Computing the Fréchet Distance between Piecewise Smooth Curves

Günter Rote[*]

November 24, 2004

Abstract

We consider the Fréchet distance between two curves which are given as a sequence of $m + n$ curved pieces. If these pieces are sufficiently well-behaved, we can compute the Fréchet distance in $O(mn \log(mn))$ time. The decision version of the problem can be solved in $O(mn)$ time. The results are based on an analysis of the possible intersection patterns between circles and arcs of bounded curvature.

1 Introduction

The Fréchet distance is a distance measure between curves.

Definition 1 (Fréchet distance).
Let $f\colon I = [l_I, r_I] \to \mathbb{R}^2$ and $g\colon J = [l_J, r_J] \to \mathbb{R}^2$ be two planar curves, and let $\|\cdot\|$ denote the Euclidean norm. Then the *Fréchet distance* $\delta_F(f, g)$ is defined as

$$\delta_F(f,g) := \inf_{\substack{\alpha\colon [0,1]\to I \\ \beta\colon [0,1]\to J}} \max_{t\in[0,1]} \|f(\alpha(t)) - g(\beta(t))\|.$$

where α and β range over continuous and non-decreasing reparameterizations with $\alpha(0) = l_I$, $\alpha(1) = r_I$, $\beta(0) = l_J$, $\beta(1) = r_J$.

In other words, we are looking for a common parameterization of f and g such that the maximum distance at any time t is as small as possible. In contrast to other common distance measures like the Hausdorff distance, the Fréchet distance respects the one-dimensional structure of the curves and doesn't just treat them as a point set.

[*]Freie Universität Berlin, Institut für Informatik, Takustraße 9, 14195 Berlin, Germany. rote@inf.fu-berlin.de.
Partially supported by the IST Programme of the EU as a Shared-cost RTD (FET Open) Project under Contract No IST-2000-26473 (ECG - Effective Computational Geometry for Curves and Surfaces).

The study of the Fréchet distance from a computational point of view has been initiated by Alt and Godau [2]. The *decision problem* is the problem to decide, for a given ε, whether the Fréchet distance between two curves is at most ε. The *optimization problem* is the problem to find the optimal ε, i. e., to *compute* the Fréchet distance.

Alt and Godau [2] treated the case of two polygonal curves. For two curves of m and n pieces, respectively, they showed how to solve the decision problem in $O(mn)$ time and the optimization problem in $O(mn\log(mn))$ time. Some related problems have also been considered, like minimizing the Fréchet distance under translations [3], or a generalized Fréchet distance between a curve and a *graph* [1]. In all cases, however, the objects are piecewise linear.

In this paper, we explore the Fréchet distance between more general curves. We assume that each input curve is given as a sequence of smooth curve pieces that are "sufficiently well-behaved", such as circular arcs, parabolic arcs, or some class of spline curves. (We will be more precise later.) Our algorithm will perform certain operations on these curves, like intersecting them with a circle, or constructing offset curves.

We will show that the *combinatorial complexity*, i. e., the number of steps, for solving the decision problem is not larger than for polygonal paths, $O(mn)$. The complexity of the individual operations (the *algebraic complexity*) depends of course on the nature of the curves. Under the stronger assumption that the curves consist of algebraic pieces whose degree is bounded by a constant, we can solve the optimization problem in $O(mn\log(mn))$ time, thus matching the running time for the polygonal case. The elementary operations, however, are algebraic operations of higher degree.

2 Assumptions

We assume that each curve is given as a sequence of pieces which are connected at their endpoints. Every piece is a smooth curve of class C^2, i. e., the curvature is defined everywhere and varies continuously within a piece. Usually a piece will be given in some convenient way, as an instance of some particular class of curves (for example, circular arcs) with certain parameters, by an equation, as a cubic spline, or in explicit parametric form. However, we will not make any assumptions about how the curves are given; it is only important that the necessary geometric operations can be carried out. Only in Section 8, for solving the optimization problem, we will assume that each piece is a smooth piece of an algebraic curve whose degree is bounded by a fixed constant. (We refer to this as the case of algebraic curves of "bounded degree".)

We can only work with curve pieces that do not "turn too much". Let f be a continuously differentiable curve. The direction of all tangent vectors of the curve sweep out a connected interval of the unit circle. If this angular range does not cover the whole circle, its length is called the *turning angle* of f. If f is convex, then the turning angle is the difference of directions between the initial and the final tangent direction, with appropriate sign.

Lemmas 2 and 3 below, which are crucial for our algorithm, hold for curves whose turning angle is bounded by π. Curves of larger bounding angle must be subdivided, for example by cutting them at every point of vertical tangency. If the pieces are algebraic of bounded degree, then each piece is cut into $O(1)$ subpieces. Thus, from now on we assume that the turning angle of each piece is at most π. Some further cuts will be necessary in Section 4 when we solve the decision problem.

3 Preliminaries

3.1 The free space diagram

The main tool of the algorithm is the *free space diagram* which was introduced in [2]. It is a two-dimensional representation of all pairs of points on the two curves, together with the identification of those pairs which are closer than ε.

Definition 2. Let $f\colon I \to \mathbb{R}^2$, $g\colon J \to \mathbb{R}^2$ be two curves, $I, J \subseteq \mathbb{R}$. The set

$$F_\varepsilon(f,g) := \{\, (s,t) \in I \times J : \|f(s) - g(t)\| \le \varepsilon \,\}$$

denotes the *free space* of f and g. the partition of $I \times J$ into the free space and its complement is called the *free space diagram*.

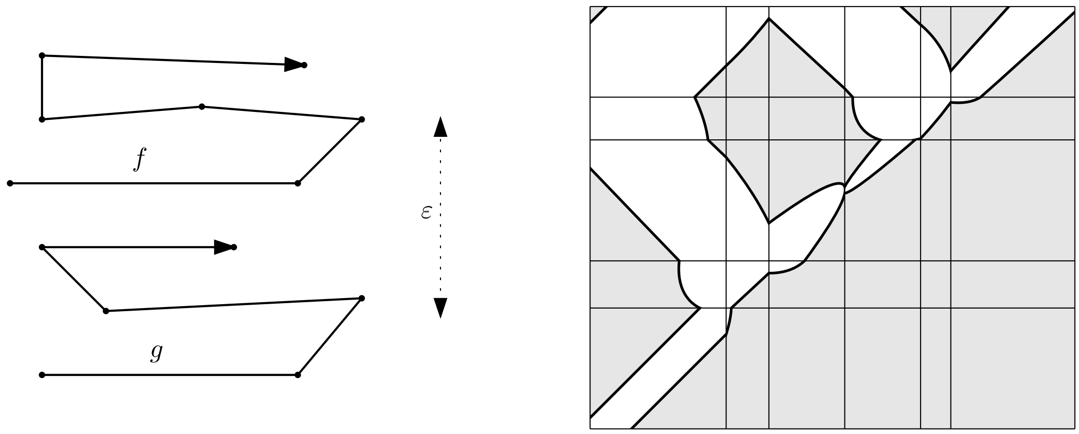

Figure 1: Two polygonal curves and their free space diagram. The scale of the free space diagram is reduced by 50% with respect to the curves.

Points in F_ε are called *feasible* or *free*, and they are usually drawn in white. The other points are called *forbidden points*. See Figure 1 for an illustration. We also view the regions of forbidden points as *obstacles*. The obstacle boundaries are formed by parameter values (s,t) for which $\|f(s) - g(t)\| = \varepsilon$. The following simple observation from [2] is crucial.

Lemma 1. *Let $f\colon I = [l_I, r_I] \to \mathbb{R}^2$, $g\colon J = [l_J, r_J] \to \mathbb{R}^2$ be two curves. Then $\delta_F(f,g) \leq \varepsilon$ if and only if there exists a curve within $F_\varepsilon(f,g)$ from (l_I, l_J) to (r_I, r_J) which is monotone in both coordinates.* $\square$

As f and g consist of several pieces, the free space diagram decomposes naturally into a grid of rectangular *cells*. There is one cell for each pair of pieces from f and g.

3.2 Auxiliary results

The following elementary lemma connects minimum and maximum curvature with some global properties of a curve [6], see Figure 2 for an illustration.

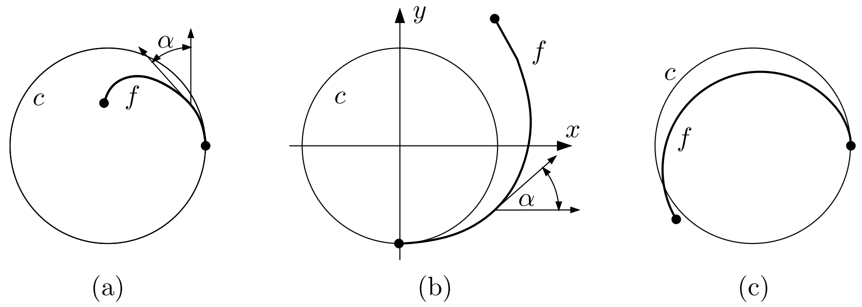

Figure 2: (a,b) Illustration of Lemma 2. (c) The bound of π on the turning angle in Lemma 2a is necessary.

Lemma 2. *Let f be a twice differentiable curve of turning angle at most π. Let c be a circle of radius r which is tangent to f at one of its endpoints.*

(a) *If f is convex and the curvature of f is at least $1/r$ everywhere, and f and c bend in the same direction at their initial point of tangency, the curve cannot enter the exterior of c.*

(b) *If the curvature of f is at most $1/r$ everywhere, the curve cannot enter the interior of c.*

Proof. (a) Let us assume without loss of generality that c is the unit circle ($r = 1$) and f starts at the rightmost point $(1, 0)$ of c in the upward direction. When we choose the arc length s as a parameter for f, the unit tangent vector of f can be written in the form

$$\dot{f}(s) = \begin{pmatrix} -\sin\alpha(s) \\ \cos\alpha(s) \end{pmatrix},$$

where the function α increases monotonically from $\alpha(0) = 0$ to a maximum $\alpha_{\max}$ of at most π. The curvature is given by the derivative of α, and by assumption we have $\dot{\alpha}(s) \geq 1$. We may choose α as a parameter, and we get

$$f(s) = f(0) + \int \begin{pmatrix} -\sin\alpha(s) \\ \cos\alpha(s) \end{pmatrix} ds = f(0) + \int_{\alpha=0}^{\alpha_{\max}} \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix} \frac{ds}{d\alpha}\, d\alpha,$$

Suppose that we want to find a point $f(s)$ on the curve which is extreme in the direction β, $0 \le \beta \le 2\pi$, i. e., we want to maximize $\binom{\cos\beta}{\sin\beta} \cdot f(s)$. This can be written as follows

$$\begin{aligned}
\begin{pmatrix}\cos\beta\\ \sin\beta\end{pmatrix} \cdot f(s) &= \begin{pmatrix}\cos\beta\\ \sin\beta\end{pmatrix} \cdot \left(f(0) + \int_{\alpha=0}^{\alpha_{\max}} \begin{pmatrix}-\sin\alpha\\ \cos\alpha\end{pmatrix} \frac{ds}{d\alpha}\, d\alpha \right)\\
&= \cos\beta + \int_{\alpha=0}^{\alpha_{\max}} \sin(\beta-\alpha) \frac{ds}{d\alpha}\, d\alpha
\end{aligned}$$

We try to maximize this expression under the constraints $0 \le ds/d\alpha \le 1$ and $0 \le \alpha_{\max} \le \pi$. The maximum is obtained if we choose $ds/d\alpha = 1$ whenever $\sin(\beta-\alpha) \ge 0$ and $ds/d\alpha = 0$ otherwise. That is, for $0 \le \beta \le \pi$, we choose $ds/d\alpha = 1$ for $0 \le \alpha \le \beta$ and obtain

$$\begin{pmatrix}\cos\beta\\ \sin\beta\end{pmatrix} \cdot f(s) \le \cos\beta + \int_{\alpha=0}^{\beta} \sin(\beta-\alpha)\, d\alpha = 1.$$

For $\pi \le \beta \le 2\pi$, we choose $ds/d\alpha = 1$ for $\beta - \pi \le \alpha \le \pi$ and obtain again

$$\begin{pmatrix}\cos\beta\\ \sin\beta\end{pmatrix} \cdot f(s) \le \cos\beta + \int_{\alpha=\beta-\pi}^{\pi} \sin(\beta-\alpha)\, d\alpha = 1.$$

Thus, for any direction β we have $\binom{\cos\beta}{\sin\beta} \cdot f(s) \le 1$. The intersection of the half-planes $\binom{\cos\beta}{\sin\beta} \cdot x \le 1$ is the unit circle c. This means that f cannot leave c.

(b) Let us assume without loss of generality that c is the unit circle and f starts at the lowest point $(0,-1)$ of c towards the right. Again, we choose the arc length s as a parameter. We denote the angle of the tangent by $\alpha(s)$, and as above, we obtain

$$f(s) = \begin{pmatrix}x(s)\\ y(s)\end{pmatrix} \quad\text{and}\quad \dot f(s) = \begin{pmatrix}\dot x(s)\\ \dot y(s)\end{pmatrix} = \begin{pmatrix}\cos\alpha(s)\\ \sin\alpha(s)\end{pmatrix}.$$

By assumption the derivative $\dot\alpha := \dot\alpha(s) := d\alpha/ds$ of α is bounded: $|\dot\alpha| \le 1$.

We first prove the lemma under the restriction that $0 \le \alpha(s) \le \pi$. This means that the curve cannot move down: $\dot y = \cos\alpha \ge 0$, and we have $y(s) \ge -1$ throughout. We have

$$\dot y = \sin\alpha \ge \sin\alpha \cdot \dot\alpha.$$

Integration gives

$$y \ge -\cos\alpha. \tag{1}$$

If y lies in the range $-1 < y < 1$, this can be written as $\arccos(-y) \ge \alpha$. The cotangent function $\cot u$ is continuous and decreasing in the interval $0 < u < \pi$, and therefore we obtain $\cot \arccos(-y) \le \cot\alpha$, or

$$\frac{-y}{\sqrt{1-y^2}} \le \frac{\cos\alpha}{\sin\alpha}.$$

Assuming $0 < \alpha < \pi$ and $-1 < y < 1$, we obtain

$$\dot{x} = \cos\alpha \geq \frac{-y \cdot \sin\alpha}{\sqrt{1-y^2}} = \frac{-y \cdot \dot{y}}{\sqrt{1-y^2}}. \tag{2}$$

This inequality also holds without the explicit restriction $0 < \alpha < \pi$, because $\alpha = \pi$ is excluded by the condition $y < 1$ and (1), and the case $\alpha = 0$ can be checked directly. The function $y(s)$ has the property that there exists some threshold $A \geq 0$ such that $y(s) = -1$ (and consequently $\alpha(s) = 0$ and $\dot{x} = 1$) for $0 \leq s \leq A$, and $y(s) > -1$ for $s > A$. Thus, by integrating (2), we obtain, for $s \geq A$:

$$x(s) \geq \sqrt{1 - y(s)^2} + A,$$

as long as $-1 < y < 1$. It follows that f cannot enter the circle $x^2 + y^2 < 1$.

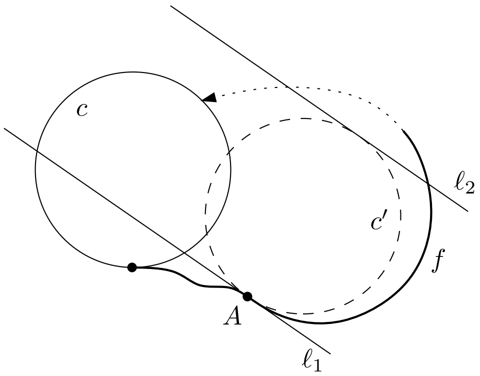

Figure 3: The general case of Lemma 2b.

Now we relax the initial assumption that $0 \leq \alpha(s) \leq \pi$. Let us assume that f terminates as soon as it intersects c and is about to enter the interior of c. Thus, the initial point (possible followed by a piece where f runs along c) and the endpoint of f are the only intersection points with c. Let A be a point on f where $\alpha(s)$ achieves its minimum value, see Figure 3. Let c' be the unit circle touching f at A from the left. Let ℓ_1 be the tangent at A, and let ℓ_2 be other tangent of c which is parallel to ℓ_1. By the definition of A, the starting point of f must lie on the side of ℓ_1 opposite to c', and hence c and c' lie on the same side of ℓ_2. We can apply the above proof to the piece of f after A and conclude that f cannot enter c'; thus, in order to reach c again, it must wind around c' and intersect ℓ_2 twice. This is not possible without creating a turning angle larger than π. $\square$

Without the bound of π on the turning angle, Lemma 2 would not hold, see for example Figure 2c for the case of part (a).

For solving the decision problem with parameter ε, we will cut the curve into pieces at all points where the curvature is $1/\varepsilon$. The following lemma, whose proof

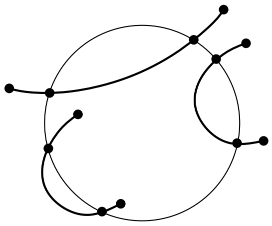

Figure 4: Illustration of Lemma 3. An arc of small curvature and two arcs of large curvature.

is based on Lemma 2, is then crucial for getting a grip on the complexity of the free space diagram, see Figure 4.

Lemma 3. *Let f be a smooth curve of turning angle at most π, and let c by a circle of radius r.*

(a) *If the curvature of f is at most $1/r$ everywhere, the curve can intersect c at most twice. If it intersects c twice, then its endpoints lie outside c or on the boundary, and the middle piece between the two intersections lies inside c.*

(b) *If f is convex and the curvature of f is at least $1/r$ everywhere, the curve can intersect c at most twice.*

Proof. If f is tangent to c at some point, the statement follows immediately from Lemma 2, applied to each piece starting from the tangent point in either direction. Thus, we can assume that each intersection point is a proper crossing where f crosses from inside to outside c or vice versa.

(a) For an intersection point A, let us look at the piece f' of f that continues outside c. It follows from Lemma 2b that f' cannot enter the interior of any of the two circles c' and c'' tangent to f at A, see Figure 5. Therefore, f cannot meet c again without making a turning angle larger than π. The statement follows.

(b) Consider three consecutive intersection points A, B, C along the curve. By reversing the orientation of f if necessary, we can assume that f crosses from outside c to inside at A, to outside at B, and back to inside at C, see Figure 6. By Lemma 2a, f is confined within some circle c' which is equal to c and tangent to f at A. Then it is easy to see that the tangent direction of f must sweep an angle larger than π between A and C, a contradiction. ☐

4 Preprocessing the input

In addition to the bound of π on the turning angle, for solving the decision problem with parameter ε, we need that in each piece of the curve, the curvature is either uniformly bigger than $1/\varepsilon$ or uniformly smaller than $1/\varepsilon$. To ensure this,

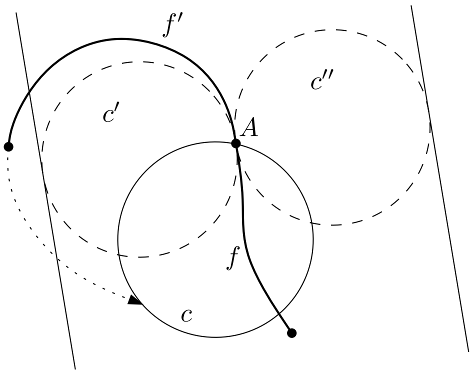

Figure 5: Proof of Lemma 3a

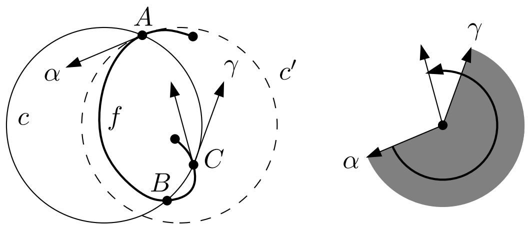

Figure 6: Proof of Lemma 3b

we subdivide the curve at all points where the curvature is $1/\varepsilon$. For convenience, we assume at this point that the curve does not contain a circular arc of radius ε. This is a technical assumption for simplifying the discussion. We will discuss later in Section 7.4 how to treat circular arcs as well.

If the pieces are algebraic curves of bounded degree, then the number of subpieces of each piece is bounded by a constant. From now on till the end of Section 7, m and n will denote the total number of resulting pieces of the two curves.

5 Critical points

We regard as *critical points* on the boundary of F_ε those points which are local extrema in the horizontal or vertical direction. A horizontal or vertical segment as part of the boundary of F_ε is excluded by assumption: it would mean that all points along some piece of f have distance ε from some fixed point of g (or vice versa). Thus f or g would contain a circular arc of radius ε. There are eight classes of critical points, shown in Figure 7.

In terms of the curves f and g, these points correspond to situations where a circle c of radius ε around a point of one curve is tangent to the other curve, see Figure 8. For example, a critical point of type W^+ occurs in the following situation: Let a point x move forward on f. The parts of g which lie within a

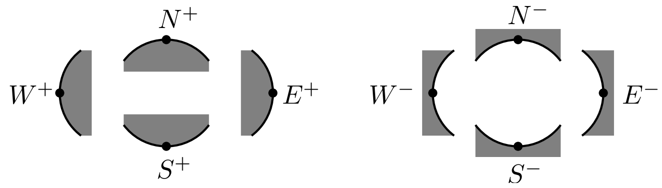

Figure 7: The eight types of critical points. N, S, E, W refers to the direction in which the point is extreme, and the superscript tells whether the area in this direction is feasible $(+)$ or forbidden $(-)$.

radius ε of x are the free points on the vertical line in the free space diagram. As this line sweeps forward to the right, an interval of forbidden points appears at the critical point W^+. This means that g touches c from inside. As x proceeds further away from g, a portion of g begins to stick out from c.

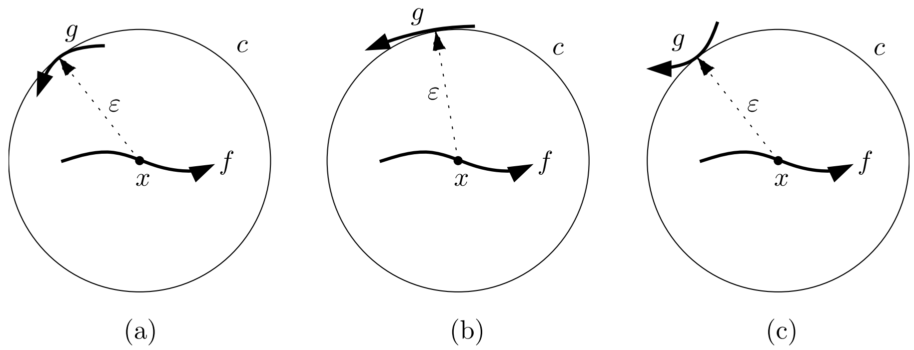

Figure 8: (a) How a critical point of type W^+ arises. (b,c) Critical points of type W^-.

A critical point of type E^+ corresponds to the opposite movement of x: The point x starts far away from g and moves closer as x proceeds along f. Critical points of type W^- and E^- occur when the curve g touches the circle c from *outside*, either because the curvature of g is bigger than $1/\varepsilon$ (Figure 8b) or because g curves away from c (Figure 8c). The critical points of type N and S correspond to the case when the roles of f and g are exchanged.

We can view these situations differently. Consider again the situation for W^+. At the critical point x has distance ε from g, and x moves away from g along f. In other words, f crosses the border which runs at fixed distance ε from g, the *offset curve* of g, see Figure 9.

The offset curves of a given curve g at distance ε are obtained as the locus of points which lie on the normal to g at any point y of g, ε away from y on both sides of g [5, 7]. If r is the radius of curvature at y, the radius of curvature at the corresponding point of the offset curve is $|r \pm \varepsilon|$. For $r > \varepsilon$, the offset curves move in the same direction as y (Figure 9b), whereas for $r < \varepsilon$, one of the curves

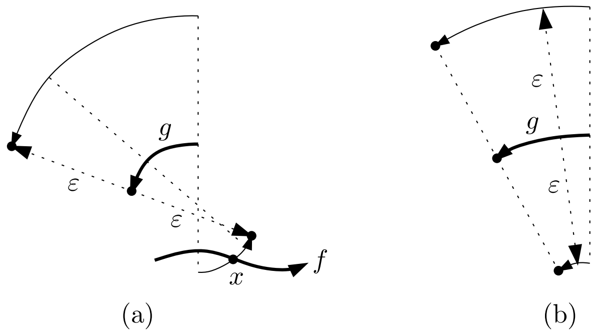

Figure 9: (a) Intersections with an offset curve. The curvature is larger than $1/\varepsilon$. (b) An offset curve when the curvature is smaller than $1/\varepsilon$.

moves in the opposite direction (Figure 9a). For $r = \varepsilon$, there is a criticality. However, since we consider pieces with $r \geq \varepsilon$ and with $r \leq \varepsilon$ separately, we need not worry about this case.

6 The structure of a single cell

The free space may be arbitrarily complicated even inside a cell. For example, if ε is very small, F_ε will contain isolated islands of free space for all intersections between f and g. However, we will show that the reachable points can be computed in a constant number of elementary geometric operations.

We have subdivided the curves, and consequently, the parameter intervals I and J into pieces. We denote by m and n the number of pieces. Correspondingly, we cut the rectangle $I \times J$ into mn cells. On the boundaries of these cells, we compute all points which are *reachable* from the lower left corner (l_I, l_J) of the rectangle by a path in free space which is monotone in both directions. We do this incrementally, starting with the lower left cell and ending in the uppermost right cell, for example in row-major order. For each cell, we assume that we know the reachable points on its left edge and on the lower edge and we transfer this information to the right edge and to the top edge.

A vertical line in the free-space diagram corresponds to a fixed point $f(s)$ on f. The points in F_ε on this line correspond to the points of g which lie inside a circle c of radius ε around $f(s)$. The boundary of F_ε corresponds to the intersections of c with g, and hence we can apply Lemma 3.

Lemma 4. (a) *Inside a cell, a vertical or horizontal line intersects the boundary of F_ε at most twice.*

A vertical tangent line trough a critical point of type E or W or a horizontal tangent line trough a critical point of type N or S does not cross the boundary of F_ε in any other point.

(b) *Moreover, if the curvature of g is smaller than $1/\varepsilon$, there can only be one interval of feasible points on a vertical line, and there are no critical points of type E^+ and W^+. Analogously, if the curvature of f is smaller than $1/\varepsilon$, there can only be one interval of feasible points on a horizontal line, and there are no critical points of type N^+ and S^+.*

Proof. The first statement follows directly from Lemma 3. A vertical line trough a critical point can be perturbed so that it intersects the boundary twice in the vicinity of this point, and thus any further intersection is excluded. (Note that the statement of Lemma 3 would permit the vertical tangent to *touch* the boundary of F_ε in another point. This is not counted as a *crossing*.) Part (b) follows from Lemma 3a and the same perturbation argument as in Part (a). □

Lemma 5. *A curve forming a component of the boundary of the free space inside a cell can contain at most four critical points. There are the following possibilities.*

1. *A closed circular loop with critical points N^-, E^-, S^-, W^- enclosing a component of free space;*
2. *a closed circular loop with critical points N^+, E^+, S^+, W^+ enclosing a component of forbidden points;*
3. *a subsequence of the above cyclic sequences containing between zero and three critical points, connecting two boundary points of the rectangle;*
4. *an "s-shaped" path between between the left edge and the right edge of the rectangle, containing two critical points S^+ and N^- or S^- and N^+;*
5. *an "s-shaped" path between between the bottom edge and the top edge of the rectangle, containing two critical points E^+ and W^- or E^- and W^+.*

Proof. If the curve always bends in the same direction (to the left or to the right) at successive critical points, it must fit one of the first three cases. Consider the "s-shaped" case of two successive critical points where the curve bends in opposite directions, as in Figure 10a. We have, say, N followed by S. (It does not matter whether it is N^+ followed by S^- or N^- followed by S^+.) One can see that there can be no further critical points after S: After another N, one can find a horizontal line with three intersections. After an E, one is also stuck: The curve cannot continue to the boundary without any further critical points, and critical points of type N or W are also not possible. □

When a component of the boundary of the free space is not a simple curve, as in Figure 10b, this can be resolved by a small perturbation of ε, giving rise to simple curves, and then the lemma can be applied.

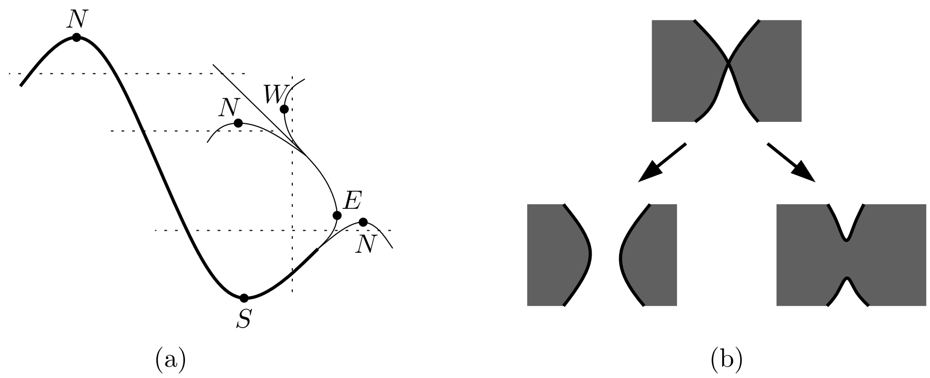

Figure 10: (a) Proof of Lemma 5. The dotted vertical and horizontal lines intersect the curve three times. (b) Eliminating degenerate situations.

7 Processing a cell

We are given the reachable points on the left and bottom edge, and we compute the points on the right edge and on the top edge which are reachable from there.

On each edge of the rectangle there are at most two intervals of free points, by Lemma 4. Inside each interval of free points, there is only a single interval of reachable points because from every free point, everything which is to the right or to the top in the same free interval is reachable directly.

Thus, in each free interval on the left and right edge, we just need to know the *lowest* reachable point, and in each free interval on the bottom and top edge, we need the *leftmost* reachable point. On the left and bottom edge, these data are given, and on the right and top edge they have to be computed.

We split the task into subproblems as follows: For each pair of a free "source" interval X on the left or bottom edge and a free "target" interval Y on the right or top edge, we compute the left-most or bottom-most point U in Y reachable from the given left-most or bottom-most reachable point B in X.

Without loss of generality, we discuss only the case where X is a free interval on the left edge. We are given the lowest reachable point B in it. There are two cases for Y: it can lie on the top edge or on the right edge.

We will describe our procedure in terms of geometric operations in the free space diagram, like finding the right-most point in a component of forbidden points. We will later discuss what these operations mean in terms of the curves f and g.

7.1 Reaching a free interval on the top edge

The upper end of X may be the upper left corner of the rectangle, or it may be a forbidden point which belongs to a component O of forbidden points. Similarly, the left endpoint F of Y may be part of a component of forbidden points, which we denote by O_2. (O and O_2 are not necessarily different, see Figure 11a.)

Lemma 6. *The leftmost point U in Y reachable from X depends only on the presence and the relative locations of O and O_2 and the horizontal line through B.*

Proof. We have to show that any other "obstacles" of forbidden points do not play any role in this question. We show this by giving an algorithm for constructing U in all cases.

If the horizontal line through B intersects O or O_2, it is clear that one cannot reach Y, see for example the interval X_1 in Figure 11a or the interval X_2 in connection with Y_2 in Figure 11b. Otherwise, we claim that the desired point U lies directly above the rightmost point of O or of O_2, whichever is further to the right. (If O or O_2 extend to the right edge of the rectangle, then again, no point of Y can be reached.)

It is clear that the monotone path from X to Y has to pass to the right of O and O_2. Thus, no point in Y left of U is reachable from X. To see that U is reachable, consider first the case that O exists, see the example of the interval X_1 in Figure 11b. Let A be the rightmost point of O. A can lie on the upper edge, or it can be a critical point of type E^+.

Assume first that A is a critical point of type E^+. The vertical line a through A lies completely in the free space, by Lemma 4, and O is the only obstacle left of a. By assumption, the horizontal line b from the lowest reachable point B in X does not intersect O before reaching a, and there are no other obstacles in this range. Thus, A is reachable from B, and the upper end A' of a is the leftmost reachable point on the top edge. If it lies in Y, we can take it as our point U, and we are done. (This is the case for the intervals X_1 and Y_1 in Figure 11b.) If Y lies left of a, we are done as well, as no points in Y are reachable from X. So let us deal with the only remaining case that Y lies to the right of a, and a is separated from Y by the obstacle O_2.

The lowest point D of O_2 must lie above O, and the horizontal line through D intersects a, which is reachable. Therefore D is reachable. From D we can reach the rightmost point C of O_2, which is either a critical point of type E^+ or the point F. In either case, we can indeed reach the point C' vertically above C as the leftmost point U.

The argument in the last paragraph also works in the case that O does not exist, because by assumption, the horizontal line through D intersects the left edge above B.

We are left with the case that the rightmost point A of O lies on the upper edge. In that case it must coincide with the left end F of Y, and the obstacles O and O_2 are the same, as in Figure 11a. Either the boundary of O is monotone and the lowest point of O lies on the left edge, or the lowest point D of $O = O_2$ is a critical point of type S^+, then we can argue similarly as above, using the horizontal line through D to show that F is reachable. □

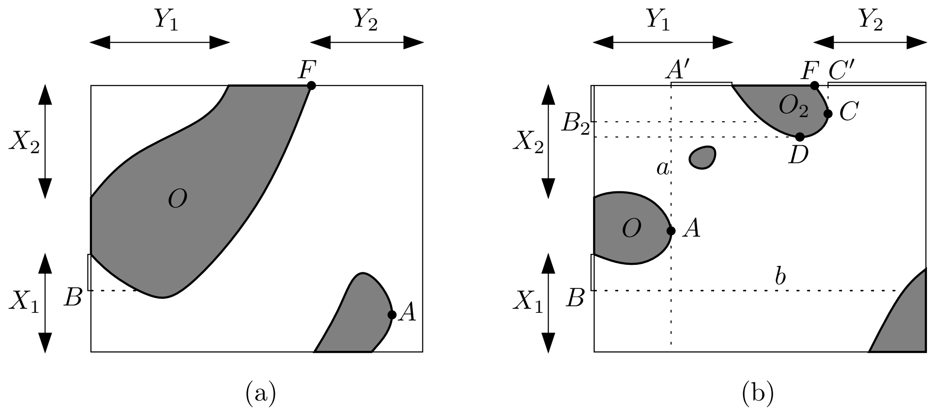

Figure 11: Determining the reachable points on the top edge

7.2 Reaching a free interval on the right edge

Again, we denote by O the component of forbidden points incident to the upper end of X. We denote by O_2 the component of forbidden points incident to the *lower* end of Y, if such a component exists. And we denote by O_1 the first (leftmost) component which is hit by the horizontal line through B, if it exists, see Figure 12a.

Lemma 7. *The lowest point U in Y reachable from X depends only on the presence and the relative locations of O, O_1 and O_2 and the horizontal line b through B.*

Proof. As in Lemma 6, it may happen that b hits O, and nothing is reachable. Otherwise, let A be the highest point of O_1. As in Lemma 6, one can argue that A is reachable from B, and that the lowest reachable point in Y is as high as A or the highest point C of O_2, whichever is higher. The cases when O_1 or O_2 don't exist are similar. ☐

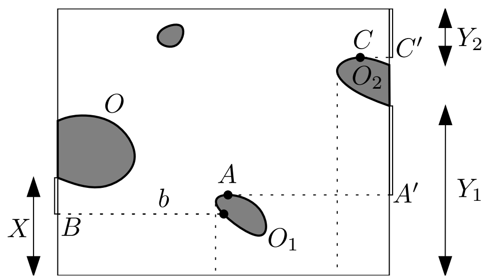

Figure 12: Determining the reachable points on the right edge

7.3 The primitive operations

We will now show how to carry out the operations that have been described above in terms of the free space diagram.

The easy operation is finding the free points on a given horizontal or vertical line, or finding the first intersection of such a line with an obstacle of forbidden points. This simply amounts to intersecting f with a circle of radius ε around a given point of g, or vice versa.

The other task is to find the rightmost, bottom-most, etc., point on the boundary of an obstacle. This is either a critical point or a point where the obstacle hits the cell boundary. These cases can be distinguished by a conceptual sweep over the free-space diagram.

The points where an obstacle boundary hits a cell boundary can be determined beforehand. There are at most eight of these points. For each boundary point, we can also find out the quadrant into which the obstacle boundary moves. For example, consider the point F on the upper edge in Figure 11a. The free space is on the right side of F. As the portion directly below F on the vertical line through F lies also in the free space, we can conclude that the boundary of O moves away to the left from F.

To show how an extreme point of an obstacle is identified, consider for example the obstacle O of forbidden points which touches the left edge in Figure 11. We want to know the rightmost point of O. As we sweep a vertical line over the cell, starting at the left edge, we keep track of the obstacle boundaries. The number of obstacle boundaries changes only when we pass a boundary point on the top or bottom edge (and in this case we know whether an obstacle boundary starts or ends there) or at a critical point A of type E, where two obstacle boundaries merge and disappear, or at a critical point of type W, where two obstacle boundaries appear. If the obstacle O has not already ended (like in the rightmost point F in Figure 11a), the point A is the desired rightmost point of O. We can stop the sweep as soon as the rightmost point is reached. Then, the sweep line cannot hit another obstacle of forbidden points, by Lemma 4, (and a critical point of type W is in fact impossible in this case).

We summarize the above arguments in the following theorem.

Theorem 1. *Given the reachable points on the bottom edge and the left edge of a cell, the reachable points on the top edge and the right edge of the cell can be computed in a constant number of the following operations:*

- *Intersecting a circle of radius ε with one of the curves*
- *Finding the first intersection of one curve with an offset curve of the other curve at distance ε.*

In both cases, we must be able to find the parameter values on the respective curves, corresponding to the points that we have computed. □

If the curves are algebraic curves whose degree is bounded by a constant, then these primitive are algebraic operations, whose degree is bounded by another (higher) constant. If we make no such assumptions on the curves, then there is no bound on the number of intersections between one curve and the offset curve of the other curve. Therefore it is important to say that we compute the *first* intersection (or the first intersection that comes after some specified parameter value) in the statement of the theorem.

We state another consequence of our considerations in the following lemma

Lemma 8. *Let X be a free interval on a horizontal line through the free space diagram. Then the leftmost reachable point in X is either the left endpoint of X or it lies vertically above the rightmost point of some obstacle of forbidden points.*

Similarly, let Y be a free interval on a vertical line through the free space diagram. Then the lowest reachable point in Y is either the lower endpoint of Y or it lies at the same height as the topmost point of some obstacle of forbidden points.

Proof. This easily follows by induction, using the way how the leftmost or lowest reachable point is constructed according to Lemma 6 and Lemma 7. □

7.4 Circular arcs

Circular arcs as part of the curve f and g require some special care but are not difficult to handle. In Lemma 3, if f shares a circular arc with c, this must be counted as a single intersection. If f contains a circular arc of radius ε and g passes through the center of this arc, then the boundary of F_ε may contain a horizontal segment. The question whether the boundary moves away towards the top or towards the bottom must be decided on the end of the circular arc, where the curvature starts to be different from $1/\varepsilon$. (If g does not pass through the center or the radius is different from $1/\varepsilon$, the circular arc requires no special treatment at all.)

7.5 Solving the decision problem

The following theorem summarizes the result that we have proved.

Theorem 2. *Given a parameter ε and two curves consisting of m and n pieces, respectively, where each piece has a turning angle at most π and curvature $\geq 1/\varepsilon$ or $\leq 1/\varepsilon$ throughout, we can decide $O(m+n)$ space and in $O(mn)$ primitive operations of the type described in Theorem 1 whether their Fréchet distance is at most ε.* □

In order to bring the input into the form required by the theorem, we may have to use additional primitive operations: Cutting a curve at all points of a given tangent direction, and at all points of curvature $1/\varepsilon$.

8 The Minimization Problem

We now solve the minimization problem of computing the Fréchet distance. However, we will make some stronger assumptions on the curves: We assume that all pieces of the curve are algebraic of degree bounded by a constant. This is needed in the analysis of the algorithm to ensure that only a limited number of obstacles can appear in a cell.

We apply Megiddo's parametric search technique [8]. We will closely follow the approach of [2] for polygonal curves, except that the technical details are a little bit more involved.

By Theorem 2, we know that we can solve the decision problem for a given threshold ε it we know all obstacles and their extreme points in all four directions in every cell, as well as the endpoints of the free intervals on the cell boundaries. There is a slight technical problem in applying parametric search, since obstacles may appear and merge together as we vary ε. Also, the cell boundaries when we cut the curves at the points of curvature $1/\varepsilon$ are not stationary.

We solve these problems in a preprocessing phase. We first ensure that each piece of the curve has turning angle at most π. As mentioned above, this multiplies the number of pieces at most by a constant. Then we consider the partition of each piece into subpieces at the points of curvature $1/\varepsilon$, for varying ε. At certain *critical values* of ε, the combinatorial structure of this partition changes: These values are the curvatures at the endpoints of the piece and the local maxima and minima of the curvature along the curve. This gives an initial set of $O(m+n)$ critical values for all pieces of f and g. Within one interval between two such critical values of ε, we consider the partition of a cell of the free space diagram into subcells by the points of curvature $1/\varepsilon$. The number of subcells is fixed, but the boundaries move. By applying Lemmas 6 and 7 inductively to these subcells, we know that we can compute the reachable points on the upper and right edge of the cell if we know the reachable points on the lower and left edge of the cell and the extreme points of all obstacles in all subcells, and where they intersect the subcell boundaries. At some values of ε, the combinatorial structure of the partition into free space and obstacles changes: A new obstacle may appear in the middle of a cell or at a cell boundary, two obstacles may grow and merge into one, of an obstacle may become tangent to the cell boundary. There is only a constant number of these critical events per subcell.

Thus, by a binary search among $O(mn)$ critical values of ε, we can narrow down the interval of possible values of the Fréchet distance ε^* so that we know the precise number of obstacles in each subcell, and whether and where they touch the subcell boundary.

Each binary search is a run of the decision algorithm, for a total time of $O(mn\log(mn))$.

By Lemmas 6, 7 and 8, we know that we can solve the decision problem if we know the sorted order of all extreme points and boundary points of all obstacles in all four directions. We apply parametric search, using a parallel sorting algorithm for the coordinates of these critical points. This narrows down the interval the interval of possible values for ε^* such that in the whole interval, no two critical points switch positions in the x- or y-order; So we know that the lower endpoint of this interval must be the Fréchet distance ε. By utilizing Cole's variant of parametric search [4] we obtain a running time of $O(mn\log(mn))$.

Theorem 3. *Given two curves consisting of m and n pieces, respectively, of smooth algebraic curves of fixed maximum degree we can compute their Fréchet distance in $O(nm)$ space and in $O(mn\log(mn))$ algebraic operations of bounded degree.* □

Here, an algebraic operation refers to a comparison between two real solutions of algebraic equations.

References

[1] H. Alt, A. Efrat, G. Rote, and C. Wenk, *Matching planar maps*, Journal of Algorithms **49** (2003), 262–283.

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

[3] H. Alt, C. Knauer, and C. Wenk, *Matching polygonal curves with respect to the Fréchet distance*, In: Proc. 18th Int. Symp. Theoret. Aspects Comput. Sci., STACS 2001 (A. Ferreira and H. Reichel, eds.), Lect. Notes Comp. Sci., vol. 2010, Springer-Verlag, 2001, pp. 63–74.

[4] R. Cole, *Slowing down sorting networks to obtain faster sorting algorithms*, J. Assoc. Comput. Mach. **34** (1987), no. 1, 200–208.

[5] R. T. Farouki and C. A. Neff, *Analytic properties of plane offset curves*, Comput. Aided Geom. Design **7** (1990), 83–99.

[6] O. Gonzalez, J. H. Maddocks, and J. Smutny, *Curves, circles, and spheres*, In: Physical Knots: Knotting, Linking, and Folding Geometric Objects in $\mathbb{R}^3$ (J. A. Calvo et al., ed.), Contemporary Mathematics, vol. 304, American Mathematical Society, 2002, pp. 195–215.

[7] M. Held, *Voronoi diagrams and offset curves of curvilinear polygons*, Computer-Aided Design **30** (1998), 287–300.

[8] N. Megiddo, *Applying parallel computation algorithms in the design of serial algorithms*, J. Assoc. Comput. Mach. **30** (1983), 852–865.