Computing the Fréchet Distance Between Folded Polygons*

Atlas F. Cook IV† Anne Driemel‡ Jessica Sherette§ Carola Wenk¶

March 30, 2012

UTSA CS Technical Report CS-TR-2012-007

Abstract

Computing the Fréchet distance for surfaces is a surprisingly hard problem and the only known algorithm is limited to computing it between flat surfaces. We study the problem of computing the Fréchet distance for a class of non-flat surfaces called folded polygons. We present a fixed-parameter tractable algorithm for this problem. Next, we present a polynomial-time approximation algorithm. Finally, we present a restricted class of folded polygons for which we can compute the Fréchet distance in polynomial time.

1 Introduction

The *Fréchet distance* is a similarity metric for continuous shapes such as curves and surfaces. In the case of computing it between two curves there is an intuitive explanation of the Fréchet distance. Suppose a man walks along one curve, a dog walks along the other, and they are connected by a leash. They can vary their relative speeds but cannot move backwards. The Fréchet distance of the curves is the minimum leash length required for the man and dog to walk along these curves. Although less intuitive, the idea is similar for surfaces. There is substantial work on how to compute or approximate this distance in various settings, see [AG00, AERW03, AKW04, Alt09, DHW10, HR11] and references therein.

While the Fréchet distance between polygonal curves can be computed in polynomial time [AG95], computing it between surfaces is much harder. In [God99] it was shown that even computing the Fréchet distance between a triangle and a self-intersecting surface is NP-hard. This result was extended in [BBS10] to show that computing the Fréchet distance between 2d terrains as well as between polygons with holes is also NP-hard. Furthermore, while in [AB10] it was shown to be upper semi-computable, it remains an open question whether the Fréchet distance between general surfaces is even computable.

On the other hand, in [BBW06] a polynomial time algorithm is given for computing the Fréchet distance between two (flat) simple polygons. This was the first paper to give any algorithm for computing the Fréchet distance for a nontrivial class of surfaces and remains the only known approach. Our contribution is to generalize their algorithm to a class of non-flat surfaces we refer to as folded polygons. Given that theirs is the only known approach it is of particular importance to explore extending it to new classes of surfaces. The major problem we encountered was that

*This work has been supported by the National Science Foundation grant NSF CAREER CCF-0643597. A preliminary version of this work has appeared as [CDH+11].

†Department of Information and Computing Sciences, University of Utrecht, Netherlands, atlas@cs.uu.nl
‡Department of Information and Computing Sciences, University of Utrecht, Netherlands, driemel@cs.uu.nl
§Department of Computer Science, University of Texas at San Antonio, USA, jsherett@cs.utsa.edu
¶Department of Computer Science, University of Texas at San Antonio, USA, carola@cs.utsa.edu

the mappings between the folded polygons which need to be considered are less restricted than those between simple polygons. We address three different methods to resolve this problem. In Section 4, we outline a fixed-parameter tractable algorithm. In Section 5, we describe a polynomial-time approximation algorithm to compute the Fréchet distance between folded polygons within a constant factor. In Section 6, we describe a nontrivial class of folded polygons for which the original algorithm presented in [BBW06] will compute an exact result.

2 Preliminaries

The Fréchet distance is defined for two k-dimensional hypersurfaces $P, Q : [0,1]^x \to \mathbb{R}^d$, where $x \leq d$, as

$$\delta_F(P,Q) = \inf_{\sigma: A \to B} \sup_{p \in A} \|P(p) - Q(\sigma(p))\|,$$

where σ ranges over orientation-preserving homeomorphisms that map each point $p \in P$ to an image point $q = \sigma(p) \in Q$. Lastly, $\|\cdot\|$ is the Euclidean norm but other metrics could be used instead.

Let $P, Q : [0,1]^2 \to \mathbb{R}^d$ be connected polygonal surfaces for each of which we have a convex subdivision. We assume that the dual graphs of the convex subdivisions are acyclic, which means that the subdivisions do not have any interior vertices. We will refer to surfaces of this type as *folded polygons*. We refer to the interior convex subdivision edges of P and Q as *diagonals* and *edges* respectively. Let m and n be the complexities of P and Q respectively. Let k and l be the number of diagonals and edges respectively. Assume without loss of generality the number of diagonals is smaller than the number of edges. Let $T_{matrixmult}(N)$ denote the time to multiply two $N \times N$ matrices. We abuse notation such that $a \in d$ denotes that a is a point on the line segment d.

3 Core Algorithm

In this section we develop the *core algorithm* which we extend in subsequent sections to compute the Fréchet distance between a more general class of surfaces. We begin by reviewing the simple polygons algorithm which ours extends, see Section 3.1. We then describe a possible extension of the simple polygons algorithm to one for folded polygons, see Section 3.2. Next, we identify a difficult problem with such an extension, see Section 3.3. Lastly, we prove that this extension can be computed in polynomial time, see Section 3.4.

3.1 Simple Polygons Algorithm Summary

Buchin *et al.* [BBW06] show that, while the Fréchet distance between a convex polygon and a simple polygon is the Fréchet distance of their boundaries, this is not the case for two simple polygons. They present an algorithm to compute the Fréchet distance between two simple polygons P, Q. The idea is to use a convex subdivision of P and map each of the convex regions of it continuously to distinct parts of Q such that taken together the images account for all of Q.

The decision problem $\delta_F(P,Q) \leq \varepsilon$ can be solved by (1) mapping ∂P onto ∂Q such that $\delta_F(\partial P, \partial Q) \leq \varepsilon$ and (2) mapping each diagonal d in the convex subdivision of P to a shortest path $f \subseteq Q$ such that both endpoints of f lie on ∂Q and such that $\delta_F(d, f) \leq \varepsilon$.

In order to solve subproblem (1) they use the notion of a free space diagram. For open polygonal curves $f, g : [0,1] \to \mathbb{R}^d$ with complexities m_f and n_g respectively the ***free space*** is defined as

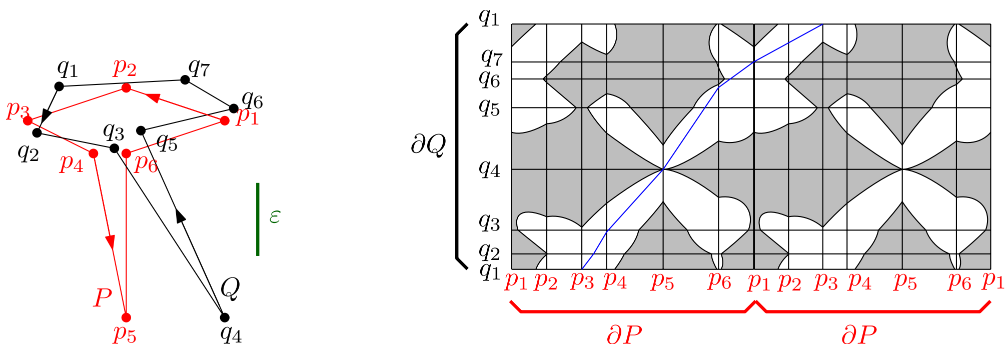

Figure 1: The white points in the free space diagram are those in the free space. The surfaces are within Fréchet distance ε since there is a monotone path in the white region starting at the bottom of the free space diagram and ending at the top which maps every point on the boundary of P exactly once. This figure was generated using an ipelet created by Günter Rote.

$FS_\varepsilon(f, g) = \{(x, y) \mid x, y \in [0, 1], ||f(x) - g(y)|| \leq \varepsilon\}$ where $\varepsilon \geq 0$. The free space diagram of f and g can be represented as a rectangle $[0, m_f] \times [0, n_g]$ consisting of m_f columns and n_g rows with a total of $m_f n_g$ cells. The bottom boundary of the free space diagram corresponds to the parameter space of f and the left boundary corresponds to the parameter space of g. A point in the free space diagram of f and g is white if it is in the free space of f and g, otherwise the point is black. There exists a monotone path through the white region of free space diagram starting at the bottom left corner of the free space diagram going to the top right if and only if the curves are within Fréchet distance ε. As shown in [AG95], this can be extended to closed curves by concatenating two copies of the free space diagram to create a ***double free space diagram*** and searching for a monotone path which covers every point in P exactly once, see Figure 1. This algorithm can be used to determine whether $\delta_F(\partial P, \partial Q) \leq \varepsilon$ and find the particular mapping(s) between ∂P and ∂Q. In turn such a mapping between the polygon boundaries defines a *placement* of the diagonals, i.e., a mapping of the endpoints of the diagonals in P to endpoints of the corresponding image curves in Q.

Subproblem (2) is solved by only considering paths through the free space diagram that map a diagonal d onto an image curve $f \subseteq Q$ such that $\delta_F(d, f) \leq \varepsilon$. Naturally, the particular placement of the diagonals determined in subproblem (1) will affect whether this is true. Therefore, they must check this for many paths in the free space diagram. Fortunately they can show that it is sufficient to only consider mapping a diagonal to an image curve which is the shortest path in Q between the end points determined by the placement.

Due to the continuous nature of the free space diagram there are an infinite number of distinct paths through it. Fortunately, as shown in [AG95], the information of a free space diagram can be encoded in a reachability graph. Such a graph can be computed using a divide and conquer algorithm. For the simple polygons algorithm this graph is stored as a adjacency matrix. Merging the information of two such graphs is accomplished by multiplying their adjacency matrices. Thus, their algorithm for computing the Fréchet distance of two simple polygons is dependent on the time complexity, $T_{matrixmult}(N)$, of multiplying two $N \times N$ matrices.

Solving these subproblems generates a mapping between P and Q for ε. This mapping might not be a homeomorphism but the authors show by making arbitrarily small perturbations of image curves the mapping can be made into one. The mapping between P and Q for ε is the limit of the homeomorphisms you get by perturbing. Thus, because the Fréchet distance is the infimum of all homeomorphisms on the surfaces, the Fréchet distance is ε. For simplicity we will refer to these generated mappings as homeomorphisms even though these may contain limits of homeomorphisms which are not homeomorphisms themselves. By performing a binary search on a set of critical values they can use the above algorithm for the decision problem to compute the Fréchet distance of P and Q in time $O(kT_{matrixmult}(mn)\log(mn))$.

3.2 Shortest Path Edge Sequences

As mentioned in Section 1 we extend an existing algorithm for computing the Fréchet distance between simple polygons to a more general set of surfaces. We refer to this set as *folded polygons* which we define as follows. Let $P : [0, 1]^2 \to \mathbb{R}^d$ be a connected polygonal surface for which we have a convex subdivision. P is a folded polygon if and only if the dual graph of the convex subdivision of it is acyclic. More intuitively, the set of folded polygons are all polygonal surfaces which do not have any interior vertices.

Note that the set of simple polygons is a proper subset of the set of folded polygons. Likewise, the set of folded polygons is a proper subset of the set of triangulated polygonal surfaces. Thus this is a set of surfaces between simple polygons (for which a polynomial time algorithm exists to compute the Fréchet distance) and triangulated polygonal surfaces (for which it is known to be NP-hard to compute the Fréchet distance [BBW06, BBS10]).

The idea of the algorithm is to subdivide one surface, P, into convex regions and pair those with corresponding regions in the other surface, Q. Those regions of Q are now folded polygons rather than just simple polygons. The authors of [BBW06] show that the Fréchet distance of a convex polygon and a simple polygon is just the Fréchet distance of their boundaries. Using essentially the same argument we prove the following lemma.

Lemma 3.1 *The Fréchet distance between a folded polygon P and a convex polygon Q is the same as that between their boundary curves.*

Proof: Assume that the Fréchet distance between the boundary curves is ε. Now consider mapping the diagonals in the convex decomposition of P to curves in Q. Because Q is convex, we can connect the images of the endpoints of each of the diagonals with a line segment. The line segments are with Fréchet distance ε to their respective diagonals since a mapping between them realizing Fréchet distance ε can be found using linear interpolation. These line segments subdivide Q into a set of convex polygons each of which are paired with a convex polygon in P. The boundaries of the paired polygons must be within Fréchet distance ε of each other, consequently the paired convex polygons must be in Fréchet distance ε.

Combining all of these mappings, we have a homeomorphism between P and Q. Thus the Fréchet distance of P and Q must also be $\leq \varepsilon$. ∎

For the simple polygon algorithm it suffices to map diagonals onto shortest paths in Q between two points on ∂Q. By contrast, there are folded polygons where a homeomorphism between the surfaces does not exist when diagonals are mapped to shortest paths but does exist when the paths are not restricted, see Figure 2. We must therefore consider mapping the diagonals to more general paths. Fortunately, we can show that these more general paths still have some nice properties for folded polygons.

Lemma 3.2 *Let u and v be points such that $u, v \in \partial Q$, $E = \{e_0, e_1, \ldots, e_s\}$ be a sequence of edges in the convex subdivision of Q, and d be a line segment. Given $\varepsilon > 0$, a curve f in Q that follows the edge sequence E from u to v such that $\delta_F(d, f) \leq \varepsilon$ can be computed, if such a curve exists, in $O(s)$ time.*

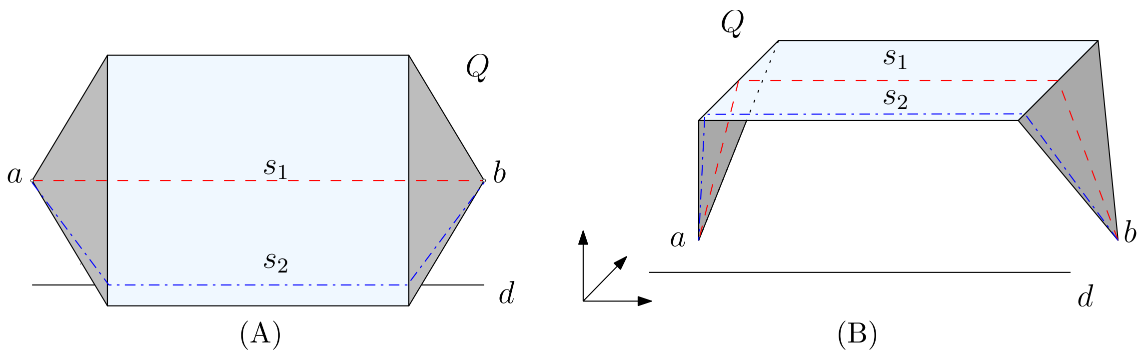

Figure 2: The curve s_1 is the shortest path in Q between the points a and b but the curve s_2 has smaller Fréchet distance to d than s_1 has. (A) Overhead view. (B) Sideview.

Proof: We construct a series $F = FS_\varepsilon(e_0, d), FS_\varepsilon(e_1, d), \ldots, FS_\varepsilon(e_s, d)$, of 2-dimensional free space diagrams. Any two edges e_i and e_{i+1} are on the boundary of the same convex polygon in the convex subdivision of Q so we can assume without loss of generality that f consists of straight line segments between the edge intersections. This is similar to the shortcutting argument used to prove Lemma 3 in [BBW06]. Thus, we only need to check the points where f crosses an edge of Q. For $\delta_F(d, f) \leq \varepsilon$ to be true, the preimages of those crossing points must be monotone along d. Let $FS'_\varepsilon(e_i, d)$ be the projection of $FS_\varepsilon(e_i, d)$ onto d. Let $F' = FS'_\varepsilon(e_1, d), FS'_\varepsilon(e_2, d), \ldots, FS'_\varepsilon(e_s, d)$.

To verify that the preimage points on d can be chosen such that they are monotone, we check the intervals of F'. Specifically, for $i < j$, the point on d mapped to e_i must come before the one mapped to e_j. This can be checked by greedily scanning left to right and always choosing the smallest point on d which can be mapped to some edge. A search of this form takes $O(s)$ time. ∎

The dual graph of the faces of Q is acyclic. This implies that there is a unique sequence of faces through which a shortest path from u to v, where $u, v \in \partial Q$, must pass. Necessarily, there must also be a unique edge sequence that the shortest path follows. We refer to this as the *shortest path edge sequence.*

Lemma 3.3 *Let Q be a folded polygon and d be a diagonal. If there is a curve $f \subseteq Q$ with $\delta_F(d, f) \leq \varepsilon$ then there is a curve $g \subseteq Q$ which follows the* shortest path edge sequence *such that $\delta_F(d, g) \leq \varepsilon$.*

Proof: Let E_f and E_g be the edge sequences of f and g respectively. By definition the dual graph of the faces of Q is acyclic, so E_g must be a subsequence of E_f. E_g induces a sequence of free space intervals. If there is a monotone path in the free space interval sequence induced by E_f, we can cut out some intervals and have a monotone path in the free space for E_g. ∎

From Lemma 3.3, we just need to consider paths that follow the shortest path edge sequence. We refer to paths that follow this edge sequence and consist of straight line segments between edges as *pseudo shortest paths* . In addition, s in Lemma 3.2 is bounded by the number of edges along the shortest path edge sequence between u and v. This implies the following theorem.

Theorem 3.4 *Let Q be a folded polygon, u and v be points such that $u, v \in \partial Q$, and d be a line segment. Given $\varepsilon > 0$, we can in $O(l)$ time find a curve f in Q from u to v such that $\delta_F(d, f) \leq \varepsilon$ if such a curve exists.*

Suppose we have a homeomorphism between ∂P and ∂Q. The endpoints of the image curves must appear on ∂Q in the same order as their respective diagonal endpoints on ∂P. The homeomorphism also induces a direction on the diagonals in P and on the edges in Q. Specifically, we consider diagonals and edges to start at their first endpoint along ∂P or ∂Q respectively in a counterclockwise traversal of the boundaries. We denote by D_e the set of diagonals whose associated shortest path edge sequences contain an edge $e \subseteq Q$. Observe that pairwise non-crossing image curves must intersect an edge e in the same order as their endpoints occur on ∂Q. We refer to this as the *proper intersection order* for an edge e.

3.3 Diagonal Monotonicity Test and Untangleability

We now define a test between two folded polygons P and Q which we call the *diagonal monotonicity test* . The diagonal monotonicity test serves as our core algorithm which we extend in later sections. Similar to the simple polygons decision algorithm, for a given ε the diagonal monotonicity test returns true if the following two conditions are true. Same as before, (1) $\delta_F(\partial P, \partial Q) \leq \varepsilon$. Given a fixed placement of the diagonals, (2) for every diagonal d_i in the convex subdivision of P, a corresponding pseudo shortest path f_i in Q has $\delta_F(d_i, f_i) \leq \varepsilon$. Condition (2) is the same as before except it uses pseudo shortest paths instead of the shortest paths.

Unfortunately, because the image curves of the diagonals are no longer shortest paths, they may cross each other and we will no longer be able to generate a mapping between the folded polygons which is a homeomorphism. Thus the diagonal monotonicity test might return true when in fact a homeomorphism does not exist. We must explicitly ensure that the image curves of all diagonals are non-crossing. We have identified cases where a pair of folded polygons pass the diagonal monotonicity test for ε but they are not within Fréchet distance ε, see Figure 3. The folded polygon P is divided into convex regions by the diagonals d_1 and d_2. These diagonals are mapped to image curves d_1' and d_2' in the folded polygon Q. P and Q pass the diagonal monotonicity test for $\varepsilon = 1$ but the image curves d_1' and d_2' are forced to cross. Such a case yields the following lemma.

Lemma 3.5 *There exist simple polygons P and Q such that P and Q pass the diagonal monotonicity test for ε but $\delta_F(P, Q) > \varepsilon$*

Proof: Let P and Q be the surfaces shown in Figure 3 (a) and (b) with $\varepsilon = 1$.

The surfaces are the same except rotated by 90 degrees. Each surface consists of two parallel 2 by 3 rectangles which are connected together by a smaller slanted rectangle. One of the large rectangles is at height 1 while the other is at height 0. Figure 3(c) shows how the two surfaces are overlaid.

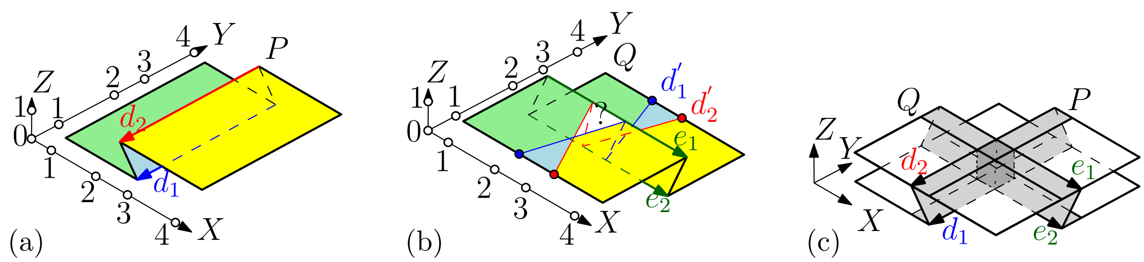

Figure 3: (a) the folded polygon P, (b) the folded polygon Q with image curves d_1' and d_2' mapped from diagonals d_1 and d_2 in P. Please note that the image curves are shown for illustrative purposes and are not with Fréchet distance 1 of their respective diagonals. (c) P and Q are overlaid as shown.

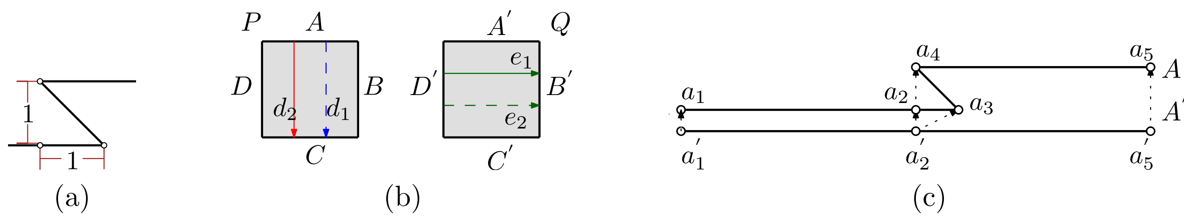

Figure 4: (a) Sideview of the surfaces, (b) Top view of P and Q showing which sides are mapped to which, (c) A side of P is mapped onto a side of Q. a_1 and $a_1^{'}$ as well as a_2 and $a_2^{'}$ occupy the same points in space and a separation is shown for clarity.

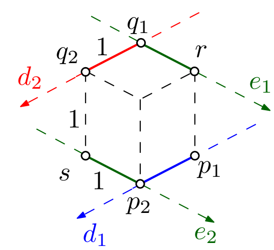

Figure 5: The cube has sides of length one. The marked points must be used in the mappings for $\varepsilon = 1$ but they can be shown to be mutually exclusive.

We can prove that the boundary curves have Fréchet distance 1 to each other. Pair up each side of P with one in Q as indicated in Figure 4(b). These paired portions of the boundary have Fréchet distance 1 to each other. The case of mapping A to $A^{'}$ is given in Figure 4(c). For the beginning and ending of the sides the curves either overlap or are exactly distance 1 apart with one directly above the other, so the mapping used is obvious. Matching the flat middle part on $A^{'}$ to the zigzag in A is a bit harder. The key insight is to observe that the point $a_2^{'}$ is within distance 1 to a_2, a_3, and a_4. The entire zigzag can be mapped to this point. The portion after the point $a_4^{'}$ can then be mapped straight down.

Consider the intersection between P and Q. At the center of the folded polygons is a cube with side length 1, see Figure 3(c) and Figure 5. The point p_1 on d_1 can map both to the point p_2 on edge e_2 and to the point r on edge e_1 within distance 1. Thus, there exists an image curve f_1 such that $\delta_F(d_1, f_1) \leq 1$. Likewise, the point q_2 on d_2 can map both to the point s on edge e_2 and to the point q_1 on edge e_1 within distance 1, so there exists an image curve f_2 such that $\delta_F(d_2, f_2) \leq 1$. Therefore, P and Q pass the diagonal monotonicity test for 1.

Notice that p_1 is the only point on d_1 which can map to e_1. Given the direction of d_1, f_1 starts at the top of Q and must cross edge e_2 then e_1. Thus no points after p_1 on d can mapped to e_2 so p_2 is the only reachable point on e_2. Likewise, s is the only point on e_2 the diagonal d_2 can map to. s comes before p_2 along e_2 so the image curve of d_2 must cross e_2 before the image curve of d_1. This is different from the proper intersection order that comes from the boundary. Therefore, even though P and Q passed the diagonal monotonicity test for 1, $\delta_F(P, Q) > 1$. ∎

We refer to a set of image curves $F = \{f_1 \ldots f_k\}$ as *untangled* if and only if the curves of F are pairwise non-crossing. We can add to the second condition of the diagonal monotonicity test an additional restriction that the set of image curves that the diagonals are mapped to is untangled. Together with condition (1) this forms what we refer to as the *enhanced diagonal monotonicity test*. We can prove the following theorem.

Lemma 3.6 *$\delta_F(P,Q) \le \varepsilon$ if and only if P and Q pass the enhanced diagonal monotonicity test for ε.*

Proof: ($\Rightarrow$): If $\delta_F(P,Q) \le \varepsilon$ then there is a mapping between ∂P and ∂Q realizing this distance. Likewise there is mapping of each diagonal to an image curve in Q. By Lemma 3.3 these image curves can be assumed to be pseudo shortest paths. In addition the image curves must be pairwise non-crossing since the mapping realizing this distance is a homeomorphism. Thus, by definition, P and Q pass the enhanced diagonal monotonicity test for ε.

($\Leftarrow$): If P and Q pass the enhanced diagonal monotonicity test for ε, then the image curves of P must subdivide Q into a set of folded polygons. Since the diagonals form a convex decomposition of P, each of these folded polygons is uniquely paired with a convex polygon in P. For any such pair the Fréchet distance between their boundary curves must be less than or equal to ε. By Lemma 3.1, the Fréchet distance of such a pair must then be less than or equal to ε. The paired regions are unique and cover P and Q. Thus, $\delta_F(P,Q) \le \varepsilon$. ∎

Unfortunately, unlike the diagonal monotonicity test, it is not clear how to compute the enhanced diagonal monotonicity test in polynomial time. In Section 4, we examine more closely the problem of deciding if an untangled set of image curves exists. In the following section, we consider optimizing the ε for which two folded polygons pass the diagonal monotonicity test.

3.4 Critical Values and Optimization

In this section we show that we can optimize the ε for which a pair of folded polygons pass the diagonal monotonicity test in polynomial time.

Let P and Q be folded polygons. As noted in Section Section 3.2, a pseudo shortest path between points $u, v \in \partial Q$ is defined by a sequence of intersection points, $f = \{a_0, \dots, a_s\}$, on the edges along a shortest path edge sequence $E = \{e_0, e_1, \dots, e_s\}$ in P. Let d be a diagonal of P the endpoints of which are mapped to u and v. Likewise, let $F = FS_\varepsilon(e_0,d), FS_\varepsilon(e_1,d), \dots, FS_\varepsilon(e_s,d)$ as described in Section 3.2. Let $FS'_\varepsilon(e_i,d)$ be the projection of $FS_\varepsilon(e_i,d)$ onto d. Let $F' = FS'_\varepsilon(e_1,d), FS'_\varepsilon(e_2,d), \dots, FS'_\varepsilon(e_s,d)$.

There are two types of critical values which we consider. First is the minimum distance, ε, between a diagonal, d, and an edge, e. This is the minimum ε where $FS'_\varepsilon(e,d)$ is nonempty. More formally, this is the minimum distance ε such that there exists points $a \in d$ and $b \in e$ with $\varepsilon = \|a-b\|$. There are $O(kl)$ such critical values.

Second, we consider the critical values associated with finding a monotone path through a pair of edges. This is the minimum ε such that, for edges e_i and e_j in Q and diagonal d in P, there exists points $a \in d$, $b_i \in e_i$, and $b_j \in e_j$, where $\varepsilon \ge \|a - b_i\|$ and $\varepsilon \ge \|a - b_j\|$. This is the minimum ε such that $FS'_\varepsilon(e_i,d) \cap FS'_\varepsilon(e_j,d)$ is nonempty. There are $O(kl^2)$ such critical values.

The minimum value where d is within Fréchet distance to some pseudo shortest path is one of the two above types of critical values. A diagonal d can be mapped to some pseudo shortest path if a monotone sequence of points $a = (a_1, \dots, a_s)$ along d can be found where $a_i \in FS'_\varepsilon(e_i,d)$. By Lemma 3.2, if d cannot be mapped to a pseudo shortest path it must be the case that for some i and j, the $FS'_\varepsilon(e_i,d)$ is empty or $FS'_\varepsilon(e_i,d) \cap FS'_\varepsilon(e_j,d)$ is empty. These are exactly the critical values we test for.

Since they are of constant complexity, an instance of either of these critical values can be computed in constant time. Therefore, all of the critical values can be computed in $O(kl^2)$ time.

As shown in Theorem 3.4, computing pseudo shortest paths instead of shortest paths does not increase the asymptotic run time. To optimize this ε we can perform a binary search on the $O(kl^2)$

set of critical values. So, by following the paradigm set forth by [BBW06], we arrive at the following theorem:

Theorem 3.7 *The minimum ε for which two folded polygons P and Q, pass the diagonal monotonicity test can be computed in time $O(kT_{matrixmult}(mn)\log(kl))$.*

4 Fixed-Parameter Tractable Algorithm

In Section 3.3 we introduced the enhanced diagonal monotonicity test and proved that two folded polygons P and Q pass the enhanced diagonal monotonicity test for ε if and only if $\delta_F(P,Q) \le \varepsilon$. In this section we outline an algorithm to decide for a fixed placement of diagonals whether the set of diagonals can be mapped within distance ε to an untangled set of image curves. From this we create a fixed-parameter tractable algorithm for computing the Fréchet distance between a pair of folded polygons.

4.1 Untangleability Space

We begin by first considering untangling image curves on a single edge. Let P and Q be folded polygons. Assume we are given a placement of the diagonals of P in Q. Let e be an edge in Q which is crossed by the image curves of h diagonals, $d_1, \ldots, d_h$. We assume without loss of generality that the image curves of the diagonals cross e in proper intersection order if, for all $1 \le i, j \le h$ where $i < j$, the image curve of d_i crosses e before the image curve of d_j crosses e. Let the ***untangleability space*** $U_\varepsilon(e)$ contain all h-tuples of points on the diagonals which can be mapped to crossing points on the edge e within distance ε and such that the crossing points are in the proper intersection order along e. We can define $U_\varepsilon(e)$ more formally as follows. As before, $a_i \in d_i$ denotes that a_i is a point on the line segment d_i. For points $p_i, p_j \in e$ let $p_i \le_e p_j$ denote that the point p_i is no further than the point p_j along the directed edge e. Let $D_e = d_1 \times d_2 \times \cdots \times d_h$.

$$\begin{aligned} U_\varepsilon(e) &= \{(a_1, \ldots, a_h) \in D_e \mid \\ &\exists\ p_1, \ldots, p_h \in e\ \forall\, 1 \le i \le j \le h \ :\ \|a_i - p_i\| \le \varepsilon \wedge p_i \le_e p_j\}. \end{aligned}$$

$U_\varepsilon(e)$ can be shown to be convex with simple arguments based on linear interpolation yielding the following lemma.

Lemma 4.1 *$U_\varepsilon(e)$ is convex.*

Proof: If $U_\varepsilon(e)$ is not convex then there are points $a = (a_1, \ldots, a_h), c = (c_1, \ldots, c_h) \in U_\varepsilon(e)$ such that there exists a point $b = (b_1, \ldots, b_h)$ on the line segment between a and c where $b \notin U_\varepsilon(e)$, see Figure 6 (a). The point b is on the line segment between a and c so there exists an x between 0 and 1 such that $b = xa + (1-x)c$. Likewise for each point $b_i \in b$, $b_i = xa_i + (1-x)c_i$. Let p_i and r_i be the points on edge e mapped to by points a_i and c_i respectively. Let $q_i = xp_i + (1-x)r_i$ be the image of b_i on e. Thus, b_i and its image q_i are defined by linearly interpolating between the original points and their images on e.

Since $a \in U_\varepsilon(e)$, $\|a_i - p_i\| \le \varepsilon$. Since $c \in U_\varepsilon(e)$. $\|c_i - r_i\| \le \varepsilon$. We can rewrite these as follows:

$$\|xa_i - xp_i\| \le x\varepsilon \quad (1)$$

$$\|(1-x)c_i - (1-x)r_i\| \le (1-x)\varepsilon. \quad (2)$$

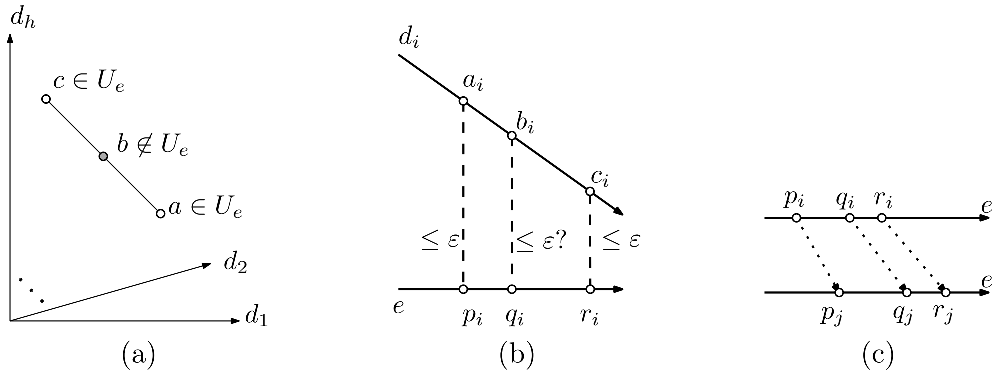

Figure 6: (a) $a, c \in U_\varepsilon(e)$ but $b \notin U_\varepsilon(e)$. (b) both a_i and c_i being within distance ε of their respective points on e necessitates that b_i is within distance ε of its own. (c) The crossing points of diagonals d_i and d_j being in order along e for both points a and c necessitates that they are also in order for b

By adding Eq. (1) and Eq. (2) we prove that $\|b_i - q_i\| \leq \varepsilon$, see Figure 6 (b). Details are given below.

$$\begin{aligned} &\|(1-x)c_i - (1-x)r_i\| + \|xa_i - xp_i\| \leq (1-x)\varepsilon + x\varepsilon \\ &\implies \|(xa_i + (1-x)c_i) - (xp_i + (1-x)r_i)\| \leq \varepsilon - x\varepsilon + x\varepsilon \\ &\implies \|b_i - q_i\| \leq \varepsilon. \end{aligned} \tag{3}$$

Thus each point b_i can be mapped to q_i within distance ε. Next, using a similar argument, we show for $1 \leq i \leq j \leq h$ that $q_i \leq_e q_j$, see Figure 6 (c). Since $a \in U_\varepsilon(e)$ and $c \in U_\varepsilon(e)$, $p_i \leq_e p_j$ and $r_i \leq_e r_j$. From this we get the following.

$$q_i = xp_i + (1-x)r_i \leq_e q_j = xp_j + (1-x)r_j \tag{4}$$

Thus, for all i, j, where $0 \leq i \leq j \leq h$, Equations Eq. (3) and Eq. (4) hold. From this it follows by definition that $b \in U_\varepsilon(e)$. Therefore $U_\varepsilon(e)$ is convex. ■

4.2 Fixed-Parameter Tractable Algorithm

Let P and Q be folded polygons. Let k and l be the complexities of the convex subdivisions of P and Q respectively. Assume we have a placement of the diagonals in Q and some fixed $\varepsilon > 0$. we can decide whether for the diagonals can be mapped within distance ε to an untangled set of image curves using the untangleability spaces of the edges. We first choose an edge, e_r, in Q to act as the root of the *edge tree*, T, that encodes the adjacency of edges of Q. Specifically, the children of an edge $e \in T$ are those which share a face of P with e and are not the parent or sibling of e, see Figure 7 (a). We propagate constraints imposed by each untangleability space up the tree to the root node to determine if, for the given placement, the diagonals can be mapped within distance ε to an untangled set of image curves.

Note that the definition of untangleability space can be extended to include diagonals which do not cross an edge e in $U_\varepsilon(e)$. Such a diagonal poses no additional restrictions on $U_\varepsilon(e)$ and this is similar to lifting $U_\varepsilon(e)$ in this extra dimension representing the diagonal. Thus we can assume for all edges that their untangleability space is k dimensional.

The untangleability space of an edge e, $U_\varepsilon(e)$, by definition contains exactly those sets of points on the diagonals in D_e which can be mapped to the edge in the proper intersection order. The point

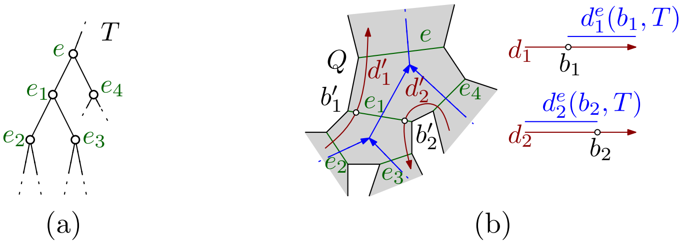

Figure 7: (a) the adjacency of the edges in Q is represented as a tree T, (b) the direction an image curve crosses an edge determines the restriction imposed on the diagonal.

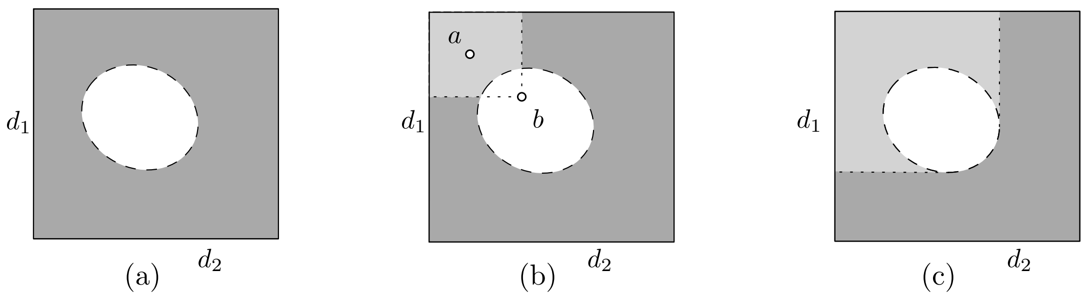

Figure 8: (a) $U_\varepsilon(e)$ is shown in white, (b) a point a is contained in $C(U_\varepsilon(e))$ because there exists a point $b \in U_\varepsilon(e)$ which is monotonically further back along each of the diagonals, (c) $C(U_\varepsilon(e))$ is the union of the white and light gray portions.

chosen in $U_\varepsilon(e)$ imposes a constraint on what points may be chosen in other untangleability spaces. In particular, the corresponding points on all of the diagonals must be monotone with respect to their edge sequence. We define $C(U_\varepsilon(e))$ to capture these constraints along the diagonals. The direction of this constraint depends on the direction the image curve of a diagonal is going when it crosses the edge. Let $d = \overline{st}$ be a diagonal, a be a point on d, and e be an edge whose parent in T is f. We define a function $d^e(a, T)$ to return the monotone portion of d relative to the point a, see Figure 7 (b). Let $d^e(a, T) = \overline{sa}$ if the direction of d is consistent with the propagation direction of e to f. Otherwise, let $d^e(a, T) = \overline{at}$. We define $C(U_\varepsilon(e))$ as follows, see Figure 8.

$$C(U_\varepsilon(e)) = \{(a_1, \ldots, a_h) \in D \mid \exists (b_1, \ldots, b_h) \in U_\varepsilon(e) \; \forall\, 1 \le i \le h \; : \; a_i \in d_i^e(b_i, T)\}$$

We recursively define for every edge e a k-dimensional *propagation space* $P_\varepsilon(e)$. If e is a leaf in the tree, then $P_\varepsilon(e) = U_\varepsilon(e)$. Otherwise, define

$$P_\varepsilon(e) = \{a \in D \mid a \in U_\varepsilon(e) \wedge a \in C(P_\varepsilon(e_1)) \wedge \ldots \wedge a \in C(P_\varepsilon(e_j))\}$$

where e is the parent of the edges $e_1, e_2, \ldots, e_j$ in T. The idea of the algorithm is that by computing $P_\varepsilon(e_r)$ we can account for the constraints encoded in all of the untangleability spaces of all of edges in Q. In particular we prove that $P_\varepsilon(e_r)$ is empty when, for a particular placement of the diagonals of P, condition (2) of the enhanced diagonal monotonicity test is passed for ε.

Theorem 4.2 *Let P and Q be folded polygons. Assume we are given a placement of the diagonals of P in Q. Let T be the edge tree of Q rooted at an edge e_r.*

Condition (2) of the enhanced diagonal monotonicity test is passed for ε if and only if $P_\varepsilon(e_r) \neq \emptyset$

Proof: ($\Leftarrow$): If $P_\varepsilon(e_r) \neq \emptyset$, then one can trace back through the propagation diagrams starting at $P_\varepsilon(e_r)$ and going down to the leaves to get a point $a_e \in P_\varepsilon(e)$ for each edge e. These points together serve as a witness for the mapping.

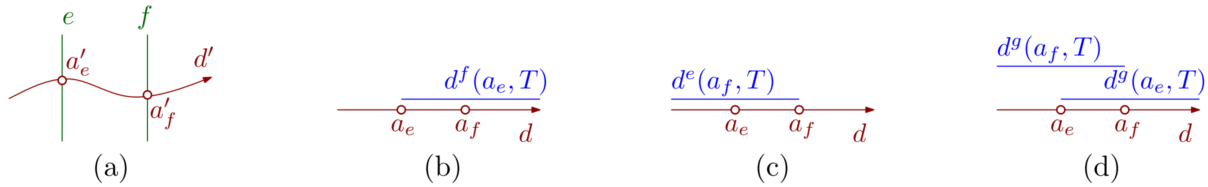

Figure 9: (a) let d' cross edges e and f, (b) a_f is contained in $d^f(a_e, T)$, (c) a_e is contained in $d^e(a_f, T)$, (d) the intersection of $d^g(a_e, T)$ and $d^g(a_f, T)$ is nonempty.

By the definition of propagation space, for any edge e, $P_\varepsilon(e) \subseteq U_\varepsilon(e)$. Thus, a point $a = (a_1, \ldots, a_h) \in U_\varepsilon(e)$ represents a set of points on the diagonals $d_1, \ldots, d_h$ such that they can be mapped in the proper intersection order along the edge e. This is half of the requirements of condition (2). Now, we must show that each diagonal is within Fréchet distance ε of its image.

Let d be a diagonal in P which is mapped to some image curve d' in Q by the witness. Let e and f be two adjacent edges in Q that are crossed by d' at points a'_e and a'_f respectively. Let a_e and a_f be the preimage points of a'_e and a'_f on d, see Figure 9 (a). Assume without loss of generality that d' crosses e before f. It must then be the case that a_e occurs before a_f along d. There are two distinct cases to consider.

Suppose f is the parent of e in T. In this case $d^f(a_e, T)$ must contain a_f, see Figure 9 (b). If it did not then by definition of propagation space $P_\varepsilon(f) = \emptyset$ and, thus, $P_\varepsilon(e_r) = \emptyset$. The case when e is the parent of f in T is similar, see Figure 9 (c).

Next, suppose e and f are siblings of a parent edge g in T. The intersection of $d^g(a_e, T)$ and $d^g(a_f, T)$ must be nonempty, see Figure 9 (b). While the image curve d' does not cross the edge g, a point in $P_\varepsilon(g)$ must have a valid coordinate in the d dimension. If the intersection is empty no such coordinate exists.

Thus, the preimages of each of the intersection points of d' must occur monotonically along d and $\delta_F(d, d') \leq \varepsilon$. Thus, for the given placement of the diagonals, condition (2) of the enhanced diagonal monotonicity test is passed for ε.

($\Rightarrow$): Suppose condition (2) of the enhanced diagonal monotonicity test is passed for ε. By definition, there exists a mapping of the diagonals to image curves which cross each edge in the proper intersection order. For a particular edge e there is a set of preimage points, $a = (a_1, \ldots, a_h)$, on the diagonals $d_1, \ldots, d_h$ such that those points on the diagonals were mapped in the *proper intersection order* along the edge e by the given mapping. Therefore, for each edge e, $a = (a_1, \ldots, a_h) \in U_\varepsilon(e)$. For simplicity assume for each edge e that $a = (a_1, \ldots, a_h)$ is the only point in $U_\varepsilon(e)$. We show that $P_\varepsilon(e_r) \neq \emptyset$ even under this more restrictive condition..

We prove that for every edge e its propagation space $P_\varepsilon(e) \neq \emptyset$ using an inductive argument. Clearly in the base case, where e is a leaf of T, $P_\varepsilon(e) = U_\varepsilon(e)$ and is thus nonempty. Next consider the inductive step. By definition $P_\varepsilon(e) = U_\varepsilon(e) \cap C(P_\varepsilon(e_1)) \cap \ldots \cap C(P_\varepsilon(e_j))$. By our inductive hypothesis $P_\varepsilon(e_i)$ is nonempty for $1 \leq i \leq j$. Thus $P_\varepsilon(e_i)$ also contains the single point in $U_\varepsilon(e_i)$. There are two cases that could cause $P_\varepsilon(e)$ to be empty.

There may exist an edge f which is a child of e in T where a constraint of f prevents using the point in $U_\varepsilon(e)$. Let d be a diagonal whose image curve, d', crosses e at a'_e and f at a'_f. Let a_e and a_f be the preimages of a'_e and a'_f respectively on d used by the original mapping. For the constraints of the diagonal d to cause $P_\varepsilon(e)$ to be empty it must be the case that $d^e(a_f, T)$ does not contain a_e, but, by assumption, a_e and a_f are the preimages of a'_e and a'_f in the given mapping between d and d' as well so they must occur in order along d. Thus $d^e(a_f, T)$ must contain a_e.

There may exist edges f and g which are children of e in T where a constraint between f and g prevents using the point in $U_\varepsilon(e)$. Let d be a diagonal whose image curve, d', crosses f at a'_f

and g at a'_g. Let a_f and a_g be the preimages of a'_f and a'_g respectively on d used by the original mapping. For the constraints of the diagonal d to cause $P_\varepsilon(e)$ to be empty it must be the case that the intersection of $d^e(a_f, T)$ and $d^e(a_g, T)$ is empty, but, by assumption, a_e and a_f are the preimages of a'_e and a'_f in the given mapping between d and d' as well so they must occur in order along d. Thus the intersection of $d^e(a_f, T)$ and $d^e(a_g, T)$ cannot be empty.

Therefore, we conclude that $P_\varepsilon(e) \neq \emptyset$ for all e in T and that $P_\varepsilon(e_r)$ is nonempty. ∎

We next give a fixed-parameter tractable algorithm to compute the Fréchet distance of P and Q. Specifically, we assume k and l are constant, where k and l are the complexities of the convex subdivisions of P and Q respectively.

Theorem 4.2 applies for a fixed mapping between ∂P and ∂Q and a fixed ε. Of course, to compute the Fréchet distance of P and Q we want to optimize ε. To do this we use the global optimization algorithm for semi-algebraic sets described by Basu *et al.* [BPR06, Section 14.4]. This requires $P_\varepsilon(e_r)$ to be described by a quantifier free logic formula. We can generate a formula for $P_\varepsilon(e_r)$ by recursively expanding the definitions for the propagation space. The recursion down the tree terminates when it reaches the leaves. The definition of the untangleability space for an edge e contains an existential quantifier. These quantifiers can be removed using quantifier elimination as described by [BPR06] in Section 14.2. The time to remove the existential quantifiers and then minimize ε using the global optimization algorithm depends on k and l which are assumed to be constants. Specifically, let $F(k, l)$ be the time complexity of computing this.

Next, consider two different mappings between ∂P and ∂Q. These determine different placements of the diagonals. If all of the image curves of all of the diagonals have the same shortest path edge sequence in both of the mappings the global optimization of ε will yield the same result. The l edges of Q divide ∂Q into $2l$ regions. Mappings of the diagonals are distinct only if the diagonals are mapped to a different sequence of regions on ∂Q. Since there are k diagonals it follows that there are at most $O((2l)^{2k})$ distinct mappings. For each distinct placement of the diagonals we compute the Fréchet distance of ∂P and ∂Q with restrictions on the diagonal placement as well as the minimum ε such that $P_\varepsilon(e_r) \neq \emptyset$. Because we are computing the Fréchet distance of P and Q we take the maximum of these two values. This gives us a table of size $O((2l)^{2k})$ values. The minimum of these values, ε, will be the Fréchet distance of P and Q. By Lemma 3.6, if $\delta_F(\partial P, \partial Q) \leq \varepsilon$ and the diagonals can be mapped within distance ε to image curves which are untangled then $\delta_F(P, Q) \leq \varepsilon$. Likewise, no other mapping can achieve a smaller ε since either $\delta_F(\partial P, \partial Q) > \varepsilon$ or the diagonals cannot be mapped within distance ε to image curves which are untangled. This yields the following Theorem.

Theorem 4.3 *We can compute the Fréchet distance of two folded polygons in time* $O((2l)^{2k} \cdot (F(k, l) + kT_{matrixmult}(mn) \log(kl)))$.

5 Constant Factor Approximation Algorithm

In this section we present an approximation algorithm which avoids the problem of finding an untangled set of image curves altogether. In particular we show that, if a pair of folded polygons pass the diagonal monotonicity test for ε then they also pass the enhanced diagonal monotonicity test for 9ε. We begin with some simple definitions.

Let P and Q be two folded polygons which are oriented counterclockwise. Let p_i and p_j be two points on ∂P. Let $\partial P_{p_i,p_j}$ denote the portion of ∂P between the points starting at p_i and ending at p_j. We refer to a diagonal $d = \overline{p_i p_j}$ as an *ear diagonal* if $\partial P_{p_i,p_j}$ contains no other endpoints of diagonals in P. Likewise, we refer to the polygon of P bounded by d and $\partial P_{p_i,p_j}$ as an *ear* of P, see Figure 10.

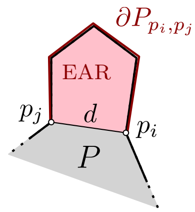

Figure 10: $d = \overline{p_i p_j}$ is an ear diagonalof P.

Consider the diagonals of P and their image curves in Q found by performing the diagonal monotonicity test. In order to pass the diagonal monotonicity test for ε there must be a homeomorphism between ∂P and ∂Q. Therefore, the endpoints of the diagonals will be mapped in the proper order along the boundary of Q. Let d be an ear diagonalof P. Let d' be its corresponding image curve in Q. It is possible, that other image curves cross d', prohibiting a homeomorphism between P and Q. Consider the arrangement of d' together with the image curves in Q which cross d' and let $a_0, \ldots, a_k$ denote the arcs of the image curves that form the top level of this arrangement, ordered along their appearance, see Figure Figure 11.

Lemma 5.1 *The Fréchet distance between d and the concatenation of the arcs $\alpha = a_0, \ldots, a_k$ is at most 9ε.*

Proof: First, we argue that for any arc a_i there exists a subsegment of d, such that the Fréchet distance between a_i and this subsegment is less than 3ε. This can be seen as follows. Let e' be the supporting image curve of a_i. Consider the two intersections of e' with d' just before and just after a_i, let them be denoted with p and q, see Figure 11. Let $\overline{e'}$ denote the segment of the image curve between the points p and q. Let p_d (respectively q_d) denote the preimage of p (respectively q) on d and let p_e (respectively q_e) be the preimage on the diagonal e. The image curve d' is in Fréchet distance ε to d and the same holds true for e' and e. By the triangle inequality, it follows that $\|p_d - p_e\| \leq 2\varepsilon$ and $\|q_d - q_e\| \leq 2\varepsilon$. Since $\overline{p_d q_d}$ and $\overline{p_d q_d}$ are straight line segments, again by the triangle inequality we have that

$$\delta_F(\overline{e'}, \overline{p_d q_d}) \leq \delta_F(\overline{e'}, \overline{p_e q_e}) + \delta_F(\overline{p_e q_e}, \overline{p_d q_d}) \leq 3\varepsilon.$$

It follows that any such arc a_i can be matched to some subsegment of d within Fréchet distance 3ε. Let ψ_i denote a mapping from points on a_i to this subsegment, that realizes this distance.

Now, consider the free space diagram for 3ε between α and d. Inside a column, corresponding to a single arc a_i, there exists a monotone path which is contained in the white area that stretches across the domain of a_i, but its endpoints do not necessarily connect to the paths in the neighboring columns, see Figure 12. Furthermore, it can happen, that two such paths overlap in the parameter of d. Let a_i and a_{i+1} be two consecutive arcs. Let p' be their intersection and let $p_i = \psi_i(p)$ and $p_{i+1} = \psi_{i+1}(p)$. Since $\overline{p_i p_{i+1}}$ is a straight line segment and p' is in distance 3ε to both p_i and p_{i+1}, it holds that $\delta_F(p', \overline{p_i p_{i+1}}) \leq 3\varepsilon$. Therefore, we can connect the endpoints of consecutive paths in the free space diagram by vertical line segments and we retrieve a connected path π through the white area of the free space for 6ε. However, π is not necessarily monotone across columns.

We claim that the right envelope π' of π is contained in the white area of the free space diagram for 9ε, which would imply the claim, that there exists such a path that is monotone, see Figure 12.

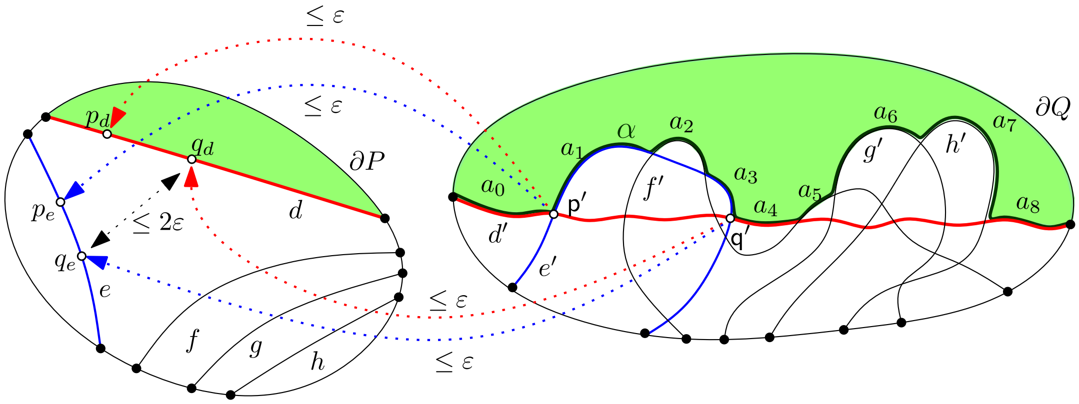

Figure 11: We remap the ear diagonal$d$ from d' to $\alpha = a_0, \dots, a_k$.

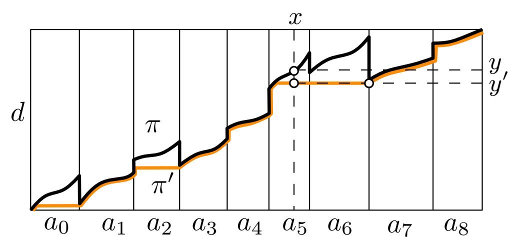

Figure 12: Consider the free space diagram of d and α.

Consider a vertical line at x through the parametric space, and its intersection points (x, y) and (x, y') with π and π'. Let p' be the point on α that corresponds to the parameter x. Consider the rightmost point on π' that intersects the horizontal line at y'. Let the corresponding point on α be q'. Let $p = d(y)$ and let $q = d(y')$. We denote with $\overline{f'}$ and $\overline{h'}$ the portions of the supporting image curves between the first intersection point with d' before p (respectively q) and after, see Figure 13. By the earlier analysis, there exist mappings for both $\overline{f'}$ and $\overline{h'}$ to subsegments of d that match to points within distance 3ε. Let $\psi_{f'}$ and $\psi_{h'}$ denote these mappings. By construction, we have that $p = \psi_{f'}(p')$ and $q = \psi_{h'}(q')$. We argue that $\|p - q\| \leq 6\varepsilon$. Clearly, if $y = y'$, this holds true. Otherwise, we have that $y > y'$, which means that q appears before p on d. Since π' is the right envelope of π, there can be no other case.

Let a' be the first intersection of $\overline{h'}$ with d' before q' and let $a = \psi_{h'}(a')$. By the monotonicity of the mapping, it holds that a appears before q on d. Similarly, we have that $\psi_{f'}$ matches the first intersection point between $\overline{f'}$ and d' after p' to a point which appears after p on d. Therefore, this point must appear after a' on d'. Let b' denote this intersection point. Since p' appears before q' on α, it must be that the curve segment of $\overline{f'}$ between p' and b' and the curve segment of $\overline{h'}$ between a' and q' intersect. Let s' be one such intersection point. The preimage $s_h = \psi_{h'}(s')$ appears after p and its preimage $s_f = \psi_{f'}(s')$ appears before q. We have, by the triangle inequality, that $\|s_f - s_h\| \leq 6\varepsilon$, since both s_f and s_h are in distance 3ε to s'. Since p and q are contained in $\overline{s_f s_h}$, we have that $\|p - q\| \leq 6\varepsilon$, as claimed. Therefore, we have that $\|p' - q\| \leq 9\varepsilon$, by the triangle inequality. Since this holds for any point p' on α, with respect to the point matched by π', the claim is implied. ■

Suppose some folded polygons P and Q pass the diagonal monotonicity test for ε. Let d be an ear diagonalof P and P_d be the ear of P associated with d. Since d is a diagonal of the convex decomposition of P, P_d is a convex polygon. By Lemma 5.1 an ear diagonal$d$ in P can be remapped

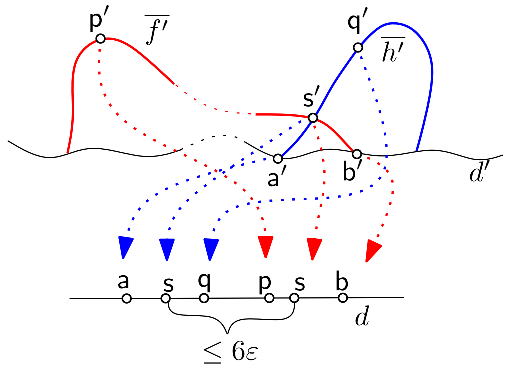

Figure 13: Using a point s where $\overline{f'}$ and $\overline{h'}$ intersect we show that the free space diagram contains a monotone path for 9ε

to an image curve d' in Q which does not cross any other image curves where $\delta_F(d, d') \leq 9\varepsilon$. This yields a folded polygon $Q_{d'} \subset Q$ between d' and ∂Q. By Lemma 3.1, $\delta_F(P_d, Q_{d'}) \leq 9\varepsilon$.

Remove P_d from P and $Q_{d'}$ from Q and then choose another ear in P. We can repeat the above arguments to remove this new ear and its corresponding folded polygon in Q. The dual graph of P is a tree. Each time we repeat this argument we are removing a leaf from the tree. Eventually, the tree will contain only a single node which corresponds to some convex polygon of P which we map to the remainder of Q.

The above argument and Theorem 3.7 yield the following theorem.

Theorem 5.2 *We can compute a 9-approximation of the Fréchet distance of two folded polygons in time $O(kT_{matrixmult}(mn)\log(mn))$.*

6 Axis-Parallel Folds and L_∞ Distance

In this section we outline a special class of surfaces for which using the L_∞ metric allows us to avoid the problem of finding an untangled set of image curves. Specifically, if all of the line segments in the convex subdivision of the surfaces are parallel to the x-axis, y-axis, or z-axis, we show that it is sufficient to use shortest paths instead of pseudo shortest paths. Since shortest paths never cross we can use the simple polygons algorithm [BBW06] to compute the Fréchet distance of the surfaces. We first prove the following lemma.

Lemma 6.1 *Let R be a half-space such that the plane bounding it, ∂R, is parallel to the xy-plane, yz-plane, or xz-plane. Given a folded polygon Q with edges parallel to the x-axis, y-axis, or z-axis and points $a, b \in Q \cap R$, let f be a path in Q, which follows the shortest path edge sequence between a and b. If f is completely inside of R so is the shortest path f' between a and b.*

Proof: For the lemma to be false there must exist a Q, R, and f which serve as a counter example. There must be at least one edge e_j in Q such that $f \cap e_j \in R$ and $f' \cap e_j \notin R$. In particular, let e_j be the first edge where this occurs along the shortest path edge sequence. First consider a Q where all of the edges of it are perpendicular to ∂R. A line segment in the shortest path f' connects the endpoints of two edges in Q. Let e_i and e_k be the edges that define the line segment in f' that passes through e_j. We now consider several cases in how those edges are positioned.

Case (I) occurs when e_k is completely outside of R, see Figure 14 (a). While this does force f' to cross e_j outside of R, there is no f which can pass through e_k while remaining inside of R. Because

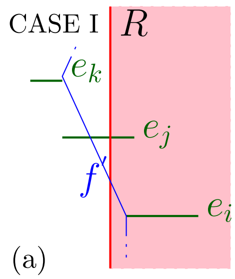

Figure 14: Figures (a) and (b) are examples of case (I) and case (II), respectively.

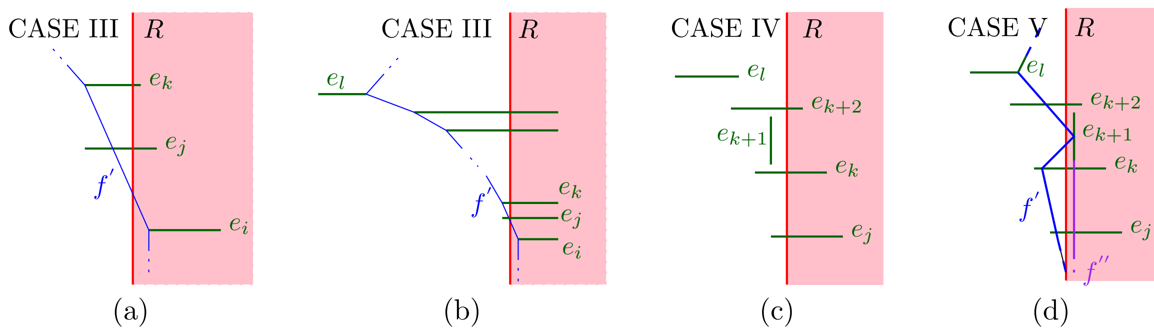

Figure 15: (a) the path can continue to curve away from R. (b) in this case, while the shortest path goes out of R eventually there must be an edge e_l where it reverses direction since f' ends at b which is in R. Such an edge corresponds to case (I). (c) the edge e_{k+1} forces f' out of R on e_j, but for this to happen e_{k+1} must be completely outside of R, thus no f can cross e_{k+1}. (d) f' is not forced out of R on e_j but rather can follow the path f''.

Q is a folded polygon any path between a and b must path through the edges in the shortest path edge sequence including e_k. Thus no f can exist entirely within R.

Case (II) occurs when part of e_k is in R and f' crosses it in the part in R, see Figure 14 (b). In this case f' does not cross e_j outside of R.

Case (III) occurs when part of e_k is in R and f' crosses it in the part outside of R, see Figure 15 (a). The path f can cross the part of e_k inside R. This case can be repeated many times but eventually, f' ends at b so it needs to cross back into R. To do this it must cross an edge on the opposite end point. If that end point is in R the entire path can be shortcut similar to case (II) and if the point is not in R then as in case (I) there does not exist an f which can cross e_l and remain in R see Figure 15 (b). Thus case (III) also leads to a contradiction.

Next we consider the two cases that arise from adding a perpendicular fold e_{k+1}. Specifically, we will consider where the first perpendicular fold after e_j is placed. Such a fold runs parallel with ∂R so it must be either entirely inside R or entirely outside.

Case (IV) occurs if the edge e_{k+1} is outside of R. In this case there does not exist an f which can pass through e_{k+1} and be entirely in R, see Figure 15 (c).

Case (V) occurs if the fold e_{k+1} is inside of R. In this case the shortest path f' will no longer be forced outside of R on the edge e_{j+1}. Instead, it goes through R to the edge e_{k+1}, see Figure 15 (d). The only way that f' could not pass through R would be if one of the edges between e_j and e_k were entirely outside R and that is ruled out by case (I).

Thus no such Q, R, and f can exist. ∎

Theorem 6.2 *The Fréchet distance between two surfaces, P and Q, both with only diagonals and edges parallel to the x-axis, y-axis, or z-axis, can be computed in time $O(kT_{matrixmult}(mn)\log(mn))$.*

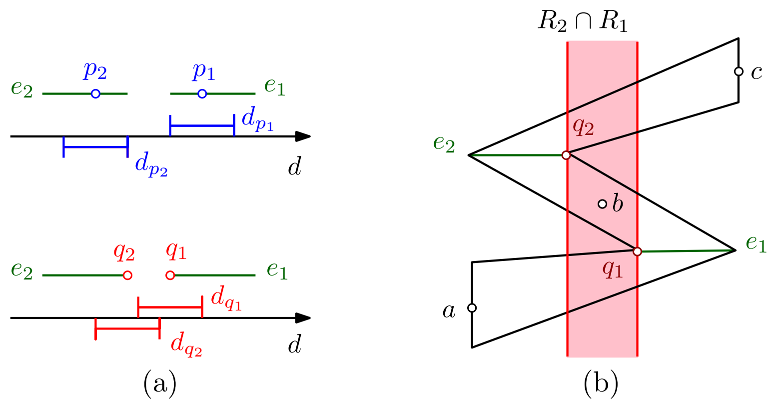

Figure 16: (a) is an example of intervals for the two different paths.(b) together the edges e_1 and e_2 cause a monotonicity constraint.

Proof: Let d be a diagonal and $f^{'}$ be the shortest path between points a and c on ∂Q. Using Lemma 6.1 we prove that if there exists a pseudo shortest path f between points a and c such that $\delta_F(d, f) \leq \varepsilon$, then $\delta_F(d, f^{'}) \leq \varepsilon$.

Minkowski Sum Constraints. Since we are using the L_∞ distance, the unit ball is a cube. The Minkowski sum of a diagonal d in P and a cube of side length ε yields a box. Points in the diagonal d can only map to points in this region. It can be defined by the intersection of 6 half-spaces; all of these have boundaries parallel to either the xy-axis, the xz-axis, or the yz-axis. Thus, from Lemma 6.1, we know that if any path through Q is completely within this box, then the shortest path $f^{'}$ will be, too. This means that for each edge e_i on the shortest path edge sequence $f^{'} \cap e_i$ is within distance ε of some non-empty interval of d.

Monotonicity Constraints. For the shortest path $f^{'}$ between the boundary points to have $\delta_F(d, f^{'}) > \varepsilon$, at least two of these intervals must be disjoint and occur out of order along d, see Figure 16 (a). Such a case introduces a monotonicity constraint on ε. If no such intervals existed then we could choose a monotone sequence of points along d such that each point is within distance ε of an edge and the sequence of edges they map to would have the same order as the shortest path edge sequence showing that $\delta_F(d, f^{'}) \leq \varepsilon$.

Let e_1 and e_2 be two edges along the shortest path edge sequence for which such bad intervals occur. Let p_1 and p_2 be points on the shortest path where it intersect edges e_1 and e_2 respectively. Let q_1 and q_2 be the same for f. Finally, let d_r denote all of the points on d which are within distance ε of the point r. Since $\delta_F(d, f) \leq \varepsilon$, d_{q_1} and d_{q_2} must overlap or occur in order along d.

Let R_1 be the half-space whose bounding plane contains q_1 and is perpendicular to d. likewise let R_2 be the half-space whose bounding plane contains q_2 and is perpendicular to d, see Figure 16(b). Let R_1 extend to the left along d and R_2 extend to the right along d. ∂R_2 must occur before ∂R_1 along d or the edges are in order and no monotonicity constraint is imposed. Assume R_1 encloses all of $f^{'}$ between a and e_2. If it does not we can choose a new edge between a and e_2 to use as e_1 for which this is true. Doing so only increases the monotonicity constraint. Likewise we can assume R_2 encloses all of $f^{'}$ between e_2 and c.

Assume a, b, and c lie on $f^{'}$. Specifically, let a and c be the end points of $f^{'}$ on ∂Q. Naturally, a shortest path must exist between a and c and it must contain at least one point in $R_1 \cap R_2$ which we call b. f follows the shortest path edge sequence between a and c, so it must also cross all of the edges in the shortest path edge sequence between a and b. Therefore, to show that p_2 is inside of R_1 we can directly apply Lemma 6.1 to the points a and b. A similar method can be used for e_2 with points b and c to show p_1 is inside R_2. Since d_{q_1} and d_{q_2} overlap or are in order, d_{p_1} and d_{p_2} must as well. Therefore, $\delta_F(d, f') \leq \varepsilon$ and shortest paths can be used for this variant of folded polygons instead of pseudo shortest paths. Because we are using shortest paths we can just use the simple polygons algorithm. This yields Theorem 6.2. ■

7 Conclusions and Possible Extensions

We adapt the simple polygon algorithm to create one for computing the Fréchet distance for a class of non-flat surfaces which we call folded polygons [BBW06]. Unfortunately, the original algorithm cannot be extended directly. We present three different methods to adapt it. The first of which is a fixed-parameter tractable algorithm. The second is a polynomial-time approximation algorithm. Finally, we present a restricted class of folded polygons for which we can compute the Fréchet distance in polynomial time.

The constant factor approximation outlined in Section 5 can likely still be improved. Specifically, we consider only the worst case for each of the out-of-order mappings which may not be geometrically possible to realize. In addition, we currently approximate the Fréchet distance by mapping image curves one-by-one to the top of the arrangement of other image curves. It would of course be more efficient to untangle image curves by mapping them to some middle curve rather than forcing one to map completely above the others.

Finally, while the problem of finding a set of untangled image curves seems hard, it is also possible that a polynomial-time exact algorithm could exist. The acyclic nature of our surfaces seems to limit the complexity of our mappings. The methods used to prove that computing the Fréchet distance between certain classes of surfaces is NP-hard in [BBS10] are not easy to apply to folded polygons.

8 Acknowledgements

The authors wish to thank Sariel Har-Peled for helpful discussions and suggestions. In particular, for his help in developing the constant factor approximation algorithm described in Section 4.

References

[AB10] H. Alt and M. Buchin. Can we compute the similarity between surfaces? *Discrete Comput. Geom.*, 43(1):78–99, 2010.

[AERW03] H. Alt, A. Efrat, G. Rote, and C. Wenk. Matching planar maps. *J. Algorithms*, 49:262–283, 2003.

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

[AG00] H. Alt and L. J. Guibas. Discrete geometric shapes: Matching, interpolation, and approximation. In Jörg-Rüdiger Sack and Jorge Urrutia, editors, *Handbook of Computational Geometry*, pages 121–153. Elsevier, 2000.

[AKW04] H. Alt, C. Knauer, and C. Wenk. Comparison of distance measures for planar curves. *Algorithmica*, 38(1):45–58, 2004.

[Alt09] H. Alt. The computational geometry of comparing shapes. In *Efficient Algorithms: Essays Dedicated to Kurt Mehlhorn on the Occasion of His 60th Birthday*, pages 235–248. Springer-Verlag, 2009.

[BBS10] K. Buchin, M. Buchin, and A. Schulz. Fréchet distance of surfaces: Some simple hard cases. In *Proc. 18th Annu. European Sympos. Algorithms*, pages 63–74, 2010.

[BBW06] K. Buchin, M. Buchin, and C. Wenk. Computing the Fréchet distance between simple polygons in polynomial time. In *Proc. 22nd Annu. ACM Sympos. Comput. Geom.*, pages 80–87, 2006.

[BPR06] S. Basu, R. Pollack, and M. F. Roy. *Algorithms in Real Algebraic Geometry*. Algorithms and Computation in Mathematics. Springer, 2006.

[CDH+11] A. F. Cook, A. Driemel, S. Har-Peled, J. Sherette, and C. Wenk. Computing the Fréchet distance between folded polygons. In *Proc. 12th Workshop Algorithms Data Struct.*, pages 267–278, 2011.

[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.

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

[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/`.