initial segments, $O(k^2)$ shortest paths, and $O(nk)$ final segments to obtain $O(mnk^4)$ paths in $O(mnk^4)$ time. □

Proposition 4.3 is not a complete characterization of minimal relative homotopy classes. It is easy to find direct geodesics of type (2) whose relative homotopy classes are not minimal. However, the next lemma shows that every type (1) direct geodesic determines a minimal homotopy class.

Lemma 4.7. *The relative homotopy class of any line segment is minimal.*

Proof. Let σ be a line segment from $A(s)$ to $B(t)$, and let h be the relative homotopy class of σ. For any relative homotopy class $h' \neq h$, the shortest path $\sigma_{h'}(s,t)$ must be longer than $\sigma = \sigma_h(s,t)$, which implies that $h' \not\preceq h$. We conclude that h is minimal. □

There are input instances that admit $\Omega(mnk^2)$ distinct minimal relative homotopy classes. For example, Figure 3 shows such an example with $k/3$ triangular obstacles. If the triangles are sufficiently small, the line through any two obstacle vertices intersects a constant fraction of the edges of both A and B, defining $\Omega(mnk^2)$ type-(iii) extremal line segments. Lemma 4.7 implies that the homotopy classes of these extremal line segments are minimal. A constant fraction of these minimal homotopy classes contain at most four extremal segments.

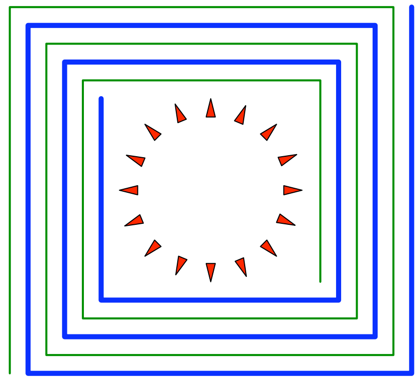

Figure 3: Curves and obstacles that admit $\Omega(mnk^2)$ minimal relative homotopy classes and $\Omega(mnk^4)$ relative homotopy classes of direct geodesics.

Moreover, this example admits $\Omega(mnk^4)$ relative homotopy classes of type-(2) direct geodesics. Consider the direct geodesics whose first and last obstacle vertices lie on the convex hull of the obstacles. There are $\Omega(k^2)$ choices for the first and last obstacle vertices; for each such choice, there are $\Omega(mk)$ choices for the initial line segment and $\Omega(nk)$ choices for the final line segment. Thus, any improvement in this portion of the algorithm will require a finer characterization of minimal relative homotopy classes.

4.4. Point Obstacles

Our previous structural results also apply to degenerate obstacles, such as points or line segments, with little modification, by replacing them with sufficiently small or thin triangles. Corollary 4.6 then implies a bound of $O(mnk^4)$ on the number of minimal homotopy classes for such inputs. The goal of this section is to provide a complete characterization of the minimal homotopy classes when all the obstacles are points, which yields a better bound of $O(mnk^2)$ on their number.

Thus, in this section, let P be a set of obstacle points in general position. Because the obstacles are now *closed* sets, there are pairs of points in $\mathcal{E} = \mathbb{E}^2 \setminus P$ that have no shortest path between them; more generally, there are homotopy classes of paths in $\mathcal{E}$ that contain no geodesics. In this setting, the distance between two points a and b (within any homotopy class) is properly defined as the *infimum* of the lengths of all paths (in that homotopy class) from a to b. For simplicity of exposition (and computation), we extend the definition of 'geodesic' to include any path in $\mathbb{E}^2$ that arises as the limit of a converging sequence of paths in $\mathcal{E}$ in the same homotopy class (with fixed endpoints), whose lengths converge to the distance between the endpoints within that homotopy class. Geometrically, geodesics in $\mathcal{E}$ are now polygonal paths in $\mathbb{E}^2$ whose internal vertices are obstacle points. (This extension is implicit in the works of Efrat *et al.* [16] and Bespamyatnikh [4], who describe algorithms to compute 'shortest' paths homotopic to a given path, in the plane minus a set of points.)

However, in order to uniquely identify the relative homotopy class of a geodesic, some additional information is now required in addition to its geometry. Specifically, we associate a ***turning angle*** with each obstacle point that the geodesic touches. Consider a geodesic γ that passes through an obstacle point p. Let C_ε be a circle centered at p with radius $\varepsilon > 0$, small enough to exclude every other obstacle in P. A turning angle of θ at an obstacle point p

indicates that replacing the portion of γ inside C_ε with an arc of length $\varepsilon|\theta|$ around C_ε, which goes counterclockwise around C_ε if $\theta > 0$ and clockwise if $\theta < 0$, yields a new path homotopic to γ. See Figure 4. A path could meet the same obstacle point more than once; we associate a different turning angle with each incidence. If γ is a geodesic, none of its turning angles is in the range $(-\pi, \pi)$, since otherwise γ could be locally shortened.

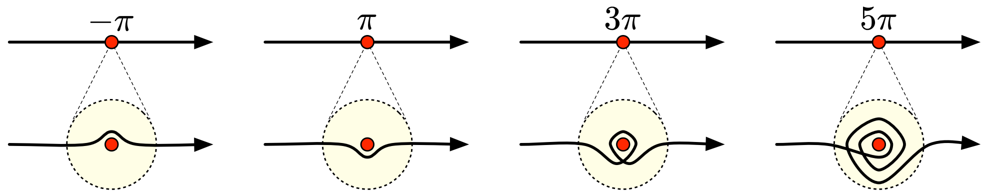

Figure 4: Four turning angles determine four different homotopy classes.

Proposition 4.8. *In the case of point obstacles, a relative homotopy class is minimal if and only if it contains a line segment.*

Proof. One direction of the proof is straightforward: Let h be the relative homotopy class of the line segment σ from $A(s)$ to $B(t)$. By our non-degeneracy assumption, up to slightly moving σ, we may assume that it touches no obstacle point and apply Lemma 4.7.

To prove the opposite implication, we consider a minimal relative homotopy class h. As in the proof of Proposition 4.3, we define $\widetilde{\pi}_h$ to be the intersection of the shortest paths between $\widetilde{A}(s)$ and $\widetilde{B}(t)$, for all $s, t \in [0, 1]$. If $\widetilde{\pi}_h = \varnothing$, then the proof of Lemma 4.4 already implies that h contains a line segment, so we assume that $\widetilde{\pi}_h \neq \varnothing$. In particular, there is a lifted obstacle point $\widetilde{p}$ such that for any $s, t \in [0, 1]$, the shortest path $\widetilde{\sigma}_h(s, t)$ passes through $\widetilde{p}$. Let $\theta(s, t)$ denote the turning angle of $\widetilde{\sigma}_h(s, t)$ at $\widetilde{p}$.

For all s and t, the path $\widetilde{\sigma}_h(s, t)$ is a shortest path, so $\theta(s, t)$ must lie outside the open interval $(-\pi, \pi)$. This turning angle is a continuous function of s and t, so we can assume without loss of generality that it is always at least π. In other words, we assume that every path $\widetilde{\sigma}_h(s, t)$ winds counterclockwise around $\widetilde{p}$. Recall that no three vertices of the input are collinear by our non-degeneracy assumption, so the minimum of $\theta(s, t)$ is not a multiple of π, and can therefore be written as $2\pi x + y$ for some integer x and some angle $y \in (-\pi, \pi)$.

Now $\widetilde{p}$ is a lift of some obstacle point p, and $\widetilde{\sigma}_h(s, t)$ similarly projects to a geodesic $\sigma_h(s, t)$. For each s and t, let $\tau(s, t)$ denote the path meeting the same obstacles in $\mathcal{E}$ in the same order and with the same turning angles as $\sigma_h(s, t)$, except that the turning angle at p is reduced by $2\pi x$. All paths $\tau(s, t)$ belong to a single relative homotopy class, which we denote by h'.

For every s and t, the paths $\tau(s, t)$ and $\sigma_h(s, t)$ have precisely the same length. This implies that $\sigma_{h'}(s, t)$ is never longer than $\sigma_h(s, t)$; thus, $h' \preceq h$.

Now let s and t be parameters such that $\theta(s, t)$ is minimized, and write $\theta(s, t) = 2\pi x + y$ for some integer x and some $y \in (-\pi, \pi)$. By construction, the turning angle of $\tau(s, t)$ at p equals y. Thus $\tau(s, t)$ is not a geodesic, so $\sigma_{h'}(s, t)$ is strictly shorter than $\tau(s, t)$, which has the same length as $\sigma_h(s, t)$. With the previous paragraph, this proves $h' \prec h$, contradicting our initial assumption that h is minimal. □

Corollary 4.9. *In the case of point obstacles, we can enumerate a superset of the minimal relative homotopy classes of size $O(mnk^2)$ in $O(mnk^2)$ time.*

Proof. In the proof of Corollary 4.6, we saw how to enumerate the $O(mnk^2)$ relative homotopy classes of line segments in $O(mnk^2)$ time; every minimal homotopy class contains a line segment by Proposition 4.8. □

We emphasize that Proposition 4.8 does not imply that the optimal *leash map* contains a line segment. At first glance, it may seem natural to conjecture that the optimal leash map must also contain a line segment; surprisingly, this conjecture is actually false.

Lemma 4.10. *There is a pair of polygonal curves and a set of point obstacles such that no optimal leash map contains a line segment.*

Proof. Consider the instance shown in Figure 5(a). The vertices of A have coordinates $(-2, 2)$, $(-2, 4)$, $(2, 4)$, and $(2, 2)$, in that order; the vertices of B have coordinates $(-2, -2)$, $(-2, -4)$, $(2, -4)$, and $(2, -2)$, in that order; and the obstacle points have coordinates $(1, 2)$, $(-1, 2)$, $(-1, -2)$, and $(1, -2)$. (This instance is highly degenerate, but it can easily be perturbed into general position without affecting the result.)

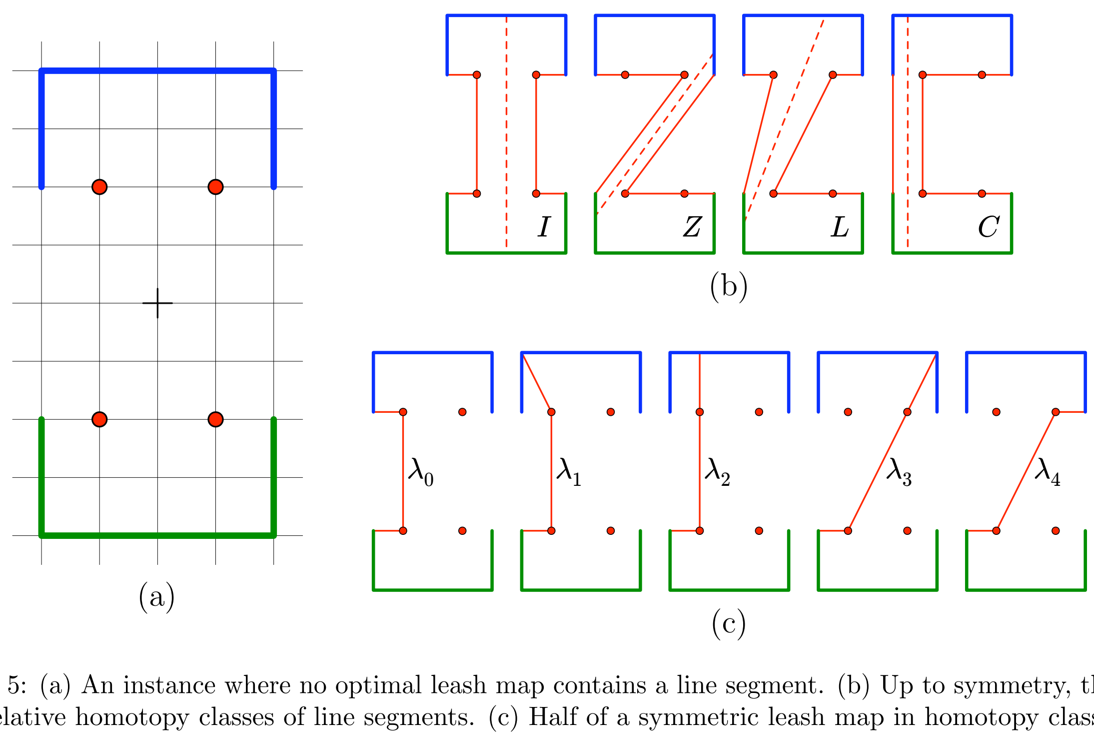

Figure 5: (a) An instance where no optimal leash map contains a line segment. (b) Up to symmetry, the only four relative homotopy classes of line segments. (c) Half of a symmetric leash map in homotopy class I.

Up to rotations and reflections, there are only four relative homotopy classes of segments with one endpoint on each curve. Figure 5(b) shows one line segment (dashed) and the initial and final leashes (solid) in each relative homotopy class. As suggested by the figure, we call these four classes I, Z, L, and C. We claim that class I is the only optimal relative homotopy class, and that the optimal leash map does not contain a line segment.

Figure 5(c) shows the first half of a leash map in class I, in which one endpoint of the leash traverses A completely before the other endpoint moves at all. The figure shows five critical leashes λ_0, λ_1, λ_2, λ_3, λ_4; between any two critical leashes, the length of the leash is either monotonically increasing or monotonically decreasing. The longest critical leash is λ_3, which has length $1 + 3\sqrt{5} \approx 7.708$; this is also the length of the leash map. The final leashes in classes Z, L, and C have lengths 8, $4 + 2\sqrt{5} \approx 8.472$, and 10, respectively. Within each relative homotopy class, the length of the final leash is a lower bound on the length of any leash map. Thus, I is the unique optimal homotopy class. On the other hand, the shortest line segment in class I has length 8, which is longer than λ_3. This completes the proof. □

4.5. Non-polygonal Obstacles

Our proof of Proposition 4.3 can be extended to non-polygonal obstacles with only minor modifications; the obstacles need not be convex or have smooth boundaries. In this more general setting, the initial and final segments of a direct geodesic must be tangent to the obstacles at their endpoints; that is, these segments can be made slightly longer without intersecting any obstacle. The algorithmic results in Sections 5 and 6 similarly extend to non-polygonal objects, provided one can efficiently compute the *visibility graph* of the obstacles [27, 28]; the running time of the resulting algorithm obviously depends on the exact representation of the objects. Further details of this extension are described by Chambers [7].

5. Computing Homotopic Fréchet Distance

Finally, we describe our algorithm to compute the homotopic Fréchet distance between two polygonal curves in the plane with polygonal obstacles. Our approach is to compute a set of relative homotopy classes that includes at least one optimal class, as described by Corollary 4.6, and then compute the Fréchet distance $\mathcal{F}_h(A, B)$ within each homotopy class h in this set. Our algorithm to compute $\mathcal{F}_h(A, B)$ is a direct adaptation of Alt and Godau's algorithm for computing the classical Fréchet distance between polygonal paths in the plane [2].

Henceforth, to simplify notation, we consider that the polygonal chain A, whose ordered sequence of vertices is $a_0, a_1, \ldots, a_m$, is parameterized over the interval $[0, m]$, instead of $[0, 1]$ as in previous sections, so that $A(i) = a_i$ for each integer i between 0 and m. Similarly, we parameterize B over the interval $[0, n]$. As in the previous section, for any $s \in [0, m]$ and $t \in [0, n]$, let $\sigma_h(s, t)$ denote the shortest path from $A(s)$ to $B(t)$ in homotopy class h, and let $\text{dist}_h(s, t) = \text{len}(\sigma_h(s, t))$. For any $\varepsilon > 0$, let $F_\varepsilon \subseteq [0, m] \times [0, n]$ denote the ***free space*** $\{(s, t) \mid \text{dist}_h(s, t) \leqslant \varepsilon\}$. Our goal is to compute the smallest value of ε such that F_ε contains a monotone path from $(0, 0)$ to (m, n); this is precisely the Fréchet distance $\mathcal{F}_h(A, B)$.

The parameter space $[0, m] \times [0, n]$ decomposes naturally into an $m \times n$ grid; let $\square_{i,j} = [i - 1, i] \times [j - 1, j]$ denote the grid cell representing paths from the ith edge of A to the jth edge of B.

5.1. Geodesic Distance Is Convex

In this section, we prove the following proposition, required by our generalization of Alt and Godau's algorithm:

Proposition 5.1. *The restriction of the function $dist_h$ to any grid cell $\square_{i,j}$ is convex.*

We first recall the following elementary classical properties of the Euclidean norm (denoted $\|\cdot\|$).

Lemma 5.2. *Let o be a fixed point in the plane, and let $\varphi_o \colon \mathbb{R}^2 \to \mathbb{R}$ be the function $p \mapsto \|\overrightarrow{op}\|$. The gradient of φ_o at any point $p \neq o$ is $\overrightarrow{op}/\|\overrightarrow{op}\|$. The function φ_o is convex everywhere, and of class C^1 everywhere except at o.*

Let $\alpha, \beta \colon [0,1] \to \mathcal{E}$ be affine functions with (constant) derivatives $\vec{a}$ and $\vec{b}$ respectively, and let h be a relative homotopy class. For each $t \in [0,1]$, let $\sigma(t)$ be the shortest path from $\alpha(t)$ to $\beta(t)$ in relative homotopy class h, and let $d(t)$ denote the length of $\sigma(t)$.

Fix $t \in [0,1]$. The shortest path $\sigma(t)$ is a polygonal curve. Let $\vec{u}(t)$ be the unit vector representing the direction of the first line segment of $\sigma(t)$ (at its initial point $\alpha(t)$). Similarly, we denote by $\vec{v}(t)$ the unit vector representing the direction of the last line segment of $\sigma(t)$.

Recall that, as t increases, the shortest path $\sigma(t)$ encounters a finite number of *events*. Between every two consecutive events, the sequence of obstacle vertices at which $\sigma(t)$ bends is the same.

Lemma 5.3. *Between any two consecutive events, d is convex and of class C^1. In particular, $d'(t) = \vec{b} \cdot \vec{v}(t) - \vec{a} \cdot \vec{u}(t)$, where $\cdot$ denotes the inner product.*

Proof. Fix two consecutive events t_0 and t_1.

Assume first that for all t between t_0 and t_1, the path $\sigma(t)$ is *not* a line segment. Then for every $t \in [t_0, t_1]$, $\sigma(t)$ is the concatenation of a line segment from $\alpha(t)$ to a fixed obstacle vertex p, a geodesic from p to another fixed obstacle vertex q, and the line segment from q to $\beta(t)$. It follows that $d(t)$ equals a constant plus $\|\overrightarrow{p\alpha(t)}\| + \|\overrightarrow{q\beta(t)}\|$. Our result is now a consequence of Lemma 5.2. Specifically, d is the sum of two convex functions, and is therefore convex. Since α and β do not meet obstacle vertices, the function d is C^1 in the interval $[t_0, t_1]$. The chain rule implies the claimed expression for d'. Specifically,

$$\frac{d}{dt}\|\overrightarrow{p\alpha(t)}\| = \frac{d}{dt}\varphi_p(\alpha(t)) = \vec{\nabla}\varphi_p(\alpha(t)) \cdot \frac{d}{dt}\alpha(t) = \frac{\overrightarrow{p\alpha(t)}}{\|\overrightarrow{p\alpha(t)}\|} \cdot \vec{a} = -\vec{u}(t) \cdot \vec{a}.$$

A similar derivation implies that $\frac{d}{dt}\|\overrightarrow{q\beta(t)}\| = \vec{v}(t) \cdot \vec{b}$.

If $\sigma(t)$ is a line segment whenever $t_0 \leqslant t \leqslant t_1$, then $d(t) = \|\overrightarrow{\alpha(t)\beta(t)}\|$. Since the function $t \mapsto \overrightarrow{\alpha(t)\beta(t)}$ is affine, Lemma 5.2 also implies that d is convex and of class C^1, and that

$$d'(t) = (\vec{b} - \vec{a}) \cdot \frac{\overrightarrow{\alpha(t)\beta(t)}}{\|\overrightarrow{\alpha(t)\beta(t)}\|}.$$

Finally, we observe that

$$\vec{u}(t) = \vec{v}(t) = \frac{\overrightarrow{\alpha(t)\beta(t)}}{\|\overrightarrow{\alpha(t)\beta(t)}\|},$$

which completes the proof. ∎

Lemma 5.4. *The function d is convex.*

Proof. Lemma 5.3 implies that between consecutive events the function d' is continuous and non-decreasing; indeed, $\vec{a} \cdot \vec{u}(t) = \|\vec{a}\| \cos\theta(t)$ where $\theta(t)$ is the angle between $\vec{a}$ and $\vec{u}(t)$. The angle $\theta(t)$ is constant if point p and segment $\alpha([0,1])$ are collinear; otherwise, $|\theta(t)|$ is strictly increasing (from 0 to π if $\alpha([0,1])$ was an infinite line). Thus, $-\vec{a} \cdot \vec{u}(t)$ is non-decreasing; a similar argument implies that $\vec{b} \cdot \vec{v}(t)$ is non-increasing.

Let t_0 be an arbitrary event. Since the functions $t \mapsto \vec{u}(t)$ and $t \mapsto \vec{v}(t)$ are continuous at t_0, Lemma 5.3 implies that d' is also continuous at t_0. Thus, d' is non-decreasing over the entire interval $[0,1]$, which implies that d is convex. □

We now conclude:

Proof of Proposition 5.1. Let $\Box_{i,j}$ be an arbitrary grid cell. Choose parameters $s, s' \in [i-1, i]$ and $t, t' \in [j-1, j]$. We claim that the function $\psi \colon [0,1] \to \mathcal{E}$ defined by setting

$$\psi(\lambda) := \mathrm{dist}_h\left((1-\lambda)s + \lambda s',\ (1-\lambda)t + \lambda t'\right)$$

is convex. Let $\alpha, \beta : [0,1] \to \mathcal{E}$ be the unique affine reparameterizations of $A|_{[s,s']}$ and $B|_{[t,t']}$, respectively. Then ψ is exactly the function d that is proved convex in Lemma 5.4.

We conclude that the restriction of dist_h to any line segment in $\Box_{i,j}$, which is a map of the form ψ above, is convex. This completes the proof that the restriction of dist_h to any grid cell $\Box_{i,j}$ is convex. □

5.2. Preprocessing for Distance Queries

The only significant difference between our algorithm and Alt and Godau's is that we require additional preprocessing to compute several *critical distances* and an auxiliary data structure to answer certain *distance queries*. (If there are no obstacles, each critical distance can be computed, and each distance query can be answered, in constant time.)

There are three types of critical distances:

- ***endpoint distances*** $\mathrm{dist}_h(0,0)$ and $\mathrm{dist}_h(m,n)$,
- ***vertex-edge distances*** $\mathrm{dist}_h(i,[j-1,j]) = \min\{\mathrm{dist}_h(i,t) \mid t \in [j-1,j]\}$ for all integers $i \in [0,m]$ and $j \in [1,n]$, and
- ***edge-vertex distances*** $\mathrm{dist}_h([i-1,i],j) = \min\{\mathrm{dist}_h(s,j) \mid s \in [i-1,i]\}$ for all integers $i \in [1,m]$ and $j \in [0,n]$.

Given integers i and j and any real value ε, a ***horizontal distance query*** asks for all values of $t \in [j-1,j]$ such that $\mathrm{dist}_h(i,t) = \varepsilon$, and a ***vertical distance query*** asks for all values of $s \in [i-1,i]$ such that $\mathrm{dist}_h(s,j) = \varepsilon$. The convexity of dist_h within any grid cell implies that any distance query returns at most two values.

We first describe how to preprocess a single vertical edge in the parameter grid to answer distance queries; critical values are automatically computed during the preprocessing. Obviously a similar result applies to horizontal grid edges.

Lemma 5.5. *Suppose we are given a point p and a line segment $\ell = \overline{xy}$, parameterized over $[0,1]$, as well as the geodesic $\sigma_h(p,x)$ and its length $dist_h(p,x)$. In $O(k \log k)$ time, we can build a data structure of size $O(k)$ such that for any ε, all values $t \in [0,1]$ such that $dist_h(p, \ell(t)) = \varepsilon$ can be computed in $O(\log k)$ time. We also report the critical vertex-edge distance $dist_h(p,\ell)$, the path $\sigma_h(p,y)$, and its length $dist_h(p,y)$.*

Proof. We first compute a constrained Delaunay triangulation [9] of the obstacles P, the segment ℓ, and point p in time $O(k \log k)$. This triangulation includes ℓ and the edges of polygons in P as edges.

We apply the following observations used in the *funnel* algorithm for computing shortest homotopic paths [8, 23, 21]. The shortest homotopic paths $\sigma_h(p,x)$ and $\sigma_h(p,y)$ may share a common subpath and then split at some vertex v; this vertex is then the apex of two concave chains that form a funnel with base xy. Each concave chain has complexity at most k and intersects a given edge of the triangulation at most twice.

The geodesic from p to x may have complexity greater than $O(k)$, but (as observed above) the concave chain from v to x has at most $O(k)$ segments. Our goal is to find a vertex w on $\sigma_h(p,x)$ such that the path from w to x contains v. In other words, the chain from w to x along $\sigma_h(p,x)$ has complexity $O(k)$ and contains the concave funnel path.

To find w, walk along the geodesic from x to p. If we find a vertex where the chain is not concave, we must have passed v, so we mark the non-concave vertex as w. If we ever re-cross a segment of the triangulation a second time, we again must have passed the funnel apex v so we can mark the second crossing as w. (We walked along $O(k)$ edges of the chain to find w.) Let π_h be the portion of $\sigma_h(p,x)$ between p and w, and τ_1 be the portion of $\sigma_h(p,x)$ between w and x.

We know that π_h is contained in $\sigma_h(p,y)$, since w is before the apex of the funnel v. Let τ_2 be the portion of $\sigma_h(p,y)$ between w and y; this can be computed in $O(k)$ time using the funnel algorithm. Given τ_2, we can then find the apex of the funnel v in $O(k)$ time.

Imagine extending each line segment on the concave chains until it intersects ℓ, the line connecting x and y. Between the two concave chains, the combinatorial description of the distance function changes only at points where the extended lines meet ℓ. To answer distance queries, we record the $O(k)$ intersections of the extended lines with ℓ. For each of the resulting intervals, record the (fixed) length of the geodesic up to the first vertex in the extended line, as well as the equations of the two lines that bracket the interval. In constant time per interval, we can also compute and store the value $t^* \in [0,1]$ such that $\text{dist}_h(p, \ell(t^*))$ is minimized; this gives the desired value $\text{dist}_h(p,\ell)$.

The funnel data structure requires $O(k)$ space to store the $O(k)$ combinatorial changes to the leash as its endpoint sweeps $\ell = \overline{xy}$.

Now given this data structure, we answer distance queries as follows. If the distance queried is smaller than $\text{dist}_h(p,\ell)$, we return the empty set. If it is equal to $\text{dist}_h(p,\ell)$, we return $\ell(t^*)$. If it is larger than $\text{dist}_h(p,\ell)$, we do two binary searches, one on the intervals between x and $\ell(t^*)$ and the other on the intervals between $\ell(t^*)$ and y. □

Lemma 5.6. *Given any shortest path in the minimal relative homotopy class h, we can compute all critical distances, in $O(mnk \log k)$ time and using $O(mnk)$ space, as well as build a data structure of size $O(mnk)$ that can answer any horizontal or vertical distance query in $O(\log k)$ time.*

Proof. We preprocess each edge of the parameter grid as described in Lemma 5.5. We start from the vertex (i, j) that is our given input, either a straight line segment or a direct geodesic. We then walk on the edges of the grid, visiting each edge at least once and at most twice. During this walk, at each current vertex (i, j), we maintain the shortest homotopic path $\sigma_h(i, j)$ and its length $\text{dist}_h(i, j)$. Each time we walk along an edge, we apply Lemma 5.5 to preprocess it and to compute the shortest homotopic path corresponding to the target vertex of that edge. Each step takes $O(k \log k)$ time, and there are $O(mn)$ edges, whence the running time. As we walk along an edge of the parameter grid, we use a deque to push and pop the obstacle vertices along the leash in constant time per operation. Since at most k vertices are pushed onto the deque for each grid edge, the total size of the deque is $O(mnk)$. □

5.3. Decision Procedure

Like Alt and Godau, we first consider the following *decision problem*: Is $\mathcal{F}_h(A, B)$ at least some given value ε? Equivalently, is there a monotone path in the free space F_ε from $(0, 0)$ to (m, n)? Our algorithm to solve this decision problem is identical to Alt and Godau's, except for the $O(\log k)$-factor penalty for distance queries; we briefly sketch it here for completeness.

For any integers i and j, let $h_{i,j}$ denote the intersection of the free space F_ε with the horizontal edge $([i-1, i], j)$, and let $v_{i,j}$ denote the intersection of F_ε with the vertical edge $(i, [j-1, j])$. In the first phase of the decision procedure, we compute $h_{i,j}$ and $v_{i,j}$ for all i and j, using one distance query (and $O(\log k)$ time) for each edge of the parameter grid.

In the second phase of the decision procedure, we propagate in lexicographic order from $\square_{1,1}$ to $\square_{m,n}$ and determine which $h_{i,j}$ and $v_{i,j}$ are reachable via a monotone path from $\square_{1,1}$. Since the free space in each $\square_{i,j}$ is convex, we can propagate through each cell in constant time.

Our decision algorithm returns true if and only if there is a monotone path that reaches (m, n). The total running time of our decision procedure is $O(mn \log k)$.

5.4. Optimization

Finally, we describe how to use our decision procedure to compute the minimum value ε^* of ε such that the free space F_ε contains a monotone path from $(0, 0)$ to (m, n); this is the Fréchet distance $\mathcal{F}_h(A, B)$.

We start by computing critical distances and the distance-query data structure in time $O(mnk \log k)$, as described in Lemma 5.6. We then sort the $O(mn)$ critical distances. Using the decision procedure, we can compare the optimal distance ε^* with any critical distance ε in $O(mn \log k)$ time. By binary search, we can, repeating this step $O(\log mn)$ times, compute an interval $[\varepsilon^-, \varepsilon^+]$ that contains ε^* but no critical distances.

We then apply Megiddo's *parametric search* technique [25]; see also [10, 31]. Parametric search combines our decision procedure with a 'generic' parallel algorithm whose combinatorial behavior changes at the optimal value ε^*. Alt and Godau observe that one of two events occurs when $\varepsilon = \varepsilon^*$:

- For some integers i, i', j, the bottom endpoint of $v_{i,j}$ and the top endpoint of $v_{i',j}$ lie on the same horizontal line.
- For some integers i, j, j', the left endpoint of $h_{i,j}$ and the right endpoint of $h_{i,j'}$ lie on the same vertical line.

Thus, it suffices to use a 'generic' algorithm that sorts the $O(mn)$ endpoint *values* of all non-empty segments $h_{i,j}$ and $v_{i,j}$, where the value of an endpoint (s, j) of $h_{i,j}$ is s, and the value of an endpoint (i, t) of $v_{i,j}$ is t.

We use Cole's parallel sorting algorithm [11], which runs in $O(\log N)$ parallel steps on $O(N)$ processors, as our generic algorithm. Each parallel step of Cole's sorting algorithm needs to compare $O(mn)$ endpoints. The graph of an endpoint, considered as a function of ε, is monotone and made of $O(k)$ hyperbolic arcs; see the proof of Lemma 5.5). It follows that the sign of a comparison between two endpoints may change at $O(k)$ different values of ε, which can be computed in $O(k)$ time. Applying the parametric search paradigm requires the following operations for each parallel step of the sorting algorithm:

- Compute the $O(mnk)$ values of ε corresponding to the changes of sign of the $O(mn)$ comparisons. This can be done in $O(mnk)$ time and $O(mnk)$ space.
- Apply binary search to these values by median finding, calling the decision procedure to discard half of them at each step of the search. This takes $O(mnk + T_d \log(mnk))$ time, where $T_d = O(mn \log k)$ is the running time of our decision procedure. We obtain this way an interval for ε where each of the $O(mn)$ comparisons has a determined sign.
- Deduce in $O(mn \log k)$ time the sign of each of the $O(mn)$ comparisons within the previously computed interval.

Since the underlying sorting algorithm requires $O(\log mn)$ parallel steps, the resulting parametric search algorithm runs in time $O(mn \log(mn)(k + \log k \log(mnk)))$.

The distance query data structure requires $O(mnk)$ space. We require $O(mnk)$ additional space to simulate sequentially each parallel step of the sorting algorithm; we can re-use this space for subsequent parallel steps. Therefore, the total space complexity of our algorithm is $O(mnk)$.

Lemma 5.7. *Given a direct geodesic in a minimal relative homotopy class h, the Fréchet distance $\mathcal{F}_h(A, B)$ can be computed in $O(N^3 \log N)$ time and $O(N^3)$ space, where $N = m + n + k + 2$ is the total input size.*

Cook and Wenk propose a more efficient and practical randomized alternative to parametric search in their algorithm to compute geodesic Fréchet distance [12, 13]. A direct application of their technique reduces the time to search for the optimal distance ε^* to $O(N^2 \log^2 N)$ with high probability. Unfortunately, because our decision procedure relies on data structures built in the preprocessing stage, which requires $O(N^3 \log N)$ time and $O(N^3)$ space, Cook and Wenk's randomized technique does not improve the overall running time of our optimization algorithm.

5.5. Putting Everything Together

Finally, to compute the homotopic Fréchet distance $\overline{\mathcal{F}}(A, B)$ in the plane minus a set of point obstacles, we enumerate the $O(N^4)$ minimal homotopy classes and compute $\mathcal{F}_h(A, B)$ for each minimal homotopy class h. Similarly, for polygonal obstacles, we construct a set of $O(N^6)$ relative homotopy classes that includes at least one optimal class, and then for each class h in that set, we compute $\mathcal{F}_h(A, B)$.

Theorem 5.8. *The homotopic Fréchet distance between two polygonal curves in the plane minus a set of points can be computed in $O(N^7 \log N)$ time and $O(N^3)$ space, where $N = n + m + k + 2$ is the total input complexity.*

Theorem 5.9. *The homotopic Fréchet distance between two polygonal curves in the plane minus a set of polygonal obstacles can be computed in $O(N^9 \log N)$ time and $O(N^3)$ space, where $N = n + m + k + 2$ is the total input complexity.*

5.6. Reduction to Geodesic Fréchet Distance?

Finally, we sketch a promising approach to a randomized algorithm which is faster with high probability than that of Section 5.5, using the recent algorithm of Cook and Wenk [12, 13] for computing the geodesic Fréchet distance between two polygonal curves inside a simple polygon. In this setting, the geodesic and homotopic Fréchet distances are identical—because the interior of a simple polygon is simply connected, there is only one relative homotopy class of leashes. Cook and Wenk's algorithm runs in $O(p + M^2 \log(pM) \log M)$ time with high probability, and in $O(p + M^3 \log(pM))$ time in the worst case, where $M = m + n$ is the total complexity of the curves and p is the complexity of the enclosing polygon.

Recall that $\mathcal{F}_h(A, B)$ is the equivalent to the classical Fréchet distance between A and B, except that distances are measured by shortest paths in relative homotopy class h. Let $\widetilde{A}$ and $\widetilde{B}$ be lifts of A and B to the universal cover $\widetilde{\mathcal{E}}$, chosen so that the shortest path $\widetilde{\sigma}_h(s, t)$ between any two points $\widetilde{A}(s)$ and $\widetilde{B}(t)$ is a lift of a geodesic in homotopy class h. Then $\mathcal{F}_h(A, B)$ is also equal to the *geodesic* Fréchet distance between $\widetilde{A}$ and $\widetilde{B}$ in $\widetilde{\mathcal{E}}$. However, the universal cover $\widetilde{\mathcal{E}}$ is infinite, so we cannot pass it as input to Cook and Wenk's algorithm.

Recall that we are given a direct geodesic $\sigma_h(s, t)$ in class h. Let $\widetilde{\Delta}$ denote the infinite triangulation of $\widetilde{\mathcal{E}}$ defined by Hershberger and Snoeyink [21]. Finally, let $\widetilde{\Pi}$ denote the union of triangles in $\widetilde{\Delta}$ that intersect either the lifted curves $\widetilde{A}$ or $\widetilde{B}$ or the shortest path $\widetilde{\sigma}_h(s, t)$. Because the dual graph of $\widetilde{\Delta}$ is a tree, the shortest path from *any* point on $\widetilde{A}$ to *any* point on $\widetilde{B}$ must be contained in $\widetilde{\Pi}$. It is not hard to prove that the region $\widetilde{\Pi}$ has complexity $O(mk + nk)$, and it can be computed in $O(mk + nk)$ time.

At this point, it is tempting to invoke Cook and Wenk's algorithm with the curves $\widetilde{A}$ and $\widetilde{B}$ and the region $\widetilde{\Pi}$ as input. However, we face a major technical hurdle: $\widetilde{\Pi}$ *is not a simple polygon*. Although $\widetilde{\Pi}$ is simply connected, has a locally Euclidean metric, and is bounded by straight line segments, it cannot be isometrically embedded in the plane without self-intersection.

We conjecture that Cook and Wenk's algorithm can be generalized to this setting with no loss of performance; this would require also generalizing the shortest-path query data structures of Guibas and Hershberger [19, 20] on which Cook and Wenk's algorithm relies. Specifically, we believe these algorithms can be modified to accept arbitrary simply connected *boundary-triangulated 2-manifolds*, as defined by Hershberger and Snoeyink [21], instead of simple polygons. If our conjecture is correct, this approach would imply an algorithm to compute $\mathcal{F}_h(A, B)$ in $O(N^2 \log^2 N)$ time with high probability, and still in $O(N^3 \log N)$ in the worst case, using only $O(N^2)$ space.

6. Extensions

6.1. Closed Curves

Formally, a ***closed curve*** in a topological space S is a continuous function from the circle $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ to S. Let A and B be two closed curves in S. A ***free homotopy*** between A and B is a continuous function $h\colon \mathbb{S}^1 \times [0,1] \to S$, such that $h(\cdot, 0) = A$ and $h(\cdot, 1) = B$; unlike a homotopy between paths, a free homotopy between closed curves does not keep any point fixed. If such a function exists, the two closed curves are ***freely homotopic***. A closed curve A is ***contractible*** if it is freely homotopic to a constant function (that is, a single point). Equivalently, A is contractible if there is a continuous function from the unit disk into S whose restriction to the disk boundary is A. A ***reparameterization of*** $\mathbb{S}^1$ is a continuous monotone surjection $\alpha\colon \mathbb{S}^1 \to \mathbb{S}^1$ of index 1. A ***reparameterization of a closed curve*** is the composition of this closed curve with a reparameterization of $\mathbb{S}^1$.

Now homotopic Fréchet distance can be defined almost exactly as it is for paths. Specifically, a leash map between A and B is a free homotopy between some reparameterization of A and some reparameterization of B; the length of a leash map ℓ is the maximum length of any leash $\ell(t, \cdot)$; and the homotopic Fréchet distance is the infimal length of any leash map:

$$\overline{\mathcal{F}}(A, B) := \inf_{\text{leash map } \ell\colon \mathbb{S}^1 \times [0,1] \to S} \left(\max_{t \in \mathbb{S}^1} \ \mathrm{len}(\ell(t, \cdot)) \right).$$

If there is no leash map between A and B—that is, if A and B are not homotopic—then we define $\overline{\mathcal{F}}(A, B) = \infty$. Our algorithm will automatically detect this situation.

In this section, we show how to adapt our algorithm to compute homotopic Fréchet distance between paths to compute the homotopic Fréchet distance between closed curves. Our derivation is complicated by the fact that different geodesics with the same endpoints can be homotopic relative to the closed curves A and B.

For notational convenience, we define $\mathbb{S}^1$ as $\mathbb{R}/\mathbb{Z}$, or equivalently, as the unit interval $[0, 1]$ with its endpoints identified. A closed curve A ***shifted by s*** is the closed curve denoted by $A + s$ where $(A + s)(t) = A((s + t) \bmod 1)$. For any closed curve A, let $A_{/}\colon [0, 1] \to \mathcal{E}$ be the corresponding ordinary curve, defined by setting $A_{/}(t) = A(t)$ for all $t \in [0, 1]$.

Consider a leash map ℓ between A and B, that is a free homotopy between some reparameterizations $A \circ \alpha$ and $B \circ \beta$. For every parameter u in $\mathbb{S}^1$, ℓ can be 'cut' along the leash $\ell(u, \cdot)$ to form a leash map $\ell_{u/}$ between the paths $(A + \alpha(u))_{/}$ and $(B + \beta(u))_{/}$. Formally,

$$\forall (s, t) \in [0, 1] \times [0, 1], \qquad \ell_{u/}(s, t) := \ell((u + s) \bmod 1, t).$$

The cut leash map $\ell_{u/}$ and the original leash map ℓ have equal length. Also note that the initial and final leashes in the cut leash map coincide: $\ell_{u/}(0, .) = \ell_{u/}(1, .)$. Conversely, for any parameters s and t in $\mathbb{S}^1$, a leash map $\ell_{/}$ between the paths $(A + s)_{/}$ and $(B + t)_{/}$ can be 'glued' to form a leash map between the closed curves A and B, if and only if $\ell_{/}(0, .) = \ell_{/}(1, .)$.

The following lemma further characterizes leashes that can appear in leash maps between closed curves. Suppose λ is a leash with endpoints $A(s)$ and $B(t)$. Let $C(\lambda)$ denote the closed curve obtained by concatenating $(A + s)_{/}$, followed by λ, followed by the reversal of $(B + t)_{/}$, followed by the reversal of λ. We call a leash λ *valid* if $C(\lambda)$ is contractible; see Figure 6.

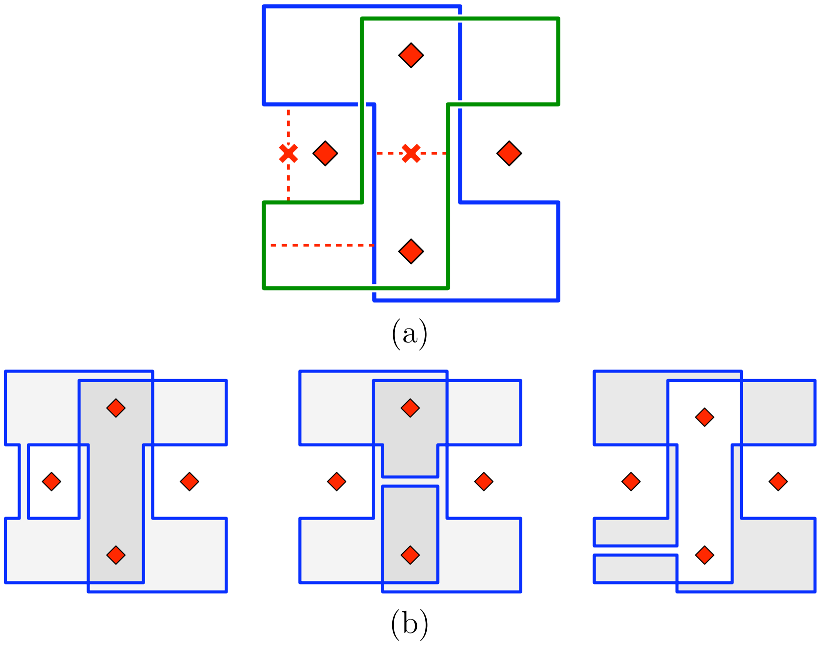

Figure 6: (a) Two invalid leashes and a valid leash between two homotopic closed curves. (b) The cycles $C(\lambda)$ corresponding to each leash λ; only the last cycle is contractible.

Lemma 6.1. *Let A and B be closed curves in some topological space S. In every leash map between A and B, every leash is valid. Conversely, any valid leash belongs to at least one leash map between A and B.*

Proof. Let ℓ be a leash map between A and B, and let $\ell(u, .)$ be a leash in ℓ. The cut leash map $\ell_{u/}$ is a continuous function defined on the topological disk $[0,1] \times [0,1]$. By definition, the restriction C of $\ell_{u/}$ to the boundary of this disk (which we can identify to $\mathbb{S}^1$) is contractible. Thus, since C is just a reparameterization of $C(\ell(u, .))$, $C(\ell(u, .))$ is freely homotopic to C and contractible.

Conversely, consider any valid leash λ, with endpoints $A(s)$ and $B(t)$. Because λ is valid, the cycle $C(\lambda)$ is contractible. Thus, there is a continuous map h from the unit disk into S whose restriction to the boundary of the disk is $C(\lambda)$. By identifying the unit disk to $[0,1] \times [0,1]$, it appears that h is actually a leash map between $(A+s)_{/}$ and $(B+t)_{/}$ with the property that $h(0, .) = h(1, .) = \lambda$. It follows that h can be glued to form a leash map between A and B. □

If two leashes are homotopic relative to $(A+s)_{/}$ and $(B+t)_{/}$, then either both leashes are valid or neither leash is valid; we say that a leash map or a relative homotopy class is *valid* if its elements are valid leashes. Let $\overline{\mathcal{F}}^v((A+s)_{/}, (B+t)_{/})$ denote the infimum length of a *valid* leash map between the ordinary curves $(A+s)_{/}$ and $(B+t)_{/}$. Cutting every leash map along its leash between some $A(s)$ and $B(0)$ and using the previous lemma, we can express the homotopic Fréchet distance between A and B as:

$$\overline{\mathcal{F}}(A, B) = \inf_{s \in \mathbb{S}^1} \overline{\mathcal{F}}^v((A+s)_{/}, B_{/})$$

Lemma 3.1 implies that $\overline{\mathcal{F}}^v((A+s)_/, B_/)$ is the length of a geodesic leash map in some valid homotopy class (relative to $(A+s)_/$ and $B_/$). The proofs of Lemmas 4.1 and 4.2 extend to *valid* relative homotopy classes without modification; thus,

$$\overline{\mathcal{F}}^v((A+s)_/,\, B_/) = \inf_{\text{valid minimal homotopy class } h} \mathcal{F}_h((A+s)_/,\, B_/).$$

Finally, Propositions 4.8 and 4.3 also extend to *valid* relative homotopy classes. The extension uses the fact that replacing a common non-empty subpath of all the leashes of a *valid* leash map yields another *valid* leash map. The two propositions imply that any valid minimal homotopy class of leashes between $(A+s)_/$ and $B_/$ contains a direct geodesic.

For any valid direct geodesic γ and any shift value s, let $[\gamma]_s$ denote the homotopy class of γ relative to $(A+s)_/$ and $B_/$. The preceding discussion implies that

$$\overline{\mathcal{F}}(A, B) = \min_{\text{valid direct geodesic } \gamma} \overline{\mathcal{F}}_\gamma(A, B),$$

where

$$\overline{\mathcal{F}}_\gamma(A, B) := \min_{s \in \mathbb{S}^1} \mathcal{F}_{[\gamma]_s}((A+s)_/,\, B_/).$$

Alt and Godau extend their algorithm to compute the classical Fréchet distance to closed curves [2]. Their algorithm solves the decision problem, whether the Fréchet distance is at most a given ε, by concatenating two copies of the free space diagram from Section 5 and therefore has $2m \times n$ cells. The decision problem can be rephrased as follows: Is there a shift s, such that the free space diagram contains a monotone path from $(s, 0)$ to $(m+s, n)$? Alt and Godau augment their representation of the free space diagram with additional pointers to answer this question efficiently.

We build an analogous free space diagram to determine, for any valid direct geodesic γ and threshold $\varepsilon > 0$, whether $\overline{\mathcal{F}}_\gamma(A, B) \leqslant \varepsilon$. Aside from the preprocessing described in Section 5.2, our approach is identical to that of Alt and Godau.

First, we modify our decision algorithm from Section 5.3 to determine whether there is a valid leash map in homotopy class h whose length is at most ε. The modified decision algorithm is more expensive by a factor of $O(\log mn)$ because it constructs a data structure analogous to that of Alt and Godau which helps to determine whether there exists a shift s, such that the $2m \times n$ free space diagram contains a monotone path from $(s, 0)$ to $(m+s, n)$. The running time of the new decision procedure is $O(mn \log k \log(mn)) = O(N^2 \log^2 N)$. The space complexity remains $O(mnk) = O(N^3)$ because the space required for storing the pointers in the data structures is $O(mnk)$.

Then, we modify our parametric search which determines the smallest ε for which there exists a valid leash map of length at most ε. As observed by Alt and Godau, we need to consider additional critical values of ε, but those also are the critical distances between a vertex of one curve and an edge of the other curve. Therefore, there are $O(mnk)$ such critical values of ε, as argued in Section 5.2. Hence, the running time of the optimization procedure is now $O(mn \log(mn)(k + \log k \log(mn) \log(mnk))) = O(N^3 \log N)$. We emphasize that the increased cost of the decision procedure does not lead to a similar increase in the overall running time of our optimization procedure.

Because each direct geodesic has complexity $O(k) = O(N)$, we can test whether a direct geodesic λ is valid in $O(N^{3/2} \log N)$ time using an algorithm of Cabello et al. [6] to decide whether $C(\lambda)$ is contractible. Thus, we can compute a set of valid minimal homotopy classes that contains an optimal one in $O(N^{11/2} \log N)$ time for point obstacles, or in $O(N^{15/2} \log N)$ for polygonal obstacles. Finally, to compute the homotopic Fréchet distance, we compute the optimum leash map in the relative homotopy class of each valid direct geodesic.

Theorem 6.2. *The homotopic Fréchet distance between two closed polygonal curves in the plane minus a set of obstacles can be computed in $O(N^7 \log N)$ time and $O(N^3)$ space if the obstacles are points, or in $O(N^9 \log N)$ time and $O(N^3)$ space if the obstacles are polygons.*

6.2. Variants of Fréchet Distance

Finally, we briefly consider two natural variants of homotopic Fréchet distance that can also be computed using our techniques.

The *weak Fréchet distance* is a variant of the classical Fréchet distance without the requirement that the endpoints move monotonically along their respective curves—the dog and its owner are allowed to backtrack to keep the leash between them short. Alt and Godau [2] gave a simpler algorithm for computing the weak Fréchet distance, using a graph shortest-path algorithm instead of parametric search. A similar simplification of our algorithm computes the *weak homotopic Fréchet distance* between curves in the plane minus polygonal obstacles in polynomial time. The proofs of Lemmas 4.1 and 4.2, as well as Propositions 4.8 and 4.3, extend to the weak variant of homotopic Fréchet distance without modification. We reiterate that Propositions 4.8 and 4.2 do not imply that the optimal (weak) *leash map* contains a direct geodesic, only the optimal relative homotopy class. Thus, we can define and compute the weak homotopic Fréchet distance as the minimum, over all relative homotopy classes h that contain a direct geodesic, of the weak Fréchet distance with respect to shortest path lengths in h.

The *discrete homotopic Fréchet distance*, also called the *coupling distance*, is an approximation of the Fréchet metric for polygonal curves defined by Eiter and Mannila [17]. The discrete Fréchet distance considers only positions of the leash where its endpoints are located at vertices of the two polygonal curves and never in the interior of an edge. This special structure allows the discrete Fréchet distance to be computed by an easy dynamic programming algorithm. As usual, Propositions 4.8 and 4.3 imply that we can define and compute the discrete homotopic Fréchet distance as the minimum, over all relative homotopy classes h that contain a direct geodesic, of the discrete Fréchet distance with respect to shortest path lengths in h.

7. Conclusion

In this paper, we introduced a natural generalization of the Fréchet distance between curves to more general metric spaces, called the homotopic Fréchet distance. We described a polynomial-time algorithm to compute the homotopic Fréchet distance between polygonal curves in the plane with point or polygon obstacles.

Improving the running time of our algorithms is the most immediate outstanding open problem. We described one promising approach in Section 5.6, which would require generalizing Cook and Wenk's algorithm for geodesic Fréchet distance to more general simply-connected spaces. We also conjecture that the running time can be improved by optimizing leash maps in every minimal homotopy class simultaneously. Since shortest paths between the same endpoints but belonging to different homotopy classes are related, we expect to (partially) reuse the results of shortest path computations in one homotopy class when we consider other homotopy classes. Finally, for polygonal obstacles, an exact characterization of minimal homotopy classes would almost certainly lead to a significantly faster algorithm.

It would be interesting to compute homotopic Fréchet distance in spaces more general than those considered in the current paper. In particular, we are interested in computing the homotopic Fréchet distance between two curves on a convex polyhedron, generalizing the algorithm of Maheshwari and Yi for classical Fréchet distance [24]. The vertices of the polyhedron are 'mountains' over which the leash can pass only if it is long enough. Shortest paths on the surface of a convex polyhedron do *not* vary continuously as the endpoints move, because of the positive curvature at the vertices, so we cannot consider only geodesic leash maps.

Finally, it would also be interesting to consider the homotopic Fréchet distance between higher-dimensional manifolds; such problems arise with respect to surfaces in configuration spaces of robot systems. Ordinary Fréchet distance is difficult to compute in higher dimensions, although the *weak* Fréchet distance between two triangulated surfaces can be computed in polynomial time [1].

Acknowledgments

This research was initiated during a visit to INRIA Lorraine in Nancy, made possible by a UIUC-CNRS-INRIA travel grant. Research by Erin Chambers and Jeff Erickson was also partially supported by NSF grant DMS-0528086; Erin Chambers was additionally supported by an NSF graduate research fellowship. Research by Shripad Thite was partially supported by the Netherlands Organisation for Scientific Research (NWO) under project number 639.023.301 and travel by INRIA Lorraine. We would like to thank the anonymous referees for their careful reading of the paper and their numerous suggestions for improvement. Finally, we thank Hazel Everett and Sylvain Petitjean for useful discussions, and Kira and Nori for great company and several walks in the woods.

References

[1] Helmut Alt and Maike Buchin. Semi-computability of the Fréchet distance between surfaces. In *Proc. 21st European Workshop on Computational Geometry*, pages 45–48, March 2005.

[2] Helmut Alt and Michael Godau. Computing the Fréchet distance between two polygonal curves. *Int. J. Comput. Geom. Appl./*, 5(1–2):75–91, 1995.

[3] Boris Aronov, Sariel Har-Peled, Christian Knauer, Yusu Wang, and Carola Wenk. Fréchet distance for curves, revisited. In *Proc. 14th European Symp. Algorithms*, volume 4168 of *Lect. Notes Comput. Sci.*, pages 52–63. Springer-Verlag, 2006.

[4] Sergei Bespamyatnikh. Computing homotopic shortest paths in the plane. In *Proc. 14th Annu. ACM-SIAM Sympos. Discrete Algorithms*, pages 609–617, 2003.

[5] Sergei Bespamyatnikh. Encoding homotopy of paths in the plane. In *Proc. LATIN 2004: Theoretical Infomatics*, volume 2976 of *Lect. Notes Comput. Sci.*, pages 329–338. Springer-Verlag, 2004.

[6] Sergio Cabello, Yuanxin Liu, Andrea Mantler, and Jack Snoeyink. Testing homotopy for paths in the plane. *Discrete Comput. Geom.*, 31(1):61–81, 2004.

[7] Erin Wolf Chambers. *Finding Interesting Topological Features*. Ph.D. thesis, Dept. Comput. Sci., Univ. Illinois Urbana-Champaign, July 2008.

[8] Bernard Chazelle. A theorem on polygon cutting with applications. In *Proc. 23rd Annu. IEEE Sympos. Found. Comput. Sci.*, pages 339–349, 1982.

[9] L. Paul Chew. Constrained Delaunay triangulations. *Algorithmica*, 4(1):97–108, 1989.

[10] Richard Cole. Slowing down sorting networks to obtain faster sorting algorithms. *JACM*, 34(1):200–208, January 1987.

[11] Richard Cole. Parallel merge sort. *SIAM Journal on Computing*, 17(4):770–785, 1988.

[12] Atlas F. Cook IV and Carola Wenk. Geodesic Fréchet and Hausdorff distance inside a simple polygon. Tech. Rep. CS-TR-2007-004, University of Texas at San Antonio, 2007.

[13] Atlas F. Cook IV and Carola Wenk. Geodesic Fréchet distance inside a simple polygon. In *Proc. 25th Symp. Theoretical Aspects of Computer Science (STACS)*, pages 193–204, 2008.

[14] Atlas F. Cook IV and Carola Wenk. Geodesic Fréchet distance with polygonal obstacles. Tech. Rep. CS-TR-2008-0010, University of Texas at San Antonio, 2008.

[15] Alon Efrat, Leonidas J. Guibas, Sariel Har-Peled, Joseph S. B. Mitchell, and T. M. Murali. New similarity measures between polylines with applications to morphing and polygon sweeping. *Discrete Comput. Geom.*, 28:535–569, 2002.

[16] Alon Efrat, Stephen G. Kobourov, and Anna Lubiw. Computing homotopic shortest paths efficiently. *Comput. Geom. Theory Appl.*, 35(3):162–172, 2006.

[17] Thomas Eiter and Heikki Mannila. Computing discrete Fréchet distance. Technical Report CD-TR 94/64, Christian Doppler Laboratory for Expert Systems, TU Vienna, Austria, 1994.

[18] Dima Grigoriev and Anatol Slissenko. Polytime algorithm for the shortest path in a homotopy class amidst semi-algebraic obstacles in the plane. In *Proc. Internat. Sympos. Symbolic and Algebraic Computation*, pages 17–24, 1998.

[19] Leonidas J. Guibas and John Hershberger. Optimal shortest path queries in a simple polygon. *J. Comput. Syst. Sci.*, 39(2):126–152, 1989.

[20] John Hershberger. A new data structure for shortest path queries in a simple polygon. *Inf. Proc. Letters*, 38(5):231–235, 1991.

[21] John Hershberger and Jack Snoeyink. Computing minimum length paths of a given homotopy class. *Computational Geometry: Theory and Applications*, 4:63–67, 1994.

[22] John Hershberger and Subhash Suri. An optimal algorithm for euclidean shortest paths in the plane. *SIAM Journal on Computing*, 28(6):2215–2256, 1999.

[23] Der-Tsai Lee and Franco P. Preparata. Euclidean shortest paths in the presence of rectilinear barriers. *Networks*, 14:393–410, 1984.

[24] Anil Maheshwari and Jiehua Yi. On computing Fréchet distance of two paths on a convex polyhedron. In *Proc. 21st European Workshop on Computational Geometry*, pages 41–44, 2005.

[25] Nimrod Megiddo. Applying parallel computation algorithms in the design of serial algorithms. *J. ACM*, 30:852–866, 1983.

[26] James Munkres. *Topology (2nd ed.)*. Prentice Hall, 2000.

[27] Michel Pocchiola and Gert Vegter. Computing the visibility graph via pseudo-triangulations. In *Symposium on Computational Geometry*, pages 248–257, 1995.

[28] Michel Pocchiola and Gert Vegter. Minimal tangent visibility graphs. *Computational Geometry: Theory and Applications*, 6, 1996.

[29] Raimund Seidel. The nature and meaning of perturbations in geometric computing. *Discrete Comput. Geom.*, 19:1–17, 1998.

[30] E. Sriraghavendra, Karthik K., and Chiranjib Bhattacharyya. Fréchet distance based approach for searching online handwritten documents. In *Proc. 9th Int'l. Conf. Document Analysis and Recognition (ICDAR)*, pages 461–465, 2007.

[31] René van Oostrum and Remco C. Veltkamp. Parametric search made practical. *Computational Geometry: Theory and Applications*, 28:75–88, 2004.