How to Walk Your Dog in the Mountains with No Magic Leash

Sariel Har-Peled* Amir Nayyeri† Mohammad Salavatipour‡

Anastasios Sidiropoulos§

December 3, 2011

Abstract

We describe a $O(\log n)$-approximation algorithm for computing the homotopic Frechét distance between two polygonal curves that lie on the boundary of a triangulated topological disk. Prior to this work, algorithms where known only for curves on the Euclidean plane with polygonal obstacles.

A key technical ingredient in our analysis is a $O(\log n)$-approximation algorithm for computing the minimum height of a homotopy between two curves. No algorithms were previously known for approximating this parameter. Surprisingly, it is not even known if computing either the homotopic Frechét distance, or the minimum height of a homotopy, is in NP.

1 Introduction

Comparing the shapes of curves – or sequenced data in general – is a challenging task that arises in many different contexts. The ***Frechét distance*** and its variants (e.g. dynamic time-warping [KP99]) have been used as a similarity measure in various applications such as matching of time series in databases [KKS05], comparing melodies in music information retrieval [SGHS08], matching coastlines over time [MDBH06], as well as in map-matching of vehicle tracking data [BPSW05, WSP06], and moving objects analysis [BBG08a, BBG+08b]. See [Fre06, AB05, AG95] for algorithms for computing the Frechét distance.

Informally, for a pair of such curves $f, g : [0,1] \to \mathcal{D}$, for some ambient space $(\mathcal{D}, \mathsf{d})$, their Frechét distance is the minimum length leash needed to traverse both curves in sync. To this end, imagine a person traversing f starting from $f(0)$, and a dog traversing g starting from $g(0)$, both traveling along these curves without ever moving backwards. Then, the Frechét distance is the infimum over all possible traversals, of the maximum distance between the person and the dog. More precisely, the Frechét distance between f and g is defined to be

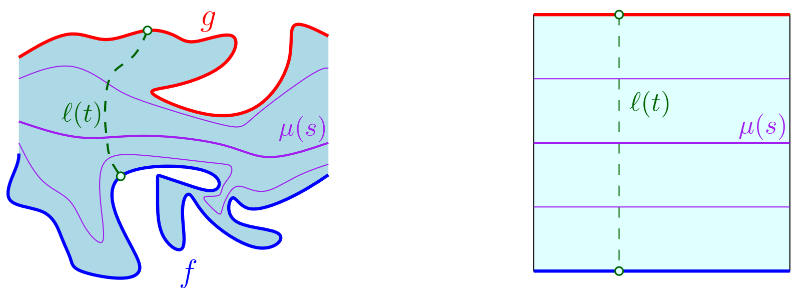
$$\delta(f,g) = \inf_{\phi:[0,1]\to[0,1]} \sup_{x\in[0,1]} \mathsf{d}\Big(f(x), g(\phi(x))\Big),$$

*Department of Computer Science, University of Illinois, Urbana-Champaign; `sariel@cs.uiuc.edu`. Work on this paper was partially supported by a NSF AF award CCF-0915984.

†Department of Computer Science, University of Illinois, Urbana-Champaign; `nayyeri2@illinois.edu`

‡Department of Computing Science, University of Alberta, and Toyota Technological Institute at Chicago; `mreza@cs.ualberta.ca`

§Toyota Technological Institute at Chicago; `tasos@ttic.edu`



(i) (ii)

Figure 1: (i) Two curves f and g, and (ii) the parametrization of their homotopic Frechét distance.

where ϕ ranges over all orientation-preserving homeomorphisms.

While this distance makes complete sense when the underlying distance is the Euclidean metric, it becomes less useful if the distance function is more interesting. For example, imagine walking a dog in the woods. The leash might get tangled as the dog and the person walk on two different sides of a tree. Since the Frechét distance cares only about the distance between the two moving points, the leash would "magically" jump over the tree.

Homotopic Frechét distance. To address this shortcoming, a natural extension of the above notion called ***homotopic Frechét distance*** was introduced by Chambers *et al.* [CdVE+08a]. Informally, revisiting the above person-dog analogy, we consider the infimum over all possible traversals of the curves, but this time, we require that the person is connected to the dog via a leash, i.e. a curve that moves continuously over time. Furthermore, one keeps track of the leash during the motion, where the purpose is to minimize the maximum leash length needed. To this end, consider a continuous mapping $\psi : [0,1]^2 \to \mathcal{D}$, where $\psi(\cdot, 0)$ and $\psi(\cdot, 1)$ are reparameterizations of f and g. Furthermore, let $\ell(t) = \psi(t, \cdot)$ be the leash at time t. Then, the homotopic Frechét distance between f and g is

$$\delta_H(f,g) = \inf_{\psi:[0,1]^2 \to \mathcal{D}} \max_{t \in [0,1]} |\ell(t)|,$$

where $|\cdot|$ denotes the length of a curve.

Clearly, $\delta_H(f,g) \geq \delta(f,g)$; in fact, $\delta_H(f,g)$ can be arbitrary larger than $\delta(f,g)$. We remark that $\delta_H(f,g) = \delta(f,g)$ for any pair of curves in the Euclidean plane, as we can always pick the leash to be a straight line segment between the person and the dog. In other words, the map ψ in the definition of δ_H can be obtained from the map ϕ in the definition of δ via an appropriate affine extension. However, this is not true for general ambient spaces, where the leash might have to pass over obstacles, hills, and such.

Another measure of interest is the ***homotopy height*** of such a mapping. Formally, given ψ as above, for $s \in [0,1]$, let $\mu(s) = \psi(\cdot, s)$ be the ***morph*** of f to g at time s. Here, $\mu(s)$ is a curve that starts as $\mu(0) = f$ and continuously deforms as s increases, and ends up as $\mu(1) = g$. Under this interpretation, the homotopic Frechét distance is refereed to as the ***morphing width*** of f and g, and it bounds how far a point on f has to travel to its corresponding point in g under the morphing of ψ [EGH+02]. The length of $\mu(s)$ is the ***height of the morph at time*** s, and the ***height*** of such a morphing is $\mathsf{height}(\mu) = \max_{s \in [0,1]} |\mu(s)|$. The homotopy height between f and g bounded by $\ell(0)$ and $\ell(1)$ is

$$\mathsf{h}(f, g, \ell(0), \ell(1)) = \inf_{\mu} \mathsf{height}(\mu)\,,$$

where μ varies over all possible morphings between f and g, such that each curve in μ has one end on $\ell(0)$ and one end in $\ell(1)$. See Figure 1 for an example. Note that if we do not constraint the endpoints of the curves during the morphing to stay on $\ell(0)$ and $\ell(1)$, the problem of computing the minimum height homotopy is trivial. One can contract f to a point, send it to g from the shortest (f, g)-path, and then expand it to g. In this paper, we first consider the problem for the special case that f and g have common endpoints. In Section 3 we show that our algorithm can be extended to the general case with the same approximation ratio. To keep the notation simple, we use $\mathsf{h}(f, g)$ when f and g have common endpoints.

Intuitively, the homotopy height measures how long the curve has to become as it deforms from f to g, and it was introduced by Chambers and Letscher [CL09, CL10]. Observe that if we are given the starting and ending leashes $\ell(0)$ and $\ell(1)$ then the homotopy height of f and g, is the homotopic Frechét distance between $\ell(0)$ and $\ell(1)$.

In this paper, we are interested in the problems of computing the homotopic Frechét distance and the homotopic height between two simple polygonal curves that lie on the boundary of an arbitrary triangulated topological disk.

Why are these measures interesting? For the sake of the discussion here, assume that we know the starting and ending leash of the homotopy between f and g. The region bounded by the two curves and these leashes, form a topological disk, and the mapping realizing the homotopic Frechét distance is a mapping of the unit square to this disk $\mathcal{D}$. This mapping specifies how to sweep over $\mathcal{D}$ in a geometrically "efficient" way (especially if the leash does not sweep over the same point more than once), so that the leash (i.e., the sweeping curve) is never too long [EGH+02]. As a concrete example, consider the two curves as enclosing several mountains between them on the surface – computing the homotopic Frechét distance corresponds to deciding which mountains to sweep first and in which order.

Furthermore, this mapping can be interpreted as surface parameterization [Flo97, SdS00] and can thus be used in applications such as texture mapping [BVIG91, PB00]. In the texture mapping problem, we wish to find a continuous and invertible mapping from the texture, usually a two-dimensional rectangular image, to the surface.

Another interesting interpretation is when f is a closed curve, and g is a point. Interpreting f as a rubber band in a 3d model, the homotopy height between f and g here is the minimum length the rubber band has to be so that it can be collapsed to a point (here, the rubber band stays on the surface as this is happening). In particular, a short closed curve with large homotopic height to any point in the surface is a "neck" in the 3d model.

To summarize, these measures seem to provide us with a fundamental understanding of the structure of the given surface/model.

Previous work. Chambers *et al.* [CdVE+08a] gave a polynomial time algorithm to compute the homotopic Frechét distance between two polygonal curves on the Euclidean plane with polygonal obstacles. Chambers and Letscher [CL09, CL10] introduced the notion of minimum homotopy height, and proved structural properties for the case of a pair of paths on the boundary of a topological disk. We remark that in general, it is not known whether the optimum homotopy has polynomially long description. In particular, it is not known whether the problem is in NP.

The problem of computing the (standard) Frechét distance between curves has been considered by Alt and Godau [AG95], who gave a polynomial time algorithm. Eiter and Mannila [EM94] studied the easier discrete version of this problem. Computing the Frechét distance between surfaces [Fre24], appears to be a much more difficult task, and to the time of this writing, its complexity is poorly understood. The problem has been shown to be NP-hard by Godau [God99], while the best algorithmic result is due to Alt and Buchin [AB05], who showed that it is upper semi-computable.

Efrat *et al.* [EGH$^+$02] considered the Frechét distance inside a simple polygon as a way to facilitate sweeping it efficiently. They also used the Frechét distance using the underlining geodesic metric as a way to get a morphing between two curves. For recent work on the Frechét distance, see [CW10, DHW10, HR11, CDHP$^+$11] and references therein.

Our results. In this paper, we consider the problems of computing the homotopic Frechét distance and the homotopy height between two simple polygonal curves that lie on the boundary of a triangulated topological disk $\mathcal{D}$ that is composed of n triangles.

We give a polynomial time $O(\log n)$-approximation algorithm for computing the homotopy height between f and g. Our algorithm to compute an approximate homotopy between f and g is via a simple, yet delicate divide and conquer algorithm.

We use the homotopy height algorithm as an ingredient for a $O(\log n)$-approximation algorithm for the homotopic Frechét distance problem. Intuitively, our algorithm for homotopic Frechét distance works as follows. We first approximately guess the optimum (i.e. δ_H). Using this guess, we classify parts of $\mathcal{D}$ as "obstacles", i.e. regions over which a short leash cannot pass. Let D' be the punctured disk obtained from $\mathcal{D}$ after removing these obstacles. The isotopy class of any leash is determined by the set of punctures that are on its left side. Observe that the leashes of the optimum solution belong to the same isotopy class. We describe a greedy algorithm to pick an isotopy class out of exponential number of choices, s.t. the homotopic Frechét distance constrained inside it is a constant factor of the homotopic Frechét distance in $\mathcal{D}$. Then, we use an extended version of our homotopy height algorithm to compute the homotopic Frechét distance.

The $O(\log n)$ factor shows up in the homotopic Frechét distance algorithm only because it uses the homotopy height as a subroutine. Thus, any constant factor approximation algorithm for the homotopy height problem implies a constant factor approximation algorithm for the homotopic Frechét distance.

As a warm-up exercise and in order to simplify the presentation we first consider the discrete version of the homotopy height problem in Section 2.1. This is how Chambers and Letscher formulated the problem. Later, in Section 2.2, we describe an algorithm to approximately find the shortest homotopy in continuous settings. In Section 3, we address the homotopic Frechét distance, both discrete and continuous. Later, in Section 4, we show that the homotopic Frechét distance between two curves on a boundary of any orientable surface of constant genus can be $O(\log n)$-approximated in the same asymptotic running time. Basic definitions are provided in Appendix A.

2 Approximating the height of a homotopy

In this section we give an approximation algorithm for finding a homotopy of minimum height in a topological disc, $\mathcal{D}$, whose boundary is defined by two walks L and R that share their end-points

s, t. We start with the discrete case, i.e. when the disk is a triangulated edge-weighted planar graph and then generalize it to the continuous case. We then use this algorithm as a subroutine (in the next section) in our algorithm for the minimum homotopic Frechét distance problem.

2.1 The discrete case

To start, let us assume we are given an embedded planar graph G all of whose faces (except possibly the outer face) are triangles. Let $s, t \in \partial G$ and L and R be the two non-crossing (s, t)-walks on ∂G in counter-clockwise and clockwise order, respectively. We use $\mathcal{D}$ to denote the topological disk enclosed by $L \cup R$. We abuse the notation and refer to vertices of G (inside of $\mathcal{D}$) as vertices of $\mathcal{D}$. Our goal is to find a minimum height homotopy from L to R of non-crossing walks. Informally, the homotopy is defined by a sequence of walks, where every two consecutive walks differ by either by a triangle or by an edge (being traversed twice). For a formal definition, see Appendix A.4.

Lemma 2.1 *Let x and y be vertices of G that are at distance ρ. Then any homotopy between L and R has height at least ρ.*

Proof: Fix a homotopy of height δ. This homotopy contains an s-t walk P that passes through x, and an s-t walk Q that passes through y. We have $\rho \leq |P[s, y]|+|Q[s, y]|$, and $\rho \leq |P[x, t]|+|Q[y, t]|$. Therefore, $\rho \leq (|P| + |Q|)/2 \leq \delta$, as required. ■

Suppose d_1 is the maximum distance of a vertex of G from either of L or R, d_2 is the largest edge weight, and let $d = \max\{d_1, d_2\}$.

Lemma 2.2 *Suppose $\mathcal{D}$, L, R, and d are defined as above. Then any homotopy between L and R has height at least d.*

Proof: For every homotopy between L and R, and for every edge e, there exists a walk in the homotopy that passes through e. Therefore, the height of the homotopy is at least d_2. Moreover, the height is at least d_1 by Lemma 2.1. ■

Here we present an algorithm which finds an (L, R)-homotopy of height at most $|L| + |R| + O(d \log n)$.

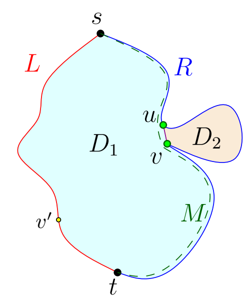
Lemma 2.3 *Let $\mathcal{D}$ be an edge-weighted triangulated topological disk with n faces such that its boundary is formed by two walks L and R that share endpoints s and t. Then, one can compute, in $O(n \log n)$ time, a homotopy from L to R of height at most $|L| + |R| + O(d \log n)$, where d is the largest among (a) maximum distance of a vertex of $\mathcal{D}$ from either of L and R, and (b) maximum edge weight, .*

Proof: Let $f(|L| + |R|, d, n)$ denote the maximum height of such a homotopy. We will show that $f(u, d, n) \in O(u + d \log n)$.

The base case $n = 0$ is easy. Indeed, if we have two edges (u, v) and (v, u) consecutive in R (or in L) we can retract these two edges. By repeating this we arrive at both L and R being identical, and we are done. The case $n = 1$ is handled in a similar fashion. After one face flip, the problem reduces to the case $n = 0$. As such, $f(|L| + |R|, d, 1) \leq |L| + |R| + d$.

For $n > 1$, compute for each vertex of G its shortest path to L, and consider the set of edges E used by all these shortest paths. Clearly $L \cup E$ form a tree. We consider each edge of R to be "thick" and have two sides (i.e., we think about these edges as being corridors with thickness). If E uses an edge of R then it uses the inner copy of this edge, while R uses the outer side. Similarly, we will consider vertices of R to be two vertices (one inside and other one on the boundary R). The set E would use only the inner vertices of R, while R would use only the outer vertices. To keep the graph triangulated we also arbitrarily triangulate inside each thick edge of R by adding corridor edges. Observe that, each corridor edge either connects two copies of a single vertex (thus has weight zero) or copies of two neighbors on R (and so has the same weight as the original edge).

Clearly, if we cut the plane along the edges of E, what remains is a simple triangulated polygon (it might have "thin" corridors along the edges of R). One can find a diagonal uv such that each side of the diagonal contains at least $\lceil n/3 \rceil$ triangles of G (and at most $(2/3)n$). (Here, we count only the "real" triangles of G – we consider the faces of the thin corridors of the edges of R to have weight 0.) Observe that, because the faces inside corridors have weight zero, we can ensure that if uv is a corridor edge then u and v are copies of a same vertex. We use this property in the following case analysis.



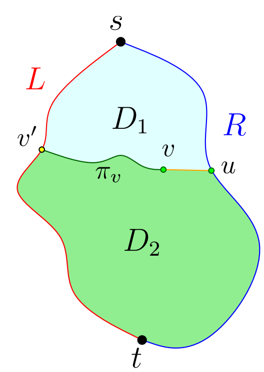
(A) Consider the case that u and v are both vertices of R. In this case, let $R[u, v]$ be the portion of R in between u and v, and let D_2 be the disk having $R[u, v] \cup uv$ as its outer boundary. Let D_1 be the disk $D \setminus D_2$. Let $M = R[s, u] \cup uv \cup R[v, t]$.

Clearly, the distance of any vertex of D_1 from L is at most d. By induction, there is a homotopy of height $f(|L| + |M|, d, (2/3)n)$ from L to M. Similarly, the distance of any vertex of D_2 from uv is smaller than its distance to L. As such, by induction, there is a homotopy between uv and $R[u, v]$ of height at most $f(|R[u, v]| + d, d, (2/3)n)$. Clearly, we can extend this to a homotopy of M to R of height $|R[s, u]| + f(|R[u, v]| + d, d, (2/3)n) + |R[v, t]|$.

Putting these two homotopies together results in the desired homotopy from L to R.

(B) If u and v are both vertices of L. The symmetric argument to the above applies.

(C) If uv is a corridor edge then, v is a vertex of E and u is an outer vertex of R (recall that we can assume that when uv is a corridor edges, v and u are inner and outer copies of a same vertex of R). Let π_v be the shortest path in $\mathcal{D}$ from v to L, and let v' be its endpoint on L.



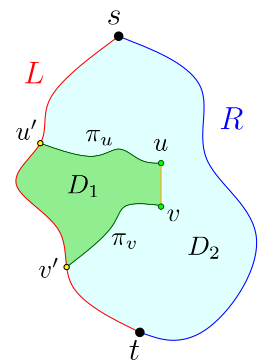
Consider the disk D_1 having the "left" boundary $L_1 = L[s, v'] \cup \pi_v \cup vu$ and $R_1 = R[s, u]$ as its "right" boundary. This disk contains at most $(2/3)n$ triangles, and by induction, it has a homotopy of height $f(|L_1| + |R_1|, d, (2/3)n)$. To see why we can apply the recursion, observe that u and v are copies of a vertex. That is, all shortest paths of vertices inside D_1 to L are completely inside D_1. Particularly, the distance of all vertices in D_1 to L_1 are at most d.

Similarly, the topological disk D_2 with the left boundary $L_2 = uv \cup \pi_v \cup L[v', t]$ and the right boundary $R_2 = R[u, t]$ has a homotopy of height $f(|L_2| + |R_2|, d, (2/3)n)$.

Starting with L, extending a tendril from v' to v, from v to u, and then applying the homotopy to first half of this walk (i.e., L_1) to move to R_1, and then the homotopy of D_2 to the second part, results in a homotopy of L to R of height

$$\max(|L| + 2d,\ f(|L_1| + |R_1|, d, (2/3)n) + |L_2|\,, |R_1| + f(|L_2| + |R_2|, d, (2/3)n))\,.$$

If the first number is the maximum, we are done. Otherwise, the above value is at most $f(|L| + |R| + d, d, 2/3n)$.



(D) Here we handle the case that u and v are both vertices of $L \cup E$. Then, as before, let u' and v' be the closest points on L to u and v, respectively. Similarly, let π_u (resp. π_v) be the shortest path from u (resp. v) to u' (resp. v').

Consider the disk D_1 having $L_1 = L[u', v']$ as left boundary, and $R_1 = \pi_u \cup uv \cup \pi_v$ as right boundary. This disk contains between $n/3$ and $2n/3$ triangles of the original surface. The distance of any vertex of D_1 to L_1 (when restricted to D_1) is at most d, and as such by induction, there is a homotopy from L_1 to R_1 of height $\alpha = f(|L_1| + |R_1|, d, (2/3)n) \leq f(|L[u', v']| + 3d, d, (2/3)n)$. This yields a homotopy of height $\alpha_1 = |L[s, u']| + \alpha + |L[v', t]|$, from L to $L_2 = L[s, u'] \cup \pi_u \cup uv \cup \pi_v \cup L[v, t]$. It is straight forward to check that $\alpha_1 \leq f(|L| + 3d, d, (2/3)n)$.

Next, let D_2 be the disk with its left boundary being L_2 and its right boundary being $R_2 = R$. Observe, that as before, the maximum distance of any vertex of D_2 to L_2 is at most d. As before, by induction, there is a homotopy form L_2 to R_2 of height $\alpha_2 = f(|L_2| + |R_2|, d, (2/3)n)$. Since $|L_2| \leq |L| + 3d$, we have $\alpha_2 \leq f(|L| + |R| + 3d, d, (2/3)n)$.

In all cases the length of the homotopy is at most

$$f(|L| + |R| + 3d, d, (2/3)n)$$

Now, it is easy to verify that the solution to the recursion $f(u, d, n)$ that complies with all the above inequalities is $f(u, d, n) = u + O(d \log n)$, as desired.

We can compute the shortest path tree in linear time using the algorithm of Henzinger *et al.* [HKRS97]. The separating edge can also be found in linear time using DFS. So, the running time for a graph with n faces is $T(n) = T(n_1) + T(n_2) + O(n)$, where $n_1 + n_2 = n$ and $n_1, n_2 \leq 2/3n$. It follows that $T(n) = O(n \log n)$. ∎

Remark 2.4 In the algorithm of Lemma 2.3, it is not necessary that we have the shortest paths from the L to all the vertices of $\mathcal{D}$. Instead, it is sufficient if we have a tree structure that provides paths from any vertex of $\mathcal{D}$ to L of distance at most d in this tree. We will use this property in the continuous case, where recomputing the shortest path tree is relatively expensive.

2.2 The continuous case

In this section we extend the arguments to the continuous case. Here we are given a piecewise linear triangulated topological disk, $\mathcal{D}$, with n triangles. The boundary of $\mathcal{D}$ is composed of two paths L and R with shared endpoints s and t. Observe that the distance of any point x in $\mathcal{D}$ from L and R is not longer than the homotopy height; there is a (s, t)-path that contains x. Here, we build a homotopy of height $|L| + |R| + O(d \log n)$, where d is the maximum distance of any point in $\mathcal{D}$ from either L or R.

Remember that the shortest path from a set of $O(n)$ edges to the whole surface can be computes in $O(n^3 \log n)$ time. Also, remember that a shortest path (a geodesic) intersects a face along a segment and it locally looks like a segment if the adjacent faces are rotated to be coplanar.

2.2.1 Homotopy height if edges are short

Here, we assume that the longest edge in $\mathcal{D}$ has length at most $2d$, where d is the maximum distance for any point of $\mathcal{D}$ from either L or R.

As in the discrete case, let E be the union of all the shortest paths from the vertices of $\mathcal{D}$ to L (as before, we treat the edges and vertices of R as having infinitesimal thickness). For a vertex v of $\mathcal{D}$, its shortest path π_v is a polygonal path that crosses between faces (usually) in the middle of edges (it might also go to a vertex, merge with some other shortest paths and then follow a common shortest path back to L). In particular, each such shortest path might intersect a face of $\mathcal{D}$ along a single segment. As such, the polygon resulting from cutting $\mathcal{D}$ along E, call it P, is a polygon that has complexity $O(n^2)$. A face of P is a hexagon, a pentagon, a quadrilateral, or a triangle. However, it has at most 3 edges that are portions of the edges of $\mathcal{D}$. We say the degree of a face is i if it has i edges that are portions of the edges of $\mathcal{D}$. Observe that, each triangle of $\mathcal{D}$ is now decomposed into a set of faces. Obviously, each triangle of $\mathcal{D}$ contains at most 1 face of degree 3. Overall, there are $O(n)$ faces of degree 3 in P.

Now consider C^*, the dual of the graph that is inside the polygon (ignore the edges on the boundary). Precisely, C^* has a vertex corresponding to each face inside the polygon P. Two vertices of C^* are adjacent if and only if their corresponding faces share a portion of an edge of $\mathcal{D}$ (a diagonal after cutting). Since the maximum degree of C^* is 3, there is an edge that is a good separator. We use this edge in a similar fashion to the proof of Lemma 2.3, except that in the recursion we avoid recomputing the shortest paths (i.e., we use the old shortest paths and distances computed in the original disk), see Remark 2.4. So, we compute the shortest paths once in the beginning in $O(n^3 \log n)$ time. Then, in each step we can find the separator in $O(n^2)$ time; remember there are $O(n^2)$ faces after cutting. That is the total time required to find the separators in all steps is $T(n) = T(n_1) + T(n_2) + O(n^2)$, where $n_1 + n_2 = n$ and $n_1, n_2 \leq 2/3n$; that is $T(n) = O(n^2 \log n)$ which is dominated by the time required to compute the shortest paths.

The proof of Lemma 2.3 then goes through literally in this case. Since all the edges have length at most $2d$, by assumption, we get the following.

Lemma 2.5 *Let $\mathcal{D}$ be a topological disk with n faces where every face is a triangle (here, the distance between any two points on the triangle is their Euclidean distance). Furthermore, the boundary of $\mathcal{D}$ is formed by two walks L and R (that share two endpoints). Let d be the maximum distance of any point of $\mathcal{D}$ from either L or R. Furthermore, assume that all edges of $\mathcal{D}$ have length at most $2d$. Then, one can compute a continuous homotopy from L to R of height $\leq |L| + |R| + O(d \log n)$ in $O(n^3 \log n)$ time.*

2.2.2 Homotopy height if there are long edges

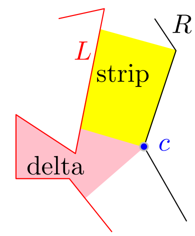
For any two points in $\mathcal{D}$ consider a shortest path π connecting them. Remember that the signature of π is the ordered sequence of edges (crossed or used) and vertices used by π.

For a point $p \in R$, let $\mathsf{s}_L(p)$ denote the signature of the shortest path from p to L. The $\mathsf{s}_L(p)$ is well defined in R except for a finite set of *medial* points, where there are two (or more) distinct shortest paths from L to p. In particular, let Π_R be the set of all shortest paths from any medial point on R to L. Observe that, the medial points are the only points that the signature of the shortest path from R to L changes in any non-degenerate triangulation.

Lemma 2.6 *The number of paths in Π_R is $O(|V(\mathcal{D})|)$, where $V(\mathcal{D})$ is the set of vertices of $\mathcal{D}$.*

Proof: Let $\{p_1, p_2, \ldots, p_k\}$ be the paths in Π_R sorted from top to bottom. That is for any $1 \leq i < k$, there is a vertex $v \in p_i$, such that $v \notin p_{i+1}$ and v is in the disc with boundary $L[s,a] \cup p_{i+1} \cup R[s,b]$, where a and b are endpoints of p_{i+1}.

Now, consider the set of elements (vertices and edges) that appear in $s_L(p_i)$, as i increases. In each two consecutive steps, at least one element is added or removed from this set. Further, whenever an element is removed from the set it cannot reappear. It follows that the cardinality if Π_R is at most the total number of vertices and edges in $\mathcal{D}$, which is $O(|V(\mathcal{D})|)$. ■



Cutting $\mathcal{D}$ along the paths of Π_R breaks $\mathcal{D}$ into ***corridors***. We say that a corridor C is a ***strip*** if all the shortest paths to L from the points interior the segment $C \cap R$ have the same signature.

Each corridor portion on R is either a point, in which case we say that it is a ***delta***, or alternatively, the corridor portion on R is a segment and in this case it is a strip. Note that the left side of a strip is either a single vertex or a segment of L.

We further break each delta into strips. So consider such a delta C with an apex c (i.e., the point of R on the boundary of C). For a point $x \in L \cap C$, we define its signature (in relation to C), to be the signature of the shortest path from x to c (restricted to lie inside C). Again, we partition $L \cap C$ into maximum intervals that have the same signature. Clearly, this breaks the delta into strips. Observe that the distance of any point of $L \cap C$ to c is at most $2d$. Applying this to all the deltas results in breaking $\mathcal{D}$ into strips. Arguing as in Lemma 2.6, we have that the total number of strips generated is $O(|V(\mathcal{D})|)$.

The shortest paths from R to L can be computed in $O(n^3 \log n)$ time. The shortest paths inside a delta to its apex can be computed in $O(n^2 \log n)$. Since there are linear number of deltas, the total running time for building the strips is $O(n^3 \log n)$.

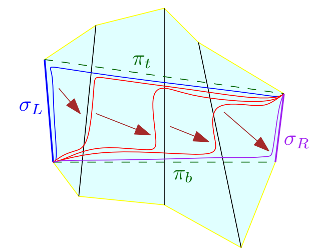
Lemma 2.7 *Consider a strip S generated by the above partition of $\mathcal{D}$. Let $\sigma_L = L \cap S$ and $\sigma_R = R \cap S$. Let π_t and π_b be the top and bottom paths forming the two sides of S that do not lie on R or L.*
(A) We have $|\pi_b| \leq 2d$ and $|\pi_t| \leq 2d$.
(B) If $|\sigma_L| > 0$ or $|\sigma_R| > 0$ then there is no vertex of $\mathcal{D}$ in the interior of S.
(C) If $|\sigma_L| > 0$ or $|\sigma_R| > 0$ then there is a homotopy from $\pi_t \cup \sigma_L$ to $\sigma_R \cup \pi_b$ of height $\max(|\sigma_L|, |\sigma_R|) + 4d$. This homotopy can be computed in linear time.

Proof: (A) If the strip was generated by the first stage of partitioning then the claim is easy.

Otherwise, consider a delta C with an apex c. For any point $x \in L \cap C$, consider the shortest path π_x from x to R in $\mathcal{D}$. If this path goes to c the claim holds immediately. Otherwise, the shortest path (that has length at most d) must cross either the top or bottom shortest path forming the boundary of C that are emanating from p. We can now modify π_x, so that after its intersection point with this shortest path, it follows it back to c. Clearly, the resulting shortest path has length at most $2d$ and lies inside the resulting strip.

(B) Indeed, the boundary paths π_t and π_b have the same signature (formally, they are the limit paths of same signature). Since $\mathcal{D}$ is non-degenerate, if there was any vertex in the middle, then the path on one side of the vertex, and the path on the other side of the vertex can not possibly have same signature.

(C) Assume that $|\sigma_L| > 0$ and $|\sigma_R| > 0$ (the other cases are degenerate instances of this case). The strip S has no vertex of $\mathcal{D}$ in its interior, and as such it is formed by taking planar quadrilaterals and gluing them together along common edges. Observe that by the triangle inequality, all such edges of any of these quadrilaterals are of length $\leq \max(|\sigma_L|, |\sigma_R|) + 4d$. It is now easy to check that we can collapse each such quadrilateral in turn to get the required homotopy. Since each of π_t and π_b is composed of two shortest paths there are linear number of such quadrilaterals, and each collapse can be done in constant time. See figure for an example. ∎



The above suggests a natural homotopy from L to R – we compute the strips, and then we order the strips according to their order along L, and transform each one of them at time, such that starting with L we end up with R.

To this end, a strip S that has only a single point on both L and R is a ***pocket***. We still need to show how to compute a homotopy of a pocket. In our case, a pocket has perimeter $4d$, and there is a point on its boundary, such that the distance of any point in it to this base point is at most $2d$ (this follows by arguing in a similar fashion to Lemma 2.7 (A)). The following is a consequence of the triangle inequality.

Observation 2.8 *Consider a topological disk $\mathcal{D}$, such that all the points of $\mathcal{D}$ are in distance at most $2d$ from some point c. Then the longest edge in $\mathcal{D}$ has length at most $4d$.*

As such, all the edges inside a pocket can not be longer than $4d$. We can now apply Lemma 2.5 to such a pocket. This results in the desired homotopy. This completes the proof of the following.

Theorem 2.9 *Given a triangulated piecewise linear surface with the topology of a disk, such that its boundary is formed by two walks L and R. Then, there is a continuous homotopy from L to R of height $\leq |L| + |R| + O(d \log n)$. This homotopy can be computed in $O(n^3 \log n)$ time.*

3 Homotopic Frechét Distance

In this section, fix $\mathcal{D}$ to be a triangulated topological disk with n faces. Let the boundary of $\mathcal{D}$ be composed of A, R, B and L, four internally disjoint walks appearing in clockwise order along the boundary. Also, let $a_0 = L \cap A$, $b_0 = L \cap B$, $a_1 = R \cap A$ and $b_1 = R \cap B$. We use the same notation to argue about both the discrete problem and the continuous problem.

Let δ be the regular Frechét distance and δ_H be the homotopic Frechét distance between A and B. Clearly, $\delta \leq \delta_H$. The following lemma implies that the Frechét distance can be approximated within a constant factor.

Lemma 3.1 *Let $\mathcal{D}$ be a triangulated topological disk with n faces, and A and B be two internally disjoint walks on the boundary of $\mathcal{D}$. Then, one can compute a reparameterizations of A and B that approximates the Frechét distance between A and B. The computed reparameterizations have width $\leq 2\delta$ in the continuous case, and $\leq 3\delta$ in the discrete case, where δ is the Frechét distance between A and B. The running time is $O(n^3 \log n)$ in the continuous case and $O(n)$ in the discrete case.*

Proof: First, we prove the continuous case. Let Π be the set of shortest paths from all points of A to the curve B. Similar to Lemma 2.6 we define Π_A to be the set of all shortest paths from medial points on A to B. The set Π_A is composed of linear number of paths; the proof is similar to Lemma 2.6. The paths in Π_A do not cross and so partition $\mathcal{D}$ into a set of regions. Each region is bounded by a portion of A, a portion of B and two paths in Π_A. A region is a *delta* if the two paths of Π_A in its boundary share a single endpoint (on A), it is a pocket if they share two endpoints (one on A and one on B), and it is strip if they share no endpoints.

Obviously the endpoints of paths in Π cover all A. Observe that they also cover all B except the bases of deltas. Now, for each delta we compute the set of all shortest paths from its base to its apex inside the delta. Let Π_B be the set of all such paths in all deltas. Clearly, the union of Π and Π_B is a set of non-crossing paths whose endpoints cover both A and B.

The shortest path from any point of A to B is at most δ. So, all paths in Π have length at most δ. Similarly, the shortest path from a point of B to A is at most δ. Now, consider a delta C with apex c. Let b be a point on the base of C (and so on B). The shortest path π_b from b to A has length at most δ. Let x be the first point that π_b intersects a boundary path of C, π_C. Now, $\pi_b[b, x] \cdot \pi_C[x, c]$ has length at most 2δ and it is inside C. We conclude that all paths in Π_B have length at most 2δ and the proof is complete.

In $O(n^3 \log n)$ time, we can compute all shortest paths from A to the whole surface. Then, we need $O(n^2 \log n)$ time to compute the shortest paths inside each of the linear number of deltas. It follows that the total running time is $O(n^3 \log n)$.

In the discrete case, we first compute the set of shortest paths, $\Pi = \{\pi_1, \pi_2, \cdots, \pi_k\}$, from vertices of A to the path B. Now, let π_i and π_{i+1} be two consecutive paths, that is the endpoints of π_i and π_{i+1}, a_i and a_{i+1}, are adjacent vertices on A. For all $1 \leq i < k$, we add the paths $\pi_i^+ = (a_i, a_{i+1}) \cdot \pi_{i+1}$ to the set Π to obtain Π^+. Observe that each path in Π^+ has length at most 2δ; it is composed of zero or one edge of A and a shortest path from a vertex of A to B. Further, Π^+ partitions the graph into regions, similar to the continuous case. Now for each vertex of B that is not an endpoint of a path in Π^+, we compute the shortest path inside its region to A. Because the region is bounded by paths of length at most 2δ, the length of such a shortest path is at most 3δ. If Π_B is the set of all such shortest paths, then $\Pi^+ \cup \Pi_B$ is a leash sequence of height at most 3δ.

We use the algorithm of Henzinger *et al.* [HKRS97] to compute the shortest paths from A in linear time. Since all regions are disjoint, we can compute all the shortest paths inside different regions in $O(n)$ time, as well. ∎

The following lemma implies that if all the vertices in $\mathcal{D}$ are not too far from the two curves, then one can transform the Frechét distance into the continuous variant (i.e., without jumps in the leash).

Lemma 3.2 *Let $\mathcal{D}$ be a triangulated topological disk with n faces, and A and B be two internally disjoint walks on the boundary of $\mathcal{D}$. Further, assume for all $p \in \mathcal{D}$, p's distance to both A and B is $O(\delta_H)$. Then, one can compute a continuous (A, B)-leash function of height $O(\delta_H \log n)$. The running time is $O(n^4 \log n)$ for the continuous case and $O(n \log n)$ for the discrete case.*

Proof: Using the algorithm of Lemma 3.1 we compute a (not necessarily continuous) leash function s of height $2\delta \in O(\delta_H)$.

Note that the leashes are not required to deform continuously in s. For a given time $t \in [0, 1]$, let $s^-(t) = \lim_{t' \to t^-} s(t')$ and $s^+(t) = \lim_{t' \to t^+} s(t')$. By definition, s is discontinuous at t if and only if $s^-(t) \neq s^+(t)$. In the discrete case let $s^-(t) = s(t)$ and $s^+(t) = s(t+1)$. In this case, we say s is discontinuous at t if $s^-(t)$ and $s^+(t)$ are more than one flip operation apart. In both cases, we also say that s has a gap at t.

Observe that a gap at time t can be filled by attaching an $(s^-(t), s^+(t))$-homotopy to s at t. Observe that all the vertices inside the disc with boundary $s^-(t) \cup s^+(t)$ have distance $O(\delta_H)$ to A and B and so to $s^-(t)$ and $s^+(t)$, so Lemma 2.9 and Lemma 2.3 imply that an $(s^-(t), s^+(t))$-homotopy with height $O(\delta_H \log n)$ can be computed.

Suppose that s has a gap at t. In the discrete case, the disc with boundary $s^-(t) \cup s^+(t)$ contains at least one face. In the continuous case, that disc contains at least one vertex. Thus, in both cases, there are $O(n)$ gaps.

We start with s and fill in all linear number of gaps to obtain a continuous leash sequence of height $O(\delta_H \log n)$.

In the discrete case, the initial leash sequence can be computed in $O(n)$ time. Since the gaps are disjoint, they can be filled in overall $O(n \log n)$.

In the continuous case, the initial leash sequence can be computed in $O(n^3 \log n)$ time. Each of the $O(n)$ number of gaps can be filled in $O(n^3 \log n)$ time. ∎

Now, we consider the case that vertices exist that are far from at least one of A and B. Intuitively, we like to pick the first and the last leash so that we can avoid moving the leash over such vertices.

A vertex $v \in V(\mathcal{D})$ is ***tall*** if and only if its distance to A or B is larger than δ_H. In the discrete case, we subdivide each edge in the beginning so that if an edge has length $> 2\delta_H$, then the vertex inserted in the middle of it is tall. Observe that, no leash of the optimum continuous leash sequence affords to contain a tall vertex. We use T to denote the set of all tall vertices in $V(\mathcal{D})$.

Now, let p and p' be two walks connecting A and B. We say that p and p' are ***isotopic*** in $\mathcal{D} \backslash T$ if and only if they are homotopic in $\mathcal{D} \backslash T$ after contracting A and B. Consequently, the disc with boundary $A \cup B \cup p \cup p'$ contains no tall vertices for isotopic p and p'.

Isotopy partitions the set of (A, B)-paths to equivalence classes. Let h be any isotopy class. We use $L(h)$ and $R(h)$ to denote the shortest (a_0, b_0)-walk and (a_1, b_1)-walk in h.

Let p be any walk in h and $a \in A$ and $b \in B$ be the endpoints of p. We define the ***left tall set*** of h, denote $T_l(h)$ to be the set of all tall vertices to the left of p; inside the disc with boundary $L \cup A[a_0, a] \cup p \cup B[b_0, b]$. We similarly define the ***right tall set*** of h, $T_r(h)$, to be the set of all tall vertices to the right of p; inside the disc with boundary $R \cup A[a, a_1] \cup p \cup B[b, b_1]$. Note that the sets $T_l(h)$ and $T_r(h)$ do not depend on the particular choice of p, since all paths in h are in the same isotopy class.

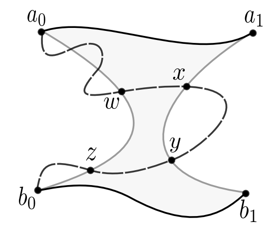
We say that h is ***extendable*** from the left (or simply extendable) if and only if $|L(h)| \leq \delta_H$ and there is an isotopy class h', such that $|L(h')| \leq \delta_H$ and $T_l(h) \subset T_l(h')$. We say h is ***saturated*** if it is not extendable and $|L(h)| \leq \delta_H$.

Lemma 3.3 *Let h be a saturated isotopy class. Then, $|R(h)| \leq 4\delta_H$.*

Proof: Let h_{opt} be the isotopy class of the leashes in the optimum solution. Remember that no leash of the optimum solution contains a tall vertex. That is, all leashes in that solution are isotopic.

Since h is saturated $T_l(h)$ is not a proper subset of $T_l(h_{opt})$. If $T_l(h) = T_l(h_{opt})$ then $h = h_{opt}$, particularly $|R(h)| = |R(h_{opt})| \leq \delta_H$.

Otherwise, the set $T_l(h) \cap T_r(h_{opt})$ is not empty. It follows that $L(h)$ crosses $R(h_{opt})$. Let $x, y \in L(h) \cap R(h_{opt})$, such that both $R(h_{opt})[a_1, x]$ and $R(h_{opt})[y, b_1]$ are internally disjoint from $L(h)$. Also, let $w, z \in L(h) \cap L(h_{opt})$, such that both $L(h)[w, x]$ and $L(h)[z, y]$ are internally disjoint from $L(h_{opt})$. In the figure, the gray paths are $L(h_{opt})$ and $R(h_{opt})$, and the dashed path is $L(h)$. Observe that the shaded area contains no tall vertex.



Clearly, the walk $A[a_1, a_0] \cdot L(h)[a_0, b_0] \cdot B[b_0, b_1]$ is isotopy to $L(h)$. Observe that the walk $A' = R(h_{opt})[a_1, x] \cdot L(h)[x, w] \cdot L(h_{opt})[w, a_0]$ is homotopic to A. Similarly, $B' = L(h_{opt})[b_0, z] \cdot L(h)[z, y] \cdot R(h_{opt})[y, b_1]$ is homotopic to B. It follows that $R(h) = A' \cdot L(h) \cdot B'$ is also isotopic to $L(h)$. Now, it is straight forward to check that $|R(h)| \leq 2\,|L(h)| + |L(h_{opt})| + |R(h_{opt})| \leq 4\delta_H$. ■

Lemma 3.4 *Let h be an isotopy class of (A, B)-leashes in $\mathcal{D} \setminus T$ such that $|L(h)|, |R(h)| = O(\delta_H)$ and $\mathcal{D}'$ be the disk with boundary $A \cdot R(h) \cdot B \cdot L(h)$. Then, all points inside $\mathcal{D}'$ are closer than $O(\delta_H)$ to both A and B in $\mathcal{D}'$.*

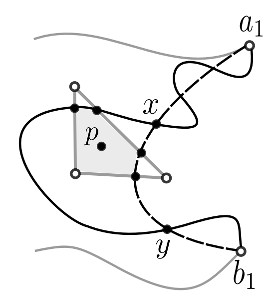
Proof: Consider a point $p \in \mathcal{D}'$ and its shortest path π_p to either of A or B in $\mathcal{D}$. Assume that $\pi_p = O(\delta_H)$ and let x be the first intersection point of π_p with either $L(h)$ or $R(h)$. Then, taking π_p from p to x and then taking the boundary path ($L(h)$ or $R(h)$) from x to A or B has length $O(\delta_H)$. So, we only need to prove that the distance of any point to both A and B is $O(\delta_H)$ in $\mathcal{D}$.

This is by construction true for the vertices of the triangulation.

In the discrete case, since each edge of length $2\delta_H$ contains a tall vertex after the subdivision, there is no edge of length $> 2\delta_H$ inside $\mathcal{D}'$. That is, the points on the edges inside $\mathcal{D}'$ have length $O(\delta_H)$ to both A and B.

In the continuous case, let $\mathcal{D}_{opt} = A \cdot R(h_{opt}) \cdot B \cdot L(h_{opt})$. If $p \in \mathcal{D}_{opt}$ then its distance to A and B is at most δ_H and we are done. Otherwise, assume that p is on the right side $R(h_{opt})$, the other case that p is on the left sue of $L(h_{opt})$ is exactly similar.

Let x and y be the intersection points of $R(h)$ and $R(h_{opt})$ such that p is inside the disk with boundary $D = R(h)[x, y] \cdot R(h_{opt})[y, x]$. Let Δ be the triangle in $\mathcal{D}$ that contains p. $\Delta \cap D$ is a polygon whose vertices are either (*not tall*) vertices of $\mathcal{D}$ that have distance at most δ_H to A and so to the boundary of D or intersection points of Δ and $R(h) \cup R(h_{opt})$. Then, triangle inequality implies that the perimeter of $\Delta \cap D$ is at most $O(\delta_H)$. Connect each vertex of Δ that is inside D to D's boundary through its shortest path (of length at most δ_H) and use the triangle inequality to bound the lengths of the sides of $\Delta \cap D$ and so its perimeter. In particular, p has distance $O(\delta_H)$ to a vertex of the polygon $\Delta \cap D$, that is p's distance to A is $O(\delta_H)$. ■



Lemma 3.5 *Let $\mathcal{D}$ be a triangulated topological disk with n faces, and A and B be two internally disjoint walks on $\mathcal{D}$'s boundary. Assuming that the homotopic Frechét distance between A and B is given, one can compute a saturated isotopy class in $O(n^7 \log n)$ time and in $O(n^3)$ in the continuous and discrete cases, respectively.*

Proof: Let δ_H denote the homotopic Frechét distance between A and B. Let h_0 be the homotopy class of the (a_0, b_0)-shortest path. That is, $L(h_0)$ is an (a_0, b_0)-shortest path. Obviously, $|L(h_0)| \leq \delta_H$.

Now, let h_i be any isotopy class such that $L(h_i) \leq \delta_H$. Then, h_i is extendable to some h_{i+1} if and only if there exists a tall vertex $v \in T_r(h_i)$, such that $L(h_{i+1}) \leq \delta_H$ and $T_l(h_{i+1}) \supseteq T_l(h_i) \cup \{v\}$. Note that $T_l(h_{i+1})$ may contain other tall vertices as well. In other words, h_i is extendable if and only if the shortest path homotopic to $A[a_0, a_1] \cdot R[a_1, b_1] \cdot B[b_1, b_0]$ in $D \backslash (T_l(h_i) \cup \{v\})$ is not longer than δ_H. Given $T_l(h_i)$ and v, we compute such a shortest path as follows.

We compute the shortest path from each vertex in $T_l(h_i) \cup \{v\}$ to L. This can be done in $O(n^3 \log n)$ time and $O(n)$ time in the continuous and discrete cases, respectively. Then, we cut the graph along all the computed shortest paths, to obtain a graph of complexity $O(n^2)$ in both cases. Finally we compute the shortest (a_0, b_0)-path in this graph. The total running time is $O(n^6 \log n)$ and $O(n^2)$ in continuous and discrete cases, respectively.

To compute a saturated homotopy class, we start with h_0. In each iteration, we extend h_i to h_{i+1} if possible. To do so, we test for all $v \in T_r(h_i)$ whether it can be used to extend h_i to h_{i+1}. If we find such a vertex we extend h_i to h_{i+1} and iterate. Otherwise, we return h_i as a saturated homotopy class. Observe that if a vertex cannot be used in some iteration it cannot be used in a later iterations either. So, we need to consider each vertex at most once.

There are $O(n)$ number of extensions. So, the total running time is $O(n^7 \log n)$ and $O(n^3)$ in the continuous and the discrete cases, respectively. ■

Theorem 3.6 *Let $\mathcal{D}$ be a triangulated topological disk with n faces, and A and B be two internally disjoint walks on the boundary of $\mathcal{D}$. Then, one can compute a continuous leash sequence between A and B of height $O(\delta_H \log n)$, where δ_H is the homotopic Frechét distance between A and B. The running time is $O(n^7 \log n \log \delta_H)$ in the continuous case and $O(n^3 \log \delta_H)$ in the discrete case.*

Proof: We use geometrically increasing values as an upper bound guess for δ_H. Given δ_H, Lemma 3.5 ensures that a saturated homotopy class, h, can be computed in polynomial time. Lemma 3.3 implies that both $L(h)$ and $R(h)$ are at most $4\delta_H$. Let $\mathcal{D}' \subseteq \mathcal{D}$ be the disc with boundary $A \cup B \cup L(h) \cup R(h)$. Since $L(h)$ and $R(h)$ are homotopic there is no tall vertices inside $\mathcal{D}'$. It follows that the shortest path inside $\mathcal{D}'$ from any vertex v to both A and B is at most $5\delta_H$. Finally, Lemma 3.2 implies that a continuous leash sequence of height $O(\delta_H \log n)$ between A and B, inside $\mathcal{D}'$, can be computed. ■

4 Homotopic Frechét distance on surfaces with positive genus

Fix X to be an orientable triangulated surface of genus g that is composed of n triangles. Let A be a (a_0, a_1)-path and B be a (b_0, b_1)-path, both on the boundary of X. We assume (without loss of generality) that X contains no other boundary except for those containing A and B; one can attach a face f_σ to every other boundary σ and then insert a very tall vertex f_σ to make sure that no leash of any approximate solution intersects f_σ.

Let $\ell_{opt}(\cdot)$ be the optimum leash function; the function that obtains the minimum homotopic Frechét distance. Indeed, all $\ell_{opt}(t)$ are isotopic in X; they are homotopic in X after contracting A and B. Intuitively, we can think of the collection of $\ell_{opt}(t)$'s, where $0 \leq t \leq 1$, as a thick edge.

Following Chambers *et al.* [CdVE+08b] and Erickson and Whittlesey [EW05], we find a greedy system of arcs $(p_1, p_2, \cdots, p_\beta)$, with the following properties.

- $\beta = 2g$ if A and B are on the same boundary and $\beta = 2g + 1$ otherwise.
- Each p_i is either a path between a_0 and b_0 or a loop that contains either a_0 or b_0.

- Each p_i is composed of two shortest paths
- If we cut X along all p_i's we obtain a topological disc P that contains 2 copies (p_i^+ and p_i^-) of each p_i on its boundary.

To compute a system of arcs in the continuous case we imitate the discrete algorithm. First, we use Mitchel *et al.* [MMP87] to compute the set of all medial points for the shortest paths from $\{a_0, b_0\}$. Observe that this set includes β cycles $\{c_1, c_2, \cdots, c_\beta\}$ whose removal makes X a topological disc; see [DL09]. By adding a single point from each c_i to the shortest paths computed from a_0 and b_0 we obtain a greedy system of arcs with the same set of properties.

Now, we assign a ***signature*** of β integers $\eta = (\eta_1, \eta_2, \cdots, \eta_\beta)$ to each path γ on X. Consider an arbitrary direction for each p_i, the integer η_i is the *algebraic intersection number* between γ and p_i; the number of times that γ crosses p_i considering the direction. Note that η_i can be a negative number.

Let C be the *universal cover* of X, i.e. the space that contains a copy of P, for any signature. For a given signature η, we use $P[\eta]$ to denote the copy of P in the universal cover corresponding to η. In addition, we use $v[\eta]$, $e[\eta]$, $A[\eta]$ and $B[\eta]$ to denote copies of a vertex v, an edge e, A or B in $P[\eta]$. The paths p_i^+ in $P[\eta]$ and p_i^- in $P[\zeta]$ are identified if and only if $\eta = (\eta_1, \eta_2, \cdots, \eta_\beta)$, $\zeta = (\zeta_1, \zeta_2, \cdots, \zeta_\beta)$, $\zeta_i - \eta_i = 1$ and $\zeta_j = \eta_j$ for all $j \neq i$.

Observe that each path in X corresponds to an infinite number of paths in the universal cover. In particular, each (u, v)-path γ with signature η in X is mapped to a path γ' in the universal cover that starts from $u[0, 0, \cdots, 0]$ and ends in $v[\eta]$. In this case, we call γ' the ***canonical*** image of γ.

Now, consider the canonical images of the set of leashes of $\ell_{opt}(\cdot)$. Since, all leashes in $\ell_{opt}(\cdot)$ are isotopic, they have the same signature η_{opt} and so their canonical images in C are paths between $A[0, 0, \cdots, 0]$ and $B[\eta_{opt}]$.

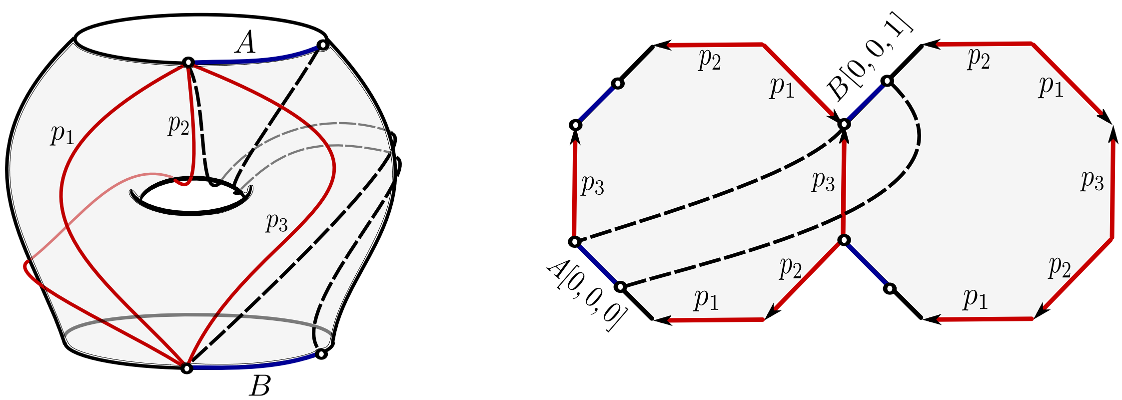


Figure 2: Left: A doubly punctured torus, a greedy system of arcs, and two canonical leashes, right: the same surface cut open along the greedy system of arcs.

The following lemma bounds the search space for η_{opt}.

Lemma 4.1 *Let π be any shortest path in X. Then, there exists an optimum leash function ℓ_{opt}, such that $\ell_{opt}(0)$ crosses π at most twice.*

Proof: Let $\ell_{opt}(\cdot)$ be the optimum solution with the minimum number of crossings on π. Assume to the contrary that π crosses $\ell_{opt}(0)$ in x, y and z in this order along π.

Figure 3: Shortcutting the leashes.

First, assume that $\ell_{opt}(0)$ crosses π in two different directions and without loss of generality at x and y; see Figure 3-left. Isotopy of the $\ell_{opt}(t)$'s implies that there is a $t' \in [0,1]$ such that for any $t < t'$, $\ell_{opt}(t)$ crosses π after x in $x(t)$ and before y in $y(t)$, in the same directions as $\ell_{opt}(0)$. For each $t < t'$ we change $\ell_{opt}(t)$ to $\ell_{opt}(t)[A, x(t)] \cdot \pi[x(t), y(t)] \cdot \ell_{opt}[y(t), B]$. Because π is a shortest path the length of no leash has been increased while the number of crossings is reduced.

Second, assume that $\ell_{opt}(0)$ crosses π in the same direction at x, y and z; see Figure 3-right. Isotopy of the $\ell_{opt}(t)$'s implies that all $\ell_{opt}(t)$'s cross π between x and y and between y and z in order. Assume that $x(t)$ and $y(t)$ are the first crossing points of $\ell_{opt}(t)$ on π after x and y alone π, respectively. For all $t \in [0,1]$ we change $\ell_{opt}(t)$ to $\ell_{opt}(t)[A, x(t)] \cdot \pi[x(t), y(t)] \cdot \ell_{opt}(t)[y[t], B]$. Again, we reduced the number of crossings on π without increasing the length of any leash. ■

The following corollary is immediate because each p_i is composed of two shortest paths.

Corollary 4.2 *Each arc in the greedy system of arcs crosses $\ell_{opt}(0)$ at most 4 times.*

We say a signature is ***feasible*** if the absolute value of each one of its corrdinates is at most 4; it follows that there are $2^{O(g)}$ feasible signatures. Corollary 4.2 implies that η_{opt} is feasible. We build Σ, a sub-space of C that only contains copies of P correspondent to feasible signatures. Observe that Σ is a topological disc composed of $2^{O(g)}n$ triangles and it has $2^{O(g)}$ copies of A and B on its boundary.

Lemma 4.1 implies that the canonical images of all the leashes in the optimum solution are completely inside Σ. If we know η_{opt} we can compute the homotopic Frechét distance between $A[0, 0, \cdots, 0]$ and $B[\eta_{opt}]$. But, there are $2^{O(g)}$ candidates for η_{opt} in Σ. For each such candidate we use Theorem 3.6 to compute the best leash function. We obtain the following:

Theorem 4.3 *Let X be a triangulated topological surface (2-manifold) of genus g that is composed of n faces, and A and B be two internally disjoint paths on X's boundary. Then, one can compute a continuous leash sequence between A and B of height $O(\delta_H \log n)$, where δ_H is the homotopic Frechét distance between A and B. The running time is $2^{O(g)} n^7 \log n \log \delta_H$ in the continuous case and $2^{O(g)} n^3 \log \delta_H$ in the discrete case.*

5 Conclusions

We presented a $O(\log n)$ approximation algorithm for approximating the homotopy height and the homotopic Frechét distance between curves on piecewise linear surfaces. It seems quite believable that the approximation quality can be further improved, and we leave this as the main open problem of our work. Since our algorithm works both for the continuous and discrete cases, it seems natural to conjecture that this algorithm should work more general surfaces and metrics.

References

[AB05] H. Alt and M. Buchin. Semi-computability of the Frechét distance between surfaces. In *Proc. 21st Euro. Workshop on Comput. Geom.*, 2005.

[AG95] H. Alt and M. Godau. Computing the Fréchet distance between two polygonal curves. *Internat. J. Comput. Geom. Appl.*, 5:75–91, 1995.

[BBG08a] K. Buchin, M. Buchin, and J. Gudmundsson. Detecting single file movement. In *Proc. 16th ACM SIGSPATIAL Int. Conf. Adv. GIS*, pages 288–297, 2008.

[BBG+08b] K. Buchin, M. Buchin, J. Gudmundsson, Maarten L., and J. Luo. Detecting commuting patterns by clustering subtrajectories. In *Proc. 19th Annu. Internat. Sympos. Algorithms Comput.*, pages 644–655, 2008.

[BPSW05] S. Brakatsoulas, D. Pfoser, R. Salas, and C. Wenk. On map-matching vehicle tracking data. In *Proc. 31st VLDB Conference*, pages 853–864, 2005.

[BVIG91] Chakib Bennis, Jean-Marc Vézien, Gérard Iglésias, and André Gagalowicz. Piecewise surface flattening for non-distorted texture mapping. In Thomas W. Sederberg, editor, *Proc. SIGGRAPH '91*, volume 25, pages 237–246, 1991.

[CDHP+11] A. F. Cook, A. Driemel, S. Har-Peled, J. Sherette, and C. Wenk. Computing the fréchet distance between folded polygons. In *Proc. 7th Workshop Algorithms Data Struct.*, pages 267–278, 2011.

[CdVE+08a] E. W. Chambers, E. C. de Verdière, J. Erickson, S. Lazard, F. Lazarus, and S. Thite. Walking your dog in the woods in polynomial time. In *Proc. 24th Annu. ACM Sympos. Comput. Geom.*, pages 101–109, New York, NY, USA, 2008. ACM.

[CdVE+08b] W. W. Chambers, E. C. de Verdière, J. Erickson, F. Lazarus, and K. Whittlesey. Splitting (complicated) surfaces is hard. *Comput. Geom. Theory Appl.*, 41:94–110, October 2008.

[CL09] E. W. Chambers and D. Letscher. On the height of a homotopy. In *Proc. 21st Canad. Conf. Comput. Geom.*, 2009.

[CL10] E. W. Chambers and D. Letscher. Erratum for on the height of a homotopy. `http://mathcs.slu.edu/~chambers/papers/hherratum.pdf`, 2010.

[CW10] A. F. Cook and C. Wenk. Geodesic fréchet distance inside a simple polygon. *ACM Trans. Algo.*, 7:9:1–9:19, 2010.

[DHW10] A. Driemel, S. Har-Peled, and C. Wenk. Approximating the Fréchet distance for realistic curves in near linear time. In *Proc. 26th Annu. ACM Sympos. Comput. Geom.*, pages 365–374, 2010. `http://arxiv.org/abs/1003.0460`.

[DL09] Tamal K. Dey and Kuiyu Li. Cut locus and topology from surface point data. In *Proceedings of the 25th annual symposium on Computational geometry*, SCG '09, pages 125–134, New York, NY, USA, 2009. ACM.

[EGH+02] A. Efrat, L. J. Guibas, S. Har-Peled, J. 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.

[EM94] T. Eiter and H. Mannila. Computing discrete Frechét distance. Tech. Report CD-TR 94/64, Christian Doppler Lab. Expert Sys., TU Vienna, Austria, 1994.

[EW05] Jeff Erickson and Kim Whittlesey. Greedy optimal homotopy and homology generators. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA '05, pages 1038–1046, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.

[Flo97] M. S. Floater. Parameterization and smooth approximation of surface triangulations. *Comput. Aided Geom. Design*, 14(4):231–250, 1997.

[Fre06] M. Frechét. Sur quelques points du calcul fonctionnel. *Rendic. Circ. Mat. Palermo*, 22:1–74, 1906.

[Fre24] M. Frechét. Sur la distance de deux surfaces. *Ann. Soc. Polonaise Math.*, 3:4–19, 1924.

[God99] M. Godau. *On the complexity of measuring the similarity between geometric objects in higher dimensions*. PhD thesis, Free University of Berlin, 1999.

[HKRS97] M. R. Henzinger, P. Klein, S. Rao, and S. Subramanian. Faster shortest-path algorithms for planar graphs. *J. Comput. Sys. Sci.*, 55:3–23, August 1997.

[HR11] S. Har-Peled and B. Raichel. The Fréchet distance revisited and extended. In *Proc. 27th Annu. ACM Sympos. Comput. Geom.*, pages 448–457, 2011. `http://www.cs.uiuc.edu/~sariel/papers/10/frechet3d/`.

[KKS05] M.S. Kim, S.W. Kim, and M. Shin. Optimization of subsequence matching under time warping in time-series databases. In *Proc. ACM symp. Applied comput.*, pages 581–586, 2005.

[KP99] E. J. Keogh and M. J. Pazzani. Scaling up dynamic time warping to massive dataset. In *Proc. of the Third Euro. Conf. Princip. Data Mining and Know. Disc.*, pages 1–11, 1999.

[MDBH06] A. Mascret, T. Devogele, I. Le Berre, and A. Hénaff. Coastline matching process based on the discrete Fréchet distance. In *Proc. 12th Int. Sym. Spatial Data Handling*, pages 383–400, 2006.

[MMP87] J. S.B. Mitchell, D. M. Mount, and C. H. Papadimitriou. The discrete geodesic problem. *SIAM J. Comput.*, 16:647–668, 1987.

[PB00] D. Piponi and G. Borshukov. Seamless texture mapping of subdivision surfaces by model pelting and texture blending. In *Proc. SIGGRAPH 2000*, pages 471–478, August 2000.

[SdS00] Alla Sheffer and Eric de Sturler. Surface parameterization for meshing by triangulation flattening. In *Proc. 9th International Meshing Roundtable*, pages 161–172, 2000.

[SGHS08] J. Serrà, E. Gómez, P. Herrera, and X. Serra. Chroma binary similarity and local alignment applied to cover song identification. *IEEE Transactions on Audio, Speech & Language Processing*, 16(6):1138–1151, 2008.

[WSP06] C. Wenk, R. Salas, and D. Pfoser. Addressing the need for map-matching speed: Localizing global curve-matching algorithms. In *Proc. 18th Int. Conf. Sci. Statis. Database Manag.*, pages 879–888, 2006.

A Some standard definitions used in the paper

A.1 Planar Graphs

An ***embedding*** of a graph G on the plane maps the vertices of G to different points on the plane and its edges to disjoint paths except for the endpoints. The faces of an embedding are maximal connected subsets of the plane that are disjoint from the image of the graph. We use ∂F to denote the boundary of a set of faces F. We abuse the notation and use ∂G to denote the boundary of outer face of G. In particular, ∂f refers to the boundary of a single face f. The term ***plane graph*** refers to a graph together with its embedding on the plane.

The ***dual*** graph G^* of a plane graph G is the (multi-)graph whose vertices correspond to the faces of G, where two faces are joined by a (dual) edge if and only if their corresponding faces are separated by an edge of G. Thus, any edge e in G corresponds to a dual edge e^* in G^*, any vertex v in G corresponds to a face v^* in G^* and any face f in G^* corresponds to a vertex f^* in G^*.

Let $G = (V, E)$ be a simple undirected plane graph with edge weights $w : E \to \mathbb{R}^+$. A ***walk*** W in G is a sequence of vertices $(v_1, v_2, \cdots, v_k)$ such that each adjacent pair $e_i = (v_i, v_{i+1})$ is an edge in G. The length of W is $|W| = \sum_{i=1}^{k-1} w(e_i)$.

Let v_i and v_j be two vertices that appear on W. By $W[v_i, v_j]$ we mean the sub-walk of W that starts from the first appearance of v_i and ends at the first appearance of v_j after v_i on W. For two walks, $W_1 = (v_1, v_2, \ldots, v_i)$ and $W_2 = (v_i, v_{i+1}, \ldots, v_j)$, we define their ***concatenation*** to be $W_1 \cdot W_2 = (v_1, v_2, \ldots, v_i, v_{i+1}, \ldots, v_j)$.

A walk with distinct vertices is called a ***path***. We use the terms (u, v)-walk to refer to a walk that starts at u and ends in v; (u, v)-path is defined similarly. A walk is closed if its first and last vertices are identical. A closed path is a ***cycle***. Two walks *cross* if and only if their images cross on the plane. That is, no infinitesimal perturbation makes them disjoint.

A.2 Piecewise Linear Surfaces and Geodesics

A ***piecewise linear*** surface is composed of finite number of Euclidean triangles by identifying pairs of equal length edges. In this paper we work with piecewise linear surfaces that can be embedded in three dimensional space such that all triangles are flat and the surface does not cross itself. Equivalently, the surface can be presented by a set of edges and three dimensional coordinates of the vertices.

We say that a triangulated surface is ***non-degenerate*** if no interior point has curvature 0, i.e. when for every non-boundary vertex x, the sum of the angles of the triangles incident to x is not equal to 2π. We can turn any triangulated surface into a non-degenerate one by perturbing all edge lengths by a factor of at most $1+\epsilon$, for some $\epsilon = O(1/n^2)$. This changes the metric of the surface by at most a factor of $O(1+1/n)$, and thus the minimum height of a homotopy. Since such a factor will be negligible for our approximation guarantee, we can assume that the input surface is always non-degenerate.

A ***path*** γ on the surface $\mathcal{D}$ is a function $\gamma : [0,1] \to X$; $\gamma(0)$ and $\gamma(1)$ are the endpoints of the path. We use $|\gamma|$ to denote the length of γ. The path γ is *simple* if and only if it maps $[0,1]$ to distinct points on X. A path is a ***geodesic*** if and only if it is locally a shortest path; i.e. it cannot be shortened by an infinitesimal perturbation. In particular, global shortest paths are geodesics. We use the term ***curve*** as an alternative for path. A path or a curve is polygonal if it is composed of a finite number of segments.

Mitchel *et al.* [MMP87] describe an algorithm to compute the shortest path distance from a single source to the whole surface in $O(n^2 \log n)$ time. The same algorithm can be adapted to compute the shortest path distance from an edge to the whole surface in the same running time. It follows that the shortest path from a set of $O(n)$ edges to the whole surface can be computed in $O(n^3 \log n)$.

The shortest path from a point in X to a set is a geodesic. So, it is a polygonal line that intersects every edge at most once at a point and passes through a face along a segment. The shortest path crossing an edge looks locally like a straight segment, if one rotates the adjacent faces so that they are coplanar. See [MMP87] for more details.

Let the source S be a set of edges of X and let π be a shortest path from a point p to S. We define the ***signature*** of π, to be the ordered set of edges (crossed or used) by π. Since π is locally a straight segment, we can rotate all faces that intersect π one by one so that π becomes a straight line. It follows that there cannot be two geodesics from p with the same signature.

A point p on the surface is a ***medial*** point if there are more than one shortest paths (with different signatures) from p to S.

A.3 Homotopy and Leash Function

Let γ and γ' be two paths with same endpoints s and t on a surface $\mathcal{D}$. A homotopy $h : [0,1]\times[0,1] \to D$ is a continuous function, such that $h(0,\cdot) = \gamma$, $h(1,\cdot) = \gamma'$, $h(\cdot,0) = s$ and $h(\cdot,1) = t$. So, for each $\tau \in [0,1]$, $h(\tau,\cdot)$ is an (s,t)-path. The ***height*** of such a homotopy is defined to be $\sup_{\tau\in[0,1]} |h(\tau,\cdot)|$.

Let A and B be two disjoint curves. A curve connecting a point in A to a point in B is called an (A,B)-***leash***. We define a (A,B)-***leash function*** to be a function f that sends every $\tau \in [0,1]$ to a leash with endpoints $a(\tau) \in A$ and $b(\tau) \in B$ such that $a : [0,1] \to A$ and $b : [0,1] \to B$ are reparametrizations of A and B, respectively. We say that a (A,B)-leash function f is *continuous* if the leash $f(\tau)$ varies continuously with τ. The *height* of a leash function f is defined to be $\sup_{\tau\in[0,1]} |f(\tau)|$. The *Frechét distance* between A and B is the height of the minimum height (A,B)-leash function. The *homotopic Frechét distance* between A and B is the height of the minimum height continuous (A,B)-leash function.

A.4 Discrete Problems

Let W_1 be an (s,t)-walk and f be a face in G. Assume that α_1 is a subwalk of W_1 and $\partial f = \alpha_1 \cup \alpha_2$, where α_1 and α_2 are walks that share endpoints u and v, such that u is closer to s on W_1. We define the ***face flip*** operation as follows. The walk $W_2 = W_1[s,u] \cdot \alpha_2 \cdot W_1[v,t]$ is the result of flipping W_1 over f. In this case, we say that W_1 and W_2 are one face flip operation apart.

Now, let W_1 be an (s,t)-walk and $e = (u,v)$ be and edge in G. Assume that $u \in W_1$. We obtain the walk $W_2 = W_1[s,u] \cdot (u,v) \cdot (v,u) \cdot W_1[u,t]$ after applying a ***spike*** operation on W_1 along e. In this case, we can obtain W_1 from W_2 by applying a ***reverse spike*** operation along e. We say that W_1 and W_2 are a spike operation apart. In general, we say that W_1 and W_2 are one operation apart if we can transform one to the other using a single face flip, spike, or reverse spike. Letscher and Chambers introduce the same set of operations with the names: face lengthening, face shortening, spike and reverse spike.

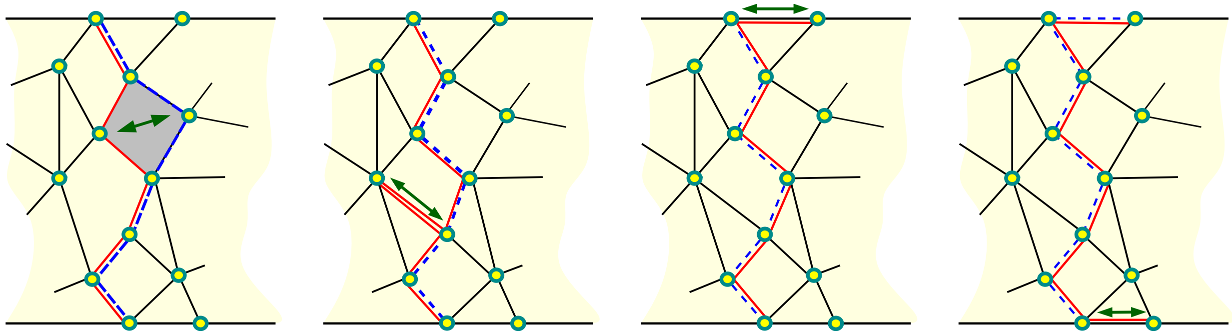


Figure 4: From left to right: face-flip, spike/reverse spike, man-move and dog-move.

Let L and R be two (s,t)-walks on the outer face of G. We define the sequence of walks $(L = W_0, W_1, \ldots, W_m = R)$ to be a (L,R)-***discrete homotopy*** if and only if for all $1 \le i \le m$, W_i and W_{i-1} are one operation apart. We may use the word homotopy as a short form of discrete homotopy when it is clear from context. A homotopy is ***monotonic*** (or equivalently it avoids backward moves) if W_{i-1} is inside the disc with boundary $L \cup W_i$ for every $1 \le i \le m$. The height of the homotopy is defined to be length of the longest walk in its sequence. The homotopy height between L and R, is the height of the shortest possible (L,R)-homotopy.

Let $A = (a_0, a_1, \ldots, a_k)$ and $B = (b_0, b_1, \ldots, b_l)$ be walks of G. The walk $W_1 = (a_i = w_1, w_2, \ldots, w_k = b_j)$ changes to the walk $W_2 = (a_{i+1}, a_i = w_1, w_2, \ldots, w_k)$ after a ***man move***. Similarly, the walk $W_1 = (a_i = w_1, w_2, \ldots, w_k = b_j)$ changes to the walk $W_2 = (w_1, w_2, \ldots, w_k = b_j, b_{j+1})$ after a ***dog move***. We say that the walk W_1 changes to the walk W_2 by a ***move*** if there is a dog move or a man move that changes W_1 to W_2. A ***leash operation*** is a move, a face flip, a spike or a reverse spike.

An (A,B)-walk is a walk that has one endpoint on A and one endpoint on B. A sequence of (A,B)-walks, $(W_1, W_2, \ldots, W_q)$ is called an (A,B)-***leash sequence*** if W_1 is a (a_0,b_0)-walk, W_q is a (a_k,b_l)-walk and for all $1 \le i < q$, W_i changes to W_{i+1} by a set of leash operations that contains at most one move. The height of a leash sequence is the length of its longest walk. The ***discrete Frechét distance*** of A and B is the height of the minimum height (A,B)-leash sequence. The leash sequence $(W_1, W_2, \ldots, W_q)$ contains no gap if W_i changes to W_{i+1} by exactly one leash operation. The ***homotopic discrete Frechét distance*** of A and B is the height of the minimum height (A,B)-leash sequence that contains no gap.