

Curve Reconstruction in Arbitrary Dimension and the Traveling Salesman Problem^{*}

Joachim Giesen

Institut für Theoretische Informatik, ETH Zürich, CH-8092 Zürich, Switzerland

Abstract. Given a finite set of points sampled from a curve, we want to reconstruct the ordering of the points along the curve. Every ordering of the sample points can be defined by a polygon through these points. We show that for simple, regular curves Traveling Salesman Paths give the correct polygonal reconstruction, provided the points are sampled densely enough. In this case the polygonal reconstruction is part of the Delaunay Triangulation of the sample points. We use this observation to design an efficient algorithm for the reconstruction problem.

Keywords. Curve Reconstruction, Traveling Salesman Problem, Minimum Spanning Tree

1 Introduction

Reconstruction of an object from an unorganized set of sample points is a fundamental problem in many areas of science. Here we address the problem to reconstruct a curve in arbitrary dimension via a finite set of sample points.

Following Amenta et al. [2], we state the curve reconstruction problem as follows

Given a curve γ in \mathbf{R}^d and a finite set of sample points $S \subset \gamma$. A polygonal reconstruction $P(S)$ of γ from S is a graph that connects all points adjacent along γ .

Thus an algorithm for curve reconstruction should reconstruct the order of the sample points along γ . That seem to be a natural claim because the main geometric characteristics of curves, length and total curvature, are defined via a limit over inscribed polygons [1]. The polygonal reconstruction is the inscribed polygon through the sample points. For dense samplings the length and the total curvature of the polygonal reconstruction approximate the length and the total curvature of γ . This is no longer true if one considers polygons, as in Attali [3], that are only near to γ in the Hausdorff or Fréchet metric. Given the sample points and no other information the polygonal reconstruction is in general the best way to estimate the length and the total curvature of γ .

Of course there is no algorithm that can reconstruct any curve from any set of samples. We have to introduce restrictions on γ and S . All known algorithms

^{*} This work was supported by grants from the Swiss Federal Office for Education and Science (Projects ESPRIT IV LTR No. 21957 CGAL and No. 28155 GALIA)

demand dense sampling. That is, these algorithms guarantee to compute a polygonal reconstruction, if they do it at all, only for sufficiently dense samples. The most common restrictions on γ are that γ is a smooth, simple (closed), plane curve.

The plane version of the curve reconstruction problem attracted some attention in recent years. O'Rourke et al. [10] published the first heuristic for this problem. Later on many algorithms and geometric graphs were designed to solve the reconstruction problem. But only very recently two algorithms were found which provably connect the sample points in the right order, provided the points were sampled densely enough: The algorithm of Bernadini and Bajaj [5] is based on α -shapes, a geometric structure introduced by Edelsbrunner [7] to model the shape of point sets. Amenta et al. [2] showed that the crust, a geometric graph defined in their article, and a β -skeleton, another geometric graph introduced by Kirkpatrick and Radke [8], solve the reconstruction problem for plane curves.

For some applications it is also interesting to reconstruct curves in higher dimension. An example is given by iso-line extraction in data from numerical solutions of partial differential equations. Furthermore the construction of Bezier or B-Spline curves through a set of sample points needs an ordering of these points. Such an ordering is induced by the ordering along the polygonal reconstruction. We show that Traveling Salesman Paths and Minimum Spanning Trees solve the reconstruction problem for regular curves in arbitrary dimension, provided the points are sampled densely enough. This might seem obvious, but there are quite well behaved curves for which this is not true. That is interesting from a theoretical point of view, because in 1930 Menger [9] suggested to base the geometry of curves and surfaces on minimum area triangulations.

We have implemented our reconstruction algorithm, an algorithm for the computation of Minimum Spanning Trees and a factor-2-approximation algorithm [6] for the Traveling Salesman Problem based on Minimum Spanning Trees. It turned out that the approximation algorithm is not suitable for the curve reconstruction problem. These implementations can be found as a JAVA applet (a JAVA 1.1 capable browser is needed) at <http://www.inf.ethz.ch/personal/giesen/>.

This article is organized as follows: First, we give the definitions of regularity and samples. Second, we show that Traveling Salesman Paths and Minimum Spanning Trees solve the reconstruction problem. Third, we present a simple and efficient algorithm that also solves the reconstruction problem. This algorithm is better suited for practical purposes, where in general no guarantee is given that the points are sampled densely enough. The proof of correctness for this algorithm uses the same techniques as used for the Traveling Salesman result. Finally, we discuss how to estimate the necessary sampling density.

2 One-Manifolds, Regularity and Samples

We want to consider regular embeddings of compact one-manifolds in arbitrary dimension. The restriction on embeddings of compact one-manifolds is equiva-

lent to the restriction on simple (closed) curves, because every simple curve is homeomorphic to either an interval or the unit circle. The unit circle \mathbf{S}^1 and the closed unit interval $[0, 1]$ are the only compact one-manifolds in the sense, that all compact one-manifolds are homeomorphic to one of them. In the following we write \mathbf{M}^1 for an one-manifold if we think of just one of \mathbf{S}^1 and $[0, 1]$.

First we want to specify our regularity assumptions. The following definition of regularity depends only on the image of γ and is equivalent to the definition of regularity used in differential geometry.

Definition 1. Let $\gamma : [0, 1] \rightarrow \mathbf{R}^d$ be an embedding, $T = \{(t_1, t_2) : t_1 < t_2, t_1, t_2 \in [0, 1]\}$ and

$$\tau : T \rightarrow \mathbf{S}^{d-1}, (t_1, t_2) \mapsto \frac{\gamma(t_2) - \gamma(t_1)}{|\gamma(t_2) - \gamma(t_1)|}.$$

The embedding γ is called regular at $\gamma(t_0)$ with tangent $\mathfrak{t}(\gamma(t_0))$ if for every sequence (ξ_n) in T which converges to (t_0, t_0) in $\text{closure}(T)$ the sequence $\tau(\xi_n)$ converges to $\mathfrak{t}(\gamma(t_0))$. We call γ regular if it is regular in all points $\gamma(t), t \in [0, 1]$.

This definition of regularity carries over to embeddings of \mathbf{S}^1 , since such embeddings locally look like embeddings of $[0, 1]$. Our definition of regularity has a pure metric interpretation which is fundamental for our proofs.

Lemma 1. Let γ be an embedding of \mathbf{M}^1 , which is regular in $p \in \gamma$. Let $(p_n), (q_n)$ and (r_n) be sequences of points from γ , that converge to p , such that $p_n < q_n < r_n$ for all $n \in \mathbf{N}$ in an order locally around p along γ . Then the sequence of angles (α_n) converges to π , where α_n is the angle at q_n of the triangle with corner points p_n, q_n and r_n .

Proof. Since γ is homeomorphic the sequences $(\gamma^{-1}(p_n)), (\gamma^{-1}(q_n))$ and $(\gamma^{-1}(r_n))$ converge to $\gamma^{-1}(p)$. Thus by our definition of regularity asymptotically the three secants

$$\text{conv}p_n, q_n, \text{conv}q_n, r_n \text{ and } \text{conv}p_n, r_n$$

have to point in the direction of the tangent $\mathfrak{t}(p)$. That is, $\lim_{n \rightarrow \infty} \alpha_n = \pi$. \square

Second we want to clarify the notion of a sample S , give a notion of a local order of a sample and introduce a measure $\varepsilon(S)$ for the density of a sample.

Definition 2. Let γ be an embedding of \mathbf{M}^1 .

1. A sample S of γ is a finite set $S = \{p^1, \dots, p^n\}$ where $p^i \in \gamma$. We assume that the sample points p^i are ordered according to the order of the $t^i = \gamma^{-1}(p^i)$, i.e. t^i and t^{i+1} have to be adjacent along \mathbf{M}^1 .
2. If $\mathbf{M}^1 = [0, 1]$ we write $i \text{ C } j$ if $t^i < t^j$. Otherwise we choose an orientation along γ and write $i \text{ C } j$ if $i \neq j$ and

$$\text{length}(\gamma(p^i : p^j)) \leq \text{length}(\gamma(p^j : p^i)),$$

where $\gamma(p^i : p^j) \subset \gamma$ is the arc connecting p^i and p^j in the orientation along γ . We write $i \text{ E } j$ if we want to include the possibility that $i = j$.

3. To every sample the number $\varepsilon(S)$ is defined as

$$\varepsilon(S) = \sup_{x \in \gamma} \min\{|p^i - x| : i = 0, \dots, n\}.$$

The notion of local order for embeddings of \mathbf{S}^1 is well defined, because regular embeddings have always finite length. A proof of this can be found in the book of Aleksandrov and Reshetnyak [1].

3 Geometric Graphs and Polygonal Reconstruction

Let S be a finite point set in \mathbf{R}^d . $TSP(S)$ is a shortest path through the points S , $MST(S)$ is a tree of minimal length on the points S and $TST(S)$ is a Hamilton cycle of minimal length of the points S . In general these geometric graphs need not to be unique.

In this section we show that in the case of regular embeddings of $[0, 1]$ Traveling Salesman Paths (TSP) and Minimum Spanning Trees (MST) solve the reconstruction problem for sufficiently dense sampling. In the case of regular \mathbf{S}^1 embeddings the reconstruction problem is solved by a closed Traveling Salesman Tour (TST). Furthermore we show that for sufficiently dense samples all edges of the polygonal reconstruction are edges of a Delaunay Triangulation of the sample points. The Delaunay Triangulation is a “nice” triangulation which can be computed efficiently, see for example [4] and [7]. But first of all we give an example which shows that the statements above are not as obvious as they might seem.

Consider the following non-regular embedding (Figure 1) $\gamma_c : [0, 1] \rightarrow \mathbf{R}^2$ given by

$$\gamma_c |_{[b_j, b_{j+1}]} : t \mapsto (b_{j+1} - t)/(b_{j+1} - b_j) \cdot \gamma_c(b_j) + (t - b_j)/(b_{j+1} - b_j) \cdot \gamma_c(b_{j+1})$$

and

$$\gamma_c(1) = (1, 2).$$

Here is

1. $b_{4i} = a_i, \gamma_c(b_{4i}) = (2a_i, a_i),$
 $b_{4i+1} = a_i + 2^{-i-3}, \gamma_c(b_{4i+1}) = (a_i + a_{i+1}, a_i),$
 $b_{4i+2} = a_i + 2^{-i-2}, \gamma_c(b_{4i+2}) = (2a_{i+1}, a_i)$ and
 $b_{4i+3} = a_i + 3 \cdot 2^{-i-3}, \gamma_c(b_{4i+3}) = (a_i + a_{i+1}, a_{i+1}).$
2. $a_i = 1 - 2^{-i}$. That is, $[0, 1] = \bigcup_{i=0}^{\infty} [a_i, a_{i+1}] \cup \{1\}$.

Note 1. Consider the embedding γ_δ one gets from a contraction by a factor $\delta \in (0, 1)$ in the y-direction from γ_c . For every $\varepsilon > 0$ there exists a sample S of γ_δ with $\varepsilon(S) < \varepsilon$ and $P(S) \neq TSP(S), MST(S)$.

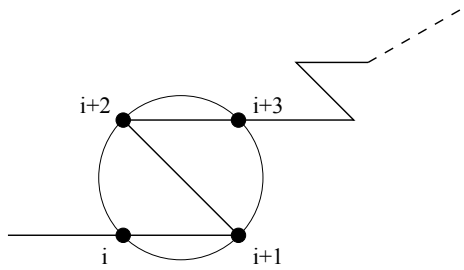


Figure 1.

Proof. Choose samples S_i of γ_δ such that $\varepsilon(S_i) < 1 - a_i$ and

$$S_i \cap \gamma_\delta([b_{4i+1}, b_{4(i+1)}]) = \{\gamma_\delta(b_{4i+1}), \gamma_\delta(b_{4i+2}), \gamma_\delta(b_{4i+3}), \gamma_\delta(b_{4(i+1)})\}.$$

Replacing the edges

$$\text{conv}\gamma_\delta(b_{4i+1}), \gamma_\delta(b_{4i+2}), \text{conv}\gamma_\delta(b_{4i+2}), \gamma_\delta(b_{4i+3}) \text{ and } \text{conv}\gamma_\delta(b_{4i+3}), \gamma_\delta(b_{4(i+1)})$$

in $P(S_i)$ by

$$\text{conv}\gamma_\delta(b_{4i+1}), \gamma_\delta(b_{4i+3}), \text{conv}\gamma_\delta(b_{4i+3}), \gamma_\delta(b_{4i+2}) \text{ and } \text{conv}\gamma_\delta(b_{4i+2}), \gamma_\delta(b_{4(i+1)})$$

leads to a shorter path through the sample points S_i . That is,

$$P(S_i) \neq TSP(S_i), MST(S_i) \text{ for all } i \in \mathbf{N}.$$

Since $\varepsilon(S_i) \rightarrow 0$ as $i \rightarrow \infty$ we are done. □

By closing γ_δ in a simple manner we get an example where $P(S) \neq TST(S)$ for arbitrarily dense samples.

Note 2. For every $\varepsilon > 0$ there exists a sample S of γ_c with $\varepsilon(S_i) < \varepsilon$ such that $P(S)$ has an edge that is not a Delaunay edge.

Proof. Choose $0 < \varepsilon_i < 2^{-i-4}$ and samples S_i of γ_c such that $\varepsilon(S_i) < 1 - a_i$ and

$$S_i \cap \gamma_c([b_{4i+1}, b_{4(i+1)}]) = \{\gamma_c(b_{4i+1} + \varepsilon_i), \gamma_c(b_{4i+2}), \gamma_c(b_{4i+3}), \gamma_c(b_{4(i+1)} - \varepsilon_i)\}.$$

Then $\text{conv}\gamma_c(b_{4i+2}), \gamma_c(b_{4i+3}) \subset P(S_i)$ cannot be a Delaunay edge by the empty ball criterion [7] for Delaunay edges (see Figure 1). Since $\varepsilon(S_i) \rightarrow 0$ as $i \rightarrow \infty$ we are done. □

Both embeddings γ_c and γ_δ have finite length. Thus rectifiability, which implies differentiability almost everywhere, is not sufficient for the Traveling Salesman path to solve the reconstruction for arbitrarily dense samples. Even worse in this example the polygonal reconstruction is not part of the Delaunay Triangulation of the sample points for arbitrarily dense samples. Our contribution is to prove that regularity is sufficient. For these proofs we need three lemmas, which we show at first. We always set $p^0 := p^{|S|}$ and $p^{|S|+1} := p^1$ if we deal with embeddings of \mathbf{S}^1 .

Lemma 2. *Let γ be an embedding of \mathbf{M}^1 and let (S_n) be a sequence of samples of γ with*

$$\lim_{n \rightarrow \infty} \varepsilon(S_n) = 0.$$

Then

$$\lim_{n \rightarrow \infty} \max\{|p_n^{i+1} - p_n^i| : p_n^i \in S_n\} = 0.$$

Proof. Assume the contrary. Then we find $c > 0$ and $p_n, q_n \in S_n$, such that $\gamma^{-1}(p_n)$ and $\gamma^{-1}(q_n)$ are adjacent along \mathbf{M}^1 and $|p_n - q_n| > 2c$. By the compactness of \mathbf{M}^1 we can assume that (p_n) converges to $p \in \gamma$ and (q_n) converges to $q \in \gamma$, otherwise we choose an appropriate subsequence. On γ we find a point r between p and q along γ with $|r - p| \geq c$ and $|r - q| \geq c$. For every $n \in \mathbf{N}$ we find $r_n \in S_n$ with $|r - r_n| < \varepsilon(S_n)$. Thus the sequence (r_n) converges to r , which is impossible for an embedding. \square

Lemma 3. *Let γ be a regular embedding of \mathbf{M}^1 . Then there exists an $\varepsilon > 0$ such that*

1. $|p^i - p^{i-1}| < |p^i - p^k|$ if $k \text{ C } i - 1$
2. $|p^i - p^{i+1}| < |p^i - p^k|$ if $k \text{ B } i + 1$

for all samples $S = \{p^1, \dots, p^{|S|}\}$ with $\varepsilon(S) < \varepsilon$ and all $i \in \{1, \dots, |S|\}$.

Proof. Assume the contrary. Then there exists a sequence (S_n) of samples with $\lim_{n \rightarrow \infty} \varepsilon(S_n) = 0$ and $p_n^i, p_n^k \in S_n$ such that

$$k \text{ C } i - 1 \text{ and } |p_n^i - p_n^{i-1}| \geq |p_n^i - p_n^k|$$

or

$$k \text{ B } i + 1 \text{ and } |p_n^i - p_n^{i+1}| \geq |p_n^i - p_n^k|.$$

By choosing a subsequence we can always assume that for all $n \in \mathbf{N}$ one of the above possibilities holds. Without loss of generality assume that this is the first one. Since γ is compact, we can also assume by choosing a subsequence that (p_n^i) converges to $p \in \gamma$. From Lemma 2 we get

$$\lim_{n \rightarrow \infty} |p_n^i - p_n^k| \leq \lim_{n \rightarrow \infty} |p_n^i - p_n^{i-1}| = 0.$$

Hence also (p_n^{i-1}) and (p_n^k) converge to p . Now look at the triangle with corner points p_n^{i-1}, p_n^i and p_n^k . From the law of cosines together with $|p_n^i - p_n^{i-1}| \geq |p_n^i - p_n^k|$, we find for the angle α_n at p_n^{i-1}

$$\cos(\alpha_n) = -\frac{|p_n^i - p_n^k|^2 - |p_n^i - p_n^{i-1}|^2 - |p_n^{i-1} - p_n^k|^2}{|p_n^i - p_n^{i-1}| |p_n^{i-1} - p_n^k|} \geq 0.$$

Thus α_n has to be smaller or equal than $\frac{\pi}{2}$, but that is a contradiction to Lemma 1. \square

Lemma 4. *Let γ be a regular embedding of \mathbf{M}^1 . Then there exists an $\varepsilon > 0$ such that*

1. $|p^i - p^m| < |p^i - p^k|$ for $k \subset m \subset i$,
 if $k \subset i - 1$ and $|p^i - p^k| \leq |p^l - p^{l+1}|$ for some $l \in \{1, \dots, |S|\}$
2. $|p^i - p^m| < |p^i - p^k|$ for $i \subset m \subset k$,
 if $k \subset i + 1$ and $|p^i - p^k| \leq |p^l - p^{l+1}|$ for some $l \in \{1, \dots, |S|\}$

for all samples $S = \{p^1, \dots, p^{|S|}\}$ with $\varepsilon(S) < \varepsilon$ and all $i \in \{1, \dots, |S|\}$.

Proof. The proof is similar to the proof of Lemma 3. □

Now we are ready to prove the theorems for embeddings of $[0, 1]$. The proof for the Traveling Salesman Path and for the Minimum Spanning Tree are done at once, because we show that asymptotically the Minimum Spanning Tree of the sample points is a unique path, which of course is a Traveling Salesman Path.

Theorem 1. *Let γ be a regular embedding of $[0, 1]$. Then there exists an $\varepsilon > 0$ such that $MST(S) = TSP(S) = P(S)$ for all samples S of γ with $\varepsilon(S) < \varepsilon$. In particular $MST(S)$ and $TSP(S)$ are unique, if $\varepsilon(S) < \varepsilon$.*

Proof. We are done if we can show that there exists an $\varepsilon > 0$ such that the polygonal reconstruction $P(S_n)$ is a Minimum Spanning Tree on the points S and that there is only one Minimum Spanning Tree for all samples S of $\gamma([0, 1])$ with $\varepsilon(S) < \varepsilon$. Assume the contrary. Then there exists a sequence (S_n) of samples $S_n = \{p_n^1, \dots, p_n^{|S_n|}\}$ with $\lim_{n \rightarrow \infty} \varepsilon(S_n) = 0$ and $MST(S_n) \neq P(S_n)$. That means, $P(S_n)$ has to contain an edge $e_n = \text{conv}p_n^i, p_n^{i+1}$ with $e_n \notin MST(S_n)$. Adding e_n to $MST(S_n)$ induces a cycle on $MST(S_n)$. All edges in this cycle different from e_n have to be of smaller or equal length than e_n , because of the minimality property of Minimum Spanning Trees. Hence there is an edge $e'_n = \text{conv}p_n^k, p_n^l \neq e_n$ in this cycle with $k \in i, i + 1 \in l$ and $|p_n^k - p_n^l| \leq |p_n^i - p_n^{i+1}|$. Lemma 4 tells us that there exists $N \in \mathbf{N}$ with

1. $|p_n^i - p_n^l| \leq |p_n^k - p_n^l|$ for all $n \geq N$. That means $|p_n^i - p_n^l| \leq |p_n^i - p_n^{i+1}|$ for all $n \geq N$.
2. $|p_n^k - p_n^{i+1}| \leq |p_n^k - p_n^l|$ for all $n \geq N$. That means $|p_n^k - p_n^{i+1}| \leq |p_n^i - p_n^{i+1}|$ for all $n \geq N$.

That is a contradiction to Lemma 3. □

Edges of an Euclidean Minimum Spanning Tree are always edges in a Delaunay Triangulation [7]. Thus the edges of the polygonal reconstruction asymptotically also have to be Delaunay edges.

The proof for embeddings of \mathbf{S}^1 is slightly more complicated. We can use Minimum Spanning Trees only indirectly for the proof.

Lemma 5. *Let γ be a regular embedding of \mathbf{S}^1 . Then there exists an $\varepsilon > 0$ such that $MST(S) \subset P(S)$ for all samples S of γ with $\varepsilon(S) < \varepsilon$.*

Proof. Assume the contrary. Then there exists a sequence (S_n) of samples $S_n = \{p_n^1, \dots, p_n^{|S_n|}\}$ with $\varepsilon(S_n) \rightarrow 0$ as $n \rightarrow \infty$ and $MST(S_n)$ with $MST(S_n) \not\subset P(S_n)$. Denote by E_n the set of edges of $P(S_n)$, by E'_n the set of edges of $MST(S_n)$ and

$$R = E_n - (E'_n \cap E_n) \text{ and } M = E'_n - (E'_n \cap E_n).$$

From our assumption we have $M \neq \emptyset$. Choose $\text{conv}p_n^k, p_n^l \in M$ with $k \subset l$. Then there exists $\text{conv}p_n^i, p_n^{i+1} \in R$ with $k \in i, i+1 \in l$, because otherwise $MST(S_n)$ has to contain a cycle, which contradicts the definition of a tree. Assume $|p_n^i - p_n^{i+1}| < |p_n^k - p_n^l|$. Removing the edge $\text{conv}p_n^k, p_n^l$ from $MST(S_n)$ decomposes $MST(S_n)$ in two connected components. Both p_n^i and p_n^{i+1} have to belong to the same component, because otherwise we get a shorter Minimum Spanning Tree by replacing $\text{conv}p_n^k, p_n^l$ by $\text{conv}p_n^i, p_n^{i+1}$ in $MST(S_n)$. Without loss of generality we can assume that $\text{conv}p_n^k, p_n^l$ connects the two components, because p_n^k and p_n^l have to belong to different components. Lemma 4 tells us that for sufficiently large n , we have $|p_n^k - p_n^i| < |p_n^k - p_n^l|$. That means, we get a shorter Minimum Spanning Tree by replacing $\text{conv}p_n^k, p_n^l$ by $\text{conv}p_n^k, p_n^i$ in $MST(S_n)$, which is a contradiction.

Hence we can assume by choosing an appropriate subsequence that $|p_n^k - p_n^l| \leq |p_n^i, p_n^{i+1}|$ for all $n \in \mathbf{N}$. Lemma 4 tells us that there exists $N \in \mathbf{N}$ with

1. $|p_n^i - p_n^l| \leq |p_n^k - p_n^l|$ for all $n \geq N$. That means $|p_n^i - p_n^l| \leq |p_n^i - p_n^{i+1}|$ for all $n \geq N$.
2. $|p_n^k - p_n^{i+1}| \leq |p_n^k - p_n^l|$ for all $n \geq N$. That means $|p_n^k - p_n^{i+1}| \leq |p_n^i - p_n^{i+1}|$ for all $n \geq N$.

That is a contradiction to Lemma 3. □

Again asymptotically a Minimum Spanning Tree of the sample points has to be a path, which consists of Delaunay edges. By adding the edge that connects the start- and the endpoint of this path we get the polygonal reconstruction. We claim that this last edge is also an Delaunay edge. To show this we use the notions of Lemma 5 and the empty ball criterion [7] for Delaunay edges. The endpoints of $MST(S_n)$ have to be p_n^i and p_n^{i+1} for some index $i \in \{1, \dots, |S_n|\}$. Assume that the open balls B_n with diameter $|p_n^{i+1} - p_n^i|$ and $\{p_n^{i+1}, p_n^i\} \subset \text{boundary}(B_n) \cap S_n$ are not empty for all $n \in \mathbf{N}$. Then there has to exist $p_n^k \in S_n$ with $k \subset i$ or $k \in i+1$ and

$$|p_n^{i+1} - p_n^k|, |p_n^i - p_n^k| \leq |p_n^{i+1} - p_n^i|.$$

That contradicts Lemma 3. Thus asymptotically all the balls B_n have to be empty and all edges of the polygonal reconstruction have to be Delaunay edges.

Theorem 2. *Let γ be a regular embedding of \mathbf{S}^1 . Then there exists an $\varepsilon > 0$ such that $TST(S) = P(S)$ for all samples S of γ with $\varepsilon(S) < \varepsilon$. Especially $TST(S)$ is unique, if $\varepsilon(S) < \varepsilon$.*

Proof. We do the proof by contradiction. Assume that there exists a sequence (S_n) of samples $S_n = \{p_n^1, \dots, p_n^{|S_n|}\}$ with $\varepsilon(S_n) \rightarrow 0$ as $n \rightarrow \infty$ and $TST(S_n) \neq P(S_n)$. Using Lemma 5 and choosing an appropriate subsequence we can assume that $MST(S_n) \subset P(S_n)$ for all $n \in \mathbf{N}$. Removing the largest edge of $TST(S_n)$ from $TST(S_n)$ gives us a path through the sample points S_n . This path cannot be a Minimum Spanning Tree, because otherwise $TST(S_n) = P(S_n)$. Furthermore this edge has to be shorter than the edge $e_n = \text{conv}p_n^i, p_n^{i+1}$ which connects the endpoints of $MST(S_n)$. For all $n \in \mathbf{N}$ we consider two cases, there exists an edge

$e'_n = \text{conv} p_n^k, p_n^l$ in $TST(S_n)$ with $k \in i, i + 1 \in l$ or such an edge does not exist. By choosing a subsequence we can assume that only one of these possibilities holds for all $n \in \mathbf{N}$. Assume that this is the first one. By construction e_n is larger than all edges of $TST(S_n)$. In particular e_n is larger than e'_n , i.e.

$$|p_n^k - p_n^l| \leq |p_n^i - p_n^{i+1}|.$$

Lemma 4 tells us that there exists $N \in \mathbf{N}$ with

1. $|p_n^i - p_n^l| \leq |p_n^k - p_n^l|$ for all $n \geq N$. That means $|p_n^i - p_n^l| \leq |p_n^i - p_n^{i+1}|$ for all $n \geq N$.
2. $|p_n^k - p_n^{i+1}| \leq |p_n^k - p_n^l|$ for all $n \geq N$. That means $|p_n^k - p_n^{i+1}| \leq |p_n^i - p_n^{i+1}|$ for all $n \geq N$.

That is a contradiction to Lemma 3. Next we consider the second case. That is, e'_n does not exist for all $n \in \mathbf{N}$. We split all $TST(S_n)$ in two paths $P_1(S_n)$ and $P_2(S_n)$, with endpoints p_n^i and p_n^{i+1} . The idea is to show that both paths alone approximate γ and especially the length of γ if e'_n does not exist. We need three facts from differential geometry. First, regular embeddings have finite length [1]. Second an inequality which follows directly from the definition of length

$$\limsup \left(\text{length}(TST(S_n)) \right) \leq \text{length}(\gamma).$$

And third an inequality which can be derived from a theorem of Menger [9] (here we omit the technical proof)

$$\liminf \left(\text{length}(P_i(S_n)) \right) \geq \text{length}(\gamma), \quad i = 1, 2.$$

On the other hand we have

$$\begin{aligned} \limsup \text{length}(TST(S_n)) &= \limsup \left(\text{length}(P_1(S_n)) + \text{length}(P_2(S_n)) \right) \\ &\geq \liminf \left(\text{length}(P_1(S_n)) \right) + \liminf \left(\text{length}(P_2(S_n)) \right) \\ &\geq 2 \text{length}(\gamma). \end{aligned}$$

That is also a contradiction. □

4 Algorithms for Polygonal Reconstruction

We are looking for an efficient algorithm which always computes a simple polygon or a collection of simple polygons through the sample points. Minimum Spanning Trees can be computed efficiently, but in general the Minimum Spanning Tree of a sample is not a polygon. We have a problem if the points are not sampled densely enough. On the other hand Traveling Salesman Paths and Traveling Salesman Tours are always simple polygons (we call a polygonal line also a polygon), but they are NP-hard to compute.

Here we describe an efficient algorithm which always computes a path or a tour. The output of this algorithm need not be a simple, but later we will discuss a variant which always computes a collection of simple polygons. We start with the description of the basis algorithm, because it is easier to analyze. The algorithm holds in every iteration of its main loop a path until all sample points are vertices of this path. Again γ denotes an embedding of \mathbf{M}^1 and $S = \{p^1, \dots, p^{|S|}\}$ a sample of γ .

Algorithm 1.

1. Choose an arbitrary point $p^i \in S$ and connect it to one of its nearest neighbors in $S - \{p^i\}$. This defines a path with two vertices v^1 and v^2 .
2. Consider the path with vertex set $V = \{v^1, \dots, v^j\}$, s.t. v^k is connected to v^{k+1} for $k = 1, \dots, j - 1$. Calculate nearest neighbors v^- of $v^1 \in V$ and v^+ of $v^j \in V$ in $S - V$. Connect v^1 to v^- if $|v^1 - v^-| \leq |v^j - v^+|$. Otherwise connect v^j to v^+ .
3. Repeat step 2 until $j = |S|$.
4. If γ is closed then connect v^1 with $v^{|S|}$.

Let $A(S)$ denote the polygon which is computed by the algorithm. Next we prove that for dense samples $A(S)$ equals the polygonal reconstruction $P(S)$

Theorem 3. *Let γ be a regular embedding of \mathbf{M}^1 . Then there exists an $\varepsilon > 0$ such that $A(S) = P(S)$ for all samples S of γ with $\varepsilon(S) < \varepsilon$.*

Proof. We look at the construction of $A(S)$. In the first step we connect an arbitrarily chosen p^i to one of its nearest neighbors. From Lemma 3 we know that this neighbor has to be one of p^{i-1}, p^{i+1} if $\varepsilon(S)$ is sufficiently small. Hence p^i is connected to p^{i-1} or p^{i+1} for sufficiently small $\varepsilon(S)$.

In the second step we connect one of the endpoints of the polygon with vertex set $V = \{v^1, \dots, v^j\}$ with one of its nearest neighbors. Let without loss of generality $v^1 = p^k$ be this endpoint and $p^{k+1} = v^2 \in V$. Assume that $v^- = p^l, l \neq k - 1$. That is $|p^k - p^l| \leq |p^k - p^{k-1}|$. Then there are the following possibilities:

1. $l \subset k - 1$. From Lemma 3 we know that this cannot happen if $\varepsilon(S)$ is sufficiently small.
2. $l \subset k + 1$. By induction we can assume that (here the indices are taken modulo n)

$$v^j = p^{k+j}, v^{j-1} = p^{k+j-1}, \dots, v^1 = p^k.$$

That means $l \notin \{k, \dots, k + j\}$ and therefore $k + j \subset l$. We distinguish two cases

- (a) $l = k + j + 1$: We have $|p^{k+j+1} - p^k| < |p^{k+j+1} - p^{k+j}|$.

That contradicts Lemma 3 in the point p^{k+j+1} for sufficiently small $\varepsilon(S)$.

- (b) $l \subset k + j + 1$: We have $|p^l - p^k| < |p^{k+j+1} - p^k|$.

That contradicts Lemma 4 in the point p^k for sufficiently small $\varepsilon(S)$.

That means p^k is connected to $v^- = p^{k-1}$ for sufficiently small $\varepsilon(S)$. If we repeat the last step until $j = |S|$ we get a path in which every vertex v^j , $j = 1, \dots, |S|-1$ is connected to the same neighbors as in $P(S)$. This remains valid if γ is closed and we connect $v^{|S|}$ with v^1 . \square

Since we start the algorithm in an arbitrary sample point it is natural to ask whether we can start in all sample points simultaneously. This leads to a variant of Algorithm 1 which does not hold just one path but a collection of vertex distinct paths in every iteration of its main loop. Again let $S = \{p^1, \dots, p^{|S|}\}$ be a sample of an embedding γ .

Algorithm 2.

1. Let $\mathcal{P} := \{\{p^1\}, \dots, \{p^{|S|}\}\}$ be a set of trivial paths. That is, all paths in \mathcal{P} consist of just one vertex.
2. Connect two paths $P_i, P_j \in \mathcal{P}$ if there exist two endpoints, one of each path, such that these endpoints are nearest neighbors* of each other. That is, the set \mathcal{P} of paths is updated as follows

$$\mathcal{P} := \left(\mathcal{P} - \{P_i, P_j\} \right) \cup \{P_i \cup P_j\}$$

3. Repeat step 2 as long as connections are possible.
4. For every path in \mathcal{P} connect its endpoints, if these endpoints are adjacent via a Delaunay edge.

* A nearest neighbor of an endpoint p of a path is another endpoint of a different path which has the shortest distance to p among all such endpoints.

This algorithm also computes the polygonal reconstruction for sufficiently dense sampling. From our earlier considerations we can restrict the search for nearest neighbors in step 2 to points adjacent via a Delaunay edge. The output of Algorithm 2 is then always a subset of the edges of the Delaunay Triangulation and its running time is bounded by the time we need to compute the Delaunay Triangulation, which is $O(|S|^{\lceil d/2 \rceil})$ resp. $O(|S| \log(|S|))$ if $d = 2$, where d is the dimension of the embedding space. The output need not consist of exactly one component, in general it is a collection of simple polygons which can be open or closed. Hence Algorithm 2 can serve as a basis for heuristics which solve the reconstruction problem for non-connected simple curves. Branching points remain difficult to handle. It seems that the best starting point for heuristics which solve the reconstruction problem with branching points is a Minimum Spanning Tree.

5 Good Samples

So far we have shown existence proofs. In this section we want to characterize a class of samples for which our theorems holds. We give a checkable condition for samples to be good.

Definition 3. Let γ be an embedding of \mathbf{M}^1 , S be a sample of γ and $B_r(p)$ the ball with radius $r > 0$ and center p . We call S a good sample, if $B_r(p^i) \cap \gamma$ is homeomorphic to the unit interval $[0, 1]$ for all

$$r < (1 + \eta) \max\{|p^{i+1} - p^i|, |p^i - p^{i-1}|\}, \text{ for all } i \in \{1, \dots, |S|\},$$

where $\eta > 0$ is an arbitrarily small constant.

The scaling with $(1 + \eta)$ prevents problems with degenerate samples, where two or more sample points exist which all have the same distance to one and the same sample point.

Good samples fulfill the properties stated in our key Lemmas 3 and 4. In the proofs of our theorems we construct contradictions to these Lemmas. Thus for regular embeddings of $[0, 1]$ and good samples S we have $P(S) = TSP(S) = MST(S) = A(S)$ and for regular embeddings of \mathbf{S}^1 and good samples S we have $P(S) = TST(S) = A(S)$.

Finally we want to show that good samples exist and that samples that contain a good sample as a subset are good samples itself.

Lemma 6. Let γ be a regular embedding of \mathbf{M}^1 . Then there exists a good sample of γ .

Proof. Set $r(p) = \sup\{r : B_{r'}(p) \cap \gamma \simeq (0, 1) \text{ for all } r' < r\}$ for all $p \in \gamma$. We show that

$$\inf_{p \in \gamma} r(p) > 0.$$

Then every sample $S = \{p^1, \dots, p^{|S|}\}$ of γ with

$$\sup\{|p^{i+1} - p^i| : i = 1, \dots, |S|\} < \frac{\inf_{p \in \gamma} r(p)}{1 + \eta}$$

is a good sample. At first we show

$$\sup\{r : B_{r'}(p) \cap \gamma \simeq (0, 1) \text{ for all } r' < r\} > 0 \text{ for all } p \in \gamma. \tag{1}$$

Assume the contrary. Then there exists $p \in \gamma$ such that for all $n \in \mathbf{N}$ there exists $0 < r_n < \frac{1}{n}$ with $B_{r_n}(p) \cap \gamma$ is not homeomorphic to $(0, 1)$. Thus we can find $p_n, p'_n \in B_{1/n}(p) \cap \gamma$ with $p \subset p_n \subset p'_n$ (or $p'_n \subset p_n \subset p$, what can be handled in a similar way) and $|p'_n - p| < |p_n - p|$. Look at the triangle with corner points p, p_n and p'_n . From the law of cosines together with $|p'_n - p| < |p_n - p|$, we find that the angle at p_n has to be smaller or equal than $\frac{\pi}{2}$. By construction (p_n) and (p'_n) converge to p , but that is a contradiction to Lemma 1. Thus inequality (1) is valid for all $p \in \gamma$. Next assume that

$$\inf_{p \in \gamma} r(p) = 0,$$

i.e. we can find a sequence (p_n) in γ with $\lim_{n \rightarrow \infty} r(p_n) = 0$. By choosing an appropriate subsequence we can assume that (p_n) converges to $p \in \gamma$, because

\mathbf{M}^1 is compact. From (1) we have $r(p) > 0$. We find $N \in \mathbf{N}$ with $p_n \in B_{r(p)}(p)$ for all $n \geq N$ and

$$B_r(p_n) \cap \gamma \subset B_{r(p)}(p) \cap \gamma$$

for all $r < r(p) - |p_n - p|$. Since $\lim_{n \rightarrow \infty} |p_n - p| = 0$ we have

$$\liminf r(p_n) > \frac{r(p)}{2} > 0.$$

That is a contradiction. □

Lemma 7. *Let γ be an embedding of \mathbf{M}^1 , S a sample of γ and $p \in \gamma - S$. Then $S \cup \{p\}$ is again a good sample.*

Proof. There has to exist an index $i \in \{1, \dots, |S|\}$ with $p^i \subset p \subset p^{i+1}$. We have

$$p \in B_{|p^{i+1} - p^i|}(p^i) \cap \gamma,$$

because otherwise S cannot be a good sample. The open balls $B_r(p)$ with radius

$$r < (1 + \eta) \max\{|p^{i+1} - p|, |p - p^i|\}$$

are all covered by the open ball $B_{|p^{i+1} - p^i|}(p)$, if η is sufficiently small. That is, the sets $B_r(p) \cap \gamma$ are open, connected subsets of $B_{|p^{i+1} - p^i|}(p^i) \cap \gamma$ and therefore are homeomorphic to $(0, 1)$. □

Acknowledgment. I want to thank my advisor Emo Welzl and Nicola Galli for helpful discussions.

References

1. A.D. Alexandrov, Yu. G. Reshetnyak *General Theory of Irregular Curves*, Kluwer Academic Publishers (1989)
2. N. Amenta, M. Bern, D. Eppstein *The Crust and the β -Skeleton: Combinatorial Curve Reconstruction*, Graphical Models and Image Processing 60/2:2, pp. 125-135 (1998)
3. D. Attali *r-Regular Shape Reconstruction from Unorganized Points*, Proc. 13th Ann. ACM Symp. on Computational Geometry 1997, pp. 248-253 (1997)
4. M. de Berg, M. van Kreveld, M. Overmars, O. Schwarzkopf *Computational Geometry*, Springer (1997)
5. F. Bernardini, C. L. Bajaj *Sampling and Reconstructing Manifolds Using Alpha-Shapes*, Proc. of the Ninth Canadian Conference on Computational Geometry 1997, pp. 193-198 (1997)
6. T. H. Cormen, C. E. Leiserson, R. L. Rivest *Introduction to Algorithms*, MIT Press (1990)
7. H. Edelsbrunner *Algorithms in Combinatorial Geometry*, Springer (1987)
8. D.G. Kirkpatrick, J.D. Radke *A framework for computational morphology*, Computational Geometry (G. Toussaint, ed.), Elsevier pp. 217-248 (1983)
9. K. Menger *Untersuchungen über eine allgemeine Metrik. Vierte Untersuchung. Zur Metrik der Kurven*, Math. Ann. 103, pp. 467-501 (1932)
10. J. O'Rourke, H. Booth and R. Washington *Connect-the-dots: A New Heuristic*, Comp. Vision, Graph. Image Proc. 39, pp. 258-266 (1987)