Curve Reconstruction and the Traveling Salesman Problem

Dissertation
zur Erlangung des Grades
des Doktors der Ingenieurswissenschaften (Dr.-Ing.)
der Naturwissenschaftlich-Technischen Fakultät I
der Universität des Saarlandes

von

Ernst Althaus

Saarbrücken
27. April 2001

Datum des Kolloqiums: 27. April 2001

Dekan der naturwissenschaftlich-technischen Fakultät I:
Professor Dr. Rainer Schulze-Pillot-Ziemen

Gutachter:
Professor Dr. Kurt Mehlhorn, MPI für Informatik, Saarbrücken
Professor Dr. Raimund Seidel, Universität des Saarlandes, Saarbrücken

Abstract

An instance of the curve reconstruction problem is a finite sample set V of an unknown collection of curves γ. The task is to connect the points in V in the order in which they lie on γ. Giesen [Gie99] showed recently that the Traveling Salesman tour of V solves the reconstruction problem for single closed curves under otherwise weak assumptions on γ and V; γ must be a single closed curve. We extend his result in several directions.

- we weaken the assumptions on the sample,
- we show that the Traveling Salesman based reconstruction also works for single open curves (with and without specified endpoints) and for collections of closed curves,
- we give alternative proofs,
- we show that in the context of curve reconstruction, the Traveling Salesman tour can be constructed in polynomial time.

Furthermore we report on experiments with a number of recent curve reconstruction algorithms.

Zusammenfassung

Die Eingabe eines Kurvenrekonstruktionsproblems ist eine endliche Menge V von Sampelpunkten auf einer unbekannten Kurve γ. Die Aufgabe besteht darin einen Graphen $G = (V, E)$ zu konstruieren, in dem zwei Punkte genau dann durch eine Kante verbunden sind, wenn die Punkte in γ benachbart sind. Giesen [Gie99] hat kürzlich gezeigt, daß die Traveling Salesman Tour durch die Punkte V das Kurvenrekonstruktionsproblem für einzelne geschlossene Kurven unter ansonsten schwachen Bedingunen löst; γ muß allerdings eine einzelne geschlossene Kurve sein. Wir erweitern dieses Ergebnis in mehrere Richtungen.

- wir schwächen die Bedingungen an das Sample ab,
- wir zeigen, daß Traveling Salesman basierte Rekonstruktionen auch für offene Kurven (mit und ohne spezifizierte Endpunkte) und für mehrere geschlossene Kurven funktionieren,
- wir geben andere Beweise,
- wir zeigen, daß die Traveling Salesman Probleme die im Zusammenhang mit Kurvenrekonstruktion auftreten, in polynomieller Zeit gelöst werden können.

Ausserdem berichten wir über Experimente mit einigen kürzlich entwickelten Kurvenrekonstuktionsalgorithmen.

Acknowledgments

This thesis couldn't have been written without the help of many people.

First of all, I thank my advisor Prof. Dr. Kurt Mehlhorn. He roused my interest in curve-reconstruction and discussed all aspects in and beyond my work in the institute. His guidance, advice, and encouragement have been invaluable throughout my master's and doctoral studies.

During the last two years, I have had many fruitful discussions with almost every member of our group. In particular, I thank Stefan Funke and Edgar Ramos, who are also interested in curve–reconstruction, and Michael Seel, who helped me in submitting the LEDA-Extention-Package. Beside the members of our institute, I had helpful discussions with Joachim Giesen and Nina Amenta about their papers.

Furthermore, I am grateful to Elmar Schömer and Mark Ziegelmann for providing interesting applications, and to Tamal Dey for providing the code of his algorithm. Special thanks to my masters student Christian Fink for his fast implementation of our ideas of the branch-&-cut algorithm for the surface reconstruction algorithm from planar contours.

In addition, I thank Maike Mellerke, Bobbye Pernice, Christian Lennerz, Volker Priebe, Sven Thiel, and Thomas Warken for helpful comments, proof reading, and correcting this thesis.

Contents

1 Introduction

The problem of reconstructing a shape from a given finite set of points has attracted much attention in the literature during the last twenty years. Its importance arises from a wide area of applications, mainly in reverse engineering. For the most important problem, the reconstruction of a surface in the Euclidean space, many algorithms have been proposed that produce good approximations of the surfaces. The drawback of these algorithms is that they provide no guarantee for the correctness of the returned solution. Recently, these reconstruction problems have been investigated from a theoretical point of view. The results are algorithms that provably solve the reconstruction problem for a certain class of shapes in a somewhat idealized setting.

A formal specification of this problem is stated in the next section. Before we mention some applications for the class of shapes we have looked at, we introduce the Traveling Salesman Problem. After that, we summarize our contribution to this problem and give an outline of the contents of the thesis.

Problem Description

In general a *shape* is defined as a subset of the Euclidean space. Given a finite set of points $S \subset \mathbb{R}^d$, called the *sample points*, the *shape reconstruction problem* asks for a shape that approximates S. We are interested in reconstruction algorithms with *guaranteed performance*, i.e. , algorithms that provably *solve* the reconstruction problem under certain assumptions on the shape and the sample set. In the case in which the shape is a *curve*, i.e. , a one–manifold, the correct solution can easily be defined as the *polygonal reconstruction*, i.e. , the graph G on S so that two points in S are connected by an edge of G iff the points are adjacent in the curve. In the case of a *surface* it is not obvious how to define what is meant by the correct solution. A possible answer is given by Amenta and Bern [AB98].

Algorithms that provably solve the problem for certain classes of curves or surfaces have been recently proposed. Most of these algorithms require a shape that is smooth, i.e. , twice differentiable and a sample that is rather dense, i.e. , that for every point p on the shape, there is a sample point of distance at most $\epsilon f(p)$, where ϵ is a constant and f is a function that describes the local complexity of the curve. Algorithms for the surface reconstruction problem are mostly extensions of similar algorithms for curves. Therefore it seems reasonable to restrict one's attention to the simpler case of curve reconstruction, to gain insight for the general case. In this thesis, we address the special case of a curve in $\mathbb{R}^2$ that is not assumed to be smooth. Figure 1 gives an example. On the left the input, i.e. , the sample points are visualized. The result of our algorithm is shown in the right. Note that a human being would probably immediately see the reconstruction. The challenge is to find algorithms that are almost as good as the human eye.

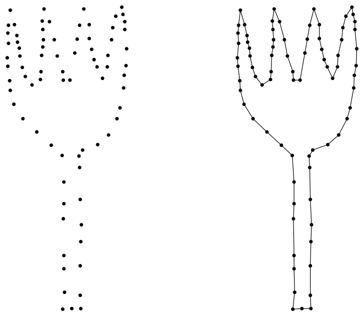

Figure 1: Part (a) shows a finite set S of points, part (b) shows the Traveling Salesman tour of the points.

The Traveling Salesman Problem

A salesman who has to visit a given set of cities and after that has to return to his home should select the order in which he is going to visit the cities, so that the total distance is minimized. The problem of ordering the cities in such a way is called the *Traveling Salesman Problem (TSP)*. The Traveling Salesman Problem is one of the oldest and most well studied problems in combinatorial optimization. Lawler, Lernstra, Rinnooy Kan, and Shmoys [LLRKS92] even call their book devoted to the TSP: *The Traveling Salesman Problem - A Guided Tour of Combinatorial Optimization.* This title demonstrates the importance of the problem. It has served as a testbed for almost every new algorithmic idea.

The traveling salesman problem has been shown to be NP–complete, even in the special case where the distance between two cities is given by their Euclidean distance [GGJ78, Pap77]. There were many investigations of special cases that are solvable in polynomial time. The first overview was given in the book of Lawler et al. [LLRKS92] and later updated by Burkard and others [BDvD$^+$98].

We follow the proposal of Giesen in using the Traveling Salesman Tour as the reconstruction for a curve. Our main contribution to this issue is the proof that the instances arising in curve reconstruction are solvable in polynomial time. This is one of the rare special cases that has important applications in practice.

Applications of Curve Reconstruction

Although the most important application of the reconstruction problem is generating a computer model of an existing three dimensional object that is scanned by some 3D scanner, there are some applications where the underlying shape is a curve. All these applications do not exactly fit our idealized setting, since there can be some noise in the input sample, the sample points can be disturbed, the sampling condition can be violated, or the curves do not necessarily belong to the class of curves that can be reconstructed with the known algorithms.

Computational Morphology

The central problem in computer vision (see [JT92]) starts with a grey–level picture of a scene and asks for a description of the scene. There has been much research on the low–level vision, i.e. , the analysis of the picture to separate objects from the background or to separate different objects. One class of methods returns so–called *dot patterns*, i.e. , a set of points describing the boundary of the shape. The shape of the corresponding object had to be found by a curve reconstruction algorithm.

There are various applications that fit into this setting. We have collected some of them.

Geographic Analysis. Assume we are given an aerial survey of an unexplored terrain, i.e. , from satellite images. Different types of ground have different textures in these images. There are sophisticated methods for extracting the boundaries of these textures. In another setting, boundaries are extracted from spatial data of an area. In both cases these boundaries are given by dot patterns and the associated polygon has to be computed.

Figures 2 gives an example. The data is from the USGS EROS Data Center [USG]. We have extracted the boundary by a simple algorithm that detects a boundary point if the image–pixel changes between water and land in a left–right scan. Since this algorithm finds some clusters of points on the boundary, we removed some points and applied our reconstruction algorithm.

Image Interpretation. There are many applications, e.g. , OCR, automatic traffic census, and many others, where some shapes of an image has to be interpreted.

One excellent descriptor for a shape is the medial axis as already proposed by Blum [Blu67] in 1967. Similarity of objects are often measured by some similarity measure on the medial axes.

Figure 3 shows a OCR–example. The boundary of the letters is extracted and reconstructed by our algorithm. Then we have computed an approximation to the medial axis.

3D Surface Reconstructions from Contours. In a very important special case of 3D surface reconstruction the image is given as a set of contours, i.e. , the image is scanned by parallel slices. Important applications of these scans are found in computer topography (CT) and magnetic resonance imaging (MRI).

There are many surface reconstruction algorithms for this special case. All assume that each slice is given as a set of polygons, i.e. , that the reconstruction problem for the single slices is already solved. The reconstruction problem for a slice can be seen as a curve reconstruction problem.

We give an example in Figure 4. The data comes from the visual human project [VHP]. For every image we try to extract the boundaries. We collect all boundaries and use an appropriate algorithm to obtain a 3D–reconstruction.

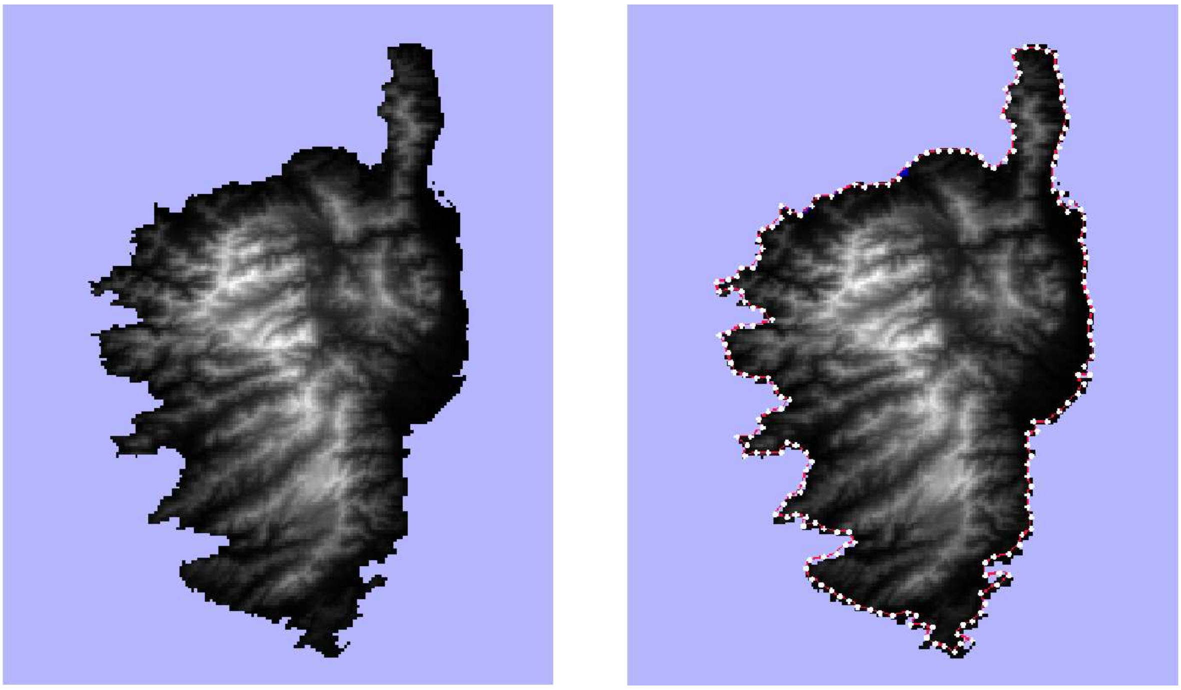

Figure 2: Part (a) shows a map of Corsica and Part (b) shows the reconstruction of our algorithm.

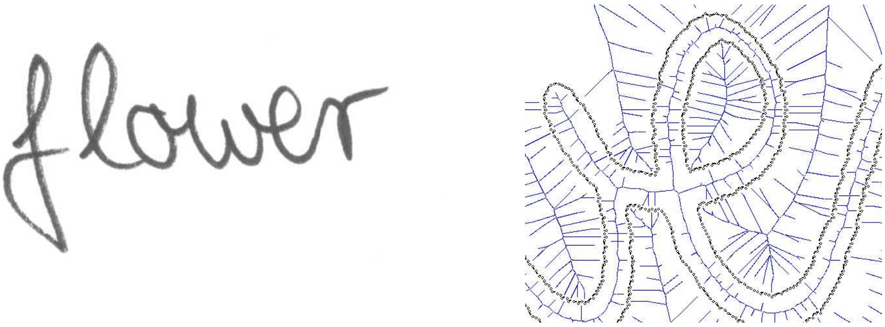

Figure 3: Part (a) shows a scan of the word “flower” and Part (b) shows the points of the boundary and an approximation of the medial axis of a part of the picture.

Plotting of Implicit Functions

We want to mention one application in which the input does not come from a scanned image.

In mathematics, some functions are given by the definition $f(x, y) = 0$. Such a definition is called an *implicit* definition. Many mathematical software packages support

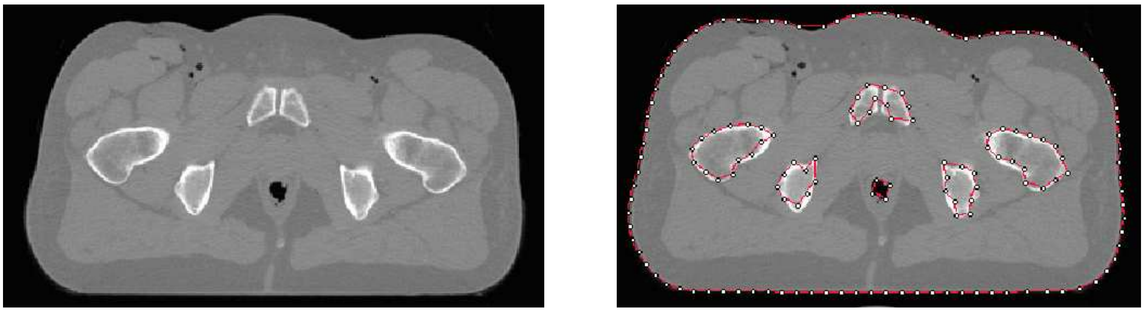

Figure 4: Part (a) shows a CT–image of a slice of a human pelvis and Part (b) a reconstruction.

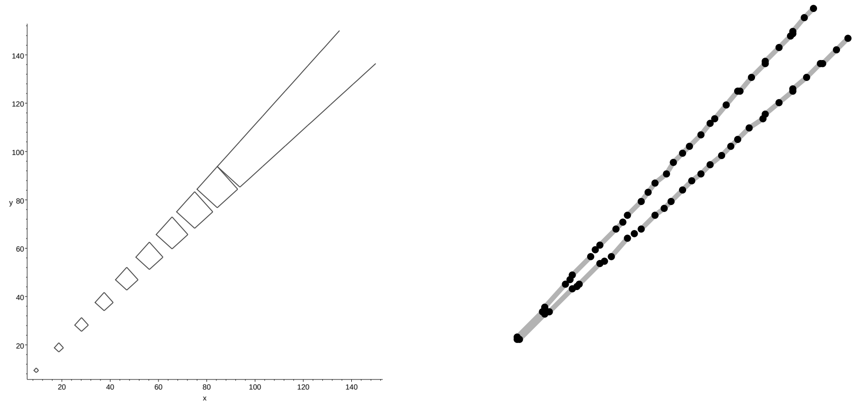

Figure 5: Part (a) shows the maple plot of the function $5|y - x| - y = 0$ and Part (b) shows the reconstruction of the sample points that were produced by maple.

the plot of such implicit functions, if they can compute for every x all values y with $f(x, y) = 0$. The result of these plots is rather poor, if the function is self-intersecting or has sharp corners. To plot such an implicit function, one computes a large set of points on this function and reconstructs it with an appropriate algorithm (see Figure 5).

Analogously, we can plot 3D implicit functions, if we take care on our sample points. By computing sample points on parallel slices, we can reconstruct these slices and than use an appropriate algorithm to compute the 3D–reconstruction.

Our Contribution

If the curve is closed, smooth, and uniformly sampled, several methods for the curve reconstruction problem are known to work ranging over minimum spanning trees [FG94], α-shapes [BB97, EKS83], β-skeletons [KR85], and r-regular shapes [Att97]. A survey of these techniques appears in [Ede98]. The case of non-uniformly sampled closed curves was first treated successfully by Amenta, Bern and Eppstein [ABE98] and subsequently improved algorithms such as [DK99, Gol99] appeared. Open non-uniformly sampled curves were treated in [DMR00]. All the papers mentioned so far require the underlying curve to be smooth.

We review the algorithms for non–uniformly sampled curves and show a variant of the algorithm of Dey and Kumar that requires the least dense sample of any known algorithm.

Giesen [Gie99] recently obtained the first result for non-smooth curves. He considered the class of *benign semi-regular curves.* An (open or closed) curve is *semi-regular* if a left and a right tangent exists in every point of the curve; the two tangents may however be different. A semi-regular curve is *benign* if the turning angle at every point of the curve is less than π. Giesen showed that the Traveling Salesman path of the sample set S solves the curve reconstruction problem for uniformly sampled benign open semi-regular curves. More precisely, he showed that for every benign semi-regular curve γ there exists a positive ϵ so that the Traveling Salesman path (tour) of S is a polygonal reconstruction provided that for every $x \in \gamma$ there is a $p \in S$ with $\|xp\| \leq \epsilon$, where $\|xy\|$ is the Euclidean distance of the two points x and y. Giesen's result is an existence result; he did not quantify ϵ in terms of properties of the curve γ. We extend Giesen's result in three directions:

- We relate ϵ to local properties of the curve γ and show that the Traveling Salesman tour (path) solves the reconstruction problem even if the sampling is non-uniform. For smooth curves our sampling condition is similar to the one used in [ABE98, DK99, Gol99, DMR00].
- We show that the Traveling Salesman tour (path) can be constructed in polynomial time if the sampling condition is satisfied.
- We give a simplified proof showing that the Traveling Salesman tour (path) solves the curve reconstruction problem.
- The TSP–algorithm reconstructs only single closed curves. We were able to extend the algorithm for collections of closed curves.

Furthermore we show the following results concerning curve reconstruction:

- In the proof of the correctness of the TSP–algorithm we introduce a sampling condition, different from those in the previous algorithms. We show that our sampling condition is implied by the other sampling conditions.
- A curve–reconstruction algorithm returns a curve it "believes" to best approximate the curve. If one adds additional sample points on the returned curve one would expect that the algorithm returns the same curve. We call an algorithm *self consistent* if it has this property. We investigate which of the algorithms are self consistent.
- All known algorithms return a subgraph of the Delaunay Diagram of the sample points. We show that the Delaunay Diagram suffices for benign semi–regular curves too.

- We relate our result to another known polynomial–solvable special case of the Euclidean Traveling Salesman problem.

Additionally we describe a testbed for curve reconstruction algorithms and report on an experimental evaluation of the curve reconstruction algorithms of Amenta, Bern, and Eppstein (ABE), Dey and Kumar (DK), Gold (Gold), Dey, Mehlhorn, and Ramos (DMR), and the TSP-algorithm. We also report on experimental results with some TSP-heuristics.

Outline

In Section 2 we collect some facts concerning linear programming and computational geometry, which we suppose to be known in the following sections. In Section 3, we review recent curve-reconstruction algorithms. The main results are proven in Section 4, where we show that the Traveling Salesman path (tour) solves the reconstruction problem in polynomial time for a large class of curves. In Section 5 we give further results concerning curve reconstruction and in Section 6 we report on experiments with recent algorithms. Finally, we state some related open problems in Section 7.

2 Preliminaries

Since we use techniques from different research areas, in particular linear programming and computational geometry, we introduce the facts we use in later sections. We start with some notations of graph theory. Then we give a short introduction to the basics of linear programming and Lagrangian relaxation. In the last section we introduce Delaunay triangulations and Voronoi diagrams.

The results in this chapter are not original. They are presented only for the sake of completeness.

2.1 Undirected Graphs

We assume that the reader is familiar with basic graph theory. The purpose of this section is to introduce some notations we use. For an introduction to basic graph theory we refer to [Meh84]. Since we only use undirected graphs, we refrain from introducing the notations for directed graphs.

A *undirected graph* $G = (V, E)$ consists of a set V of nodes and a set E of edges, where $E \subseteq \{\{u, v\} \mid u, v \in V, u \neq v\}$. If $E = \{\{u, v\} \mid u, v \in V, u \neq v\}$ we call G the complete graph over V.

For an edge $e = \{u, v\}$, abbreviated as uv, we call u and v the endpoints of e. The endpoints u and v of an edge uv are called the *incident nodes* of the edge and e is said to be *adjacent* to u and v. Two edges are called *adjacent* if they share a common endpoint.

For a subset V' of V, we call the set of edges E' that have exactly one endpoint in V' the *cut* of V' and denote this set by $\delta(V')$. The edges that have both endpoints in V' are called the *induced edges* of V', denoted by $\gamma(V')$. The graph $G = (V', \gamma(V'))$ the *induced subgraph* of V'. For a subset E' of E the induced subgraph is defined by (V', E') where $V' = \{v \in V \mid \exists uv \in E'\}$.

A *path* between two nodes u and v is an alternating tuple $(u, uw_1, w_1, w_1w_2, w_2, w_2w_3, w_3, \ldots w_kv, v)$ of nodes of V and edges of E starting with u and ending with v so that every node is adjacent to its neighboring edges and so that all nodes are pairwise distinct. We call the nodes (edges) of the tuple the nodes (edges) of the path.

A path is called a *Hamilton path* of G if the nodes of the path are all nodes of the graph. A *Hamilton cycle* is the union of an edge $uv \in E$ and a Hamilton path between u and v.

A subset E' of E is called a *spanning tree* of G if there is exactly one path between every two distinct nodes of V in the induced subgraph of E'.

If there is a bijective mapping between the nodes V of a graph and a set S of points in the plane, we call the graph an *embedded graph*. We identify the nodes of the graph with the corresponding point in the plane and edges uv with the segment between the corresponding points of u and v.

2.2 Linear Programming

We briefly introduce linear programming and cite the theorems we use in the following sections. The results are collected from the sources [NW88, Chv83, Sch86, Wol98].

A linear program (LP) is given by

$$\begin{array}{lrcl} \min & c^T x & & \\ s.t. & A_1 x & \leq & b_1 \\ & A_2 x & = & b_2 \\ & A_3 x & \geq & b_3. \end{array}$$

The term $c^T x$ is called the *objective function* of the linear program and the equalities and inequalities of the descriptions are called the *constraints* of the LP.

A linear program has either no solution, is unbounded or has a finite solution. Many optimization problems can be formulated as linear programs or integer linear programs. An integer linear program, or integer program for short, is a linear program with the additional constraint that some of the variables have to be integral. Linear programs can be solved in polynomial time, whereas finding an optimal solution of integer programs is NP-complete in general.

By replacing $A_3 x \geq b_3$ by $-A_3 x \leq -b_3$ and $A_2 x = b_2$ by $-A_2 x \leq -b_2, A_2 x \leq b_2$ we get an LP of the form

$$\begin{array}{ll} \min & c^T x \\ s.t. & Ax \leq b \end{array}$$

without changing the optimal solution or the objective function of this solution. A maximization problem $\max c^T x$ can be written as minimization problem by $\min -c^T x$ and vice versa. The optimal solution does not change, the objective function value of the optimal solution is multiplied by -1. Very common inequalities are $x \geq 0$ and $x \leq 0$ for some variable x. Variables, for which the inequality $x \geq 0$ is valid, are called *non-negative* variables. If for a variable the inequality $x \leq 0$ is valid we call the variable *non-positive*. If none of the two inequalities are valid, we call the variable *free*.

A further transformation leads to the *standard form* of an LP. First we replace each non-positive variable x by a non-negative variable x^- and substitute $x = -x^-$. For each free variable x of the LP we introduce two new non-negative variables x^+ and x^- and substitute $x = x^+ - x^-$. Then we introduce a slack variable y_i for every column of the constraint matrix and replace $a_i x \leq b_i$ by $a_i x + y_i = b_i$. In this manner we get an LP of the form

$$\begin{array}{lrcl} \min & c^T x & & \\ s.t. & Ax & = & b \\ & x & \geq & 0. \end{array}$$

The objective function value of the optimal solution of this LP is the same as the objective function value of the optimal solution of the original problem. Furthermore an optimal solution of the new LP transforms into an optimal solution of the original LP by setting $x = x^+ - x^-$ for the free variables and $x = -x^-$ for the non-positive variables. We ignore the values of the slack variables y.

2.2.1 Some Linear Algebra

For a matrix A, A_i denotes the i-th row and A^j denotes the j-th column. The $n \times n$ *unit matrix* is denoted by I_n.

A set of points $x^1, \ldots, x^k \in \mathbb{R}^n$ is *linearly independent* if the unique solution of the system $\sum_{i=1}^k \lambda_i x^i = 0$ is $\lambda = 0$. It is *affinely independent* if the unique solution of $\sum_{i=1}^k \lambda_i x^i = 0, \sum_{i=1}^k \lambda_i = 0$ is $\lambda = 0$.

Note that the maximum number of linearly independent vectors in $\mathbb{R}^n$ is n, the maximal number of affinely independent vectors in $\mathbb{R}^n$ is $n+1$. The following corollary makes a connection between linear and affine independence.

Corollary 1 *The following statements are equivalent:*

- $x^1 \ldots, x^k \in \mathbb{R}^n$ *are affinely independent.*
- $x^2 - x^1, \ldots, x^k - x^1 \in \mathbb{R}^n$ *are linearly independent.*
- $(x^1, 1), \ldots, (x^k, 1) \in \mathbb{R}^{n+1}$ *are linearly independent.*

The following corollary leads to the definition of the rank of a matrix.

Corollary 2 *If A is an $m \times n$ matrix, the maximum number of linearly independent rows of A, viewed as vectors $A_i \in \mathbb{R}^n$, equals the maximum number of linearly independent columns, viewed as vectors $A^j \in \mathbb{R}^m$.*

The maximum number of linearly independent rows of a matrix A is called the *rank* of A and is denoted by rank(A). If A is a square matrix, i.e. , the number of columns is equal to the number of rows, and has full rank, i.e. , the rank is equal to the number of rows/columns, we call the matrix *non-singular*. If the square matrix does not have full rank, we call it *singular*.

Corollary 3 *If $\{x \in \mathbb{R}^m \mid Ax = b\} \neq \emptyset$, the maximum number of affinely independent solutions of $Ax = b$ is $n + 1 - \mathrm{rank(A)}$.*

2.2.2 Polyhedra

A *polyhedron* $P \subseteq \mathbb{R}^n$ is the set of points defined by a finite number of linear inequalities, i.e. , $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$, where (A, b) is an $m \times (n+1)$ matrix. A polyhedron is said to be *rational* if there exists an $m' \times (n+1)$ matrix (A',b') with rational coefficients so that $P = \{x \in \mathbb{R}^n \mid A'x \leq b'\}$. If there is a constant U for a polyhedron P so that $P \subseteq [-U, U]^n$, the polyhedron is called *bounded polyhedron*, or *polytope*. A set C is called *convex*, if $x, y \in C$ implies $\lambda x + (1-\lambda)y \in C$ for all $0 \leq \lambda \leq 1$. The *convex hull* of a set of points S is the smallest convex set containing S.

Theorem 4 (Finite Basis Theorem) *A subset S of $\mathbb{R}^n$ is a polytope if and only if it is the convex hull of a finite number of points.*

2.2.3 Solving Linear Programs

We briefly present two different methods for solving linear programs: The simplex method, which is the most commonly used method in practice, because it is very efficient, although no polynomial bound on the running time is known, and the ellipsoid method, which is the most cited method in theory, since it has a polynomial running time, but is never used in practice, since there is no efficient implementation. We do not present interior points methods, which are practically efficient and provide a polynomial running time.

The Simplex Method

The basic primal simplex method requires a LP in standard form. Look at the LP

$$\begin{array}{rrcl} \min & c^T x & & \\ s.t. & Ax & = & b \\ & x & \geq & 0, \end{array}$$

for an $m \times n$ matrix A and an m-vector b. Assume that A has full row rank, so $rank(A) = m$. Let $B = \{B_1, \ldots, B_m\}$ a set of column indices and $N = \{1, \ldots, n\} \setminus B$. Let $A_B = (A^{B_1}, \ldots A^{B_m})$ the $m \times m$ matrix consisting of the columns given by B. A_B is called a *basis* of A, if A_B is non-singular. The solution $x_B = A_B^{-1}b$, $x_N = 0$ is called a *basic solution* of $Ax = b$. We call a basic solution *primal feasible*, if $A_B^{-1}b \geq 0$.

A point $x \in p$ is an *extreme point* of a polyhedron P, if there are no two points $y, z \in P, y \neq z$ so that $x = (y + z)/2$. We can characterize the extreme points by bases of the matrix A.

Lemma 1

- *Let x be an extreme point of a polyhedron $P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$, for a matrix A with full row rank. Then there exists a (at least one) basis A_B of A, so that $x = A_b^{-1}x_B$.*
- *If x is a primal feasible basic solution, than x is an extreme point.*

We call an extreme point, which has several bases, a *degenerate solution*.

Geometrically, the simplex method starts with a feasible extreme point x^0. It then iteratively computes a feasible extreme point x^{i+1} that provides a better objective function value. The algebraic counterpart to extreme points are the bases of the matrix A. If the actual solution is not degenerate, there is a rather simple calculation that provides the information, if the addition of an index in the actual basis leads to a basis with a better objective function value. If there is no such variable, the actual solution is also the optimal solution. A further calculation provides the indices, which we can remove of the basis to preserve primal feasibility. If there is no such variable, we know that the problem is unbounded. If we are in a degenerate basis, we perform the same algebraic steps. The problem is, that in this case, we could end up in the same extreme point.

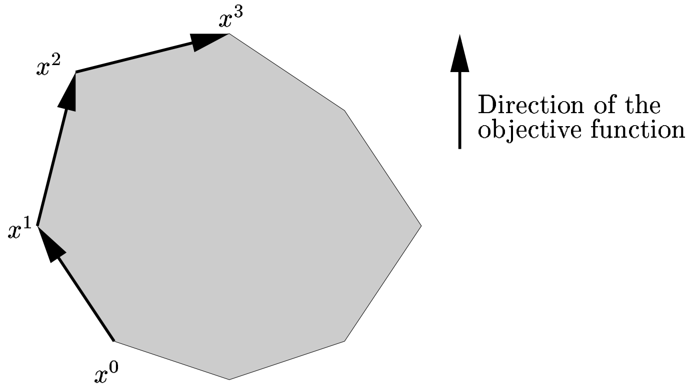

Figure 6: Geometric illustration of the simplex method: Starting at the feasible point x^0, we iteratively move segment by segment to a better point, until we end up in x^3, which we cannot improve further.

Geometrically, an exchange of an index B_i for another of a basis B corresponds to a move along the line defined by $A_{B \setminus B_i} x_B = 0$ until a further constraint is met. See Figure 6 for an illustration.

There are efficient implementations of a single iteration of the simplex method. The crucial point for an efficient implementation is to get a small number of iterations. The following questions and problems arise:

- If there are several indices that would lead to a better objective function value, we must decide which of them to choose. There are several rules leading to rather few iterations in most of the practical applications, but there is no polynomial bound for the number of iterations on any rule. Moreover, it is not known if there can be a rule, leading to a polynomial iteration bound.
- Analogously, we have to choose an index if there are several variables that can leave the basis. Most implementations choose arbitrarily one of those variables. This selection seems to be less critical for the total number of iterations.
- If we are in a degenerate basis, the algorithm could run in a cycle and hence would not terminate. There are rules governing how to select the indices entering and leaving the basis that guarantee termination. Since these rules lead to more iterations in practice, they are only applied if the actual extreme point does not change for some iterations with the rule that does not guarantee termination.

For the primal simplex method, we need a feasible basis to start with. Often, a feasible basis is known for a specific problem. A general solver can solve the *auxiliary problem*, an LP for which a feasible basis is known, and whose optimal solution is feasible for the original constraint system. Look at the optimization problem

$$\begin{array}{lrcl} \min & c^T x & & \\ s.t. & Ax & = & b \\ & x & \geq & 0, \end{array}$$

where A is a $m \times n$–matrix and without loss of generality$b \geq 0$ (if $b_i < 0$ for some row i, multiply this line by -1). We introduce m new variables $y_1 \ldots y_m$. The auxiliary problem is written as follows:

$$\begin{array}{lrcl} \min & y & & \\ s.t. & Ax + I_m y & = & b \\ & x & \geq & 0 \\ & y & \geq & 0. \end{array}$$

A feasible solution for this problem is $x = 0, y = b$. If the original problem has a solution, say x^*, the solution $x = x^*, y = 0$ is valid for the auxiliary problem. Furthermore it is optimal, since the optimal function value of this solution is 0 and every feasible solution has a non-negative objective function value, since $y \geq 0$. If there is a feasible basis for which $y = 0$, this basis is also feasible for the following problem:

$$\begin{array}{lrcl} \min & c^T x & & \\ s.t. & Ax + I_m y & = & b \\ & x & \geq & 0 \\ & y & = & 0. \end{array}$$

This problem has the same optimal solution as the original problem, except that the additional variables y have value 0. The only reason for keeping the variables y in the LP is that the basis obtained from the auxiliary problem could contain some of the y-variables (in the case of a degenerate basis) and so the basis is not a basis of the original LP. Each variable y_i which is not in the basis can be deleted at once. Similarly, if a variable y_i leaves the basis during the second optimization process, it can be deleted from the system.

We just described the *primal simplex method*. In every iteration, we have a primal feasible basis and we iterate until there is no index whose addition could improve the objective function value. There is also a *dual simplex method*. In every iteration we have a basis, so that there is no index whose addition could improve the objective function value, but it is not necessarily primal feasible, i.e. , there can be indices i with $x_i < 0$. In an iteration we move in a direction that increases the value of a variable whose value is negative, preserving the optimality of the new extreme point, if it is primal feasible.

The Ellipsoid Method

The ellipsoid method is a method for solving the *strict membership problem*, i.e. , given integers m and n, an integer $m \times n$ matrix A, and an integer m-vector b, find a point in $P_< = \{x \in \mathbb{R}^n \mid Ax < b\}$ or show that $P_< = \emptyset$. It can be extended to solve LPs.

A $n \times n$ matrix D is *positive definite*, if $x^T D x > 0$ for all $x \in R^n \setminus \{0\}$. An *ellipsoid* with center y is a set $E = \{x \in R^n \mid (x-y)^T D^{-1} (x-y) \leq 1\}$, where D is a positive definite $n \times n$ matrix. We write this ellipsoid as $E(D, y)$. The volume of an ellipsoid E is denoted by vol(E). We refrain from giving an exact definition of the volume of a subset S of $\mathbb{R}^n$.

Let T be the largest integer of A and b. We assume that $P_<$ is bounded. An important observation is that if $P_< \neq \emptyset$ then $P_< \subset E((nT)^n \cdot I_n, 0)$ is and $\mathrm{vol(P_<)} > \mathrm{n^{-n}(nT)^{-n^2(n+1)}}$.

The ellipsoid method iteratively computes ellipsoids that cover $P_<$. The volume of these ellipsoids decreases geometrically. It works as follows:

- Start with the ellipsoid $E(D^0, x^0)$, where $x^0 = 0$ and $D^0 = T^n \cdot I_n$. Let $k = 0$.
- If $\mathrm{vol(E(D^k, x^k))} < \mathrm{n^{-n}(nT)^{-n^2(n+1)}}$, we know that $P_<$ is empty, because otherwise we have a contradiction according to $n^{-n}(nT)^{-n^2(n+1)} < \mathrm{vol(P_<)} < \mathrm{vol(E(D^k, x^k))} < \mathrm{n^{-n}(nT)^{-n^2(n+1)}}$.
- If $x^k \in P_<$, we are done. Otherwise let i be an index so that $A_i x^k \geq b_i$.
- Find an ellipsoid $E(D^{k+1}, x^{k+1})$ containing the relevant part $E(D^k, x^k) \cap \{x \in \mathbb{R}^n \mid A_i x \leq b\}$ of the ellipsoid with volume at most $e^{-1/2(n+1)} \mathrm{vol(E(D^k, x^k))}$.
- Set $k = k + 1$ and iterate.

For an illustration of the ellipsoid method see Figrue 7.

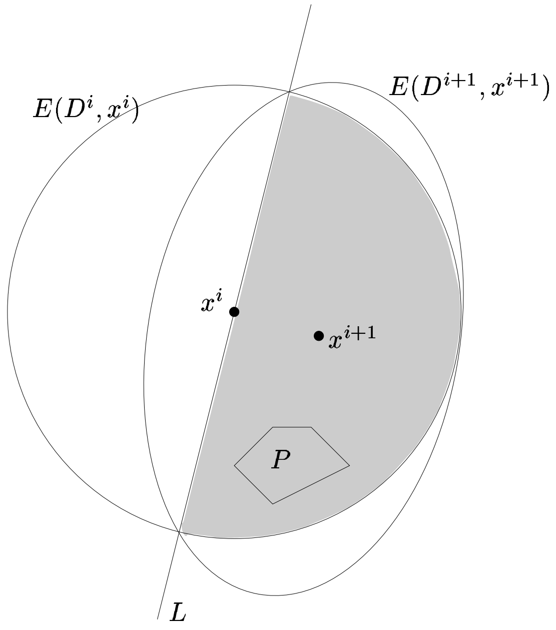

Figure 7: Illustration of the ellipsoid method: The line L separates the center of the ellipsoid $E(D^i, x^i)$ from the polytope P. So we can construct the new ellipsoid $E(D^{i+1}, x^{i+1})$ with smaller volume. Since the ellipsoid $E(D^i, x^i)$ covers the polytope P and P is completely “right” of L, the constructed ellipsoid $E(D^{i+1}, x^{i+1})$ covers P.

We refrain from giving the formula for finding the new ellipsoid $E(D^{k+1}, x^{k+1})$ from $E(D^k, x^k)$ and the index i.

The ellipsoid method above solves the strict membership problem. To use the method for linear programming, we need to extend it in two directions. We need to find an optimal point, and the ground set is a polyhedron. The details of this extension are rather complicated and omitted here. An important insight is that if x^* is an extreme point of a polyhedron, every coefficient x_j^* of x^* is a rational with denominator at most $L = (nT)^n$.

Exponential Number of Constraints

An interesting fact concerns linear programs with an exponential number of constraints.

If we look carefully at the ellisoid method, we see that the only part of the algorithm which depends on the problem is deciding whether a point x is in the polyhedron and, if not, finding a hyperplane separating x from the polyhedron. This problem is called the *separation problem*.

If the separation problem is solvable in polynomial time, the ellipoid method for the linear program runs in polynomial time. We note that the analysis of the runnning time does not depend on the number of constraints, without proving this.

But we can solve such linear programs in a practically efficient way with the simplex method. We start with a small subset of the constraints and solve the corresponding linear program with the simplex method. Let x be the optimal solution. We solve the separation problem for x. If x is in the polyhedron, it is also the optimal solution and we are done. Otherwise, we add a violated constraint to the actual subset of constraints and iterate. Note that the optimal basis of the old problem is dual feasible for the new problem, thus we can start the dual simplex method with this basis. One oberves that the dual simplex needs only very few iterations to solve the new problem.

A surprising observation is that this method terminates after a rather small number of iterations. We now mention some improvements of the basic algorithm, that are often used in practice to accelerate the algorithm.

- Instead of adding only one violated constraint in an iteration, we add several, cleverly chosen, violated constraints.
- If a constraint is not tight, i.e. , is not satisfied with equality, for a number of iterations we remove the constraint of the actual subset of constraints. We do this to keep the solved constraint set small.

2.2.4 Linear Programming Duality

We look at the maximization problem

$$\begin{array}{rrcl} \max & c^T x & & \\ s.t. & Ax & \leq & b \\ & x & \geq & 0. \end{array}$$

We call this problem the *primal problem* (P). The corresponding *dual problem* (D) has a variable for every constraint in (P) and a constraint for every variable in (P). It writes as follows

$$\begin{array}{rrcl} \min & b^Tp & & \\ s.t. & A^Tp & \geq & c \\ & p & \geq & 0. \end{array}$$

There are rules for transforming a general LP directly to its dual without transforming it to a LP as (P). The resulting dual is similar to the LP one gets if one first makes the transformations.

Lemma 2 *The dual of the dual is the primal.*

Proof: If we rewrite the dual in a form as (P) we get

$$\begin{array}{rrcl} \max & -p^Tb & & \\ s.t. & -A^Tp & \leq & -c. \\ & p & \geq & 0. \end{array}$$

The dual of this problem is

$$\begin{array}{rrcl} \min & -c^tx & & \\ s.t. & -(A^T)^Tx & \geq & -c \\ & x & \geq & 0. \end{array}$$

Rewriting this LP leads to the original LP. ▮

The following lemmas and theorems make a connection between the primal and dual problem. These facts are only a small fraction of the large number of theoretical and practical applications of the dual.

Lemma 3 (Weak duality) *If x^* is primal feasible and p^* is dual feasible than $c^Tx^* \leq b^Tp^*$.*

Proof: Since $p^{*T}A \geq c$ and $x^* \geq 0$ we have $c^Tx^* \leq p^{*T}Ax^*$. Furthermore since $Ax^* \leq b$ and $p^* \geq 0$ we have $p^{*T}b \geq p^{*T}Ax^*$. These two inequalities together form the claim of the Lemma. ▮

We now state one of the fundamental results of linear programming.

Theorem 5 (Strong duality) *If the primal or the dual LP has an optimal finite solution, then both LPs have finite solutions and the objective function values are equal.*

We can use this theorem to prove the optimality of a pair of primal and dual feasible solutions by the following theorem.

Theorem 6 (Complementary slackness) *If x^* is a primal feasible solution and p^* is a dual feasible solution then x^*, p^* are primal-dual optimal if and only if*

$$\begin{array}{cc} b_i - A_i x^* = 0 \text{ or } p_i^* = 0 & \text{for all constraints } i \\ \text{and} & \\ (A^T)_j p_j^* - c_j = 0 \text{ or } x_j = 0 & \text{for all variables } j. \end{array}$$

2.2.5 Minimum Spanning Trees and Linear Programming

Given a graph $G = (V, E)$ and edge weights c_e for all edges $e \in E$ a minimum spanning tree of (G, c) is a spanning tree of G with minimal weight. The weight of a subgraph $G' = (V, E')$ is defined as $\sum_{e \in E'} c_e$.

There are simple algorithms for solving the minimum spanning tree problem, i.e. the problem of finding a minimum spanning tree of a given graph G. We describe two well-known algorithms without proving their correctness.

Both algorithms start with the empty edge set and iteratively add edges until the current edge set forms a spanning tree of G. Kruskal's algorithm maintains a forest and inserts in every iteration the shortest edge which does not close a cycle. If there are several edges of the same length, choose one arbitrarily.

Prim's algorithm maintains a spanning tree of a subset of the nodes. It starts with the set $V_0 = \{u\}$ for any node $u \in V$. It then iteratively adds the shortest edge leaving the current node set V_i. If this edge is not unique, we can again choose one arbitrarily. The new set V_{i+1} is the union of the current set V_i and the node reached. The first algorithm can be implemented with running time $O(|E| \log |E|)$, the second with running time $O(|V| \log |V|)$.

We now look at the vectors $x = (x_e)_{e \in E} \in \mathbb{R}^{|E|}$, i.e. , the vector has a real valued entry for every edge of the graph. The incidence vector x of a subset $E' \subseteq E$ is defined by $x_e = \begin{cases} 1 & \text{if } e \in E' \\ 0 & \text{otherwise} \end{cases}$.

Lemma 4 *The incidence vectors of spanning trees of G are feasible for the following linear program. Moreover the incidence vector of a minimum spanning tree is an optimal solution of the LP.*

$$\begin{array}{rrcll} \min & \sum_{e \in E} c_e x_e & & & \\ s.t. & \sum_{e \in \gamma(S)} x_e & \leq & |S| - 1 & \text{for all } \emptyset \neq S \subset V \\ & \sum_{e \in E} x_e & = & |V| - 1 & \\ & x_e & \geq & 0 & \end{array} \tag{1}$$

Proof: We obtain a somewhat extended result. This extension is not of general interest, but we need it for a proof in section 4.4.

For any real constant r with $0 < r \leq |V| - 1$ we look at the linear program where we exchange the constraint (1) of the above system for

$$\sum_{e \in E} x_e = r.$$

For $r = |V| - 1$ we have the same LP as above. We claim that an optimal solution for the LP is obtained by the following two steps:

- Run Kruskal's algorithm. Let T be the edges chosen in the first $\lfloor r \rfloor$ iterations. Let f be the edge chosen if the $\lfloor r \rfloor + 1$st iteration, if $\lfloor r \rfloor < |V| - 1$ and an arbitrary edge which is not in T otherwise.
- Let $x = (x_e)_{e \in E}$ with $x_e = \begin{cases} 1 & \text{for all } e \in T \\ r - \lfloor r \rfloor & \text{for } e = f \\ 0 & \text{otherwise.} \end{cases}$

We call this algorithm the modified Kruskal's algorithm. Note that if r is integral, $x_f = 0$. In particular, if $r = |V| - 1$, the choice of the arbitrary edge does not matter.

We use the modified Kruskal's algorithm to construct a primal and dual solution simultaneously. The primal and dual solutions will satisfy the complementary slackness conditions (see Theorem 6).

The dual has a variable y_R for every subset $R \subseteq V$. The constraints of the dual are as follows:

$$\begin{aligned} y_V + \sum_{R \subset V | e \in \gamma(R)} y_R &\leq c_e && \text{for all } e \in E && (2) \\ y_R &\leq 0 && \text{for all } R \subset V. && (3) \end{aligned}$$

After the i-th iteration, the actual forest has i edges. The incidence vector of the actual subsolution is called x^i, the actual dual solution is called y^i. Let $P^i = \{P_1^i, \dots P_{|V|-i}^i\}$ be the current partition of the nodes (after inserting the i-th edge) in the connected subtrees. We call the sets P_j^i the active sets of phase i.

We start with $x^0 = 0$, $y^0 = 0$ and P^0 is the set of all singleton sets $\{v \mid v \in V\}$.

We maintain the following invariants:

- After every iteration i, x^i satisfies all primal constraints except $\sum_{e \in E} x_e = r$. This constraint is satisfied after the $\lfloor r \rfloor + 1$st iteration.
- For all active sets $P_j^i \in P^i$, we have $x^i(P_j^i) = |P_j^i| - 1$, except after the last iteration.
- For all edges e with $x_e^i \neq 0$ we have $y_V^i + \sum_{R \subset V | e \in \gamma(R)} y_R^i = c_e$.
- All dual solutions y^i are dual feasible.
- In an iteration i, we change the dual constraint $y_V^i + \sum_{R \subset V | e \in \gamma(R)} y_R^i$ of an edge e only for edges which are selectable in iteration i, i.e. , edges which we can add to the current primal solution without closing a cycle.

These invariants hold for the initial solution x^0, y^0.

The complementary slackness conditions are written as follows: $y^i_V + \sum_{R \subset V | e \in \gamma(R)} y^i_R = c_e$ for all edges e with $x_e \neq 0$ and $x(R) = |R - 1|$ or all $R \subset V$ with $y^i_R \neq 0$. These conditions are satisfied, if the invariants hold.

In the i-th iteration we increase y^i_V and decrease all $y_{P^i_j}$ at the same rate until a constraint gets tight. This happens for $y^i_V = c_e$, where e is the edge the modified Kruskal's algorithm selects in the i-th iteration. This can be seen as follows. As long as an edge is selectable, the corresponding constraint $y^i_V + \sum_{R \subset V | e \in \gamma(R)} y^i_R \leq c_e$ gets tighter, because none of the subsets $\{R \subset V \mid e \in \gamma(R)\}$ is in a partition P^j for any $j \leq i$. As soon as an edge is no longer selectable, exactly one of the sets $\{R \subset V \mid e \in \gamma(R)\}$ decreases its value at the same rate as y^i_V increases, so the tightness of the constraint does not change in this case. If the constraints of several edges get tight at the same time (if several edges have the same cost) select the same edge as Kruskal's algorithm and continue.

Now we set x^i as defined above and preserve the invariants above. ▮

2.2.6 TSP and Integer Programming

Given a set S of points in the plane, the *Traveling Salesman tour* is the shortest Hamiltonian cycle in the complete embedded graph with nodes corresponding to S and edge costs corresponding to the Euclidean distance between the two endpoints. For two points $a, b \in S$, the *Traveling Salesman path* between a and b is the Hamiltonian path between a and b in the same graph.

The construction of Traveling Salesman paths or tours is an NP-hard problem. A successful method for solving the Traveling Salesman problem is to formulate the problem as an integer linear program (ILP) and to use a branch-and-cut algorithm based on the LP-relaxation of the problem. We give the formulation for Traveling Salesman paths with fixed endpoints a and b. We introduce a variable x_{uv} for every edge uv between two points and describe the set of all Hamiltonian paths with endpoints a and b in the following way:

$$\sum_{v \in V} x_{uv} = 2 \text{ for all } u \in V \setminus \{a, b\} \tag{4}$$

$$\sum_{v \in V} x_{uv} = 1 \text{ for } u \in \{a, b\} \tag{5}$$

$$\sum_{u \in V', v \in V'} x_{uv} \leq |V'| - 1 \text{ for } V' \subset V,\ V' \neq \emptyset \tag{6}$$

$$x_{uv} \in \{0, 1\} \text{ for all } u, v \in V. \tag{7}$$

We refer to this program as the *Subtour-ILP for the Traveling Salesman problem with specified endpoints*. The equality constraints (4) and (5) in the Subtour-ILP are called *degree constraints*, the inequality constraints (6) are called *subtour elimination constraints*, and the constraints $x_{uv} \in \{0, 1\}$ are called the *integrality constraints*. Relaxing the integrality constraints to $0 \leq x_{uv} \leq 1$ gives the *Subtour-LP for the Traveling Salesman problem with specified endpoints*. The objective function for both programs is $\sum_{u,v \in V} \|uv\| x_{uv}$, i.e., the total Euclidean length of the edges selected.

Lemma 5 *The feasible incidence vectors of the integer program above are exactly the Hamilton path between a and b of the complete graph over V.*

Proof: Let x be a feasible solution of the above system. Since x is a 0/1–vector, x corresponds to a set of edges. Since x satisfies the constraints (4), every node has exactly two adjacent edges, except a and b, which have exactly one adjacent edge. If one adds the edge ab, every node has exactly two adjacent edges. Thus the new set of edges is a 2–factor, i.e. , a set of disjoint cycles meeting all nodes. Assume that we have more than one cycle. Then at least one cycle does not contain the edge ab, and thus the subtour elimination constraint (6) for the nodes of this cycle is violated, a contradiction. Thus the modified set of edges is a Traveling Salesman Tour and so x corresponds to a Traveling Salesman Path between a and b. ∎

It is well known that the Subtour-LP can be solved in polynomial time in the size of the bit representations of the distances by use of the ellipsoid method, see [Sch86]. A potentially exponential, but practically very efficient algorithm uses the simplex method and the cutting plane framework. One starts with the LP consisting only of the degree constraints and then solves a sequence of LPs. In each iteration one checks whether the solution of the current LP satisfies all subtour elimination constraints and, if not, one adds a violated subtour elimination constraint to the LP. We use the latter algorithm in our experimental framework for curve reconstruction.

Assuming that all degree constraints are satisfied, the test whether there is a violated subtour elimination constraint can be made by computing the minimal cut in the graph with edge weights equal to the the current LP solution for all edges different from ab and the LP solution plus 1 for the edge ab. A violated subtour elimination constraint exists if and only if this minimal cut is less than 2. This can be seen as follows. For every node v, the sum of the values of the edges adjacent to v is exactly 2. Let V' be one side of the cut not containing both a and b. Thus $2|V'| \sum_{u \in V'}(\sum_{v \in V'} x_{uv} + \sum_{v \notin V'} x_{uv}) = 2\sum_{uv \in G(V')} x_{uv} + \sum_{uv \in \delta(V')} x_{uv}$. Thus the value of all edges with both endpoints in the cut is larger than $|V'| - 1$, if the value of the cut is less than 2 and since the edge ab does not have both endpoints in V', the value is equal to the sum of the LP values of the edges. Since the computation of the minimal cut in a graph is a polynomial time algorithm, on can check whether there is a violated degree constraint or a violated subtour elimination constraint in polynomial time.

Remark: We stress the fact that we cannot optimize the Subtour-LP with one of the methods of section 2.2.3, since the Euclidean lengths are not necessarily rational.

2.2.7 Lagrangian Relaxation

Consider the following linear program:

$$\begin{array}{rrcl} \min & c^T x & & \\ s.t. & A_1 x & = & b_1 \\ & A_2 x & = & b_2 \\ & x & \geq & 0 \end{array}$$

and assume that dropping the constraint $A_2 x = b_2$ makes the problem "easy". *Lagrangian Relaxation* is a method for making use of this fact. If we dualize the problem above, we get

$$\begin{array}{rlcl} \max & y_1^T b_1 + y_2^T b_2 & & \\ s.t. & y_1^T A_1 + y_2^T A_2 & \leq & c^T. \end{array}$$

We rewrite the dual as double maximization:

$$\max_{y_2} y_2^T b_2 + (\max_{y_1} \quad y_1^T b_1 \qquad s.t. \quad y_1^T A_1 \leq c^T - y_2^T A_2)$$

and redualize the inner maximization problem

$$\max_{y_2} y_2^T b_2 + \left(\begin{array}{rrcl} min & (c^T - y_2^T)x & & \\ s.t. & A_1 x & = & b_1 \\ & x & \geq & 0 \end{array}\right).$$

None of the transformations above have changed the objective value. Thus we can formulate the following theorem.

Theorem 7 (Lagrangian Relaxation) *The following two problems have the same objective value*

- $\min c^T x$ *s.t.* $A_1 x = b_1, A_2 x = b_2, x \geq 0$.
- $\max_{y_2} y_2^T b_2 + (\min (c^T - y_2^T A_2)x$ *s.t.* $A_1 x = b_1, x \geq 0)$.

A straightforward extension to problems which are not in standard form leads to the following corollary.

Corollary 8 *The following two problems have the same objective value*

- $\min c^T x$ *s.t.* $A_1 x \leq b_1, A_2 x \leq b_2$.
- $\max_{\{y_2 | y_2 \geq 0\}} y_2^T b_2 + (\min (c^T - y_2^T A_2)x$ *s.t.* $A_1 x \leq b_1)$.

Remark: Note that in the case in which only equalities are relaxed, the order of the feasible solutions does not depend on the potential function, whereas in the general case, the order can change for different potential functions. However, for the optimal potential function, the optimal solution for the original cost functions remains optimal.

Example

We look at the linear program for the traveling salesman path problem given above. If we add the redundant constraint

$$\sum_{uv\in E} x_{uv} \;=\; n-1$$

and relax the equality constraints (4) and (5) we get a dual variable for every node of the graph. We call the value of this variable the potential of the node. In the context of the traveling salesman problem a variable for a node v is often called μ_v. Thus we reformulate the problem as follows.

$$\begin{array}{lrcll} \max\limits_{\mu} +(\min & \mu_a+\mu_b+2\sum\limits_{v\in V\setminus\{a,b\}}\mu_v+\sum\limits_{uv\in E}(\|uv\|-\mu_u-\mu_v)x_{uv} & & & \\ s.t. & \sum\limits_{uv\in E} x_{uv} & = & n-1 & \\ & \sum\limits_{u\in V',v\in V'} x_{uv} & \leq & |\,V'\,|-1 & \text{for } V'\subset V,\ V'\neq\emptyset \\ & 0\leq x_{uv} & \leq & 1 & \text{for all } uv\in E). \end{array}$$

The inner linear program is the description of the spanning trees of G, thus we can solve this linear program in time $O(|V|\log|V|)$. We will see in the next section how to use this fact to efficiently get an approximation of the value of the linear program for the traveling salesman problem.

The optimal bound one can obtain with any potential function is known as the Held-Karp bound. We have just seen that the Held-Karp bound is equal to the bound obtained by solving the Subtour-LP.

We will see other examples of a Lagrangian Relaxation later.

Solving the Lagrangian Dual

There are several methods for solving the Lagrangian dual, i.e. , to compute a potential function y which leads to the optimal or to an almost optimal bound.

The most common technique is the subgradient optimization. Let $L(y)=\min(c^T - y_2^TA_2)x$ s.t. $A_1x=b_1, x\geq 0$ be the bound obtained for the potential function y. The function $L:\mathbb{R}^m\mapsto\mathbb{R}$ has the following important properties:

- $L(y)$ is piecewise linear, and therefore differentiable except in a finite set of vectors y.
- $L(y)$ is concave, i.e. , for all $y_1,y_2\in\mathbb{R}^m$, and $\alpha\in[0,1]$ we get $L(\alpha y_1+(1-\alpha)y_2)\geq \alpha L(y_1)+(1-\alpha)y_2$.

The subgradient optimization is a method for optimizing a concave function. A subgradient of a concave function $L:\mathbb{R}^m\mapsto\mathbb{R}$ at $y^*\in\mathbb{R}^m$ is a vector $\gamma(y^*)\in\mathbb{R}^m$ so that

$$\gamma(y^*)(y^*-y)\geq L(y)-L(y^*) \text{ for all } y\in\mathbb{R}^m.$$

The set of all subgradients at y^* is called $\delta L(y^*)$. Some facts of subgradients:

- If L is continuous and differentiable at u then $\delta L(y^*)$ is the differential of L at y^*. Furthermore $\gamma(y^*)$ is a direction of increase, i.e. , there exits $\theta > 0$ such that $L(y^* + \theta\gamma(y^*)) > L(y^*)$.
- In general $\gamma(y^*)$ is not a direction of increase. But there is a justification for moving in the direction normal to a subgradient. Any point with a larger objective value than y^* is contained in the half-space $\gamma(y^*)(y^* - y) > 0$. In particular there exits a $\theta > 0$ so that $y^* + \theta\gamma(y^*)$ is closer to the optimal point than y^*.
- y^* is optimal if and only if $0 \in \delta(y^*)$.

These observations lead to the following algorithm for solving a subgradient optimization problem.

1. Start with $t = 0$ and any y_0.
2. Solve the Lagrangian dual for y_t. If $0 \in \delta L(y_t)$ the current solution is optimal.
3. Set $y_{t+1} = y_t + \theta_t\gamma(y_t)$. We explain below how to choose θ_t. Set $t \leftarrow t + 1$ and continue with step 2.

We have to select the sequence $(\theta_t)_{t=1}^{\infty}$. Similarly to the different algorithms for solving LPs, there are theoretical and practical good choices.

- A series $(\theta_t)_{t=1}^{\infty}$ with $\sum_{t=1}^{\infty} \theta_t = \infty$ and $\lim_{t\to\infty} = 0$, so a divergent series. This setting theoretically converges to an optimal point. The convergence is slow in practice.
- A geometric series $\theta_t = \theta_0\rho^t$, where $0 < \rho < 1$. This setting often leads to an algorithm which converges fast in practice. Theoretically, it converges to the optimal solution only if θ_0 and ρ are chosen sufficiently large.

It remains to show how to compute a subgradient in the case of a linear program. It can be computed directly from the optimal solution: The vector $b_1 - A_1x^t$, where x^t is an optimal solution of $L(y^t)$ is a subgradient of L at y^t. This can be seen as follows. Let $\gamma(y^t) = b - Ax^t$ and let y be any vector. Then $L(y) \leq c^Tx^t + (y^t)^T(b - Ax^t) = L(y^t) - (y^t)^T(b - Ax^t) + y^T(b - Ax^t) = L(y) - (y^t - y)^T\gamma(y^t)$.

2.3 Delaunay Triangulations and Diagrams

In this section we follow the presentation in [MN99].

The *convex hull* of a set of points S is the minimal convex set of points covering S. Recall that a set X is called convex if for any two points p and q of X the entire line segment pq is contained in X. A *triangulation* of a set of points S in the plane is a partition of the convex hull of S into triangles with three points of S as its vertices. If all points of a point set S are collinear, there is no triangulation of S. A triangulation is called *Delaunay triangulation*(DT) if the interior of the minimum enclosing disk of any

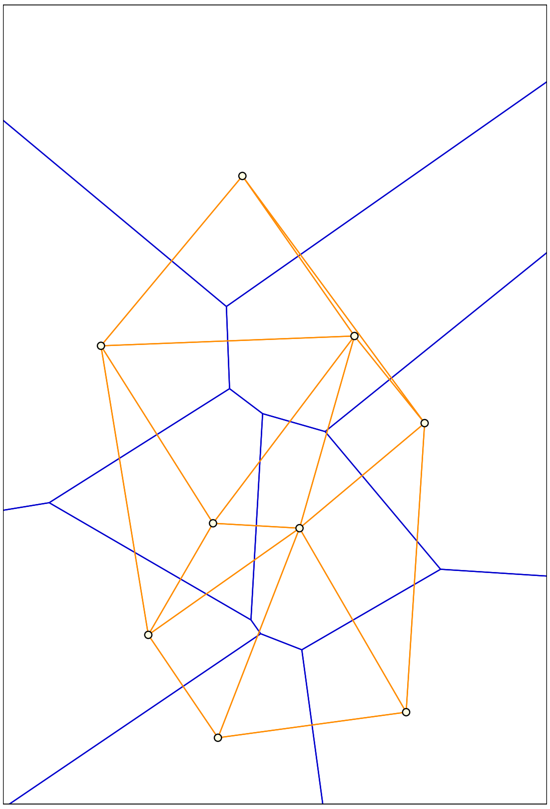

Figure 8: A Delaunay triangulation and the Voronoi diagram of a point set

triangle in the triangulation contains no point of S (see Figure 8). We show the existence of a Delaunay triangulation for every set S.

Remark: Note that in a triangulation every edge has two adjacent faces, i.e. edges of the convex hull have exactly one adjacent triangle and the other adjacent face is the unbounded face, and all other edges have two adjacent triangles.

Lemma 6 *A point set S has a Delaunay triangulation unless all points of S are collinear.*

Proof: The proof is by construction. Starting at any triangulation, we show that, if the triangulation does not have the Delaunay property, we can increase the lexicographical order of all inner angles of the triangles sorted from the smallest to the largest angle. Since there is only a finite number of triangulations this process has to terminate. Since the process only terminates if the triangulation has the Delaunay property, it follows that there is a triangulation with the Delaunay property.

We first look at the very special case of only four non-collinear points in convex position and show that the triangulation with the larger smallest angle has the Delaunay property. Then we show that every triangulation that has not the Delaunay property has a subset of four points in convex position, so that the induced subgraph is triangulated and does not have the Delaunay property.

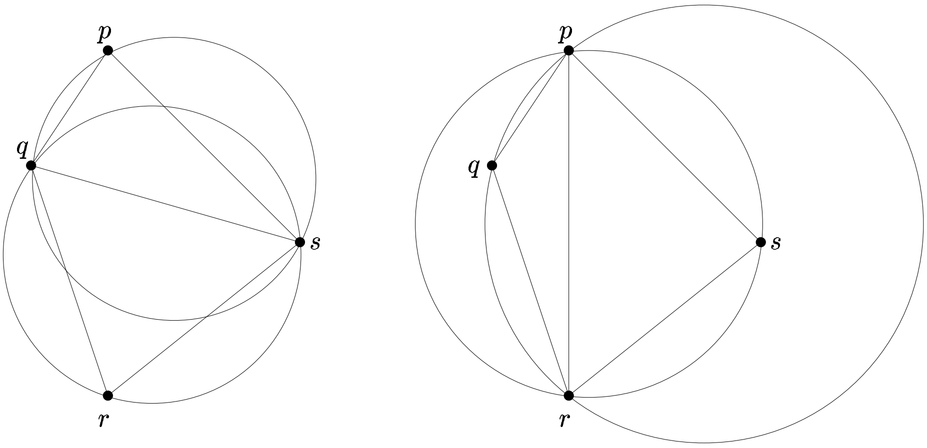

Figure 9: Four points and the two possible triangulations.

Let p, q, r, s be four points in convex position in the plane so that the order of the points in the convex hull is p, q, r, s. Look at the triangulation $\triangle p, q, r \cup \triangle r, s, p$. Assume the triangulation does not have the Delaunay property. The critical observation is the following. If the minimum enclosing disk of $\triangle p, q, r$ contains the point s, the minimum enclosing disk of $\triangle r, s, p$ contains the point q. Furthermore the point p is not contained in the minimum enclosing disk of $\triangle qrs$ and r is not contained in the minimum enclosing disk of $\triangle spq$. Thus we replace the diagonal pr by the diagonal qs and get a Delaunay Triangulation. We call this transformation a *diagonal-flip*.

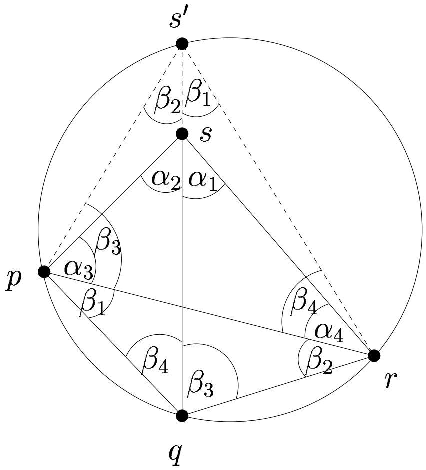

Figure 10: By the Inscribed Angle Theorem, we know that angles with the same name have the same value.

We now show that the smallest of the six angles of the two triangles increases for every diagonal-flip. Look at the intersection of the line through q and s with the circumcircle of pqr different from q, and call this point s' (see Figure 10 for an illustration). By the Inscribed Angle Theorem applied to the secant rq we know that $\angle(\overrightarrow{s'q}, \overrightarrow{s'r}) = \angle(\overrightarrow{pq}, \overrightarrow{pr})$. We call this angle β_1. Analogously we can apply the Inscribed Angle Theorem to the secants qp, ps' and $s'r$ and obtain $\angle(\overrightarrow{rp}, \overrightarrow{rq}) = \angle(\overrightarrow{s'p}, \overrightarrow{s'q}) = \beta_2$, $\angle(\overrightarrow{qs'}, \overrightarrow{qp}) = \angle(\overrightarrow{rs'}, \overrightarrow{rp}) = \beta_4$, and $\angle(\overrightarrow{pr}, \overrightarrow{ps'}) = \angle(\overrightarrow{qr}, \overrightarrow{qs'}) = \beta_3$. We show that for any inner angle of the two triangles $\triangle p, q, s$ and $\triangle q, r, s$ there is an angle of $\triangle p, q, r$ or $\triangle p, r, s$ that is smaller. We can see that $\alpha_4 + \beta_2 > \alpha_4$ and $\alpha_3 + \beta_1 > \alpha_3$, since the points are in convex position. Furthermore $\alpha_2 > \beta_2$, $\alpha_1 > \beta_1$, $\beta_3 > \alpha_3$ and $\beta_4 > \alpha_4$, since moving a point $\tilde{s}$ from s' to s increases the angles at $\tilde{s}$ and decreases the angles at p and r.

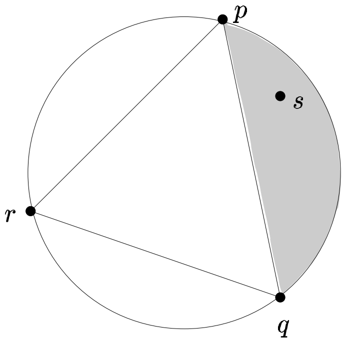

Figure 11: If the point s lies in the shaded region, but is not connected to p and q, one of the two edges pt or qt is closer to s than pq.

It remains to show that any triangulation which is not Delaunay contains two triangles $\triangle p, q, r$ and $\triangle r, s, p$, sharing an edge so that the point s is in the minimum enclosing disk of $\triangle p, q, r$ and the points are in convex position. If the triangulation is not Delaunay, there is a triangle $\triangle p, q, r$ and a point s, so that s is in the minimum enclosing disk of $\triangle p, q, r$ (see Figure 11). If $\triangle r, s, q$ is a triangle of the triangulation, we are done. Now assume otherwise. Notice that no point can lie in the interior of a triangle. Without loss of generality let s be in the part bounded by the segment p, q and the bounding arc of the minimum enclosing disk between p and q, not containing r. Let t be the third point of the second triangle adjacent to pq. The point s is in the disk of $\triangle p, q, t$, since the minimum enclosing disk of $\triangle p, q, t$ contains the region we have described. The distance from s to the triangle $\triangle p, q, t$ is shorter than the distance to the triangle $\triangle p, q, s$, since the triangle intersects the perpendicular from s to pq. We repeat the argument for the triangle $\triangle p, q, r$. After a finite number of iterations, we must arrive at the first case. The points lie in convex position, because the point lies in the minimum enclosing disk of the triangle and outside the triangle itself. ∎

Delaunay triangulations are not unique. The set of edges that are contained in every Delaunay triangulation is called the *Delaunay diagram*. We will now characterize the edges of the Delaunay diagram.

Lemma 7 *An edge pq is in the Delaunay diagram, iff there is a closed ball B with* $B \cap S = \{p, q\}$.

Remark: Note that if an edge pq has a ball B with $B \cap S = \{p, q\}$ we can choose the ball B so that p and q are on the boundary of B. Such a ball is called a *protecting ball* of pq. Its center lies on the perpendicular bisector of pq. For an edge pq the *center ball* of pq is defined as the ball centered at the midpoint of pq with radius $\|pq\|/2$.

Proof: We first prove, that every edge which has a protecting ball is indeed in every Delaunay triangulation. Assume otherwise. Let pq be an edge, B a closed ball with $B \cap S = \{p, q\}$ and T be a Delaunay triangulation not containing the edge pq. Move along the edge pq and observe that an initial part of the segment starting at p is contained in another triangle as an initial part starting at q. So the segment pq crosses a segment, say ab of the triangulation T. Let a' and b' be the intersections of ab with B. Since p, q, a' and b' are cocircular, every closed ball containing a' and b' either contains p or q. So the protecting ball of a, b contains either p or q in its interior, a contradiction.

We now come to the converse. Let st be an edge of a Delaunay triangulation DT. We show that either there is a Delaunay triangulation DT′ that does not contain st, or there is a ball B containing only s and t.

If st is not an edge of DT, we are done. If st is an edge of the convex hull, there is a ball B containing only s and t. In the remaining case, there are two triangles in DT adjacent to st, say $\triangle s, t, a$ and $\triangle s, t, b$. Let B_a and B_b be the minumum enclosing disks of $\triangle s, t, a$ and $\triangle s, t, b$. If the four points s, t, a and b are cocircular, we can flip the segments st and ab and get a Delaunay triangulation DT′ without the segment st. Otherwise look at the ball B centered at the midpoint of the centers of the balls B_a and B_b through s and t. It is completely contained in $B_a^0 \cup B_b^0 \cup \{s, t\}$, thus it contains no point and we are done. ∎

2.3.1 The Voronoi Diagram

Let S be a set of points in the plane. The *Voronoi diagram* VD of S is defined as the set of points in the plane with more than one nearest neighbor in S. Notice that VD consists of segments of the perpendicular bisectors of two points of S, i.e. , VD consists of line segments, rays, and lines. So VD is a graph-like structure with nodes in some of the intersection points of the perpendicular bisectors. The edges of VD are either segments between two nodes or rays starting at one node. The exception is, if all points are collinear, then the Voronoi diagram is formed by the complete perpendicular bisectors of "neighbored" points.

For simplicity, we assume that the points in S are not collinear. Then this structure can be obtained from a Delaunay triangulation of S with the following simple rules:

- The nodes of VD are the centers of the minimum enclosing disks of the triangles of the Delaunay triangulation. So the nodes are the common intersection of the perpendicular bisectors of the three edges of the triangle.

- The edges of VD are obtained from the edges of the Delaunay Diagram. Let e be an edge of the Delaunay Diagram.
    - If e is not an edge of the convex hull of S, the Voronoi diagram has a segment between the two nodes, defined for the two triangles containing e.
    - If e is an edge of the convex hull, the Voronoi diagram has the ray, which is part of the perpendicular bisector of s, starting at the node, defined for the triangle containing s.

Thus we have a node in VD for every triangle in DT, an edge in VD for every edge in DT, and a face in VD for every node in DT. These are called the *dual node*, *dual edge*, and *dual face*, respectively. Note that the only case in which the Voronoi diagram has a line is the one in which all points are collinear. In this case, the Voronoi diagram is the set of all perpendicular bisectors of neighbored points with respect to the common line.

Lemma 8 *The rules above construct the Voronoi diagram from a Delaunay triangulation.*

Proof: We prove the Theorem in two steps. First we show that the edges we construct belong to the Voronoi diagram and in the second step we derive that all edges of the Voronoi diagram are covered by the construction.

Let S be the set of points for which we construct the Delaunay triangulation and the Voronoi diagram. Let pq be an edge of a Delaunay triangulation. Assume first that the dual edge is a line segment, say st. Then s and t are the centers of two triangles of the Delaunay triangulation sharing one edge, say pq. The minimum enlcosing disks B_s and B_t of the two triangles are empty of points of S. Thus s and t have at least three nearest neighbors and belong to the Voronoi diagram. Look at any point x on the segment st and the ball with center x and radius $\|xp\|$. Since this ball is contained in the union of the balls B_s and B_t, its interior is empty of points of S. Thus x has at least two nearest neighbors, namely s and t and thus belongs to the Voronoi diagram (see Figure 12). Now, look at the case in which the dual edge is a ray, starting at a point s. Let x be a point on this ray. The ball B_x with center x and radius $\|xp\|$ is contained in the union of B_s and the halfspace, defined by the line through p and q not containing s. Since ps is an edge of the convex hull of S, this halfspace contains no point of S. Thus the interior of B_x is empty of points of S and x has at least two nearest neighbors (see Figure 12).

We now turn to the proof that no point of the Voronoi diagram is missing. Assume there is a point with at least two nearest neighbors in S which is not constructed. Since the set of points with at least three nearest neighbors is discrete and the Voronoi diagram and the constructed diagram is closed and connected, there is a point p with exactly two nearest neighbors, say s and t, which is not constructed S. Thus the ball with center p and radius $\|pt\|$ is a protecting ball for the edge st and thus the edge st is in every Delaunay triangulation. We show that the dual edge of st contains all points that have s and t as nearest neighbours. Assume the dual edge is terminated on the Voronoi point v and let x be the third point of the triange defining v. A point y behind v is closer to x than to s and t, since it is behind the perpendicular bisector of sx. ∎

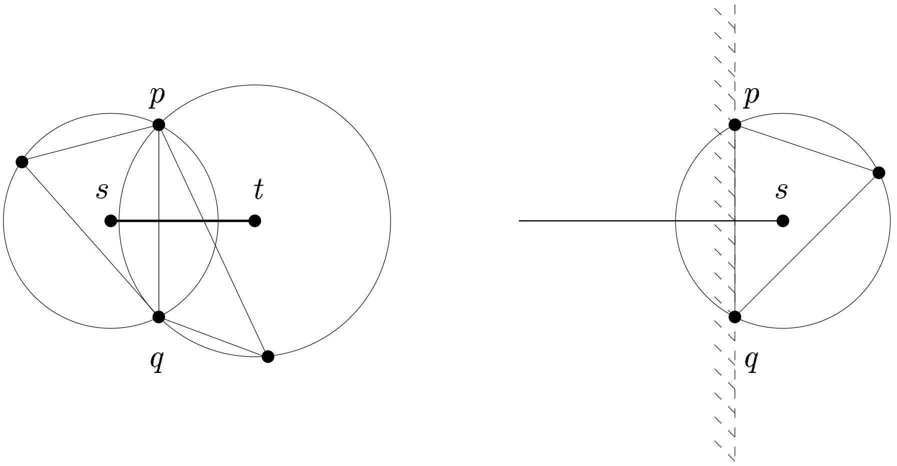

Figure 12: On the left we see the contruction of a segment of the Voronoi diagram and on the right the construction of a ray.

3 Previous Work

Many algorithms have been proposed for the curve reconstruction problem. We are interested in *reconstruction algorithms with guaranteed performance*, i.e. , algorithms which provably solve the reconstruction problem under certain assumptions on γ and S. If the curve is closed, smooth and uniformly sampled, several methods are known to work ranging over minimum spanning tree [FG94], α-shapes [BB97, EKS83], β-skeleton [KR85], and r-regular shapes [Att97]. A survey on these techniques appears in [Ede98]. The case of non-uniformly sampled closed curves was first treated successfully by Amenta, Bern and Eppstein [ABE98] and subsequently improved algorithms such as [DK99, Gol99] appeared. Open non-uniformly sampled curves were treated in [DMR00].

We review some of the algorithms which work for non-uniform sample sets. Before we explain these algorithms and prove their correctness, we introduce some notations for curves and prove some basic facts. Furthermore we introduce a variant of the algorithm by Dey and Kumar [DK99] that requires a weaker sampling density than the known algorithms.

3.1 Terminology and Basic Properties of Curves

The following definitions and Lemmas are collected from the papers [Gie99, ABE98, DK99]. A *single open curve* is given by an *embedding* $\gamma : [0,1] \to \mathbb{R}^2$ and a *single closed curve* is given by an embedding $\gamma : S^2 \to \mathbb{R}^2$, where S^2 is the unit circle.

Definition 1 ([Gie99]) *Let*

$$T = \{ (t_1, t_2) \ ; \ t_1 < t_2, t_1, t_2 \in [0,1] \}$$

and

$$\tau : T \to S^2, (t_1, t_2) \mapsto \frac{\gamma(t_2) - \gamma(t_1)}{|\, \gamma(t_2) - \gamma(t_1) \,|}.$$

The single curve γ is called left (right) regular at $\gamma(t_0)$ with left (right) tangent $\mathrm{t}_l(\gamma(t_0))$ ($\mathrm{t}_r(\gamma(t_0))$) if for every sequence (ξ_n) in T which converges to (t_0, t_0) from the left (right) in the closure of T the sequence $\tau(\xi_n)$ converges to $\mathrm{t}(\gamma(t_0))$. We call γ semi-regular if it is left and right regular in all points $\gamma(t)$, $t \in [0,1]$. We call γ regular if it is semi-regular and the left and right tangent coincide in every point of the curve.

A *curve* is a finite collection of disjoint single curves. It is called *smooth* if all single curves are regular, and it is called *closed* if all single curves are closed.

Remark: In other papers, e.g. , [ABE98] a curve is defined as a closed, compact, twice-differentiable one-manifold, without boundary, embedded in the plane. Thus in their definition, a curve is always smooth and closed.

Figure 13 shows two semi-regular curves. Tangents are unit vectors. The angle between two vectors with the same source is the smaller of the two angles between the vectors. The angle is zero if the two vectors point in the same direction and the angle is π if the

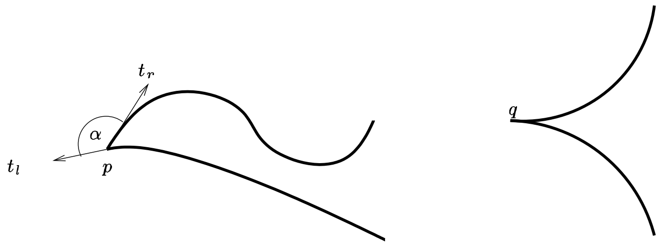

Figure 13: Two semi-regular curves, one benign and one not. In the left curve, the two tangents t_l and t_r at a point p of the curve are shown. The turning angle at p is α. In the curve on the right the turning angle at q is π as the left and right tangents at q point to opposite directions.

vectors point in opposite directions. The angle between the left and right tangent at a point $p \in \gamma$ is called the *turning angle* at p. If the curve has a tangent at p, the turning angle at p is zero. If the turning angle at p is non-zero, we call p a *singularity* of the curve. A semi-regular curve is *benign* if the turning angle is less than π at every point of the curve.

A *sample set* S of a curve γ is a finite set of points $p \in \gamma$. For every point $p \in \gamma$, the adjacent sample points are defined as the two sample points one reaches first, if one moves along γ in one of the two possible directions. The polygonal reconstruction of γ wrt. S is an embedded graph with node set S and edge set E defined by $pq \in E$ if and only if p and q are adjacent on γ.

The *medial axis* M of a curve γ is defined as the closure of all points having more than one nearest neighbour in γ. The *local feature size* $f(p)$ of a point $p \in \gamma$ is the distance of p to the medial axis. For any $\epsilon > 0$, a sample set S is called a ϵ-sample, if for all $p \in \gamma$ there is a sample $s \in S$ with $\|sp\| \leq \epsilon f(p)$. If one extends a sufficiently dense sample set by additional sample points, the new set should also be sufficiently dense. The following corollary makes this intuition formal, for all sampling conditions of the form above.

Corollary 9 *Let* $\epsilon < 1$. *If* S *is an* ϵ*–sample of a smooth curve* γ, *and* $S' \supset S$ *is a sample set of* γ, *than* S' *is an* ϵ*–sample of* γ.

We now summarize a few facts on the local feature size. The first Lemma is a simple consequence of the triangle-inequality.

Lemma 9 $f(p) \leq f(q) + \|pq\|$ *for all* $p, q \in \gamma$.

We can use this Lemma to bound the length of a segment, adjacent to some sample point $p \in \gamma$.

Lemma 10 *The length of a segment* pq *of the polygonal reconstruction of an* ϵ*-sampled smooth curve* γ *is at most* $2\epsilon/(1-\epsilon)f(p)$.

Proof: Let r be the point on γ between p and q that lies on the perpendicular bisector of p and q. Then $\|pr\| \leq \epsilon f(r)$ and $\|pr\| \geq \|pq\|/2$. By Lemma 9 we have $f(r) \leq \|pr\| + f(p)$. So $\|pr\| \leq \epsilon(\|pr\| + f(p))$. Algebraic transformations lead to $\|pr\| \leq \epsilon/(1-\epsilon)f(p)$. Thus $\|pq\| \leq 2\epsilon/(1-\epsilon)f(p)$. ∎

To prove the correctness of an algorithm, it is important to have an upper bound of the distance of a point $p \in \gamma$ to its nearest medial axis point.

Lemma 11 *Let B be a closed ball so that $B \cap \gamma$ is not connected. Then B contains a medial axis point.*

Proof: For a point x let x^* be a point on γ with minimal distance to x. Let p and q be two points of γ in different connected components with respect to B. Assume we are moving a point x from p to the center of B and than to q. In the beginning, the nearest point of γ is p and at the end, the nearest point of γ is q. Since $\|xx^*\|$ is continous, either the x^* remains in the component of p or there are at least two points on γ with distance $\|xx^*\|$ to x. This x is a medial axis point of γ. ∎

Corollary 10 *Let $p \in \gamma$ and S be a 1–sample of γ. The point in S with smallest distance to p is one of the two adjacent sample points.*

Proof: The ball around p through the sample q with shortest distance to p must intersect γ in a connected component, since the distance from p to q is less than $f(p)$. Thus q is adjacent to q. ∎

Our next goal is to bound the angle between two adjacent segments of the polygonal reconstruction. Therefore we prove the following Lemma.

Lemma 12 *Let B be a ball, tangent to γ at a point p with radius $f(p)$. Then B contains no point of γ in its interior.*

Proof: Let m be the center of B. Let B' be the largest ball, tangent to γ at p on the same side of the curve as B which contains no point of γ in its interior. Since B' is maximal and γ is compact, B' has a point q of γ, different from p in its boundary. So p and q have the same distance from the center m' of B' and since the interior of B' is empty of terminals they are both nearest neighbors of p on γ. We conclude that m' is a point of the medial axis and so $\|pm\| = f(p) \leq \|pm'\|$. Since B and B' are tangent balls at p on the same side, B is completely contained in B' and so the interior is empty of points of γ. ∎

The sampling condition states that the distance from any point of the curve to its next sample point is at most ϵ times the distance to the medial axis. A crucial point, which we often use to derive a contradiction, is the point on the curve which has the same distance to both adjacent sample points. We will see later that this point is unique.

Definition 2 ([ABE98]) *Let γ be a smooth curve and S an ϵ–sample for an $\epsilon < 1$. Let st be an edge of the polygonal reconstruction of γ. The* curve Voronoi point *of st is defined as the intersection of the perpendicular bisector of the segment st with the curve between s and t, i.e. , the part of the curve which contains no other sample point.*

Remark: An 1–sample of any curve has at least 3 sample points, thus the part between two adjacent sample points is well defined.

We are now ready to bound the angle of two adjacent segments. We do so by bounding the angle between segments and the tangent in one of its endpoints. From this it is a simple corollary to bound the angle between two adjacent segments.

Lemma 13 *Let $\epsilon \leq 1$, γ be a smooth curve and S be an ϵ–sample of γ. Let s, t be two adjacent sample points of γ and p be the curve Voronoi point of st.*

- *The angle $\angle(\vec{st}, \vec{sp})$ is at most* $\arcsin(\epsilon/2)$.
- *The angle between the tangent at s and the segment sp is at most* $\arcsin(\epsilon/2)$.

Corollary 11 *The angle between two adjacent segments of the polygonal reconstruction of an ϵ–sampled smooth curve is at least* $\pi - 4\arcsin(\epsilon/2)$.

Proof: [of Lemma 13]

Let s, t and p be defined as in the Lemma (see Figure 14).

We show the first item. By Lemma 12 we know that s is outside the two tangent circles in p with radius $f(p)$. By the sampling condition, we know that $\|ps\|$ and $\|pt\|$ is at most $\epsilon f(p)$. The angle $\angle(\vec{st}, \vec{sp})$ is maximized if $\|ps\| = \|pt\| = \epsilon f(p)$ and s and t lie on the boundary of the same tangent circle. In this case the angle $\angle(\vec{st}, \vec{sp})$ is exactly $\arcsin(\epsilon/2)$.

We turn to the second item. Let β be the angle between the tangent in s and the segment sp. Let without loss of generalityt(s) be a horizontal line and p be below t(s). Look at the largest tangent circle in s below t(s) so that its interior does not intersect γ. Let m be its center. Since the circle is maximal, it meets at least two points of γ. Thus m is a point of the medial axis.

Let $c = \|mp\|/\|ms\|$. Since $p \notin B^0(m, \|ms\|)$, we have $c \geq 1$. By the low of cosine, we have

$$\sin(\beta) = \cos(\pi/2 - \beta) = \frac{\|sp\|^2 + \|sm\|^2 - c^2\|sm\|^2}{2\|sp\|\|sm\|}$$

The right hand side increases with the distance $\|sp\|$, so it is maximized if $\|sp\|$ is maximial. Since $\|sp\| \leq \epsilon f(p) \leq \epsilon c\|sm\|$ we have

$$\sin(\beta) \leq \frac{c^2\epsilon^2\|sm\|^2 + \|sm\|^2 - c^2\|sm\|^2}{2c\epsilon\|sm\|^2} = \frac{1 + c^2(\epsilon^2 - 1)}{2c\epsilon} \leq \epsilon/2$$

The last inequality holds since the term $(1 + c^2(\epsilon^2 - 1))/(2c\epsilon)$ is decreasing in c and $c \geq 1$.

∎

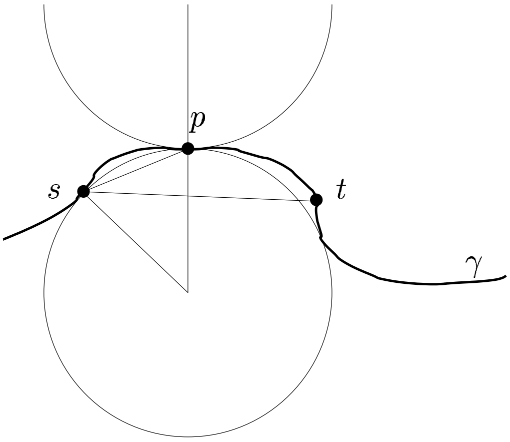

Figure 14: p is the curve Voronoi vertex between s and t. The points s and t must lie outside the tangent balls of p with radius $f(p)$.

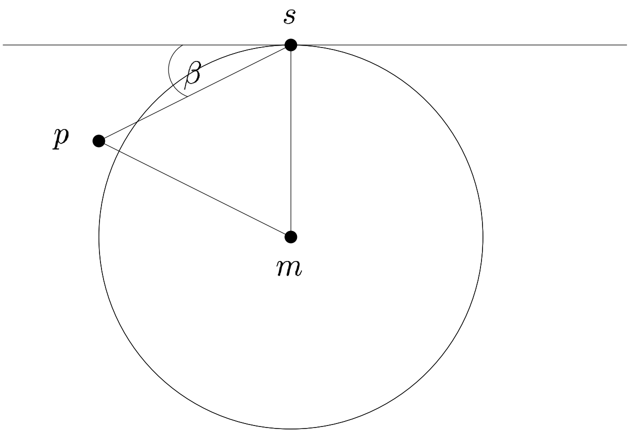

Figure 15: The point p lies outside the largest empty tangent ball of s. The angle is maximized if p lies on the boundary of this ball.

3.2 Algorithm of Dey and Kumar

Although the algorithm of Dey and Kumar was not the first algorithm, we start with this algorithm, since it is the simplest one. The algorithm works as follows:

- For every sample point s include the edge from s to one of its nearest neighbours.
- For every sample point s which has only one adjacent edge, say st, include the edge from s to one of the nearest neighbours which form an angle of more than $\pi/2$ with st.

To prove the correctness, we show that in both phases of the algorithm, only edges of the polygonal reconstruction are added to the returned graph. Since we add only correct edges in the first phase and the angle between two adjacent segments is more than $\pi/2$ by Corollary 11, we conclude that every node has at least two adjacent edges. Thus all edges of the polygonal reconstruction are added.

Lemma 14 *If S is an $1/3$–sample of a smooth closed curve γ, only edges of the polygonal reconstruction are added in the first phase of the algorithm.*

Proof: Assume otherwise. Let t be the nearest neighbour of s and st not in the polygonal reconstruction. Let p be one of the adjacent sample points of s in the polygonal reconstruction. Using Lemma 10, we conclude $\|st\| \leq \|sp\| \leq 2\epsilon/(1-\epsilon)f(s) \leq f(s)$. The center ball B of st splits γ in more than one component (since both neighbours of s are outside B) and thus Lemma 11 for B shows that there is a medial axis point on st, a contradiction. ∎

Lemma 15 *If S is an $1/3$–sample of a smooth closed curve γ, only edges of the polygonal reconstruction are added in the second phase of the algorithm.*

Proof: Assume otherwise. Let st be an edge added in the second phase that is not in the polygonal reconstruction. Let p be the nearest neighbour of s and q be the other sample point adjacent to s in the polygonal reconstruction.

Corollary 11 shows that $\angle(\vec{sp}, \vec{sq}) > \pi/2$, thus st is shorter than sq. Again the center ball B of st splits γ in more than one component. The same argument as above leads to a contradiction. ∎

We summarize the two Lemmas in the following corollary.

Corollary 12 *If S is an $1/3$–sample of a smooth curve γ, the algorithm of Dey and Kumar returns the polygonal reconstruction.*

3.3 Algorithm of Amenta, Bern, and Eppstein

Amenta, Bern, and Eppstein [ABE98] developed the notion of "non-uniform sampled" and prove the correctness of two algorithms: The β–crust, which was known to be correct for uniform samples, and the CRUST, an algorithm they present in the paper. They discovered that the β–crust finds the correct reconstruction for every 0.297–sample and the CRUST for every 0.252–sample. Gold and Snoeyink improved the analysis and came with a faster variant. They showed that a 0.42–sample suffices for both variants. We now show the correctness of both variants for $\epsilon = 1/3$.

The algorithm of Amenta, Bern, and Eppstein [ABE98] works as follows:

- Compute the Voronoi diagram V of the sample Points.
- Compute the Delaunay Triangulation T of $V \cup S$.
- Keep all edges of T which run between two sample points.

Another view of the algorithm is that the algorithm selects all edges of the Delaunay diagram which have an protecting ball empty of sample points and Voronoi points. Gold and Snoeyink [Gol99] noticed that it suffices that there is a ball with the two sample points on its boundary which has none of the two endpoints of the dual edge in its interior. This can be checked by a simple incircle test of the endpoints of the two edges. Thus the second computation of a Delaunay Triangulation can be omitted.

To prove the correctness of the algorithm, we first prove that all edges of the polygonal reconstruction are in the solution and then that no edges which are not in the polygonal reconstruction are in the solution.

Lemma 16 *If S is an $1/3$–sample of a smooth closed curve γ, the algorithm of Amenta, Bern, and Eppstein (Gold and Snoeyink) has all edges of the polygonal reconstruction in the returned graph.*

Proof: Certainly the algorithm of Gold and Snoeyink returns more edges than the algorithm of Amenta, Bern and Eppsein. So it suffices to prove the Lemma for the second algorithm.

For every segment of the polygonal reconstruction we define a protecting ball and show that it is empty of sample points, as well as of Voronoi points.

Let st be an edge of the polygonal reconstruction. Look at the ball B centered at the curve Voronoi vertex p between s and t with radius $\|ps\| \leq f(p)/3$ (see Figure 16).

Assume first that there is a sample point in B. Then B splits γ into more than one component and thus there is a medial axis point in B, which is a contradiction to the definition of local feature size.

Assume now that there is a Voronoi vertex v in B. Let R be the radius of the Voronoi circle of v. Then $R \leq \|vs\|$ and $R \leq \|vt\|$.

If V has no point of the curve in its interior, the center of V has at least three nearest neighbours on γ, thus it is a medial axis point. Otherwise there is a ball $V' \subset V$ that splits γ into more than one component. In both cases V contains a medial axis point.

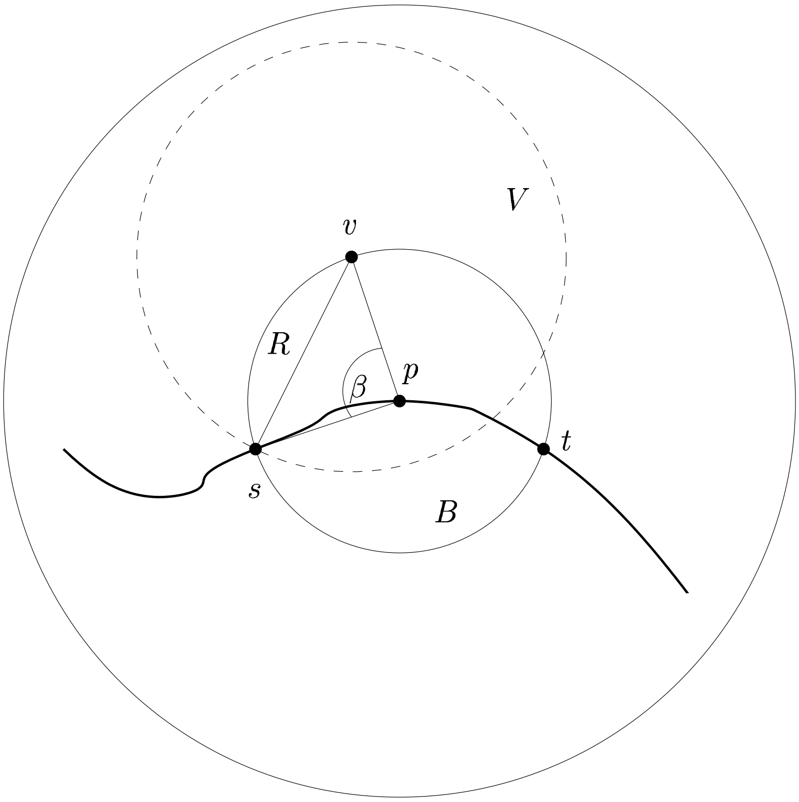

Figure 16: The construction of the protecting ball and the proof that it is empty of Voronoi vertices.

It remains to show that V is completely contained in $B(p, f(p))$. Since $\angle(\vec{ps}, \vec{pt}) \leq \pi - 2\arcsin(\epsilon/2)$, either $\angle(\vec{ps}, \vec{pv})$ or $\angle(\vec{pt}, \vec{pv})$ is at most $\beta = \pi/2 + \arcsin(\epsilon/2)$. Furthermore $\|sp\| \leq f(p)$ and $\|pv\| \leq f(p)$. Thus $R/2 \leq f(p)/3\sin(\pi/2 + \arcsin(\beta/2))$ and hence $R < f(p)/2$. ∎

Lemma 17 *If S is a $1/3$–sample of a smooth closed curve γ, the algorithm of Amenta, Bern, and Eppstein (Gold and Snoeyink) has no edges that are not in the polygonal reconstruction in the returned graph.*

Proof: We show that for an edge st not in the polygonal reconstruction, there is no ball with s and t on its boundary which has no endpoint of the dual edge of st in its interiors. This shows the Lemma for both algorithms.

Assume otherwise. Let B be such a ball, and v and v' the intersections of B with the perpendicular bisector of st. Let m be the midpoint of st (see Figure 17).

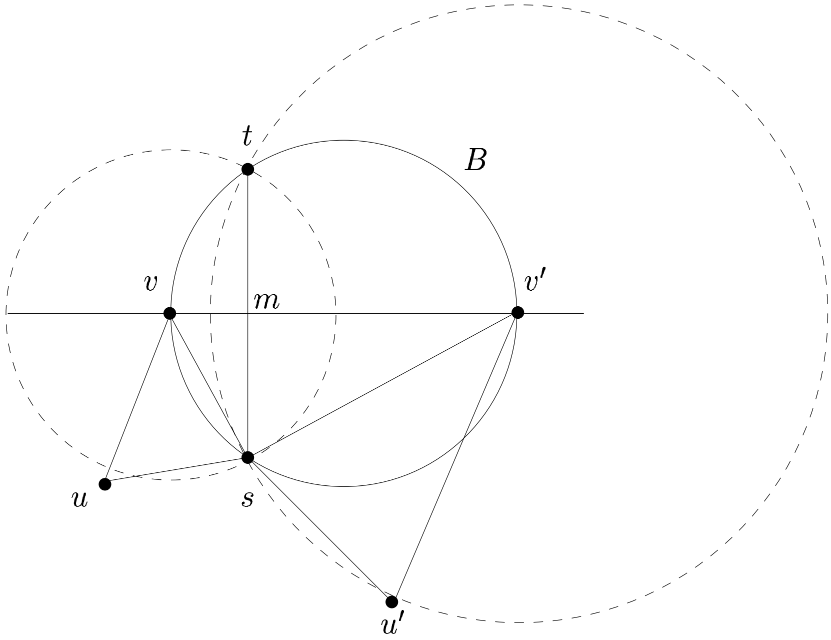

Figure 17: If the sampling condition is satisfied and u and u' are outside the two balls, the sum of the angles around s is more than 2π.

We derive a contradiction by showing that the sum of the angles around s is more than 2π.

The ball $B(v, \|vs\|)$ contains no sample point in its interior, since it is enclosed by one of the balls centered at the dual Voronoi points of the edge st through s and t. These balls have at least three sample points on their boundary and no sample point in their interior.

Thus there is a medial axis point on st (γ has at least two components in $B(m, \|ms\|)$, one containing s and another one containing t. Move along st. In s the closest curve point is s and in t the closest curve point is t. Between s and t, the closest curve point is always in $B(m, \|ms\|)$ and thus it must jump from the component of s to the component of t. At this point we have a medial axis point). Let without loss of generalitya medial axis point on sm and $\|sm\| = 1$.

Let u and u' be the sample points adjacent to s and p, p' the curve Voronoi points of su and su'. By Corollary 11, we know that $\angle(\vec{su}, \vec{su'}) \geq \pi - 4 \arcsin(\epsilon/2)$. The angle $\angle(\vec{sv}, \vec{sv'})$ is exaclty $\pi/2$, according to the Theorem of Thales.

Let $\epsilon \leq 1/3$. We bound the angle $\angle(\vec{su}, \vec{sv})$. Since $\epsilon \leq 1/3$ we have by Lemma 10 $\|su\| \leq 1$. Thus $\angle(\vec{su}, \vec{sv}) \geq \arccos(1/2\sqrt{\|sv\|^2 + 1})$. Analogously we have $\angle(\vec{su}, \vec{sv'}) \geq \arccos(1/2\sqrt{\|sv'\|^2 + 1})$. Since either $\|sv\|$ or $\|sv'\|$ is at least 1, we have $\angle(\vec{su}, \vec{sv}) + \angle(\vec{su}, \vec{sv'}) \geq \pi/3 + \arccos 1/4$.

We summarize $2\pi = \angle(\vec{su}, \vec{su'}) + \angle(\vec{sv}, \vec{sv'}) + \angle(\vec{su}, \vec{sv}) + \angle(\vec{su}, \vec{sv'}) \geq \pi - 4 \arcsin(\epsilon/2) + \pi/2 + \pi/3 + \arccos 1/4 > 2\pi$, a contradiction. ∎

3.4 A New Algorithm

We derive a variant of the algorithm of Dey and Kumar for which we can prove that a 1/2–sample suffices to guarantee the correct reconstruction.

To prove this, we use the following Lemma.

Lemma 18 *Let S be a 1/2–sample of a smooth curve γ. Let s and t be two non-adjacent sample points of S so that the center ball of st is empty of sample points. Then there is a medial axis point on the segment st with distance at most $3/4\|st\|$ from s.*

Proof: Assume otherwise. Move a point x along the segment st until it initially meets the medial axis and call this point m (see Figure 18). Lemma 11 states that there is a medial axis point on this segment.

Since none of the interior points is a medial axis point, m is in the same connected component of $\gamma \cap B(s, s + \vec{st}/2)$ as s. Thus the connected component of s touches the circle centered at m with radius $\|mt\|$. Let p be the intersection point of γ with this circle. If we add p to the sample set, the new sample set is clearly a 1/2 sample. Since the center ball of st contains no sample point and the component of γ between s and p does not leave the ball (since it is a component with respect to this ball), p is adjacent to s in the new sample set. Let u be the curve Voronoi point of sp. So $\|up\| = \|us\| \geq 1/4\|st\|$ and $f(u) \leq \|up\| + \|st\|/2$. From Corollary 10, we know that either p or s is the nearest sample point of u, but $\|up\| - f(p)/2 \geq \|up\| - \|up\|/2 - \|st\|/4 \geq 0$. This contradicts the sampling condition. ∎

We now derive the promised variant of the algorithm of Dey and Kumar.

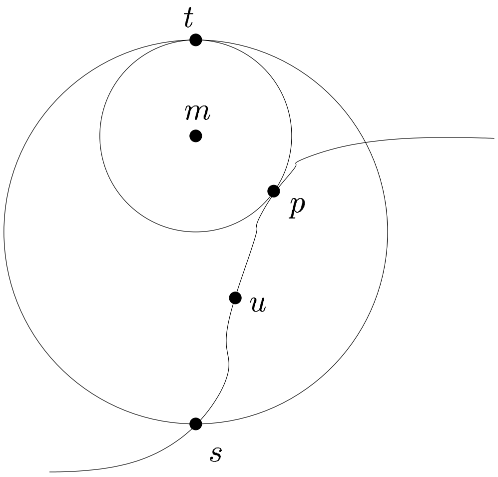

Figure 18: Adding the point p to the sample set leads to a contradiction.

- For every point $p \in S$, add the edge between p and the nearest neighbour of p.
- For every point p of degree 1, add the edge between p and the nearest neighbour that forms an angle of at least $\pi - 2\arcsin(1/4)$ with the edge adjacent to p.

Lemma 19 *If S is a 1/2–sample of a smooth closed curve γ, the modified algorithm of Dey and Kumar returns the polygonal reconstruction.*

Proof: As above, it suffices to show that the algorithm adds only edges of the polygonal reconstruction.

Assume there is an edge st, added in phase one, which is not of the polygonal reconstruction. Let without loss of generality$t$ be the nearest neighbour of s. Let p and q be the two adjacent samples of s and u and v the curve Voronoi points between sp and sq (see Figure 19). Lemma 18 states that there is a medial axis point on st with a distance of at most $3/4\|st\|$ from s and Lemma 14 states that $\angle(\vec{su}, \vec{sv})$ is at least $\pi - 2\arcsin(1/4)$. Thus either $\angle(\vec{st}, \vec{su})$ or $\angle(\vec{st}, \vec{sv})$ is at most $\pi/2 + \arcsin(1/4)$. Let without loss of generality$\angle(\vec{st}, \vec{su}) \leq \pi/2 + \arcsin(1/4)$ and $\|su\| = 1$. Then $\|st\| \leq 2$, since t is the nearest neighbour from s and $\|sp\|$ is at most 2. Thus

$$f(u) \leq \|p(s + 3/4\vec{st})\| \leq \sqrt{1^2 + (2 \cdot 3/4)^2 - 2 \cdot 3/4 \cos(pi/2 + \arcsin(1/4))} < 2,$$

a contradiction to the sampling condition.

We turn to phase two. Assume there is an edge st added in phase two, which is not of the polygonal reconstruction. Let without loss of generality$t$ be the nearest neighbour of s which forms an angle of at most $2\arcsin(1/4)$ with s. Let p be the point, adjacent to s, which is not added in phase 1 and let u be the curve Voronoi point between s and p (see Figure 20). Let without loss of generality$\|su\| = 1$. Lemma 18 states that there is

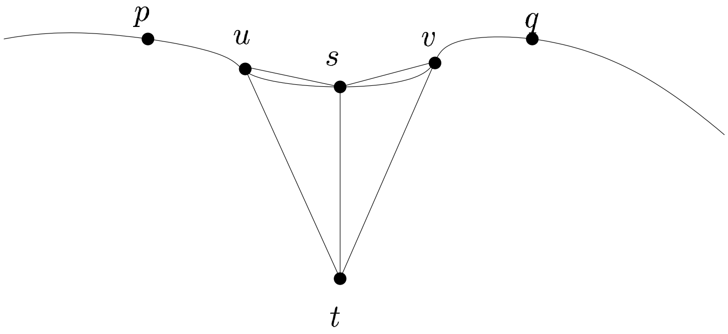

Figure 19: Either the angle $\angle(\vec{st}, \vec{su})$ or $\angle(\vec{st}, \vec{sv})$ is small and thus the distance from either u or v to the medial axis point is less than two times the distance to s.

a medial axis point and st with a distance of at most $3/4\|st\|$ from s. Furthermore the angle $\angle(\vec{st}, \vec{su})$ is at most $3\arcsin(1/4)$ and $\|st\| \leq 2$. Thus

$$f(u) \leq \|p(s + 3/4\vec{st})\| \leq \sqrt{1^2 + (2 \cdot 3/4)^2 - 2 \cdot 3/4 \cos(3\arcsin(1/4))} < 2$$

yeilds a contradiction to the sampling condition.

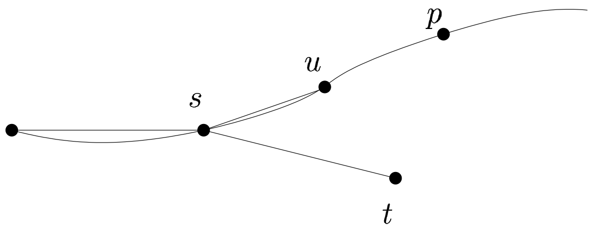

Figure 20: The angle $\angle(\vec{st}, \vec{su})$ is small and hence the distance from u to the medial axis point is less than twice the distance from u to s.

∎

3.5 The Algorithm of Dey, Mehlhorn, and Ramos

Dey, Mehlhorn, and Ramos developed the first algorithm that comes with guarantees for collections of open and closed curves. We do not present this algorithm in detail, but discuss the problems which arise when open curves are allowed and how to overcome these difficulties.

The first problem they had to overcome is defining the problem properly. Assume the sampling condition should be such that if S is a sufficiently dense sample, every superset $S' \supset S$ which is a sample set, is also sufficiently dense. Then there are open curves that have no sufficiently dense sample: Assume otherwise. Let S be a dense sample for the unit-circle and let p and q be two neighboured sample points. Let S' be a sufficiently dense sample for the unit-circle, without the arc between p and q. Then $S \cup S'$ is a dense sample for both curves.

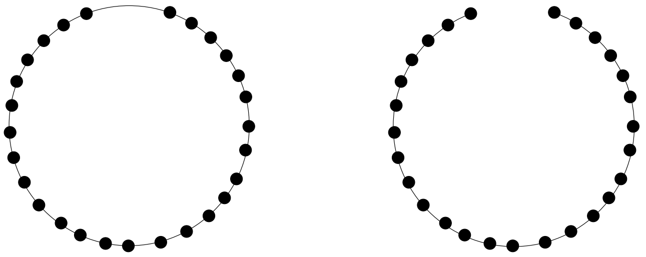

Figure 21: The same sample set is a dense sample for curves with different polygonal reconstructions

A *feasible reconstruction* G for S is a set of edges, so that there exists a curve γ, for which S is a "dense" sample and G is the polygonal reconstruction. Dey, Mehlhorn, and Ramos overcome this problem in the following way. They looked at all polygonal reconstructions, for which they can prove feasiblility and return the polygonal reconstruction which is "maximal". They prove that this reconstruction is a superset of the polygonal reconstruction of the underlying curve. So they return the maximum feasible reconstruction they can justify, i.e., beside the maximal reconstruction G, they return a curve Γ connecting the points as defined by G, so that the input sample is a dense sample for the curve. Certainly the sampling condition to guarantee that the result is a superset of the polygonal reconstruciton has to be stronger than the guaranteed sampling density of the returned curve Γ. The algorithm has an parameter ρ to adjust the required and the guaranteed sampling density. The exact result is stated as follows:

Theorem 13 (Dey, Mehlhorn, Ramos (99)) *Let S be a set of points in the plane and $\rho < 1$. The reconstruction G and the curve Γ returned by the algorithm of Dey, Mehlhorn, and Ramos with parameter ρ satisfy the following two properties:*

- *If $\rho < 1/2$ and S is a $(\rho/8)$–sample from a curve γ then the polygonal reconstruction of γ for S is a subset of G.*
- *Let $c = 13.35$. If $\rho < 1/8$ then S is a $(c\rho)$–sample from Γ and G is the polygonal reconstruction of Γ for S.*

The algorithm starts with a set of edges obtained from a known algorithm. Then they delete edges until they can justify the result. The result of the algorithm is such that, for any edge pq, the two balls with radius $\|pq\|/\rho$ are empty of Voronoi points of S and the ball centered at the midpoint of pq with radius $\|pq\|/4\rho$ contains no node with degree less than 2. Notice that no point can have degree higher than 2.

The proof that the edges of the polygonal reconstruction satisfy the properties above is very similar to the proof of Amenta, Bern, and Eppstein. Furthermore Dey, Mehlhorn, and Ramos gave a construction for the curve Γ.

4 The TSP-Algorithm

In this section we show that the Traveling Salesman path or cycle solves the curve-reconstruction problem for single open or closed curves, respectively, for sufficient dense sample sets. We derive this by showing that the Held-Karp relaxation is equal to the polygonal reconstruction. In addition, we show how this relaxation can be computed in polynomial time Furthermore, we extend the result for collections of closed curves.

4.1 Statement of the Results

Our research was motivated by the following recent result of Giesen [Gie99]. He proves that the Traveling Salesman tour solves the curve reconstruction problem for uniformly sampled benign semi-regular curves. The result was stated as follows.

Theorem 14 ([Gie99]) *For every benign semi-regular closed curve* γ *there exists an* $\epsilon > 0$ *with the following property: If* V *is a finite sample set of* γ*, so that for every* $x \in \gamma$ *there is a* $p \in V$ *with* $\|pv\| \leq \epsilon$*, the optimal Traveling Salesman tour is the polygonal reconstruction of* γ*.*

An efficient successful method for solving the Traveling Salesman problem is to use a branch-and-cut algorithm based on the Subtour-LP. In general, the optimal solution of the Subtour-LP is fractional. Our main result states that the optimal solution of the Subtour-LP is integral, whenever V is a sufficiently dense sample of a benign semi-regular curve. The reader might be interested to know that we discovered this fact in our experiments on curve reconstruction. We had implemented a branch-and-cut algorithm based on the Subtour-LP, the algorithm which solves the Subtour-LP and is then supposed to branch on a fractional variable. We observed that the algorithm never branched. After seeing this behavior in a large number of examples, we formulated it as a conjecture and set out to prove it.

Theorem 15 (Main Theorem) *Let* γ *be a closed benign semi-regular curve, let* V *be a finite set of samples of* γ*. If* V *satisfies the sampling condition given below, then:*

- *The optimal Traveling Salesman tour of* V *is a polygonal reconstruction of* γ*.*
- *The Subtour-LP for Traveling Salesman tours has an optimal integral solution and this solution is unique.*

In the case of an open curve, let a *and* b *be the first and last sample point, respectively (in the order on* γ*). The statements above hold true for the optimal Traveling Salesman path with endpoints* a *and* b *and the Subtour-LP for Traveling Salesman paths with specified endpoints; furthermore they also hold for the Traveling Salesman path with unspecified endpoints and for the Subtour-LP for Traveling Salesman paths with unspecified endpoints. The latter result required a strengthened sampling condition.*

We will prove our main theorem for open curves with specified and unspecified endpoints in Sections 4.2 and 4.3, respectively. The proof for closed curves follows in Section 4.4.

Our main theorem suggests a reconstruction algorithm for benign semi-regular curves: *Solve the Subtour-LP. If the optimal solution is integral, output it.* We briefly discuss two strategies for solving the Subtour-LP.

As stated in Section 2.2.3, there are several methods for solving the Subtour-LP. A potentially exponential, but practically very efficient algorithm uses the simplex method and the cutting plane framework. One starts with the LP consisting only of the degree constraints and then solves a sequence of LPs. In each iteration one checks whether the solution X^* to the current LP satisfies all subtour elimination constraints and, if not, one adds a violated subtour elimination constraint to the LP. The check for a violated subtour elimination constraint has been explained in Section 2.2.6. We use a simplex-based strategy in our experiments on curve reconstruction (see Section 6).

The Ellipsoid method [Sch86] solves the Subtour-LP in polynomial time in the size of the bit representations of the coefficients of the cost function. Distance values are, in general, non-rational numbers and hence the Ellipsoid method is not directly applicable in our setting. In Section 4.5, we extend our results to the situation where the position of the points and the distances between points are only approximately known and show how to obtain a polynomial time algorithm.

4.2 Open Curves

We assume that our open curve is oriented and write $p < q$ if p precedes q on γ. We use $B(p, r)$ and $B^0(p, r)$ to denote the closed and open ball with center p and radius r, respectively.

4.2.1 The Held-Karp Bound

Our proof of Theorem 15 exploits the connection between the Subtour-LP and the Held-Karp bound briefly introduced in Section 2.2.7. The purpose of this section is to review the relevant facts about the Held-Karp bound.

Let $G = (V, E)$ be an undirected graph, let a and b be two designated vertices of G, and let c be an arbitrary cost function on the edges of G. A function $\mu : V \to \mathbb{R}$ is called a *potential function*. It gives rise to a modified distance function c_μ via $c_\mu(u, v) = c(u, v) - \mu(u) - \mu(v)$. Consider now any Traveling Salesman path T with endpoints a and b. Its costs under c and c_μ are related by $c_\mu(T) = c(T) - 2\sum_{v \in V} \mu(v) + \mu(a) + \mu(b)$ since the path uses two edges incident to every vertex except for a and b. Observe that $c_\mu(T) - c(T)$ does not depend on T and hence the optimal Traveling Salesman path for endpoints a and b is the same under both cost functions. Let T_0 be an optimal Traveling Salesman path for endpoints a and b.

Let MST_μ be a minimum spanning tree with respect to the cost function c_μ and let $C_\mu = c_\mu(\mathrm{MST}_\mu)$ be its cost. Then $C_\mu \leq c_\mu(T_0)$, since a Traveling Salesman path is a spanning tree.

Fact 1 *Let μ be any potential function. If MST_μ is a Traveling Salesman path with endpoints a and b, it is an optimal Traveling Salesman path for endpoints a and b.*

Proof: From $C_\mu \le c_\mu(T_0)$, we conclude that MST_μ is an optimal Traveling Salesman path with respect to c_μ. Since the ranking of paths is the same under both cost functions, it is also optimal with respect to c. ▮

The inequality $C_\mu \le c_\mu(T_0) = c(T_0) - 2\sum_{v\in V}\mu(v) + \mu(a) + \mu(b)$ is valid for every potential function and hence

$$\max_\mu C_\mu + 2\sum_{v\in V}\mu(v) - \mu(a) - \mu(b) \le c(T_0)\ .$$

The quantity on the left is called the Held-Karp bound. The following fact which was shown in Section 2.2.7, is crucial for our proof.

Fact 2 *The Held-Karp bound is equal to the optimal objective value of the Subtour-LP.*

Proof: The fact follows by relaxing the degree constraints of the Subtour-LP in a Lagrangian fashion. For a short introduction to Lagrangian Relaxation see Section 2.2.7. ▮

We note that an optimal choice of μ in the Held-Karp bound is given by the optimal solution of the linear programming dual of the Subtour-LP; μ corresponds to the dual variables for the degree constraints. We next draw a simple consequence from the two facts above, which forms the basis of our proof.

Lemma 20 *Let μ be any potential function. If MST_μ is the unique minimum spanning tree with respect to c_μ and it is a Traveling Salesman path with endpoints a and b, then the Subtour-LP has a unique optimal solution and this solution is integral.*

Proof: If MST_μ is a Traveling Salesman path, it is optimal (Fact 1) and hence $c(\mathrm{MST}_\mu) = c(T_0)$. The Held-Karp bound is therefore equal to $c(T_0)$ and the same holds true for the optimal objective value of the Subtour-LP (Fact 2). The incidence vector of MST_μ is a feasible solution of the Subtour-LP of cost $c(\mathrm{MST}_\mu)$ and hence it is *an* optimal solution of the Subtour-LP. We will next argue that it is the unique optimal solution. Assume that there is an optimal solution of the Subtour-LP with $x_e > 0$ for some $e \notin \mathrm{MST}_\mu$. Since MST_μ is unique, there is a $\eta > 0$ so that decreasing the cost of e by η the minimum spanning tree will not change and hence the value of the Held-Karp bound will not change. However, the objective value of the Subtour-LP will decrease. This is a contradiction to the equality of the two bounds. ▮

We can now describe our proof strategy for open curves. We define a potential function μ so that MST_μ is the unique minimum spanning tree in the complete network $G = (V, V \times V, c)$, where c is the Euclidean distance function, and moreover MST_μ coincides with the polygonal reconstruction (and hence is a Traveling Salesman path with endpoints a and b, where a and b are the first and last sample point, respectively). Then MST_μ and hence the polygonal reconstruction is the unique optimal solution of the Subtour-LP. We want to stress that the definition of μ is only needed for the proof of our main theorem. The reconstruction algorithm simply solves the Subtour-LP.

4.2.2 Intuition

When will the minimum spanning tree of the sample set be the correct reconstruction?

Let $V = \{v_1, v_2, \ldots, v_n\}$ where we assume the points to be numbered according to their order on the curve. Kruskal's algorithm considers the edges $v_i v_j$ in increasing order of length and adds an edge to the spanning tree if it does not close a cycle. Kruskal's algorithm will therefore construct the path $v_1 – v_2 – \ldots – v_n$ if the potential function μ is such that

$$c_\mu(v_i, v_{i+1}) < c_\mu(v_h, v_j) \text{ whenever } h \leq i < j, j - h \geq 2 \ . \tag{8}$$

Let us consider two special situations: γ is essentially straight (any two left or right tangents to γ form an angle of less than $\pi/3$) or γ consists of a sharp corner (a point in which γ turns by at least $7\pi/24$) and two incident straight line segments (see Figure 22). We will show in Section 4.2.3 that any curve γ can be decomposed into subcurves which are either essentially straight or which consist of a sharp corner with two incident essentially straight legs.

For an essentially straight curve the minimum spanning tree will reconstruct for a large choice of potential functions. It will work without a potential function, i.e., $\mu(p) = 0$ for all p, and, more generally, it will work for any potential function that does not change too fast as a function of the position of its argument. For a point p which belongs to an essentially straight part of γ, we will essentially[1] define $\mu(p) = d(p)/3$, where $d(p)$ is maximal, such that $B^0(p, d(p)) \cap \gamma$ is connected and essentially straight and $B^0(p, r)$ denotes the open ball with center p and radius r. This choice guarantees that $\mu(p)$ changes slowly with the position of its argument (with at most one third of the change in argument) and that $\mu(p)$ depends on local properties of the curve and is large in parts of γ that are intuitively simple to reconstruct. For sharp corners, the definition above leads to a potential value of zero.

Corners with a turning angle of more than $\pi/2$ will confuse the minimum spanning tree when used without a potential function, as Figure 22 shows. One of our insights is that a simple potential function can be used to make the minimum spanning trees work. Assume that our curve consists of the two line segments $y = \pm m \cdot x$ for $0 \leq x \leq 1$ and let V be a finite set of samples. We define the potential as a function of the x-coordinates of the sample points. Fix $\pi(O)$ arbitrarily, let $p(x) = (x, mx)$ and $q(x) = (x, -mx)$ and define

---

[1] The precise definition given in Section 4.2.4 is more involved.

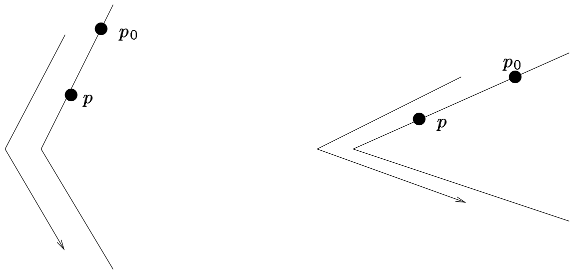

Figure 22: In the figure on the left, the Euclidean distance between p and p_0 grows as p moves away from p_0 along γ. In the figure on the right, the distance first grows and then shrinks again as p moves around the corner. The minimum spanning tree with $\mu(p) = 0$ for all p will reconstruct the curve on the left, but may fail on the curve on the right.

$\pi(p(x)) = \pi(q(x)) = \pi(O) - x$. Then $c_\pi(p(x), p(x_0)) = \sqrt{1+m^2} |\, x - x_0 \,| - 2\pi(O) + x_0 + x$ and $c_\pi(q(x), p(x_0)) = \sqrt{(x - x_0)^2 + (mx + mx_0)^2} - 2\pi(O) + x_0 + x$. It is an easy exercise in calculus to show that $c_\pi(p(x), p(x_0))$ is an increasing function of $|\, x - x_0 \,|$ and that $c_\pi(q(x), p(x_0))$ is an increasing function of x. We conclude that the minimum spanning tree for the modified distance function reconstructs. In the argument above the choice of $\pi(O)$ is arbitrary. The actual choice of $\pi(O)$ will depend on local properties of the curve. For every sharp corner s we will consider the largest open disk $B^0(s, c_s)$ such that $\gamma \cap B^0(s, c_s)$ is connected and is essentially a sharp corner with two incident straight legs. We set $\pi(O) = c_s$.

We can now sketch our definition of μ. For every point $p \in \gamma$, the first definition $\mu(p) = d(p)/3$ is applicable. It assigns potential zero to sharp corners. Near sharp corners we use the second definition, namely $\mu(p) = c_s - \|sp_s\|$, where p is near the sharp corner s and p_s is the projection of p onto the angular bisector of the two tangents at s (see Figure 27).

The analysis above suggests that with this definition of μ, the minimum spanning tree solves the reconstruction task locally, i.e., if given the points in $V \cap \gamma'$, where γ' is a subcurve of γ that is either essentially straight or a sharp corner with two incident straight legs. In other words, inequality (8) holds if v_h,v_i, and v_j belong to the same γ'.

Does it also hold globally?

Consider the situation shown in Figure 23. We have two points p and q that belong to distinct essentially straight parts of γ. We have $\max(d(p), d(q)) \leq \|pq\|$ and hence $c_\mu(p, q) > 0$. More generally, we will show in Section 4.2.3 that any edge pq, where p and q do not belong to an either essentially straight subcurve or to a sharp corner with its incident legs, has positive modified cost.

The previous paragraph suggests our sampling condition. We require that any edge of the polygonal reconstruction has non-positive reduced cost. Then (8) certainly holds when $c_\mu(v_j, v_j) > 0$. When $c_\mu(v_h, v_j) \leq 0$, v_h and v_j are guaranteed to lie in a common essentially straight subcurve or near a common sharp corner, and the local analysis applies.

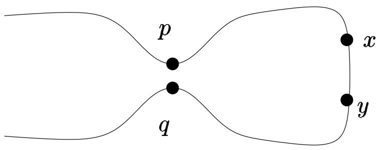

Figure 23: The edge pq does not belong to the polygonal reconstruction. Our definition of the potential function ensures that $c_\mu(p,q) > 0$. Our sampling condition that any edge in the polygonal reconstruction has non-positive reduced cost.

This ends the informal description of our proof.

4.2.3 The Sampling Condition and the Global Reasoning for Open Curves

We will give the detailed definition of our potential function in Section 4.2.4; it assigns a positive real $\mu(x)$ to every point x of γ. Define the turning angle of a subcurve γ' of γ as the opening angle of the smallest double-cone that contains all left and right tangents to points $p \in \gamma'$. We require:

Sampling Condition:

(a) For any two adjacent (on γ) samples u and v: $c_\mu(u,v) \leq 0$.

(b) For any two adjacent samples u and v: $\gamma[u,v]$ turns by less than π. For two adjacent points p, q on the curve, $\gamma[p,q]$ denotes the subcurve of γ with endpoints p and q not containing another sample point (In the case of closed curves we always have at least 3 sample points).

Condition (a) implies condition (b) in the case of open curves as we will see below. For closed curves, condition (a) is not sufficient as the example in Figure 24 indicates. Condition (a) states that adjacent sample points must be sufficiently close in a metric sense and condition (b) states that the curve must not turn too much between adjacent sample points.

The sampling condition is easily satisfied. Let $\epsilon = \inf_{x\in\gamma}\mu(x)$. Then $\epsilon > 0$ since γ is compact and hence a sample set in which γ turns by less than π between adjacent samples and in which there is at least one sample point in every curve segment of length $\epsilon/2$ satisfies the sampling condition. We want to stress that the sampling condition can also be satisfied with non-uniform sampling. In regions of γ where μ is large, the sampling may be less dense than in regions where μ is small. In Section 5.3 we will relate our sampling condition to the conditions used in other papers on curve reconstruction.

In order to show that MST_μ is the polygonal reconstruction of γ from V, we define a family Γ of (overlapping) subcurves γ' of γ so that:

(P1) Each subcurve γ' is connected and the minimum spanning tree (with respect to cost function c_μ) of the points in $V \cap \gamma'$ is unique and coincides with the polygonal reconstruction of γ'.

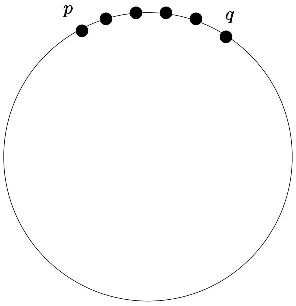

Figure 24: p and q are adjacent sample points and $c_\mu(p, q) \leq 0$ if p and q are sufficiently close. However, the sample set contradicts our intuition of what constitutes a dense sample set. Condition (b) excludes the case.

(P2) For every edge e with $c_\mu(e) \leq 0$ there is a subcurve $\gamma' \in \Gamma$ containing both endpoints of e.

We show that these conditions imply that MST_μ is the polygonal reconstruction of γ.

Lemma 21 *Conditions (P1) and (P2) imply that MST_μ is the polygonal reconstruction of γ.*

Proof: Let MST_μ be any minimum spanning tree in (V, E, c_μ) and let $e = uv$ be any edge which does not belong to the polygonal reconstruction. We show that $e \notin \mathrm{MST}_\mu$. Observe first that any edge in the minimum spanning tree has non-positive modified cost since there is a spanning tree, namely the polygonal reconstruction, in which every edge has non-positive cost. This follows from the cycle rule for minimum spanning trees. So assume $c_\mu(e) \leq 0$. Then there is a subcurve $\gamma' \in \Gamma$ containing both endpoints of e by (P2). We even have $\gamma[u, v] \subseteq \gamma'$ since γ' is connected by (P1). Moreover, the minimum spanning tree of $V \cap \gamma'$ is unique and coincides with the polygonal reconstruction of γ'. Thus $c_\mu(e') < c_\mu(e)$ for every edge e' on the part of the polygonal reconstruction between u and v. We conclude that $e \notin \mathrm{MST}_\mu$. ∎

4.2.4 The Definition of the Potential Function

In this section, we give the precise definition of our potential function. The definition depends on the thresholds $\theta_{max\_sharp}, \theta_{turn}, f_{scale}, f_{wriggle}, \theta_{wriggle}$, and f_{shrink}, whose choice is somewhat arbitrary but not completely independent. In Section 4.2.6 we summarize the conditions.

Singularities cause difficulties for most curve reconstruction algorithms; the difficulties grow with the turning angle. We call a singularity p a *sharp corner* if the turning angle at p is at least $\theta_{max\_sharp} = 7\pi/24 = 52.5°$.

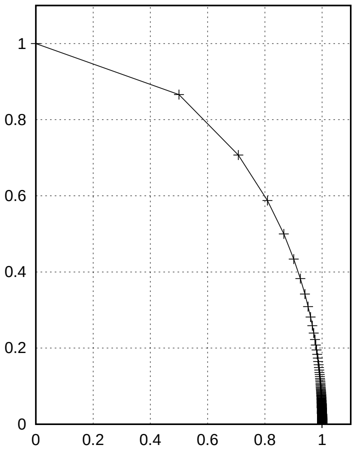

Figure 25: A semi-regular curve with an infinite number of corners

A semi-regular curve may have an infinite number of singularities. For example, the convex hull of the points $(\cos(\pi/n), \sin(\pi/n))$, $n \geq 2$, has an infinite number of singularities (see Figure 25). However, a semi-regular curve can have only a finite number of sharp corners. Assume otherwise. Then the sharp corners have an articulation point p. Let p_1, p_2, … be an increasing sequence of sharp corners converging to p. For any sharp corner p_i we can choose two points q_i and r_i in the vicinity of the corner so that the tangents at q_i and r_i form an angle of at least $\pi/6$ and so that the sequence q_1, r_1, q_2, r_2, … increases and converges to p. The sequence shows that γ has no left tangent at p.

We use S to denote the set of sharp corners of γ.

We are now ready to define our potential function. The definition consists of two parts dealing with the neighborhoods of sharp corners and curve parts "far away" from all sharp corners, respectively. We start with the latter parts.

For every point $p \in \gamma$, let $d(p)$ be maximal so that the open ball $B^0(p, d(p))$ has the following properties:

- $B^0(p, d(p)) \cap S = \emptyset$.
- $B^0(p, r) \cap \gamma$ is connected for all r with $r \leq d(p)$.
- $B^0(p, d(p)) \cap \gamma$ turns by less than $\theta_{turn} = \pi/3$.

We will define our potential in parts that are far away from sharp corners as $f_{scale} d(p)$ and choose $f_{scale} = 1/3$. We will define later what we mean exactly by "far away".

Observe that the closed ball $B(p, d(p))$ has one of the following properties: it has a point in S on its boundary, it intersects γ in more than one component, or γ turns by $\pi/3$ in the ball (see Figure 26). For sharp corners $s \in S$, we have $d(s) = 0$ and for points $p \in \gamma \setminus S$, we have $d(p) > 0$. The function $p \mapsto d(p)$ is continuous. Thus $d(p)$ will be an increasing function of p as p moves away from a sharp corner for a neighborhood of any

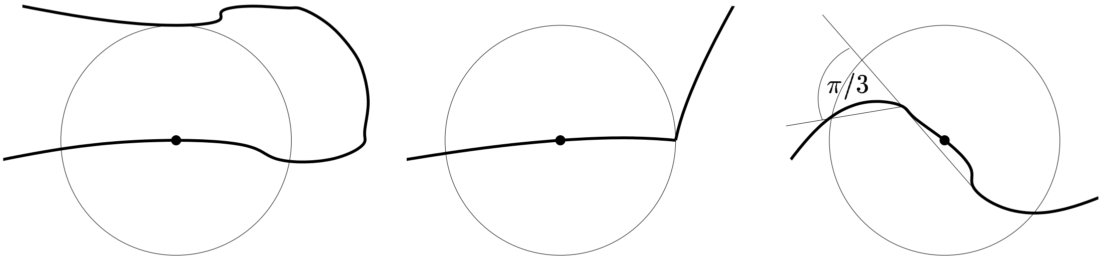

Figure 26: The closed ball intersects γ in more than one part, has a singularity on its boundary, or γ turns by $\pi/3$ in the ball.

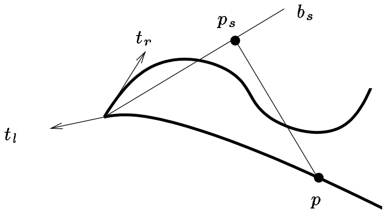

Figure 27: Illustration of the definition of b_s and p_s.

sharp corner. In Section 5.3 we will relate $d(p)$ to the distance of p to the medial axis of γ.

For sharp corners we define a quantity δ_s. For a sharp corner $s \in S$ let b_s be the bisector of the angle between the right tangent and the reversal of the left tangent, let α_s be the turning angle at s (we have $\alpha_s \geq 7\pi/24$).

For an angle α, let $\bar{\alpha} = \pi - \alpha$. For a point $p \in \gamma$, let p_s be the orthogonal projection of p onto b_s. See Figure 27 for an illustration of these definitions. For every $s \in S$, let δ_s be maximal, so that

- $\gamma \cap B^0(s, \delta_s)$ is connected. We call the two components of $(\gamma \setminus s) \cap B^0(s, \delta_s)$ the two legs of γ incident to s.
- The angle between any segment with both endpoints on one leg and the tangent in s of the same leg is less than $\min\{f_{wriggle}\bar{\alpha}_s, \theta_{wriggle}\}$. We choose $f_{wriggle} = 1/4$ and $\theta_{wriggle} = \pi/9$. The second bound guarantees that the angle between any segment and the perpendicular bisector is less than $\pi/2$.
- For either of the two legs $d(p)$ increases as p moves away from s.
- $B^0(s, 2\delta_s)$ contains no sharp corner different from s.

The last condition ensures that the balls $B^0(s, \delta_s)$, $s \in S$, are pairwise disjoint. For an illustration of the definition, see Figure 28. Clearly, $\delta_s > 0$ for all $s \in S$.

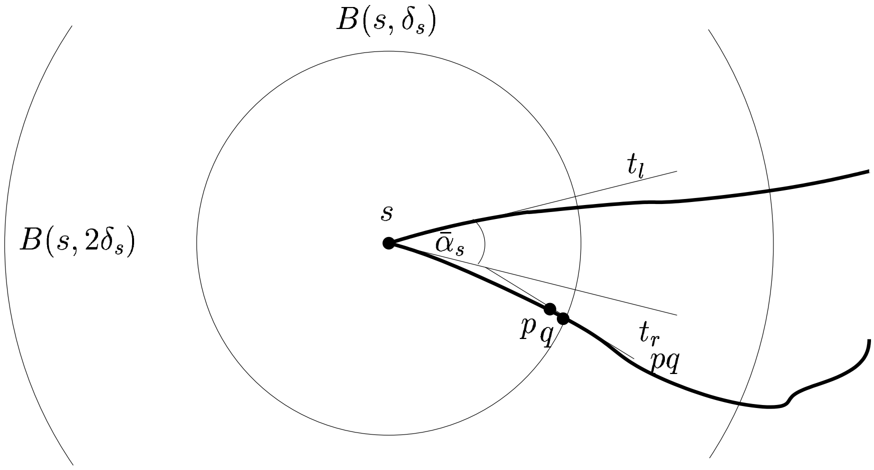

Figure 28: By the definition of δ_s, we know that γ is connected in $B^0(s, \delta_s)$, there is no sharp corner in $B^0(s, 2\delta_s)$, the angle between a segment through two points of a leg in $B^0(s, \delta_s)$, say pq, and the tangent of this leg, say t_r, is less than $\min(\alpha_s/4, \pi/9)$, and the d values are increasing in $B(s, \delta_s)$.

Recall from Section 4.2.2 that we want to define the potential near a sharp corner as $c_s - \|sp_s\|$. The change from the potential for sharp corners and smooth areas is made by choosing the maximum of the two possibilities. We use the constant c_s to specify the exact point where we change from one definition to the other. We choose c_s maximal under the restriction that all points outside the $B(s, \delta_s f_{shrink})$ ball ($f_{shrink} = 1/5$) have potential $d(p)/3$. Thus let q^1 and q^2 be the points where the two legs intersect the boundary of the circle $B(s, \delta_s/5)$ and let

$$c_s = \min\left\{ d(q^i)/3 + \|sq_s^i\| \ ; \ i = 1, 2 \right\}.$$

Then $c_s \leq 2\delta_s/5$ since $d(q^i) \leq \|sq^i\| = \delta_s/5$ and $\|q_s^i\| \leq \|sq_i\| = \delta_s/5$.

Remark: In the proofs of Section 4.2.5, we will only use the fact that $0 < c_s \leq 2\delta_s/5$, the exact value of c_s does not matter. This will become important in Section 4.4.

We are now ready to define our potential-function μ:

$$\mu(p) = \begin{cases} d(p)/3 & \text{if } p \text{ is in no } B(s, \delta_s). \\ \max\{c_s - \|sp_s\|, d(p)/3\} & \text{if } p \in B(s, \delta_s). \end{cases}$$

Observe that this definition "combines" the two cases discussed in Section 4.2.2. We use T to denote the set of points $p \in \gamma$ with $\mu(p) = d(p)/3$, i.e. , the points that are not affected by the singularities. Then $q^i \in T$ since $c_s \leq d(q^i)/3 + \|sq_s^i\|$ and hence $c_s - \|sq_s^i\| \leq d(q^i)/3$. Since $d(p)$ and $\|sp_s\|$ increase as p moves away from s (at least as long as $p \in B^0(s, \delta_s)$), we have $p \in T$ for any curve point p that is not contained in $\cup_{s \in S} B^0(s, \delta_s/5)$. This also implies that μ is a continuous function. For a point $p \in B(s, \delta_s/5)$, we have $d(p) \leq \delta_s/5$ and hence $\mu(p) \leq c_s$.

We will frequently use the following simple observation.

Lemma 22 *Let* $u \in T \cap B(s, \delta_s)$ *for some* $s \in S$ *and let* v *be a point on the other leg of* s*. Then* $d(u) \leq \|uv\|$*.*

Proof: Assume otherwise, i.e. , $\|uv\| < d(u)$. We have $d(u) \leq \|us\|$ and hence $B^0(u, d(u)) \cap \gamma$ consists of at least two components, one containing u and one containing v. ▮

4.2.5 Local Reasoning

We consider the following family Γ of subcurves:

1. $B^0(p, d(p)) \cap T \cap \gamma$ for all $p \in T$.
2. $B^0(s, \delta_s) \cap \gamma$ for all $s \in S$.

We call the subcurves of the first kind regular subcurves and the subcurves of the second kind singular subcurves.

Lemma 23 *The subcurves* $\gamma' \in \Gamma$ *are connected.*

Proof: This is obvious for singular subcurves. So consider a subcurve $\gamma' = B^0(p, d(p)) \cap T \cap \gamma$ for some $p \in T$. The subcurve $B^0(p, d(p)) \cap \gamma$ is connected by definition. If γ' is not connected, $B^0(p, d(p)) \cap \gamma$ decomposes into three non-trivial segments γ_1, γ_2, and γ_3 with $\gamma_1 \cap T \neq \emptyset$, $\gamma_2 \cap T = \emptyset$ and $\gamma_3 \cap T \neq \emptyset$. This implies that γ_2 passes through a sharp corner, a contradiction to the definition of $d(p)$. ▮

Lemma 24 *Let* u *and* v *be adjacent sample points and let* $\gamma' \in \Gamma$*. If* γ' *contains* u *and* v*, then* $\gamma[u, v] \subseteq \gamma'$*.*

Proof: γ' is connected and hence either $\gamma[u, v] \subseteq \gamma'$ or $\gamma \setminus \gamma[u, v] \subseteq \gamma'$. The latter case is impossible, since $\gamma \setminus \gamma[u, v]$ is not connected in the case of an open curve and turns by more than π according to our second sampling condition in the case of a closed curve. However γ' turns by less than π according to the definition of Γ. ▮

For open curves Lemma 24 holds true without the second sampling condition. Since curves $\gamma' \in \Gamma$ turn by less than π, the second sampling condition is implied by the first for open curves. We will next verify the properties (P1) and (P2).

Lemma 25 (Property (P2)) *Let* $e = pq$ *be an edge with non-positive modified cost. Then there is a subcurve* $\gamma' \in \Gamma$ *containing* p *and* q*.*

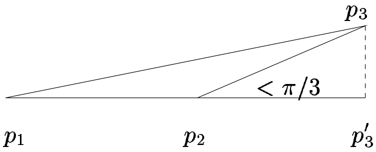

Figure 29: The situation in the proof of Lemma 27. We have $\|p_1p_3\| \geq \|p_1p_3'\| > \|p_1p_2\| + \|p_2p_3\| \cos \pi/3$.

Proof: We have $\|pq\| \leq \mu(p) + \mu(q)$ by assumption. If $p \in T$ and $q \in T$, $\mu(p) = d(p)/3$ and $\mu(q) = d(q)/3$ and hence $\|pq\| \leq (d(p)+d(q))/3 \leq 2\max(d(p), d(q))/3$. Thus $\{p,q\} \subseteq B^0(x, d(x))$ for one of the endpoints x of e.

If one of the endpoints does not belong to T, say $p \notin T$, then $p \in B^0(s, \delta_s/5)$ for some sharp corner. If $q \in B^0(s, \delta_s)$ we are done. So assume otherwise. We have $\mu(p) \leq 2\delta_s/5$ and $\|pq\| \geq 4\delta_s/5$. If $q \in T$, then $\mu(q) = d(q)/3 \leq \|sq\|/3 \leq (\|pq\| + \delta_s/5)/3$ and hence $\|pq\| \leq (\|pq\| + \delta_s/5)/3 + 2\delta_s/5$ or $2\|pq\|/3 \leq 7\delta_s/15$ or $\|pq\| \leq 21\delta_s/30$, a contradiction to $\|pq\| \geq 4\delta_s/5$. If $q \notin T$ then $q \in B^0(t, \delta_t/5)$ for some sharp corner t different from s and hence $\|pq\| > 4(\delta_s + \delta_t)/5$. But $\mu(p) \leq 2\delta_s/5$ and $\mu(q) \leq 2\delta_t/5$, a contradiction. ∎

In order to show that the MST$_\mu$ coincides with the polygonal reconstruction, we show that the modified distance between two points p and r is either non-negative, or the modified distance from p to r is greater than the modified distance from p to any point q between p and r. Since we also want to use these Lemmas if we treat the problem with finite precision arithmetic, we quantify the change of the modified distance in the distance of q and r.

Lemma 26 *Let p, q, and r be sample points on a regular subcurve with $p < q < r$. If $\{p,q,r\} \in B(t, d(t))$ for some point $t \in \gamma$ then $c_\mu(p,r) - c_\mu(p,q) \geq \|qr\|/6$.*

Proof: Let $t \in \gamma$ and $p < q < r \in B(t, d(t))$. Since the points are contained in a regular subcurve, we have $\angle(\vec{pq}, \vec{qr}) < \pi/3$ and hence $\|pr\| > \|pq\| + \|qr\| \cos \pi/3 = \|pq\| + \|qr\|/2$ (see Figure 29). Furthermore observe that $d(r) \leq d(q) + \|qr\|$. Thus

$$c_\mu(p,r) - c_\mu(p,q) = \|pr\| - \mu(p) - \mu(r) - \|pq\| + \mu(p) + \mu(r) \geq \|qr\|/2 - \|qr\|/3 = \|qr\|/6.$$

∎

Lemma 27 (Prop. (P1) for regular regions) *MST$_\mu$ coincides with the polygonal reconstruction for regular subcurves.*

Proof: The Lemma above for both possible orientations of the curve directly implies that Prim's minimum spanning tree algorithm finds the polygonal reconstruction. ∎

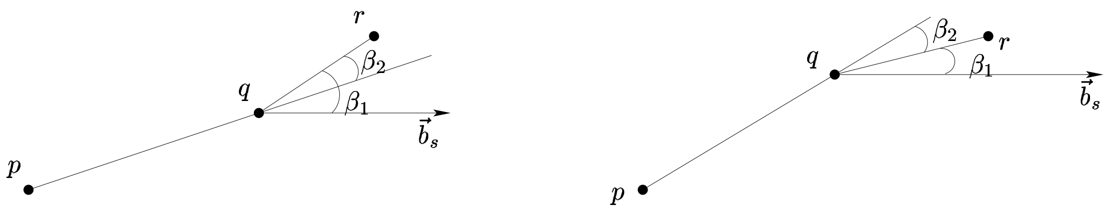

Figure 30: Illustration of the definition of β_1 and β_2, if p, q, and r are on the same leg.

Lemma 28 *Let p, q, and r be sample points of γ with $p < q < r$. If $\{p, q, r\} \in B(s, \delta_s)$ for some sharp corner $s \in \gamma$ then $c_\mu(p, r) \geq 0$ or $c_\mu(p, r) - c_\mu(p, q) \geq \|qr\|(\sin \bar{\alpha}_s/4)^2/3$.*

Proof: We first argue that it is sufficient to prove the claim for the situations where either q and r belong to T or neither of them does. Assume for the moment that those two cases have been dealt with. If exactly one of q and r belongs to T, there is a point u between q and r which belongs to the boundary of T. For this point, we have $d(u)/3 = c_s - \|su_s\|$ and hence u can be considered to be in T or outside T. The triples (p, q, u) and (p, u, r) are both in one of the special situations and hence we have $c_\mu(p, u) \geq 0$ or $c_\mu(p, u) - c_\mu(p, q) \geq \|qu\|(\sin \bar{\alpha}_s/4)^2/3$ and $c_\mu(p, r) \geq 0$ or $c_\mu(p, r) - c_\mu(p, u) \geq \|ur\|(\sin \bar{\alpha}_s/4)^2/3$. If $c_\mu(p, r) \geq 0$, we are done. Otherwise $c_\mu(p, r) < 0$ and hence $c_\mu(p, u) < 0$. Thus $c_\mu(p, r) - c_\mu(p, q) = c_\mu(p, r) - c_\mu(p, u) + c_\mu(p, u) - c_\mu(p, q) \geq (\|qu\| + \|ur\|)(\sin \bar{\alpha}_s/4)^2/3 \geq \|qr\|(\sin \bar{\alpha}_s/4)^2/3$, where the last inequality follows from the triangle inequality. From now on we may assume that either q and r belong to T or neither of them does.

We need some further case distinctions. The first distinction is according to the sign of $c_\mu(p, q)$. The case $c_\mu(p, q) > 0$ is dealt with in the last paragraph of the proof.

We start with the assumption $c_\mu(p, q) \leq 0$. We make a further case distinction according to the position of s in the sequence $p < q < r$. In all four cases we employ a common strategy. We have $c_\mu(p, r) - c_\mu(p, q) = (\|pr\| - \|pq\|) - (\mu(r) - \mu(q))$. We bound $\|pr\| - \|pq\|$ from below and $\mu(r) - \mu(q)$ from above and estimate the difference of the bounds. In all cases we also use the estimates $\bar{\alpha}_s \leq 17\pi/24$, $\sin \bar{\alpha}_s/4 \leq 0.5281$, $(\sin \bar{\alpha}_s/4)^2/3 \leq 0.1$ and $\cos 2\pi/9 \geq 2/3$.

$s \leq p < q < r$: If $\{q, r\} \cap T = \emptyset$, $\mu(r) \leq \mu(q)$, and if $\{q, r\} \subseteq T$, $\mu(r) \leq \mu(q) + \|qr\|/3$. In either case[2], $\mu(r) \leq \mu(q) + \|qr\|/2$.

Let β_2 be the angle between the vectors $\vec{pq}$ and $\vec{qr}$. Then $\|pr\| - \|pq\| \geq \|qr\| \cos \beta_2$. By the sampling condition, we have $\beta_2 \leq 2\pi/9$, since the angle between pr, respectively qr, and the left tangent at s is at most $\pi/9$. Thus

$$c_\mu(p, r) - c_\mu(p, q) \geq \|qr\|(\cos(\beta_2) - 1/2) \geq \|qr\|/6 \geq \|qr\| \sin(\bar{\alpha}_s/4)^2/3 .$$

---

[2] In Section 4.4.2 we will consider a modified potential function for which we only know $\mu(r) \leq \mu(q) + \|qr\|/2$. We want to reuse the proof there.

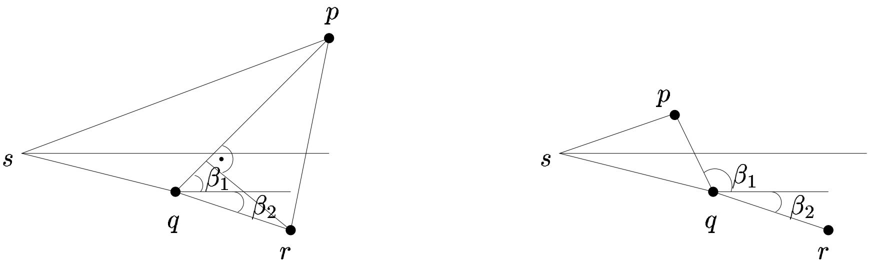

Figure 31: The case $p < s \le q < r$: In the left part q_s is closer to s than p_s and in the right part the converse is true. In both situations the definitions of β_1 and β_2 are illustrated.

$p < q < r \le s$: If $\{qr\} \subseteq T$, $\mu(r) \le \mu(q) \le \mu(q) + \|qr\|/2$. Let β_2 be the angle between the vectors $\vec{pq}$ and $\vec{qr}$. Then $\|pr\| - \|pq\| \ge \|qr\| \cos \beta_2$ and $\beta_2 \le 2\pi/9$ and hence the argument used in the case $s \le p < q < r$ applies.

Assume next that $\{q, r\} \cap T = \emptyset$. Let β_1 be the angle between the vectors $\vec{qr}$ and $\vec{b}_s$. Then $\mu(r) - \mu(q) = \|qr\| \cos \beta_1$. By the sampling condition, we have $\beta_1 \ge 3/4\bar{\alpha}_s$. Thus $\mu(r) - \mu(q) = \|qr\| \cos 3/4\bar{\alpha}_s$. Let β_2 as above be the angle between the vectors $\vec{pq}$ and $\vec{qr}$ (for an illustration, see Figure 30). Then $\beta_2 \le \bar{\alpha}_s/2$ (since the angle between pq, respectively qr, and the right tangent at s is at most $\bar{\alpha}_s/4$) and $\|pr\| - \|pq\| \ge \|qr\| \cos \beta_2 \ge \|qr\| \cos \bar{\alpha}_s/2$.

Combining bounds, we obtain

$$\begin{aligned} c_\mu(p, r) - c_\mu(p, q) &\ge \|qr\|(\cos(\bar{\alpha}_s/2) - \cos(\bar{\alpha}_s/2 + \bar{\alpha}_s/4)) \\ &\ge \|qr\|(\cos(\bar{\alpha}_s/2) - \cos(\bar{\alpha}_s/2)\cos(\bar{\alpha}_s/4) + \sin(\bar{\alpha}_s/2)\sin(\bar{\alpha}_s/4)) \\ &\ge \|qr\| \sin(\bar{\alpha}_s/4)^2. \end{aligned}$$

$p < s \le q < r$: Assume first that $\{q, r\} \cap T = \emptyset$. If q_s is at least as far from s as p_s, we have $\|pr\| - \|pq\| \ge 0$ and $\mu(q) - \mu(r) \ge \|qr\| \cos(17\pi/48 + \pi/9) = \|qr\| \cos(\pi/2 - 5\pi/144) = \|qr\| \sin(5\pi/144) > 0.1 \cdot \|qr\| \ge \|qr\|(\sin \bar{\alpha}_s/4)^2/3$. The claim follows.

If q_s is closer to s than p_s, let β_1 be the angle between the vectors $\vec{qp}$ and $\vec{b}_s$ and β_2 be the angle between the vectors $\vec{qr}$ and $\vec{b}_s$; for an illustration see the left part of Figure 31. We have $\mu(q) - \mu(r) = \|qr\| \cos \beta_2$ and $\|pq\| - \|pr\| \le \|qr\| \cos(\beta_1 + \beta_2)$. By the sampling condition, we have $\beta_2 \ge \bar{\alpha}_s/4$ and $\beta_1 \ge \angle(\vec{sp}, \vec{b}_s) \ge \bar{\alpha}_s/4$, since $\angle(\vec{qp}, \vec{b}_s) \ge \angle(\vec{sp}, \vec{b}_s)$ (moving along the line $\vec{sq}$ increases the angle). Combining bounds, we obtain:

$$\begin{aligned} c_\mu(p, r) - c_\mu(p, q) &\ge \|qr\|(\cos \beta_2 - \cos(\beta_1 + \beta_2)) \\ &= \|qr\|(\cos \beta_2 - \cos \beta_1 \cos \beta_2 + \sin \beta_1 \sin \beta_2) \\ &\ge \|qr\| \sin(\bar{\alpha}_s/4)^2. \end{aligned}$$

We come to the case that $\{q, r\} \subseteq T$: If $p \in T$, we have $c_\mu(p, r) = \|pr\| - d(p)/3 - d(r)/3 > 0$, since $d(p) \le \|pr\|$ and $d(r) \le \|pr\|$.

So assume $p \notin T$. Since $q \in T$, we have $d(q)/3 \ge c_s - \|sq_s\|$ and hence $c_s \le d(q)/3 + \|sq_s\|$. Thus[3] $0 \ge c_\mu(p, q) = \|pq\| - (c_s - \|sp_s\|) - d(q)/3 \ge \|pq\| - (\|sq_s\| - \|sp_s\|) - 2d(q)/3 \ge \|pq\|/3 - (\|sq_s\| - \|sp_s\|)$ and hence $\|sq_s\| - \|sp_s\| \ge \|pq\|/3$. In particular, q_s lies further away from s than p_s does.

Let β_1 be the angle between the vectors $\vec{qp}$ and $\vec{b}_s$ and β_2 be the angle between the vectors $\vec{qr}$ and $\vec{b}_s$; for an illustration see the right part of Figure 31. Then $\cos\bar{\beta}_1 \ge 1/3$, $\beta_2 \ge \bar{\alpha}_s/4$, and $\|pr\| - \|pq\| \ge \|qr\| \cos(\pi - \beta_1 - \beta_2)$.

Combining bounds we obtain $c_\mu(p, r) - c_\mu(p, q) \ge \|qr\|(\cos(\bar{\beta}_1 - \beta_2) - 1/3)$. If $\cos(\bar{\beta}_1 - \beta_2) - 1/3 \ge 4/27$, we are done since $4/27 \ge (\sin\bar{\alpha}_s/4)^2/3$. So assume $\cos(\bar{\beta}_1 - \beta_2) \le 1/3 + 4/27 = 13/27 \le 1/2$. Then $\bar{\beta}_1 - \beta_2 \ge \pi/3$ and hence $\bar{\beta}_1 \ge \pi/3$ and hence $(1 - \cos\bar{\beta}_1)/\bar{\beta}_1 \ge 1/2$. Since $\cos(\bar{\beta}_1 - x)$ is convex in $[0, \bar{\beta}_1]$ and is therefore above the line through the points $(0, 1)$ and $(\bar{\beta}_1, \cos\bar{\beta}_1)$, we have

$$\cos(\bar{\beta}_1 - \beta_2) - 1/3 \ge \cos\bar{\beta}_1 + (1 - \cos\bar{\beta}_1)(\beta_2/\bar{\beta}_1) - 1/3 \ge \beta_2/2 \ge (\sin\bar{\alpha}_s/4)^2/3.$$

$p < q < s < r$: We have $\|qr\| \le \|qs\| + \|sr\|$. One of the previous cases applies to the triples (p, q, s) and (p, s, r) and therefore $c_\mu(p, r) - c_\mu(p, q) = c_\mu(p, r) - c_\mu(p, s) + c_\mu(p, s) - c_\mu(p, q) \ge (\|qs\| + \|sr\|)(\sin\bar{\alpha}_s/4)^2/3 \ge \|qr\|(\sin\bar{\alpha}_s/4)^2/3$.

The discussion of the case $c_\mu(p, q) \le 0$ is now completed. So let us assume $c_\mu(p, q) > 0$. If $c_\mu(p, r) < 0$, there is a point q' between q and r with $c_\mu(p, q') = 0$. The first case applies to the triple (p, q', r) and hence $c_\mu(p, r) > c_\mu(p, q')$, a contradiction. Thus $c_\mu(p, r) \ge 0$. ∎

Lemma 29 (Prop. (P1) for singular regions) *MST_μ coincides with the polygonal reconstruction for singular subcurves.*

Proof: The Lemma above for both possible orientations of the curve directly implies that Prim's minimum spanning tree algorithm finds the polygonal reconstruction. ∎

4.2.6 Conditions on the Thresholds

In the preceding sections we showed that properties (P1) and (P2) hold, if one chooses the thresholds as in Section 4.2.4. There are other possible choices for the thresholds that make the arguments work. We now collect the conditions on the thresholds. Note that the Subtour-LP has an unique integral solution if there is a choice of the thresholds that make the MST_μ unique and equal to the polygonal reconstruction.

We introduced six thresholds:

---

[3] Recall that we work under the assumption $c_\mu(p, q) \le 0$.

θ_{max_sharp}: The minimum turning angle of a singularity that we call sharp. We have chosen $\theta_{max\_sharp} = 7\pi/24$.

θ_{turn}: The maximal turning angle in a $B(p, d(p))$ ball. We have chosen $\theta_{turn} = \pi/3$.

f_{scale}: The factor by which we scaled the $d(p)$ value for the potential. We have chosen $f_{scale} = 1/3$.

$f_{wriggle}$: The factor by which we scaled the $\bar{\alpha}_s$ as maximal angle between a tangent and a segment in a sharp corner. We have chosen $f_{wriggle} = 1/4$.

$\theta_{wriggle}$: The maximal angle between a tangent and a segment in a sharp corner. We have chosen $\theta_{wriggle} = \pi/9$.

f_{shrink}: The factor by which we shrunk the δ_s ball to define the area where we use the potential function for sharp corners. We have chosen $f_{shrink} = 1/5$.

First of all, we must guarantee that $\mu > 0$, which is equivalent to

$$\boxed{\theta_{turn} < \theta_{max\_sharp}.}$$

We start our investigations by looking at property (P2), i.e. , to guarantee that all edges which are not contained in a region have positive modified cost. Look at any edge pq with negative reduced cost. If p and q are in T, we have to show that there is a point x with $\|px\| \le d(x)$ and $\|qx\| \le d(x)$. For simplicity, we have assumed that x is either p or q. This is reasonable, since the conditions we get are weaker than the conditions we need later. Thus we require $\max(d(p), d(q)) \ge \|pq\|$. Since $\|pq\| \le f_{scale}(d(p) + d(q))$, it is sufficient that

$$\boxed{f_{scale} < 1/2.}$$

If exactly one of p, q is not in T, we require that $2f_{shrink}\delta_s + f_{scale}\|sq\| \le \|pq\|$ (the first summand is an upper bound for the potential of p, the second for the potential of q). Using $\|pq\| \ge 1 - f_{shrink}$ for all points outside the $B(s, \delta_s)$ ball, we conclude that it is sufficient that

$$\boxed{3f_{shrink} + f_{scale} < 1.}$$

Assume next that p and q are not in T. Then $p \in B(s_1, f_{shrink}\delta_{s_1})$ and $q \in B(s_2, f_{shrink}\delta_{s_2})$ for different sharp corners s_1 and s_2. Let without loss of generality$\delta_{s_1} \ge \delta_{s_2}$. It is sufficent that

$$\boxed{f_{shrink} < 1/3.}$$

Let us now turn to property (P1). In regular regions, we require that "potential change takes place more slowly than the change in distance". Thus we need

$$\boxed{\cos(\theta_{turn}) > f_{scale}.}$$

For singular regions, we have to go through all cases of Lemma 28. If $s \leq p < q < r$, we need similar to the case of regular cures

$$\boxed{\cos(\theta_{wriggle}) > f_{scale}.}$$

Analogously, if $p < q < r < s$ and q and r are in T. If $p < q < r < s$ and q and r are not in T, we require that $\cos(1 - f_{wriggle}) < \cos(2 f_{wriggle})$, which is equivalent to

$$\boxed{f_{wriggle} < 1/3.}$$

If $p < s \leq q < r$ and $q, r \notin T$, we additionally require that

$$\boxed{\theta_{max\_sharp}/2 + \theta_{wriggle} < \pi/2.}$$

If $q, r \in T$, we compute that $\cos(\bar{\beta}_1) \geq 1 - 2 f_{scale}$ and require that $\cos(\bar{\beta}_1 - f_{wriggle} \alpha_s) > 1/3$. This is equivalent to

$$\boxed{f_{scale} \leq 1/3.}$$

Note that the last requirement is the only one where the condition can also be satisfied with equality, since we subtract a positive number in the derivation. This ends our discussion.

4.3 Open Curves with Unspecified Endpoints

In the preceding sections we assumed that the first and last sample point (= endpoints of the Traveling Salesman path) are specified as part of the input. In this section, we show that the Subtour-LP can also reconstruct when the endpoints are not specified. Of course, the requirements on the sample will be stronger. The argument used in this section is a variant of the argument used in Section 4.2.

We use the following formulation of the Subtour-ILP. The goal is to select a total of $n - 1$ edges so that at most two of them are incident to any node and so that no subset V' with $V' \neq \emptyset$ is "overfilled".

$$\begin{array}{rlcll}
\min & \sum\limits_{u,v \in V} c_{uv} x_{uv} & & & \\
s.t. & \sum\limits_{v \in V} x_{uv} & \leq & 2 & \text{for all } u \in V \\
 & \sum\limits_{u,v \in V'} x_{uv} & \leq & |V'| - 1 & \text{for } V' \subset V,\ V' \neq \emptyset \\
 & \sum\limits_{u,v \in V'} x_{uv} & = & n - 1 & \\
 & x_{uv} & \in & \{0, 1\} & \text{for all } u, v \in V.
\end{array}$$

Selecting a total of $n - 1$ edges so that at most two of them are incident to any node amounts to selecting a path and a set of cycles covering all nodes. The constraint that no set can be overfilled implies that no cycles can be used and hence any solution must be a Traveling Salesman path. The Subtour-LP is obtained by replacing the integrality constraints $x_{uv} \in \{0, 1\}$ by the linear constraints $0 \leq x_{uv} \leq 1$.

As in the case of open curves with specified endpoints, we have to show that the separation problem can be solved in polynomial time. The separation algorithm works as follows:

Let x^* be the optimal solution. We assign a capacity of x_e^* to edge e for every edge e. Furthermore, we introduce an artificial vertex s and edges us with capacity $2 - \sum_{e\in\delta(u)} x_e^*$ for any node u of the graph. We compute a minimal cut in this graph and take as subset for the subtour elimination constraint the side of the cut that does not contain the artificial node s.

To see the correctness of this separation algorithm, we show that the subtour elimination constraint for S is violated iff the size of the cut is less than 2. Let $S \subset V$ be a subset of the nodes. First notice that the sum of the capacities of edges adjacent to a node is exactly 2 for every node in the graph above. Thus $2\sum_{e\in\gamma(S)} x_e^* + \sum_{e\in\delta(s)} x_e^* = 2|S|$, since for every node $u \in S$, we sum the capacities of the adjacent edges. We conclude that $\sum_{e\in\delta(s)} x_e^* < 2$ iff $\sum_{e\in\gamma(S)} x_e^* > |S| - 1$.

We consider only non-positive potential functions $\mu \le 0$ in this section. Let a and b be fixed vertices and consider any Traveling Salesman path T with endpoints a and b. Its costs under c and c_μ are related by $c_\mu(T) = c(T) - 2\sum_{v\in V}\mu(v) + \mu(a) + \mu(b)$, since the path uses two edges incident to every vertex except for a and b. Observe that $c_\mu(T) - c(T)$ does not depend on T and hence the optimal Traveling Salesman path for endpoints a and b is the same under both cost functions. However, the relative order of Traveling Salesman path with distinct endpoints is changed.

Let MST_μ be a minimum spanning tree with respect to the cost function c_μ and let $C_\mu = c_\mu(\mathrm{MST}_\mu)$ be its cost. Then $C_\mu \le c_\mu(T)$ for any Traveling Salesman path T.

Fact 3 *Let $\mu \le 0$ be any potential function. If MST_μ is a Traveling Salesman path and $\mu(a) = \mu(b) = 0$ for the endpoints of this path, it is an optimal Traveling Salesman path.*

Proof: Let T_0 be an optimal Traveling Salesman path, say with endpoints u and v. Then $C_\mu \le c_\mu(T_0)$, since T_0 is a spanning tree, $c(\mathrm{MST}_\mu) = c_\mu(\mathrm{MST}_\mu) + 2\sum_{v\in V}\mu(v) - \mu(a) - \mu(b) = c_\mu(\mathrm{MST}_\mu) + 2\sum_{v\in V}\mu(v)$, since MST_μ is a path with endpoints a and b and a and b have potential zero, and $c(T_0) = c_\mu(T_0) + 2\sum_{v\in V}\mu(v) - \mu(u) - \mu(v) \ge c_\mu(T_0) + 2\sum_{v\in V}\mu(v)$, since T_0 is a path with endpoints u and v, and since the potentials of u and v are non-positive. Thus

$$c(T_0) \ge c_\mu(T_0) + 2\sum_{v\in V}\mu(v) \ge c_\mu(\mathrm{MST}_\mu) + 2\sum_{v\in V}\mu(v) = c(\mathrm{MST}_\mu).$$

∎

The inequality $C_\mu \le c_\mu(T_0) = c(T_0) - 2\sum_{v\in V}\mu(v) + \mu(u) + \mu(v) \le c(T_0) - 2\sum_{v\in V}\mu(v)$ is valid for every non-positive potential function (the last inequality uses non-positivity) and hence

$$\max_{\mu\le 0}\,(C_\mu + 2\sum_{v\in V}\mu(v)) \le c(T_0)\ .$$

The quantity on the left is called the Held-Karp bound. The following fact is crucial for our proof.

Fact 4 *The Held-Karp bound is equal to the optimal objective value of the Subtour-LP.*

Proof: The proof follows from [CCPS98, page 259]. Relaxing the degree constraints $\sum_{v \in V} x_{uv} \leq 2$ in a Langrangean fashion, we obtain the problem

$$\begin{array}{rrcll} \max\limits_{\mu \leq 0} \min\limits_{x \geq 0} & \sum\limits_{u,v \in V} c(uv) x_{uv} + \sum\limits_{u} \mu(u)(2 - \sum\limits_{v \in V} x_{uv}) & & & \\ s.t. & \sum\limits_{u,v \in V'} x_{uv} & \leq & |V'| - 1 & \text{for } V' \subset V,\, V' \neq \emptyset \\ & \sum\limits_{u,v \in V} x_{uv} & = & n - 1 & \\ & x_{uv} & \leq & 1 & \text{for all } u, v \in V. \end{array}$$

Observe that we are only maximizing over non-positive potential functions μ. This stems from the fact that in contrast to Section 4.2, the degree constraints are now inequalities instead of equalities. The reformulation has the same objective value. The objective function of the LP can be reformulated as $\sum_{u,v \in V}(c(uv) - \mu(u) - \mu(v))x_{uv} + 2\mu(V)$. We conclude that the LP is simply a minimum spanning tree problem for the cost function c_μ. ∎

We note that the optimal choice of μ in the Held-Karp bound is given by the optimal solution of the linear programming dual of the Subtour-LP; μ corresponds to the dual variables for the degree constraints. We next draw a simple consequence from the two facts above.

Lemma 30 *Let $\mu \leq 0$ be any potential function. If MST_μ is the unique minimum spanning tree with respect to c_μ and it is a Traveling Salesman path, and $\mu(a) = \mu(b) = 0$ for its endpoints a and b, then the Subtour-LP has a unique optimal solution and this solution is integral.*

Proof: If MST_μ is a Traveling Salesman path, it is optimal (Fact 1) and hence $c(\mathrm{MST}_\mu) = c(T_0)$. The Held-Karp bound is therefore equal to $c(T_0)$ and the same holds true for the optimal objective value of the Subtour-LP (Fact 2). The incidence vector of MST_μ is a feasible solution of the Subtour-LP of cost $c(\mathrm{MST}_\mu)$ and hence it is *an* optimal solution of the Subtour-LP. We will next argue that it is the unique optimal solution. Assume that there is an optimal solution of the Subtour-LP with $x_e > 0$ for some $e \notin \mathrm{MST}_\mu$. Since MST_μ is unique there is a $\eta > 0$ so that decreasing the cost of e by η will not change the minimum spanning tree and hence the value of the Held-Karp bound will not change. However, the objective value of the Subtour-LP will decrease. This is a contradiction to the equality of the two bounds. ∎

It remains to define the appropriate potential function. We obtain it as a modification of the potential function defined in the preceding section. We use $\bar{\mu}$ to denote it. Let V be a set of sample points and let a and b be the first and the last sample point.

Let $m = \min(\bar{\mu}(a), \bar{\mu}(b))$ and set $c_s = \min(c_s, m)$ for all sharp corners. This changes $\bar{\mu}$; it makes $\bar{\mu}$ smaller for some points near sharp corners. Define new potential functions $\tilde{\mu}$ and μ by

$$\tilde{\mu}(p) = \min(\bar{\mu}(p), m) \quad \text{and} \quad \mu(p) = \tilde{\mu}(p) - m$$

for all $p \in \gamma$, i.e., first all potential values are capped at m and then m is subtracted. Then $\tilde{\mu}(p) \geq \mu(p)$ for all p and $\mu(p) \leq 0$ for all p. Also $\mu(a) = \mu(b) = 0$.

We strengthen the sampling condition and require $c_{\tilde{\mu}}(pq) = \|pq\| - \tilde{\mu}(p) - \tilde{\mu}(q) \leq 0$ for all edges in the reconstruction.

Since $c_s \leq m$ for all sharp corners, we have $\tilde{\mu}(p) = \bar{\mu}(p)$ for all $p \notin T$. Here $\bar{\mu}$ denotes the original potential function with the capped c-values.

Suppose that V satisfies the strengthened sampling condition. Then the minimum spanning tree with respect to $c_{\tilde{\mu}}$ is equal to the polygonal reconstruction. This requires a check of Lemmas 25 to 29; we leave the straightforward but tedious check to the reader. The minimum spanning tree with respect to μ is the same as the minimum spanning tree with respect to $\tilde{\mu}$, since μ and $\tilde{\mu}$ differ only by a constant. We conclude that the minimum spanning tree with respect to μ is equal to the polygonal reconstruction. Moreover it is unique. We finally observe that μ is non-positive and that $\mu(a) = \mu(b) = 0$. Thus MST_μ is the unique optimal solution of the Subtour-LP by Lemma 21.

4.4 Closed Curves

We extend the result to closed curves in two steps.

- In Section 4.4.2 we alter the potential function to μ', so that the two longest edges of the polygonal reconstruction have the same modified cost and so that the minimum spanning trees with respect to the new modified cost are precisely the polygonal reconstruction minus one of the edges of maximal modified cost.
- In Section 4.4.1 we show that the preceding sentence implies our main theorem for closed curves.

Observe in our write-up that the second step is dealt with first.

Readers familiar with the Held-Karp bound for Traveling Salesman tours may wonder why we are not arguing about 1-trees. We tried but could not get the argument to work. A 1-tree is defined as follows. An arbitrary node $v \in V$ is fixed. A 1-tree consists of the two cheapest edges incident to v plus a minimum spanning tree of $V \setminus v$. We were unable to construct a potential function for which the optimal 1-tree coincides with the polygonal reconstruction. We were able to construct a potential function where the two cheapest edges incident to v were indeed the edges to the two neighbors in the polygonal reconstruction and were able to construct a potential function where the minimum spanning tree on $V \setminus v$ coincided with the polygonal reconstruction minus the two edges incident to v. We were unable to satisfy both conditions simultaneously.

4.4.1 The Subtour-LP and the Global Reasoning

Assume that the potential function μ' has been constructed. The Subtour-LP for the traveling salesman problem can be formulated as follows:

$$\begin{array}{lrcll}
\min & \displaystyle\sum_{u,v\in V} c_{uv}x_{uv} & & & \\
s.t. & \displaystyle\sum_{v\in V} x_{uv} & = & 2 & \text{for } u \in V \\
 & \displaystyle\sum_{u\in V',v\in V'} x_{uv} & \leq & |\,V'\,| - 1 & \text{for } V' \subseteq V,\, \emptyset \neq V' \neq V \\
 & \displaystyle\sum_{u,v\in V} x_{uv} & = & |\,V\,| & \\
 & 0 \;\leq\; x_{uv} & \leq & 1 &
\end{array}$$

The last equality is redundant but helpful for our Lagrangian-relaxation. The length of the polygonal reconstruction is an upper bound for the objective value of the Subtour-LP. We relax the set of degree equalities to the objective function and obtain the following problem with the same objective value [CCPS98, pages 258–260]:

$$\begin{array}{lrcll}
\max & 2\displaystyle\sum_{u\in V} \mu(u) + \min \sum_{u,v\in V} c_\mu(uv)x_{uv} & & & \\
s.t. & \displaystyle\sum_{u\in V',v\in V'} x_{uv} & \leq & |\,V'\,| - 1 & \text{for } V' \subseteq V,\, \emptyset \neq V' \neq V \\
 & \displaystyle\sum_{u,v\in V} x_{uv} & = & |\,V\,| & \\
 & 0 \;\leq\; x_{uv} & \leq & 1 &
\end{array}$$

In this formulation the maximization is over all choices of μ. For fixed μ the inner minimization is over the choices for the x_{uv}. We will show that for $\mu = \mu'$ the polygonal reconstruction is the unique optimal solution for the minimization problem and hence the objective value of the maximization problem is at least the length of the polygonal reconstruction. It cannot be larger and hence the objective value of the maximization problem is equal to the length of the reconstruction. This proves that the polygonal reconstruction is an optimal solution of the Subtour-LP. We still need to argue uniqueness. Assume that there is another optimal solution for the Subtour-LP. Since the solution satisfies the degree constraints, it will give the same value to the inner minimization problem as the polygonal reconstruction and hence the polygonal reconstruction is not the unique optimal solution of the inner minimization problem.

It remains to prove that for $\mu = \mu'$, the polygonal reconstruction is the unique optimal solution for the inner minimization problem. Orient γ arbitrarily, let e_1 and e_2 be edges in the polygonal reconstruction which have maximal modified cost, and let u and v be the starting nodes of e_1 and e_2, respectively, and let R_1 and R_2 be the sample points from u to v respectively from v to u with respect to the order of the points along the curve (see Figure 32). Then $R_1 \cup R_2 = V$ and $|R_1 \cap R_2| = 2$. Let E_i, $i = 1, 2$, be the set of edges having both endpoints in R_i and let C be the remaining set of edges. Then $e_i \in E_i$ and any edge $e \in C$ has a modified cost larger than e_1 (and hence e_2). Otherwise, there would

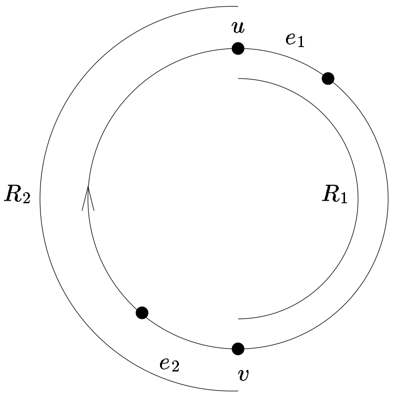

Figure 32: The notation used in the reformulations of the Subtour-LP. We have $e_i \in E_i$.

be a minimum spanning tree that is not contained in the polygonal reconstruction. In an optimal solution of the inner LP, the total weight of the edges in E_i is $|\,R_i\,| - 1 - o_i$ for some $o_i \geq 0$, $i = 1, 2$; the total weight of the edges in C is $o_1 + o_2$. Thus the inner LP is relaxed by:

$$\begin{array}{rrcl}
\min\limits_{o_1,o_2 \geq 0} \min & \sum\limits_{u,v \in V} c_{\mu'}(uv) x_{uv} & & \\
s.t. & \sum\limits_{u \in R_1', v \in R_1'} x_{uv} & \leq & |\,R_1'\,| - 1 \qquad \text{for all } R_1' \subseteq R_1 \text{, with } R_1' \neq \emptyset \\
& \sum\limits_{u,v \in R_1} x_{uv} & = & |\,R_1'\,| - 1 - o_1 \\
& \sum\limits_{u \in R_2', v \in R_2'} x_{uv} & \leq & |\,R_2'\,| - 1 \qquad \text{for all } R_2' \subseteq R_2 \text{, with } R_2' \neq \emptyset \\
& \sum\limits_{u,v \in R_2} x_{uv} & = & |\,R_2\,| - 1 - o_2 \\
& \sum\limits_{uv \in C} x_{uv} & = & o_1 + o_2 \\
& 0 \leq x_{uv} & \leq & 1
\end{array}$$

Observe that we have dropped some of the subtour elimination constraints.

Consider the inner minimization problem for fixed values of o_1 and o_2. The first two lines describe a partial minimum spanning tree for R_1 while the next two lines describe a partial minimum spanning tree for R_2. More precisely, the system is minimized if one chooses the $\lfloor |R_1| - 1 - o_1 \rfloor$ shortest edges of the minimum spanning tree and fills the fractional part with the next edge (as seen in Section 2.2.5). The same is true for R_2. The LP takes its minimum for $o_1 = o_2 = 0$, since any edge in C has higher cost than any edge in the minimum spanning tree. For $o_1 = o_2 = 0$, the system describes the minimum spanning trees for R_1 and R_2. Thus the polygonal reconstruction is the unique optimal solution of the inner minimization problem and hence of the Subtour-LP.

4.4.2 The Modified Potential Function and the Local Reasoning

We show how to alter the potential so that the two longest edges of the polygonal reconstruction have the same modified cost and so that all minimum spanning trees for V remain part of the polygonal reconstruction. Note that the new potential is defined according to a given sample set, whereas the original potential only depends on the curve.

Let e_{max} be the edge of the polygonal reconstruction with highest modified cost. We claim that one of the following cases arises:

(1) There is a sharp corner so that both endpoints of e_{max} are outside the ball $B(s, \delta_s/5)$.

(2) There is a point $v \in \gamma$, so that $\|vu\|/2 \geq \mu(u)$ for both endpoints u of e_{max} and so that v does not lie in the $B^0(s, \delta_s)$ ball of any sharp corner s.

The first case certainly arises when there are at least three sharp corners. Assume that the first case does not hold. We make a further case distinction: Either both endpoints of e_{max} lie in T (this will certainly be the case when there is no sharp corner) or one endpoint of e_{max} lies in $B(s, \delta_s/5) \setminus T$ for a sharp corner s. In the former case, the curve must leave the union of the $B(u, d(u))$ balls of the two endpoints u of e_{max},because the curve can turn by at most $2\pi/3$ within the union of these balls. As a result any curve point v outside the two balls (since $\|uv\| \geq d(u) = 3\mu(u)$ for any of the endpoints) and outside $\cup_{s\in S} B^0(s, \delta_s)$ will work; v exists since the regions $(B^0(s, \delta_s))_{s\in S}$ are pairwise disjoint and since γ turns by less than π in any such region. So assume that some endpoint of e_{max} lies in $B(s, \delta_s/5) \setminus T$ of some sharp corner s. Consider the leg of s that does not contain the other endpoint of e_{max}. It contains no endpoint of e_{max} that lies in T. Let v be the point on this leg of s with distance δ_s from s and let u be an endpoint of e_{max}. If $u \in B(s, \delta_s/5)$ then $\|uv\| \geq 4\delta_s/5$ and $\mu(u) \leq \max(c_s, d(u)/3) \leq 2\delta_s/5$; if $u \notin B(s, \delta_s/5)$, then $u \in T$ and hence u does not lie on the same leg as v does. Consequently $d(u) \leq \|uv\|$ by Lemma 22 and we are done. In either case we have shown that one of the items above holds. We also need the following Lemma.

Lemma 31 *Let s be a sharp corner and let x and y be adjacent sample points so that $s \in \gamma[x, y]$. Then $x, y \in B(s, 3\delta_s/5)$ and either x or y lies in $B(s, \delta_s/5) \setminus T$.*

Proof: Since the regions $B^0(p, d(p)) \cap T$ do not contain any sharp corner and since the regions $(B^0(t, \delta_t))_{t\in S}$ are pairwise disjoint, we have $x, y \in B^0(s, \delta_s)$ by Lemmas 24 and 25. Assume $x \notin B(s, 3\delta_s/5)$. Then

$$c_\mu(x, y) \geq c_\mu(s, x) \geq \|sx\| - 2\delta_s/5 - \|sx\|/3 \geq 2\|sx\|/3 - 2\delta_s/5 > 2\delta_s/5 - 2\delta_s/5 = 0,$$

where the first inequality follows from the third claim in the proof of Lemma 29. Assume next that $x, y \in T$. Lemma 22 implies $d(x), d(y) \leq \|xy\|$ and hence $c_\mu(x, y) \geq \|xy\|/3 > 0$.
∎

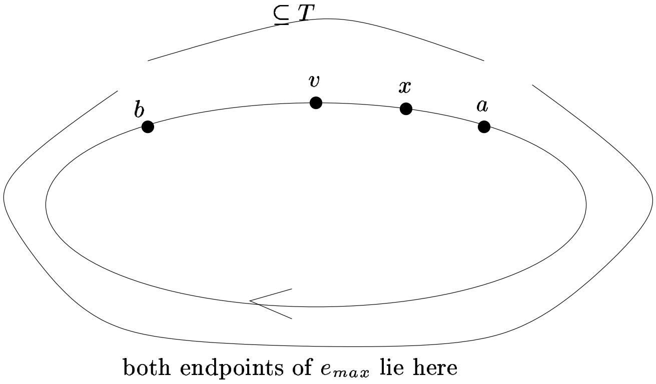

Figure 33: The construction of the potential function μ'.

We now turn to the definition of the modified potential function and the proof that all minimum spanning trees for V with respect to the modified potential function are subsets of the polygonal reconstruction.

Assume first that there is a sharp corner s, so that both endpoints of e_{max} lie outside the $B(s, \delta_s/5)$ ball. We decrease c_s continuously to zero. For $c_s = 0$, we have $\mu(x) = d(p)/3$ for all $x \in B(s, \delta_s)$. Hence the edge xy in the reconstruction connecting the two legs of s (it is existent by Lemma 31) has positive modified cost: $c_{\mu'}(x, y) = \|xy\| - \mu'(x) - \mu'(y) \geq \|xy\| - \|xy\|/3 - \|xy\|/3 > 0$, since $d(x), d(y) \leq \|xy\|$ and hence $c_\mu(x, y) \geq \|xy\|/3 > 0$ by Lemma 22. The modified cost of the edge e_{max} is not affected by the change of c_s. Thus there must be a value of c_s for which the two largest modified costs in the reconstruction are the same. All edges in the reconstruction still have non-positive modified cost (because e_{max} has non-positive modified cost) and the minimum spanning tree with respect to $c_{\mu'}$ remains the polygonal reconstruction, as noted in Section 4.2.4 that only $0 < c_s \leq 2\delta_s/5$ is used in our proofs. This completes the discussion of the first case.

In the next case, there is a point $v \in \gamma$, so that $\mu(u) \leq \|vu\|/2$ for both endpoints u of e_{max}. So, v lies outside the $B^0(s, \delta_s)$ balls of all sharp corners. We split the curve at v and orient the resulting open curve arbitrarily. For any $l \in \mathbb{R}$, $0 \leq l \leq \mu(v)$, we define a as the first point with $\mu(a) = \|va\|/2 + l$ and b as the last point with $\mu(b) = \|vb\|/2 + l$ and define a potential μ' by

$$\mu'(p) = \begin{cases} \mu(p) & \text{if } a \leq p \leq b \\ \|vp\|/2 + l & \text{otherwise} \end{cases}$$

Observe that a, b, and μ' depend on l. For simplicity of notation we do not make this dependence explicit in the notation. Figure 33 illustrates the definition of μ'.

We need to argue that a and b are existent for all choices of l, that the cost of e_{max} does not depend on l, that there is a choice of l for which the two largest modified costs are the same, and that every minimum spanning tree with respect to $c_{\mu'}$ uses only edges of the polygonal reconstruction.

We first show the existence of a and b for all choices of l. For $l = \mu(v)$ we have $a = b = v$. For $0 \leq l < \mu(v)$, we have $\mu(v) > l + \|vv\|/2$ and $\mu(u) \leq \|uv\|/2 \leq l + \|uv\|/2$ for either endpoint of e_{max}. From the continuity of μ we conclude that there is a point a between v and u, for which $\mu(a) = \|va\|/2 + l$, and a point b between u and v, for which $\mu(b) = \|vb\|/2 + l$. We have shown that a and b exist and that e_{max} does not depend on l. Furthermore, we know that $\mu'(p) \leq \mu(p)$ for all p by the definition of a and b.

We next show that $a \leq r$, where r is the first point on γ with $\|sr\| = 3\delta_s/5$ for some sharp corner s, hence $a \in T$. We have $\|rv\| \geq 2\delta_s/5$ and $\mu(r) = d(r)/3 \leq \|sr\|/3 = \delta_s/5 \leq \|rv\|/2 \leq \|rv\|/2 + l$. Continuity of μ implies that a lies between v and r. Similarly, b lies in T after the last point r on γ with $\|sr\| = 3\delta_s/5$ for some sharp corner.

We next argue that there is a choice of l for which the two largest modified costs in the reconstruction are the same. From $a \leq u \leq b$ for any endpoint u of e_{max} we conclude that the modified cost of e_{max} does not depend on l. For $l = \mu(v)$, we have $a = b = v$ and hence $\mu = \mu'$. Thus e_{max} has the maximal modified cost $c_{\mu'}$ among all edges in the reconstruction. For $l = 0$, we consider the reconstruction edge between the last and the first sample points x and y, respectively, and show that it has positive modified cost $c_{\mu'}(xy)$. There must be a subcurve γ' containing $\gamma[x, y]$ and hence v. Since $v \notin B(s, \delta_s)$ for any sharp corner, γ' is a regular subcurve and hence $\gamma[x, y] \subseteq T$. We first show that $a < x$ and $y < b$ is impossible. So assume otherwise. Then $\mu(x) \leq \mu(a) + \|ax\|/3 = \|va\|/2 + \|ax\|/3$ and $\mu(y) \leq \mu(b) + \|by\|/3 = \|vb\|/2 + \|by\|/3$ and $\|xy\| > \|xa\|/2 + \|ab\| + \|by\|/2$ since γ' turns by less than $\pi/3$. Thus $c_\mu(x, y) > 0$, a contradiction to our sampling condition. Thus either $v \leq x \leq a$ or $b \leq y \leq v$ or both. We may assume without loss of generality that $v \leq x \leq a$. If $y < b$, we have $\mu(y) \leq \mu(b) + \|by\|/3 = \|vb\|/2 + \|by\|/3$, $\|xy\| > \|xv\|/2 + \|vb\| + \|by\|/2$, and $\mu'(x) = \|vx\|/2$, hence $c_{\mu'}(x, y) > 0$. If $b \leq y$, we have $\|xy\| > \|xv\|/2 + \|vy\|/2$, $\mu'(y) = \|vy\|/2$ and $\mu'(x) = \|vx\|/2$ and hence $c_{\mu'}(x, y) > 0$. In either case we have shown that for $l = 0$ there is an edge in the reconstruction with positive modified cost. Continuity implies that there is a value of l for which the two largest modified costs in the reconstruction are the same; this completes the definition of the modified potential function in the second case.

In order to verify that all minimum spanning trees are subsets of the polygonal reconstruction, it is sufficient to show that Lemma 25 holds for the new potential function and that $c_{\mu'}(p, r) > c_{\mu'}(p, q)$ for any three points with $p < q < r$ and $\{p, q, r\} \subseteq \gamma'$ for some $\gamma' \in \Gamma$.

From $\mu'(x) \leq \mu(x)$ for all x, we conclude $c_{\mu'}(pq) \leq c_\mu(pq)$ for all edges pq. Thus Lemma 25 still holds.

If $\{p, q, r\} \subseteq B(u, d(u)) \cap T$ for some $r \in \gamma$, we have $\angle(\vec{pq}, \vec{qr}) < \pi/3$ and hence $\|pr\| - \|pq\| > \|qr\|/2$ and $\mu'(r) \leq \mu'(q) + \|qr\|/2$. Thus $c_{\mu'}(p, r) > c_{\mu'}(p, q)$.

Assume next that $\{p, q, r\} \subseteq B(s, \delta_s)$ for some sharp corner s. If $c_{\mu'}(p, r) \geq 0$ we are done. Otherwise $c_\mu(p, r) < 0$, since the modified distance with respect to μ is at most the modified distance with respect to μ'. We make the same case distinction as in the proof of Lemma 28. In this proof, we bounded $\|pr\| - \|pq\|$ from below and $\mu(r) - \mu(q)$ from above. Now we need to bound $\mu'(r) - \mu'(q)$. Since $\mu'(x) = \mu(x)$ for $x \notin T$, we only need to reconsider the case that r and q are in T. We have $\mu'(r) \leq \mu'(q) + \|qr\|/2$; hence the arguments used in the cases $p < q < r \leq s$, $s \leq p < q < r$, and $p < q < s < r$ stay valid. We only need to reconsider the case $p < s \leq q < r$.

If $\mu(p) \neq \mu'(p)$, we have $\mu(q) = \mu'(q)$ and $\mu(r) = \mu'(r)$ and are done. So assume $\mu(p) = \mu'(p)$. If $\mu(q) = \mu'(q)$ we have $c_{\mu'}(p,r) - c_{\mu'}(p,q) \geq c_\mu(p,r) - c_\mu(p,q) \geq 0$, since $\mu'(r) \leq \mu(r)$. Otherwise $\mu'(r) - \mu'(q) \leq 0 \leq \mu(r) - \mu(q)$, since r is closer to v than q and both are in $T \cap B(s, \delta_s)$. Thus $c_{\mu'}(p,r) - c_{\mu'}(p,q) \geq c_\mu(p,r) - c_\mu(p,q) \geq 0$.

4.5 Solving the Subtour-LP

The Subtour-LP has an exponential number of constraints. The ellipsoid method [Sch86] allows to solve LPs with an exponential number of constraints in polynomial time in the number of variables if the following conditions are satisfied:

- the coefficients of the variables in the constraints are polynomially bounded. This is the case for the Subtour-LP.

- the separation problem can be solved in polynomial time; i.e., given a vector x^*_{uv} of polynomially bounded values for the variables, one can decide in polynomial time whether the vector satisfies all constraints and, if not, exhibit a violated constraint. This is the case for the Subtour-LP. It is trivial to check the degree constraints and the constraint that all values lie between zero and one. In order to check the subtour elimination constraints (we discuss the case of tours, having already discussed the cas of path in Sections 4.2 and 4.3), we consider the complete network on V and assign capacity x^*_{uv} to edge uv. Consider any subset V' of V with $\emptyset \neq V' \neq V$ and observe that $2|V'| = \sum_{u \in V'} \sum_{v \in V} x^*_{uv} = 2 \sum_{u \in V', v \in V'} x^*_{uv} + \sum_{u \in V'} \sum_{v \notin V'} x^*_{uv}$. We conclude that the subtour elimination constraint for V' is satisfied iff the capacity of the cut $(V', V \setminus V')$ is at least two. Some subtour elimination constraint is satisfied if the minimum cut is less than two. A minimum cut can be computed in polynomial time.

- The coefficients in the objective function are polynomially bounded. This is not the case since the bit-representation of an Euclidean length is in principal infinite. Furthermore there are curves where the bit-representation of any point is infinite.

We show that it suffices to know the sample points and their Euclidean distances only approximately. More precisely, we show the following: Let S be a set of points. Let m be the minimal distance between any two points in S and for a point $p \in S$ let p^* be a closest point of γ. If

1. $\|pp^*\| \leq \rho m/10$ for all $p \in S$, where ρ is a constant depending on γ, and
2. the set $S^* = \{p^* \mid p \in S\}$ satisfies a strengthened sampling condition

then the Subtour-LP has a unique optimal integral solution even when the distances between sample points are only known up to an error of $m\rho/10$. Moreover, the Subtour-LP can be solved in polynomial time and the optimal solution is a tour connecting the points in S in the order in which the points S^* lie on γ. We also show how to estimate ρ from the sample set and without knowledge of γ.

Now we make precise what we mean by a approximate sample set of a curve γ. Recall that for an angle α we have defined $\bar{\alpha}$ by $\bar{\alpha} = \pi - \alpha$. Let $\alpha_0 = \max_{s \in S} \alpha_s$. We define $\bar{\alpha}_0 = 17\pi/24$, if there is no sharp corner.

Definition 3 *Let γ be a benign semi-regular curve. We call a set S of points $p \in \mathbb{Q}^2$ an* approximate sample set *of a curve γ, if for all $p \in S$*

$$\|p\gamma\| \le (\sin \bar{\alpha}_0/4)^2 m/30.$$

For the following let $\rho = (\sin \bar{\alpha}_0/4)^2 m/3$. Then $m > 7\rho$ (since $\bar{\alpha}_0 < 17\pi/24$) and $\|pp^*\| \le \rho/10$ for the points in an approximate sample. We need the following sampling condition.

Sampling Condition for Approximate Sample Sets:

(a) For any two adjacent (on γ) samples u^* and v^*: $\|u^*v^*\| \le 9/10(\mu(u^*) + \mu(v^*))$.

(b) For any two adjacent samples u^* and v^*: $\gamma[u^*, v^*]$ turns by less than π. For two adjacent points p, q on the curve, $\gamma[p, q]$ denotes the subcurve of γ with endpoints p and q not containing another sample point.

Let γ be a benign semi-regular curve and S an approximate sample set satisfying the sampling condition. For any two sample points p and q let $\|pq\|_{\approx}$ be a rational number, so that $|\|pq\| - \|pq\|_{\approx}| \le \rho/10$. Then $|\|p^*q^*\| - \|pq\|_{\approx}| \le 3\rho/10$. Note that for all $p, q \in S$ there exists a choice of $\|pq\|_{\approx}$ which has a bit representation of polynomial size in m and $\bar{\alpha}_0$. We consider the approximate Subtour-LP of the approximate sample set S.

Theorem 16 *Let γ be an open (closed) benign semi-regular curve and S be an approximate sample set of γ satisfying the sampling condition above.*

- *The approximate Subtour-LP for S has a unique optimal integral solution.*
- *The approximate Subtour-LP can be solved in polynomial time.*

Proof: We define the potential function of an approximate sample point p as $\mu(p) := \mu(p^*)$. We call the new modified cost function $c_\mu^{\approx}$. We show:

1. if p and q are adjacent sample points, then $c_\mu^{\approx}(p, q) < -3\rho/10$ and hence there is a minimum spanning tree in which each edge has length less than $-3\rho/10$.
2. if p and q are sample points which are not contained in some $\gamma' \in \Gamma$, then $c_\mu^{\approx}(p, q) \ge -3\rho/10$ and hence none of these edges belong to a minimum spanning tree.
3. if $p < q < r$ are three sample points contained in some $\gamma' \in \Gamma$ and $c_\mu^{\approx}(p, r) < -3\rho/10$, then $c_\mu^{\approx}(p, r) > c_\mu^{\approx}(pq)$, and hence the minimum spanning tree reconstructs locally.

We turn to the first item. Let p and q be adjacent sample points. We have $6\rho < m - \rho \leq \|pq\| - \rho \leq \|p^*q^*\| + 2\rho/10 - \rho \leq \|p^*q^*\| \leq \mu(p^*) + \mu(q^*)$. Thus

$$\begin{aligned} c_\mu^{\approx}(p,q) &= \|pq\|_{\approx} - \mu(p) - \mu(q) \\ &\leq \|p^*q^*\| + 3\rho/10 - 1/10(\mu(p^*) + \mu(q^*)) - 9/10(\mu(p^*) + \mu(q^*)) \\ &< \|p^*q^*\| - 3\rho/10 - 9/10(\mu(p^*) + \mu(q^*)) \\ &\leq -3\rho/10. \end{aligned}$$

For the second item, consider two points p and q that are not contained in any common subcurve $\gamma' \in \Gamma$. We have

$$\begin{aligned} c_\mu^{\approx}(p,q) &\geq \|p^*q^*\| - 3\rho/10 - \mu(p^*) - \mu(q^*) \\ &= c_\mu(p^*, q^*) - 3\rho/10 \\ &\geq -3\rho/10. \end{aligned}$$

We come to the third item. Consider three sample points $p < q < r$ that are contained in some $\gamma' \in \Gamma$ and for which $c_\mu^{\approx}(pr) < -3\rho/10$. Then $c_\mu(p^*, r^*) < 0$ and hence (using Lemmas 26 and 28 and the fact that $(\sin\bar{\alpha}_0/4)^2/3 \leq 1/6$):

$$\begin{aligned} c_\mu^{\approx}(p,r) - c_\mu^{\approx}(p,q) &\geq c_\mu(p^*, r^*) - c_\mu(p^*, q^*) - 6\rho/10 \\ &\geq \|qr\|(\sin\bar{\alpha}_0/4)^2/3 - 6\rho/10 \\ &\geq \rho - 6\rho/10 > 0. \end{aligned}$$

▮

What have we achieved at this point? We have shown that the Subtour-LP reconstructs provided that our sample set S satisfies the sampling condition for approximate sample sets *and* that we are given approximate distances of polynomial size that differ by at most $\rho/10$ from the true distances. We could compute approximate distances with the required property, if we were given ρ or alternatively α_0 as an additional input. We now show how to compute a lower bound on α_0, which leads to a polynomial precision in the input size, without any additional knowledge of the curve.

Lemma 32 *Let m and M be the minimal and maximal distance between two sample points, respectively. Then $sin(\bar{\alpha}_0/4) \geq m/(15M)$.*

Proof: If γ has no sharp corners, $\alpha_0 = 17\pi/24$ and there is nothing to show. So assume otherwise and let s be any sharp corner. We prove that there is a sample point p in $B(s, \delta_s) \cap T$ on each leg of the sharp corner and we use this fact to bound $\bar{\alpha}_s$ from below.

We look arbitrarily at one of the two orders obtained by splitting the curve at s and prove that there is a sample point behind s in $B(s, \delta_s) \cap T$. Assume otherwise. Let x be the first sample point behind s outside $B(s, \delta_s)$ and let y be the sample point preceding x. Since every edge of the polygonal reconstruction must lie in at least one subcurve $\gamma' \in \Gamma$, y must lie behind s. By assumption, y does not lie in T. Assume first that

$x \in T$. Then $\|xy\| \geq 4\delta_s/5$ and $\mu(y) \leq 2\delta_s/5$. Thus $c_\mu(x,y) = \|xy\| - \mu(x) - \mu(y) \geq \|xy\| - 2\delta_s/5 - \|xy\|/3 \geq 2\|xy\|/3 - 2\delta_s/5 \leq 8\delta_s/15 - 2\delta_s/5 > 0$, a contradiction. Assume now $x \notin T$. Let s' be the corner so that $x \in B(s', \delta_{s'})$ and assume without loss of generality$\delta_s \geq \delta_{s'}$. Then $\mu(x) \leq 2\delta_s/5$, $\mu(y) \leq 2\delta_s/5$ and $\|xy\| \geq 2\delta_s - \delta_s/5 - \delta_{s'}/5 \geq \delta_s$. Thus $c_\mu(x,y) > 0$, a contradiction.

Let p be the first sample point after s in $B(s,\delta_s) \cap T$ and let q be the adjacent sample point after p. Then $q \in T$; if $q \notin T$ then $q \in B(s', \delta_{s'})$ for some sharp corner s' and hence s' would have no sample point in $B(s', \delta_{s'}) \cap T$. The distance between p and q is at least m. Also $\|pq\| \leq d(p)/3 + d(q)/3 \leq d(p)/3 + (d(p) + \|pq\|)/3$ and so $d(p) \geq \|pq\| \geq m$. Consider the intersections of the two legs of s with the boundary of the δ_s-ball centered at s. The intersections have distances of at least m (since d-values grow along each leg) and s sees the intersections under an angle of at least $\bar{\alpha}_0/2$. Thus $\sin\bar{\alpha}_0/4 \geq m/(2\delta_s)$. Since there is at least one sample outside the ball $B(s,\delta_s)$ and at least one sample inside the ball $B(s,\delta_s/5)$, we have $M \geq 4\delta_s/5$. Thus $\sin\bar{\alpha}_0/4 \geq 4m/(10M)$. ∎

4.6 Collections of Closed Curves

In the preceding sections we showed that the Subtour-LP formulation of the Traveling Salesman problem is able to reconstruct *single* closed and open curves. In this section we extend the algorithm so that it can handle collections of closed curves. We do not know how to handle collections of open and closed curves. Please note that the algorithms [DMR00, FR01] can handle open and closed curves.

The algorithm works in rounds. The first round constructs an initial partition of the sample points and subsequent rounds merge blocks of the partition. The construction of the initial partition and the merging is done conservatively, i.e., all points in the same block provably belong to the same curve. In the first round, every point is joined to points close to it; Section 4.6.1 gives the details. In later rounds (see Section 4.6.2) we solve the Subtour-LP for each block and then analyze the solution. If the Subtour-LP fails on a block or if curves constructed for different blocks interfere, some blocks are merged.

Throughout this section we assume our set of sample points to satisfy a strengthened sampling condition. The strengthened sampling condition leads to denser sampling near sharp corners. We change μ to μ' by decreasing the δ- and c-values of sharp corners. We set $\delta'_s \leq \delta_s$, so that for any two points p and q in $B(s, \delta'_s/7)$, the angle between the segment pq and the corresponding tangent in s is at most $\pi/40$. Furthermore, we decrease the value c_s to c'_s so that $c'_s - \|sp_s\| \leq d(p)/3$ for every $p \notin B(s, \delta'_s/60)$. This enlarges the region in which the potential is defined by $d(p)/3$ from T to T'. Recall that the choice of c_s only guaranteed that the points outside $B(s, \delta_s/5)$ belonged to T, whereas T' contains all points outside the balls $B(s, \delta'_s/60)$.

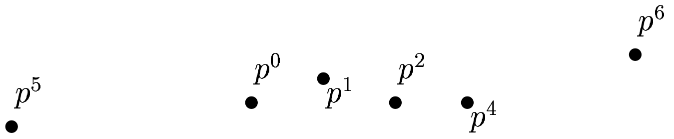

Figure 34: The points p^1 up to p^5 are joined with p^0, but p^6 is not joined.

4.6.1 The Initial Partition

We define a graph on our set of sample points. The connected components of this graph form the initial partition. For a sample point $p = p^0$, let p^1, p^2, … be the other sample points in order of increasing distance (ties are broken arbitrarily). We always join p^0 with p^1 and p^2. We join p^0 and p^i, $i \geq 3$, if $\angle \overrightarrow{p^{k-1}p^{k-2}}, \overrightarrow{p^{k-1}p^k} \geq 2\pi/3$ for all k with $2 \leq k \leq i-1$. Observe that the decision of whether p^i is joined to p^0 depends only on the points p^0 up to p^{i-1}, but not on the point p^i itself. This is essential for making connections between the points on different legs of a sharp corner but also hinders the extension to open curves. Figure 34 illustrates the definition.

Lemma 33 *If p and its two adjacent sample points are in T', then p is only joined with points in $B(p, d(p))$ and is joined with both adjacent sample points.*

Proof: Since p is in T', $\gamma \cap B(p, d(p))$ consists of a single component and turns by less than $\pi/3$. Also the two sample points q, r adjacent to p on γ lie in $B(p, d(p))$. We show this for q. We have $d(q) \leq d(p) + \|pq\|$, $\mu(q) = d(q)/3$ since $q \in T'$, and, according to our sample condition, $\|pq\| \leq d(p)/3 + d(q)/3$. Thus $\|pq\| \leq d(p)/3 + (d(p) + \|pq\|)/3$ and hence $\|pq\| \leq d(p)$.

Assume without loss of generalitythat q is considered before $r = p^i$. Orient $\gamma \cap B(p, d(p))$ so that $r < p < q$. Then $q = p^1$ and $p^1 < p^2 < \ldots < p^{i-1}$. Since $\gamma \cap B(p, d(p))$ turns by less than $\pi/3$, we have $\angle \overrightarrow{p^{k-1}p^{k-2}}, \overrightarrow{p^{k-1}p^k} \geq 2\pi/3$ for all k with $2 \leq k \leq i-1$ and $\angle \overrightarrow{p^{i-1}p^{i-2}}, \overrightarrow{p^{i-1}p^i} \leq \pi/3$. Thus p is joined with p^1 up to p^i, but not with p^{i+1}. ▮

We turn to the non-smooth parts of the curve. We first show that our sampling condition implies that each leg of a sharp corner must contain several sample points.

Lemma 34 *a) Let p be a sample point in $B(s, \delta'_s) \setminus B(s, 2\delta'_s/60)$. Then both adjacent sample points lie on the same leg as p, the one closer to s has distance at most $\|ps\|/2$ from p and the one further from s has distance at most $\|ps\|$ from p.*

b) For every leg ℓ of a sharp corner s we have at least one sample point in $B(s, 2 \cdot 2^j\delta'_s/60) \setminus B(s, 2^j\delta'_s/60)$ for $j = 1, 2, 3, 4$.

Proof: a) We have $\mu'(p) \leq \|sp\|/3$. The point q on the same leg as p with distance exactly $\|sp\|/2$ to s lies in T' and hence $\mu'(q) \leq \|sp\|/6$. Thus $c_{\mu'}(p, q) \geq \|sp\|/2 - \|sp\|/6 - \|sp\|/3 = 0$. Lemma 28 implies that $c_{\mu'}(p, x) \geq 0$ for any point x between s and q and for any point

x on the other leg. Thus there must be a sample point between p and q. We conclude that both adjacent sample points lie on the same leg as p and that the one closer to s satisfies the stated distance constraint. For the one further from p, we consider the point q on the same leg as p, further away from s than p, and having distance $\|sp\|$ from p. Then $\mu'(q) \leq 2\|sp\|/3$ and hence $c_{\mu'}(p,q) \geq \|sp\| - 2\|sp\|/3 - \|sp\|/3 = 0$. We now argue as above.

We turn to part b). Part a) implies that if one of the annuli contains a point, the adjacent annuli do also and hence all annuli do. We conclude that either all annuli contain a sample point or none does. Assume the latter and let p be the first sample point on ℓ outside $B(s,\delta'_s)$. The sample point q preceding p must lie in $B(s, 2\delta'_s/60)$. Thus $\|pq\| \geq \|ps\| - 2\delta'_s/60$, $\mu'(p) \leq \|ps\|/3$, $\|ps\| \geq \delta'_s$, and $\mu'(q) \leq \max(2\delta'_s/60, c_s) \leq 2\delta'_s/5$, a contradiction. ▮

Lemma 35 *A sample point p is only joined with points of the same curve.*

Proof: Consider a sample point p. If p and both adjacent sample points of p are in T', the claim follows from Lemma 33.

Otherwise $p \in B(s, 2\delta'_s/60)$ according to Lemma 34, part a). Let $p = p^0$, p^1, …, p^j be the sample points in $B(s,\delta'_s)$ ordered accordingly to their distance from p. Since both legs of the sharp corner contain sample points in the annuli $B(s, 2 \cdot 2^j\delta'_s/60) \setminus B(s, 2^j\delta'_s/60)$ for $j = 2,3,4$, we must have a subsequence of length at least three, so that the first and the last element of the subsequence, call them q and r, respectively, lie on the same leg as p and further away from s than p from S, and the points in between (there is at least one) lie on the other leg. Figure 35 illustrates the situation. We show that no point after r is joined to p. Assume otherwise. Let q' be the point added directly after q and let r' be the point added directly before r.

We want to bound the angle between the segments $q'q$ and $r'r$. If q' is equal to r', this angle is at least $2\pi/3$. Otherwise, the angle between $q'q$ and the segment between q' and the sample q'' added after q' is at least $2\pi/3$ and the angle between $r'r$ and the segment between r' and the sample r'' added before r' is at least $2\pi/3$ (since r is not the last point joined to p). Also q', q'', r'', and r' lie all on the same leg and hence the angle between $q'q$ and $r'r$ is at least $\pi/3 - 4\pi/40$. Here we use the strengthened sampling condition.

Let x be the intersection between the lines supporting qq' and rr'. We have $d(q) < d(r) \leq d(q) + \|qr\|$, $\|rr'\| \geq d(r)$ and $\|qq'\| \geq d(q)$ by Lemma 22, $\|qr\| \leq (d(q)+d(r))/3$ by our sampling condition, and hence $\|rx\| \geq \|rr'\| \geq d(r) \geq (d(r)+d(q))/2 \geq 3\|qr\|/2$ and $\|qx\| \geq \|qq'\| \geq d(q) \geq (d(q)+d(r)-\|pq\|)/2 \geq \|qr\|$. The application of the "theorem of cosines[4]" with $D = \|qr\|$ yields $\cos(\angle(q'q, r'r)) \geq [(3D/2)^2 + D^2 - D^2]/(2 \cdot D \cdot 3D/2) = 3/4$ and hence $\angle(q'q, r'r) < \pi/3 - \pi/10$, a contradiction. ▮

Lemma 36 *For any curve γ_j all sample points in $T' \cap \gamma_j$ belong to the same component.*

---

[4] $\|qr\|^2 = \|qx\|^2 + \|rx\|^2 - 2\|qx\| \cdot \|rx\| \cdot \cos(\angle(q'q, r'r))$

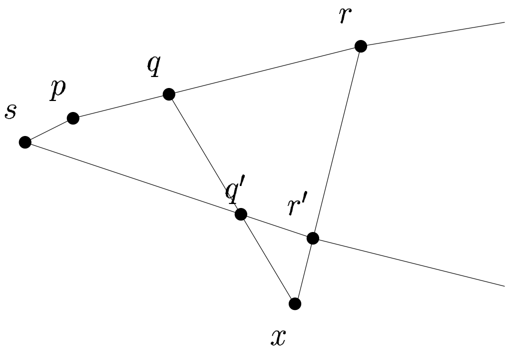

Figure 35: The angle between $q'q$ and $r'r$ must be large, otherwise the region would not be grown further. This contradicts the fact that the distance between q and r is short.

Proof: Breaking γ_j at its sharp corners gives us a collection of subcurves γ_{jk}. Consider any subcurve γ_{jk}. The sample points on γ_{jk} come in three groups: first a group of points outside T', then a group of points in T', and finally a group of points outside T'. Lemma 33 implies that all sample points in $\gamma_{jk} \cap T'$ belong to a single component. Consider now two adjacent subcurves incident to a sharp corner s. In both subcurves, all points in T' belong to the same component. Let p and q be the points in the components containing the points T' which are closest to s and do not connect to both neighbors. We claim that p and q are connected to points on the other leg. Assume otherwise, say p is not connected to a point on the other leg. Since p is not connected to both neighbors we have $p \in B(s, \delta'_s/60)$. Let $p = p^0, p^1, \ldots, p^i$ be the sample points connected to p and ordered accordingly to their distance from p. Then $s < p < p^1 < \ldots < p^i$, since p is not connected to both neighbors and since p is not connected to a sample point on the other side. We have also shown in the proof of Lemma 35 that $p^i \in B(s, \delta'_s)$ Thus $\angle(p^{i-1}p^{i-2}, p^{i-1}p^i) \geq 2\pi/3$, a contradiction to the fact that p^i is the last point joined with p.

Thus p connects to a point u on the other leg and q connects to a point v on the other leg. If either u or v belong to the component containing the points in T', we are done. So assume otherwise. Then u is closer to s than q to s and v is closer to s than p to s and hence $pvuq$ builds a convex quadrangle. The two segments pu and qv are crossing. Thus either $\|pv\| < \|pu\|$ or $\|qu\| < \|qv\|$ (since $\|pv\| + \|qu\| < \|pu\| + \|qv\|$) and hence either p or q will be joined with a sample closer to s, a contradiction. ∎

We call the component containing all sample points in $\gamma_j \cap T'$ the main component of γ_j.

Lemma 37 *The Subtour-LP applied to the main component of* γ_j *reconstructs* γ_j.

Proof: We show that the sample points in $\gamma_j \cap T'$ satisfy our original sampling condition. Consider two points p and q in $\gamma_j \cap T'$ which are adjacent along γ_j. If they are also adjacent in the full sample, we are done. Assume otherwise. Then $\{p, q\} \subseteq B(s, 2\delta'_s/60)$ for some

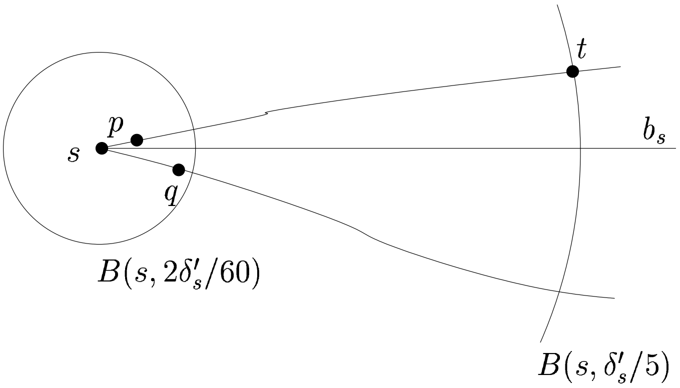

Figure 36: The definition of t.

sharp corner s. Let t be the point so that $c_s = d(t)/3 + \|st_s\|$ (see Figure 36). We have $\|pq\| \leq 2 \cdot (2\delta'_s/60) \cdot \sin 3\overline{\alpha}_s/4 + |\|sp_s\| - \|sq_s\||$, since $\angle(sp, sp_s) \leq 3\overline{\alpha}_s/4$ and similarly for q, $d(t) \geq \delta'_s/5 \sin \overline{\alpha}_s/2$, since the ball around t with radius $\delta_s/5 \sin \overline{\alpha}_s/2$ does not intersect the other leg, $\mu(p) \geq c_s - \|sp_s\| = d(t)/3 + \|st_s\| - \|sp_s\| = d(t)/3 + \|t_s p_s\| \geq \frac{\delta'_s}{15} \sin \overline{\alpha}_s/2 + \|t_s p_s\|$ and, by the same argument, $\mu(q) \geq \frac{\delta'_s}{15} \sin \overline{\alpha}_s/2 + \|t_s q_s\|$. Thus $c_\mu(p, q) \leq \frac{4\delta'_s}{60} \sin 3\overline{\alpha}_s/4 + |\|sp_s\| - \|sq_s\|| - 2\frac{\delta'_s}{15}\frac{2}{3} \sin 3\overline{\alpha}_s/4 - \|t_s p_s\| - \|t_s q_s\| \leq 0$. ∎

4.6.2 Merging Components

The initial partition may contain too many components. For every curve γ_j, there is a main component which contains all sample points in $\gamma_j \cap T'$ and maybe other components. We call them minor components. Each minor component is contained in $B(s, 2\delta'_s/60)$ for some sharp corner s. The reconstruction based on the Subtour-LP is guaranteed to succeed for the main component of every curve. For the minor components it may or may not produce a tour. In this section we describe a strategy for merging components.

We define a region R_{pq} for every edge pq of the computed reconstruction. If the region R_{pq} for an edge pq of the computed reconstruction contains a sample point from another component, we join the component containing p and q with the component of the point closest to pq and lying in another component. We continue until the components stabilize.

Before we define the regions R_{pq} we draw an important consequence of the merging rule. For an edge in a minor component the point closest to it and in another component is guaranteed to lie on the same curve. This follows from the fact that a minor component is contained in $B(s, 2\delta'_s/60)$ for some sharp corner, that the corresponding major component has a point in $B(s, 4\delta'_s/60)$, and that any point within $B(s, \delta'_s)$ belongs to the same curve.

In the arguments to follow we can therefore concentrate on edges in the reconstruction of the main component. In particular, we can use the fact that the Subtour-LP correctly reconstructs the main component. We come to the definition of the regions R_{pq}. We define R_{pq} as the union of a region R'_{pq} and the circumcircle of the segment pq. For every

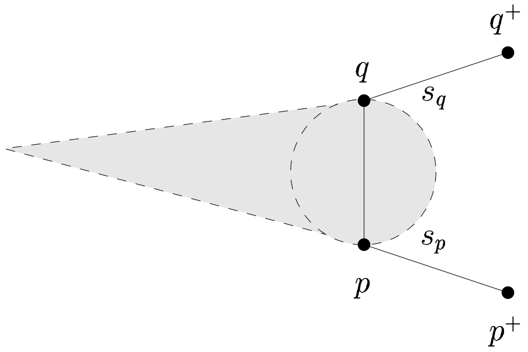

Figure 37: The definition of R_{pq}.

sample point p we define β_p to be the angle between the two segments incident to p. The following paragraph motivates our definition of the region R'_{pq}.

Assume pq is the segment in the main component connecting the two legs of a sharp corner s (we say that the edge straddles the sharp corner) and let $\bar{\alpha}_s$ be the angle between the two tangents at s. Let s_p be a segment defined by two adjacent sample points on the leg of p, let s_q be a segment defined by two adjacent sample points on the leg of q, both lying in $B(s, \delta'_s/7)$. Let θ_s be the angle formed by the segments. Since either segment forms an angle less than $\bar{\alpha}_s/4$ with the corresponding tangent at s, we have $\bar{\alpha}_s/2 \leq \theta_s \leq 3\bar{\alpha}_s/2$ or $2\theta_s/3 \leq \bar{\alpha}_s \leq 2\theta_s$, i.e. , θ_s is a good estimator for $\bar{\alpha}_s$. Since again the angle between any tangent on a leg and the appropriate tangent in the corner is at most $\alpha_s/4$, we know that the angle between the corner point and the points p and q is between $\bar{\alpha}_s/2$ and $3\bar{\alpha}_s/2$, thus between $\theta_s/3$ and $3\theta_s$.

We come to the definition of R'_{pq}. For an edge pq let p^+ and q^+ be the other neighbors of p and q, respectively, and let θ_s be the angle between the segments pp^+ and qq^+, i.e. , $\theta_s = \beta_p + \beta_q - \pi$. We define R'_{pq} as the set of all points r with

- $\angle(\overrightarrow{rp}, \overrightarrow{rq}) \geq \theta_s/3$,
- $\angle(\overrightarrow{pr}, \overrightarrow{pq}) \leq \pi - \beta_p + \min(\pi/20, \theta_s)$ and r lies on the opposite halfspace with respect to the line pq as p^+, and
- $\angle(\overrightarrow{qr}, \overrightarrow{pq}) \geq \pi - \beta_q + \min(\pi/20, \theta_s)$ and r lies on the opposite halfspace with respect to the line qp as q^+.

For an illustration of this definition see Figure 37. Note that $\pi - \beta_q + \min(\pi/20, \theta_s) \leq \pi - \beta_q + \theta_s = \beta_p \leq \pi$.

Lemma 38 *If the main component of a curve does not yet contain all sample points from the curve, it will grow.*

Proof: The main component contains all points in T'. Consider a sample point on γ_j which does not belong to the main component and let p and q be its adjacent sample

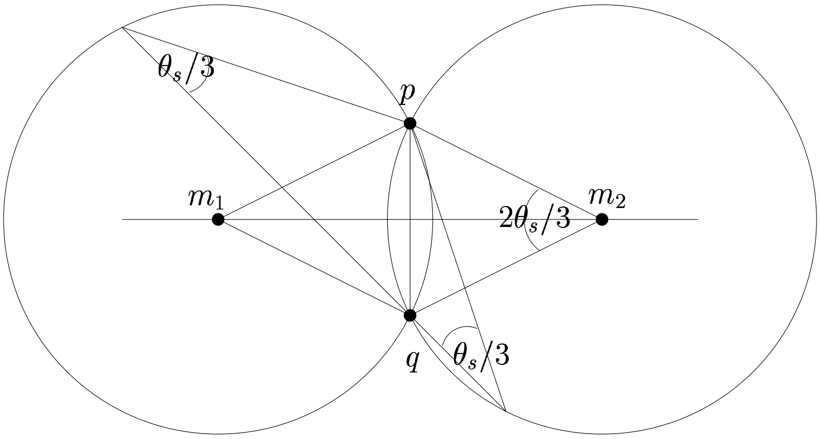

Figure 38: The Inscribed Angle Theorem: The central angle is equal to twice the inscribed angle

points in the main component. Then $\{p, q\} \subseteq B(s, 2\delta'_s/60)$ for some sharp corner. If p and q lie on the same leg of s the centerball of pq contains the subcurve between p and q and if pq straddles the sharp corner, the region R'_{pq} contains the subcurve. ∎

To prove that we do not merge components that do not belong to the same curve, we need the following three Lemmas.

Lemma 39 *Every point r in R'_{pq} has distance at most $\|pq\|/sin(\theta_s/3)$ from p and q.*

Proof: Let m_1 and m_2 be the points on the perpendicular bisector of pq with distance $\|pq\|/(2\sin(\theta_s/3))$ from p and q (see Figure 38).

By the Inscribed Angle Theorem, every point which sees pq under an angle of at least $\theta_s/3$ lies inside the union of the the balls with center m_1 or m_2 through p and q. Thus any point in the region R_{pq} has a distance of at most $\|pq\|/\sin(\theta_s/3)$ from p and q. ∎

Lemma 40 *$\theta_s \geq \alpha_s/2$.*

Proof: If the segment pq straddles the corner, $\theta_s \geq \alpha_s/2$, since pp^+ lies completely on one leg and qq^+ lies completely on the other leg. If p^+ and q^+ are on the same leg, this follows directly from the sampling condition. If the segment pp^+ straddles the corner, the angle formed by the segments pp^+ and qq^+ is smaller than the angle formed by the segments sp^+ and qq^+, which is at least $\alpha_s/2$. ∎

Lemma 41 *Let pq be a segment of the polygonal reconstruction, with $p, q \in T$ and $d(p) \geq d(q)$. Then $d(p) \geq 3\|pq\|/2$.*

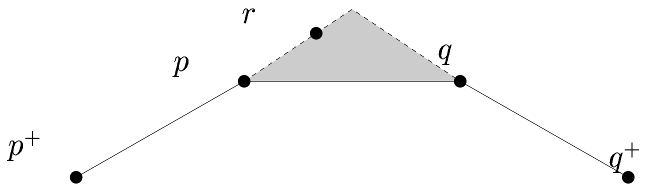

Figure 39: The segments rp and rq form a small angle with the segment pq. Thus $\|pr\|$ can not be much larger than $\|pq\|$.

Proof: Assume otherwise. Than $c_\mu(p,q) = \|pq\| - d(p)/3 - d(q)/3 \geq \|pq\| - 2d(p)/3 > \|pq\| + \|pq\| \geq 0$. ▮

We now turn to the proof that we do not merge a curve with points from another component.

Lemma 42 *The region R_{pq} of an edge pq in the polygonal reconstruction of the main component contains no sample point of another component.*

Proof: This is obvious for the center ball of pq. We turn to the region R'_{pq}.

Assume first p and q are in T' and without loss of generality$d(p) \geq d(q)$. Assume there is a point r outside the the $B(p, d(p))$ ball in R'_{pq}. Then $\|pr\| \geq d(p) \geq 3\|pq\|/2$ by Lemma 41.

We know $\angle(\overrightarrow{pq}, \overrightarrow{pp^+}) \geq 2\pi/3$ and $\angle(\overrightarrow{pp^+}, \overrightarrow{pr}) \geq 19\pi/20$. Hence $\angle(\overrightarrow{pq}, \overrightarrow{pr}) \leq (\pi - 2\pi/3) + (\pi - 19\pi/20) = 23\pi/60$, see Figure 39. Analogously $\angle(\overrightarrow{qp}, \overrightarrow{qr}) \leq 23\pi/60$, thus $\angle(\overrightarrow{rp}, \overrightarrow{rq}) \geq 7\pi/30$. So[5] $\|pr\| \leq \sin(23\pi/60)/\sin(7\pi/3)\|pq\| < 3\|pq\|/2$.

Assume next that one of p and q does not belong to T'.

We will show that $R'_{pq} \subseteq B(s, \delta'_s)$. Let r be any point in R'_{pq} and let θ be the angle under which r sees the segment pq. Then $\theta \geq \theta_s/3$ by the definition of R'_{pq}.

Assume first $\theta_s \geq \pi/6$. We know $\|pr\| \leq \|pq\|/\sin(\theta_s/3) \leq 6\|pq\|$. Thus $\|sr\| \leq 6(8\delta'_s/60) + 2\delta'_s/60 < \delta'_s$.

Assume now $\theta_s < \pi/6$. Look at the triangle $\triangle pqr$ and assume without loss of generality$\pi - \beta_p > \pi/2$ (see Figure 40). We show that the corner point s is almost as far away from p as r. The angle at r is at least $\theta_s/3$, the angle at p is at most $\pi + \theta_s - \beta_p$. Thus $\|pr\| \leq \|pq\| \sin(\pi + \theta_s - \beta_p)/\sin(\theta_s/3)$.

The sharp corner s forms an angle of at most $3\theta_s$ with points p and q, since $\bar{\alpha}_s \leq 2\theta_s$ according to Lemma 40. The angle at p of the triangle $\triangle pqs$ is between $\pi + \theta_s - \beta_p$ and $\pi - \theta_s + \beta_p$.

We have to distinguish two cases according to β_p. Assume first $\beta_p > \pi - 2\theta_s$. We conclude $\|rs\| \leq \|ps\| + \|pr\| \leq 2\delta'_s/60 + \|pq\| \sin(\pi + \theta_s - \beta_p)/\sin(\theta_s/3) \leq 2\delta'_s/60 + \|pq\| \sin(3\theta_s)/\sin(\theta_s/3) \leq 2\delta'_s/60 + 10\|pq\| \leq 2\delta'_s/60 + 40\delta'_s/60 \leq \delta'_s$.

---

[5] In this proof we make frequent use of the fact that $a/\sin\alpha = b/\sin\beta = c/\sin\gamma$ for a triangle with sides a, b, c and corresponding angles α, β, and γ.

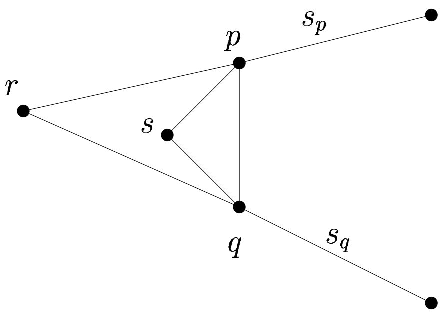

Figure 40: The two triangles pqr and pqs.

Assume now $\beta_p < \pi - 2\theta$. We conclude $\|rs\| \leq \|ps\| + \|pr\| \leq 2\delta'_s/60 + \|pq\| \sin(\pi + \theta_s - \beta_p)/\sin(\theta_s/3) \leq 2\delta'_s/60 + \|ps\|(\sin(3\theta)/\sin(\pi - \theta_s + \beta_p))(\sin(\pi + \theta_s - \beta_p)/\sin(\theta_s/3)) \leq 2\delta'_s/60 + \|ps\|(\sin(3\theta_s)/\sin(\theta_s))(\sin(3\theta_s)/\sin(\theta_s/3)) \leq 2\delta'_s/60 + 27\|ps\| \leq 2\delta'_s/60 + 52\delta'_s/60 < \delta'_s$. ∎

5 Further Results Concerning Curve Reconstruction

In this section we present some further interesting results on curve-reconstruction. For computational purposes it is desirable to restrict the search for the reconstruction to a sparse graph defined on the sample set. We show that the edges of the polygonal reconstruction are in the Delaunay Diagram for a slightly strengthened sampling condition. Curve reconstruction problems are allowed to "invent" curves, if the input does not satisfy the required sampling condition. This issue is discussed in Section 5.2. All previous algorithms use a sample condition based on the medial axis. In Section 5.3 we relate our sample condition to the medial-axis bases sample conditions. In Section 5.4, we relate our result to so-called necklace tours. Necklace tours are a polynomially solvable case of the Traveling Salesman Problem.

5.1 Curve Reconstruction and the Delaunay Diagram

Most previous curve reconstruction algorithms use sampling conditions that guarantee that the polygonal reconstruction is a subset of the Delaunay diagram. Our sampling condition does not imply that the Traveling Salesman Tour is a subgraph of the Delaunay triangulation, see Figure 41. This fact can be interpreted positively and negatively: positively, as an indication of the strength of the TSP-reconstruction, and negatively, since the optimal Traveling Salesman tour must be searched for in the complete graph on the sample set. In this section we show that a slight strengthening of our sample condition implies that the polygonal reconstruction is contained in the Delaunay Diagram.

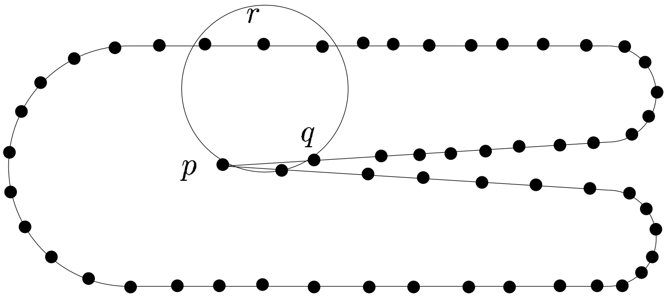

Figure 41: The edge pq does not belong to the Delaunay Triangulation; p is a sharp corner with $\delta_p = \|pr\|/2$. Also $c_p \approx 2\delta_p/5 = \|pr\|/5$. If $\|pq\| < \|pr\|/5$, the edge pq has negative reduced cost. Thus our sampling condition is satisfied.

Additional condition on the sample set: An edge pq of the polygonal reconstruction with an endpoint not in T has length of at most $4\delta_s \sin(\alpha_s/2)/5$, where s is the sharp corner with $\{p, q\} \subseteq B(s, \delta_s)$.

Lemma 43 *If the sample set V satisfies the strengthened sampling condition, the polygonal reconstruction is contained in the Delaunay diagram of V.*

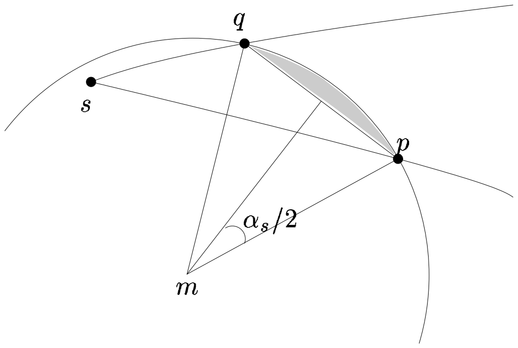

Figure 42: The sample x must lie on the upper leg behind q and in the shaded lune.

Proof: Let pq be an edge of the polygonal reconstruction. We construct a Delaunay ball B for it.

Assume first that $\{p, q\} \subseteq T$. The center ball of pq is contained in $B(p, d(p))$ and hence empty of sample points; otherwise, the curve γ would either intersect the ball in more than one component or turn by more than $\pi/3$ within the ball.

Assume next that one of the endpoints of the edge, say p, does not lie in T. Let s be the sharp corner with $p \in B(s, \delta_s)$. We distinguish the cases whether p and q are on the same leg of s or not.

Assume first that p and q lie on different legs, see Figure 42. We put the center m of B into the halfspace containing s and having p and q in its boundary and set the radius r of B to $(\|pq\|/2)/\sin(\bar{\alpha}_s/2) \leq 2\delta_s/5$; the upper bound on the radius follows from the strengthened sampling condition. Since one of p or q is contained in $B(s, \delta_s/5)$, we conclude $B(m, r) \subseteq B(s, \delta_s)$.

Assume now that there is a sample point $x \in B(m, r)$. We discuss the case in which x lies on the same leg as q and leave the other case to the reader. Since p and q are adjacent samples, x cannot lie on the segment sq and hence $\angle(\vec{sp}, \vec{qx}) = \angle(\vec{sp}, \vec{sq}) + \angle(\vec{sq}, \vec{qx}) \leq \angle(\vec{qp}, \vec{sq}) + \angle(\vec{sq}, \vec{qx}) = \angle(\vec{qp}, \vec{qx})$ (We have $\angle(\vec{sp}, \vec{sq}) < \angle(\vec{qp}, \vec{sq})$, since moving along the segment $\vec{sq}$ increases the angle). Since x lies in the lune of $B(m, r)$ defined by pq, we have $\angle(\vec{qx}, \vec{qp}) \leq \bar{\alpha}_s/2$. Thus $\angle(\vec{sp}, \vec{qx}) \leq \bar{\alpha}_s/2$. On the other hand, the angle between the two tangents at s is $\bar{\alpha}_s$ and hence $\angle(\vec{sp}, \vec{qx}) > \bar{\alpha}_s - 2\bar{\alpha}_s/4 = \bar{\alpha}_s/2$, a contradiction.

Assume next that p and q lie on the same leg, see Figure 43. The center m of B lies on the same side of the angular bisector of the cone defined by the tangents $\mathrm{t}_r(s), -\mathrm{t}_l(s)$ as p and q and sees the segment pq under an angle of $\bar{\alpha}_s$. As above, we conclude that $B(m, r) \subseteq B(s, \delta_s)$. Then x must be contained in the lune of $B(m, r)$ defined by the segment pq. Since p and q are adjacent sample points, x must lie on the other leg. Assume without loss of generalitythat p is closer to s than q. We have $\angle(\vec{pq}, \vec{px}) \leq \bar{\alpha}_s/2$ and hence $\bar{\alpha}_s = \angle(\mathrm{t}_r(s), -\mathrm{t}_l(s)) < \angle(\vec{ps}, \vec{px}) + 2(\bar{\alpha}_s/4) \leq \bar{\alpha}_s$, a contradiction. ∎

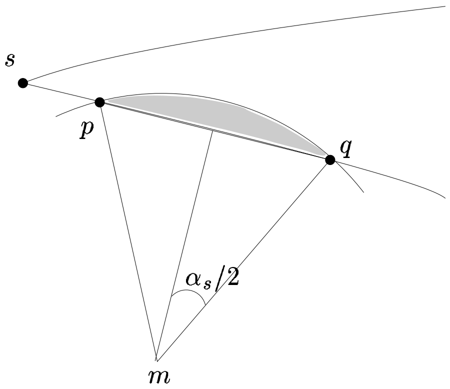

Figure 43: The sample x must lie on the upper leg and in the shaded lune.

5.2 Monotonicity

Intuitively, a larger sample set makes the reconstruction task simpler. We discuss how various sampling conditions and reconstruction algorithms behave with respect to larger sample sets.

A sampling condition is called *monotone* if any superset of a set satisfying the sampling condition also satisfies the sampling condition.

For closed curves and open curves with specified endpoints, our sampling condition is monotone. For open curves with unspecified endpoints the superset must satisfy the additional constraint that the additional points must lie in $\gamma[a, b]$, where a and b are the extreme sample points in the old sample. The sampling conditions used in the papers [ABE98, DK99, Gol99, DMR00] are also monotone; again the additional constraint is needed for open curves. The conditions in [FR01] are not monotone.

All algorithms mentioned in this thesis come with a guarantee of the form: If the curve γ is from a certain class of curves and the sample set V is sufficiently dense, the algorithm will reconstruct γ. It is not specified what the algorithm does if the hypothesis of the theorem is not satisfied. The algorithm may either fail, i.e., indicate that it could not find a curve, or "invent" a curve. From a practical point of view this situation is unsatisfactory as a user has in general no way of distinguishing reconstruction from invention. The situation is aggravated by the fact that the sampling densities required by the theorems are quite high and that the algorithms tend to work for smaller densities and hence are likely to be used in situations not covered by the theorems. *It would be nice to have algorithms that never invent curves.*

A reconstruction algorithm is called *self-consistent* if it has the following property. On an input V it either outputs `FAILURE` or `SUCCESS`. In the latter case it also outputs a curve Γ passing through V so that for any sample V' from Γ with $V \subseteq V'$, it will also output Γ. A reconstruction algorithm that is not self-consistent can change its mind if given additional sample points that seem to confirm the output of the algorithm.

Theorem 17 *The algorithm in [DMR00] and the TSP-algorithm are self-consistent, the*

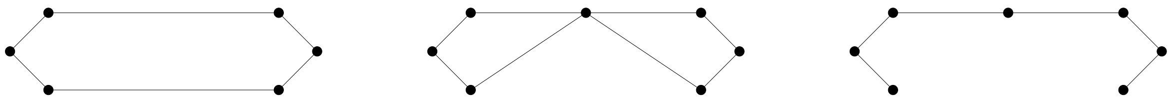

Figure 44: For the sample set in the left figure, the algorithms in [ABE98, DK99, Gol99] produce the hexagon, as shown. If one adds a sample point in the middle of one of the long segments, the algorithm in [DK99] produces the output of the middle figure, the algorithms in [ABE98, Gol99] produce the output of the right figure. Thus none of these algorithms is self-consistent.

algorithms in [ABE98, DK99, Gol99, FR01] are not self-consistent.

Proof: The algorithm in [DMR00] is constructed to be self-consistent. For the algorithms in [ABE98, DK99, Gol99, FR01] it is easy to come up with examples that show non-self-consistency (see Figure 44).

It remains to show self-consistency for the TSP-algorithm. We show that if the solution for the Subtour-LP for a set V is unique and integral then the same is true for the Subtour-LP for $V \cup \{z\}$ for any z on an edge of the integral solution of the Subtour-LP of V (and different from all points in V).

The claim is a simple consequence of the so-called splitting-off lemma (see [Lov79], Problem 6.53). Consider an optimal solution of the Subtour-LP for $V \cup \{z\}$. The value C of this solution is at most the length C_0 of the optimal tour of V (since the tour for V is also a tour for $V \cup \{z\}$). The splitting-off lemma allows the construction of a solution for the Subtour-LP for V from the solution for $V \cup \{z\}$. It implies the existence of a set of triples (e, f, r), where e and f are edges incident to z and r is a non-negative real, so that

- for each edge e incident to z the sum of the third components of all triples containing e is equal to x_e (the value of the edge e in the Subtour-LP for $V \cup \{z\}$) and
- a solution for the Subtour-LP for V can be obtained by modifying the solution for the Subtour-LP for $V \cup \{z\}$ as follows: for each edge uv with $u \neq z$ and $v \neq z$ increase x_{uv} by the sum of the third components of all triples (uz, zv, r). Delete all edges incident to z.

For any edge $e = uv$ let y_e be the increase of x_e in this construction. The cost of the solution obtained is $C + \sum_{uv} y_{uv}(c_{uv} - c_{uz} - c_{zv})$. This cost is at most C (and hence at most C_0) since $c_{uv} \leq c_{uz} + c_{zv}$ by the triangle inequality. Since the solution for the Subtour-LP for V is unique and is equal to the tour for V, the cost of the solution cannot be smaller than C_0 and hence for any edge $e = uv$ with $y_{uv} > 0$ we must have $c_{uv} = c_{uz} + c_{zv}$, i.e., z lies on the line segment $\overline{uv}$. Moreover, $y_e + x_e$ must be integral for every edge $e = uv$.

In the tour for V there is only one edge passing through z (since optimal tours are non-self-intersecting) and hence there can be only one edge uv with $y_{uv} > 0$. Thus our set of triples consists of a single triple (uz, zv, r) and since the degree constraint at z must be satisfied for the optimal solution of the Subtour-LP for $V \cup \{z\}$ we conclude that $r = 1$. We conclude that the optimal solution of the Subtour-LP for $V \cup \{z\}$ is unique and integral. ∎

5.3 Our Sample Condition and the Local Feature Size

The papers [ABE98, DK99, Gol99, DMR00] investigated the reconstruction problem for smooth curves. A curve is smooth if it is twice-differentiable. They expressed the sampling condition in terms of the so-called *local feature size*. The local feature size $f(p)$ at a curve point p is the distance of p from the medial axis of γ. The *medial axis* of a curve is the closure of the set of points in the plane which have at least two nearest (with respect to the Euclidean metric) points on the curve. They required a sampling condition of the form: For any $p \in \gamma$ there must be a sample point $v \in V$ with $\|pv\| \leq \epsilon \cdot f(p)$; here ϵ is a parameter which depends on the algorithm. All algorithms require $\epsilon \leq 1/2$.

The experimental results of Section 6 suggest that the TSP-algorithms works for sparser sample sets than the algorithms mentioned above. We do not know whether this observation is a fact and can prove only a much weaker result.

Lemma 44 *Let γ be a smooth curve and $\epsilon < 1/10$. If for any $p \in \gamma$ there is a sample point $v \in V$ with $\|pv\| \leq \epsilon \cdot f(p)$, then V satisfies our sampling condition.*

Before we prove Lemma 44, we show the following Lemma.

Lemma 45

$$f(p) < 3d(p)$$

Proof: The following fact is a reformulation of Corollary 11.

Fact 5 *Let r be a point of a smooth curve γ. Furthermore let q and s be points on γ with distance less than $f(r)$ from r with $q < r < s$. Then $\angle(\vec{rq}, \vec{rs}) > \pi/3$.*

By the definition of $d(p)$, either $B(p, d(p)) \cap \gamma$ is not connected or $B(p, d(p))$ contains three points turning by $\pi/3$.

In the first case, there is a medial axis point in $B(p, d(p))$ by Lemma 1 of [ABE98] and hence $f(p) \leq d(p)$.

We turn to the second case. Let $q < r < s \in B(p, d(p)) \cap \gamma$ forming an angle of $\pi/3$. By Fact 5 we conclude $f(r) \leq \max(\|qr\|, \|rs\|)$. Thus $f(p) \leq \|pr\| + \max(\|qr\|, \|rs\|) < 3d(p)$. ∎

Proof: [of Lemma 44]

We have to show that for an ϵ–sampled curve, with $\epsilon < 1/10$, the modified cost of the edge between two adjacent sample points is less then 0.

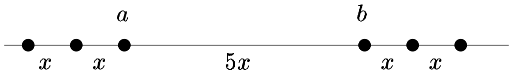

Figure 45: The underlying part of the curve is a line segment. Hence, there is no condition on the sample set. In a necklace tour the disks centered at a and b have radius at most $2x$. Thus the two disks will not intersect.

Let p and q be adjacent sample points. Let without loss of generality$d(p) \leq d(q)$. Lemma 10 states $\|pq\| \leq 2\epsilon f(p)/(1-\epsilon) = 1/5 f(p)/(1-1/10)$. Thus $c_\mu(p,q) = \|pq\| - d(p)/3 - d(q)/3 < 1/5 f(p)/(1-1/10) - f(p)/9 - f(p)/9 \leq 0$ ∎

Our sample condition depends on several parameters and we have set these parameters to particular values in Section 4.2. In Section 4.2.6 we discussed the dependency between the parameters. For smooth curves we can set f_{scale} to 1/2 and then the argument above works for $\epsilon = 1/7$

5.4 Necklace Tours

We have shown that curve reconstruction gives rise to a polynomially solvable case of the Euclidean traveling salesman problem. In this section we relate our results to a known solvable case, the so-called *necklace tours*. Let V be a set of points in the plane and assume that there is a set of disks centered at the points in V so that each disk intersects with exactly two other disks and so that the intersection graph of the disks is connected. The intersection graph of the disks defines a tour on V; the two neighbors of a point v correspond to the two disks that intersect the disk associated with v. The tour is called a necklace tour and is known to be an optimal traveling salesman tour of V; see [BDvD+98].

We cannot claim that necklace tours are a special case of our result, since there is no curve underlying a necklace tour. The optimality proof for necklace tours is a special case of our argument. We simply define the potential of any point v as the radius of the disk associated with v. Then exactly the edges in the tour have non-positive cost. Any tour which uses an edge outside the necklace tour must include edges of positive cost and can include only a subset of the edges in the necklace tour. This implies that the necklace tour is optimal.

Figure 45 shows an example of a TSP problem which is covered by Theorem 15 and whose solution is not a necklace tour.

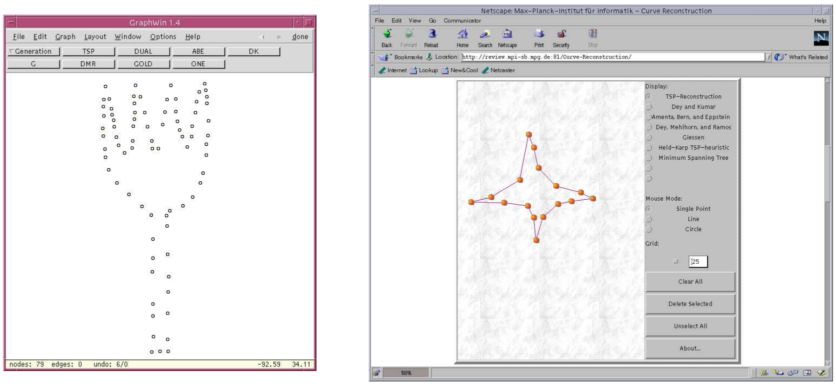

Figure 46: A screen-shot of the testbed and the Internet demo.

6 Experiments

In the preceding sections we have seen several algorithms for the curve reconstruction problem that guarantee the correct reconstruction under some assumptions on the curve γ and the sample set S. In this section we make an experimental comparison of the algorithms.

6.1 The Testbed

Reproducibility of experiments is a major concern for experimental algorithms. We therefore provide our reconstruction algorithms and problem generators as a LEDA extension package. As part of the package we provide a literate programming document describing all the implementations. The package also contains a testbed that allows one to experiment with reconstruction algorithms (not just ours) and visualizes the outcome of the reconstructions. Figure 46 shows a screen-shot of the testbed. The package will be available at `http://www.mpi-sb.mpg.de/LEDA/friends/leps.html`.

We also offer an Internet interface to our implementations. It is available at `http://review.mpi-sb.mpg.de:81/Curve-Reconstruction/`. It consists of a JAVA-applet that allows the user to construct problem instances and to select an algorithm. The applet contacts a server at the MPI to run the algorithm and displays the result of the reconstruction.

6.2 About the Implementation

All our implementations are based on LEDA [LED, MN99]. All algorithms use the graph data type, the geometry kernels, and the Delaunay diagram algorithm. ABE and Gold's algorithms also use the Voronoi diagram algorithm and DMR uses a data structure for nearest neighbor queries. The TSP algorithm uses, in addition, the min-cut and the connected components algorithm, it uses SOPLEX [Wun97] or CPLEX [ILO99] to solve

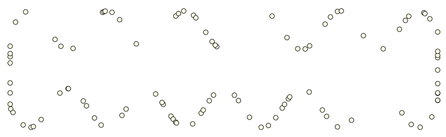

Figure 47: Our sample curve (5 periods and 300 points).

linear programs. All geometric programs can be run with either the rational or the floating point kernel of LEDA.

The implementations of the previous known algorithms are straight forward. Given the sample set S, we first compute the Delaunay Triangulation of S using the algorithm provided by LEDA. Than the filtering of the edges is done.

Also the TSP–algorithm first computes the Delaunay Triangulation. For every edge of the Delaunay triangulation, we introduce a binary variable and set up a matrix with the degree constraints. We iteratively solve the LP, using a LP–solver as black–box and look for violated subtour elimination constraints until we can not find further violated constraints. If the solution of the last LP is integral, we return the induced graph, otherwise we return the empty graph. It remains to explain how we find violated subtour elimination constraints. We run a fast heuristic, and if the heuristic fails to find a violated constraint, we run the exact algorithm to compute the minimal cut in the appropriate graph, as explained in section 4. The heuristic works as follows. Let G' be the subgraph of the Delaunay Triangulation with nodes corresponding to the sample points and edges corresponding to those edges of the Delaunay Triangulation that have a value 1 in the current LP solution. We test the subtour elimination constraints for the connected components of G' for violation. If we find violated inequalities, we add them to the LP, otherwise the heuristic fails.

6.3 The Experiments

We performed experiments that compare the reconstruction quality and the running time of our implementations. All experiments were carried out with the floating point geometry kernel of LEDA. We used several curves for our experiments. Since the results are fairly consistent over the curves, we report only about the experiments for the curve shown in Figure 47. The curve is essentially a sinusoidal curve. We use p to denote the number of periods, n to denote the number of sample points and $\rho = n/p$ to denote the number of points per period.

6.3.1 Reconstruction Quality

In order to test the reconstruction quality of the different algorithms we used a fixed curve and varied the sampling density. More precisely, we used a fixed number n of points and varied the number p of periods in our curve. For each value of p we chose n random

Figure 48: A reconstruction with the TSP-algorithm and the algorithm of Dey and Kumar.

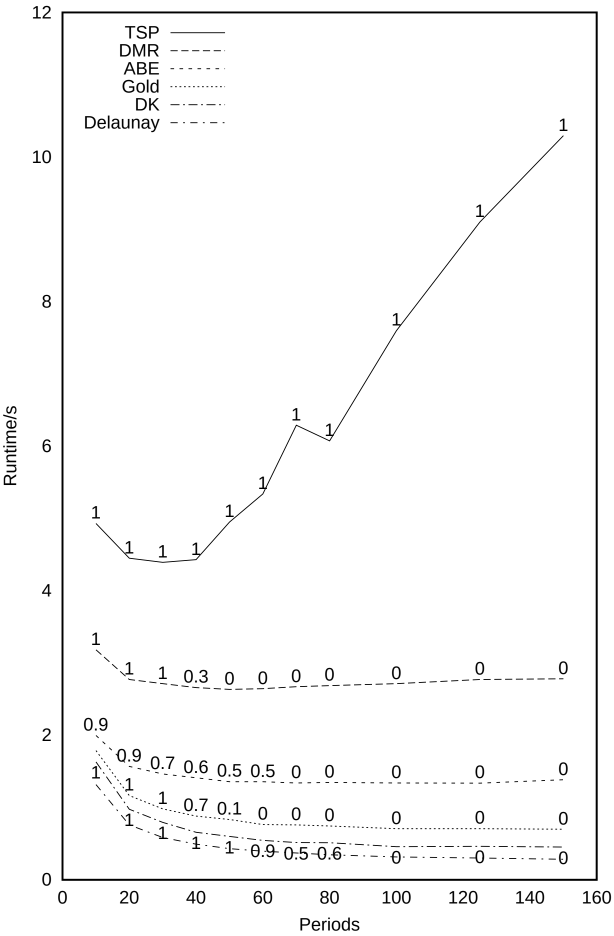

Figure 49: The abscissa shows the number of periods and the ordinate shows the running time (average over ten runs). The numbers on the curves indicate the fraction of problems solved, e.g., 0.4 indicates that 40% of the problems were solved. For all experiments we used $n = 3000$ points. The curve Delaunay indicates the time of the Delaunay diagram computation.

points on the curve. We repeated each experiment ten times. We measured the running time and in what percentage of the cases the algorithms found the reconstruction. The results are shown in Figure 49. We observe that the TSP-algorithm can work with much sparser samples than the other algorithms. Figure 48 shows a reconstruction example.

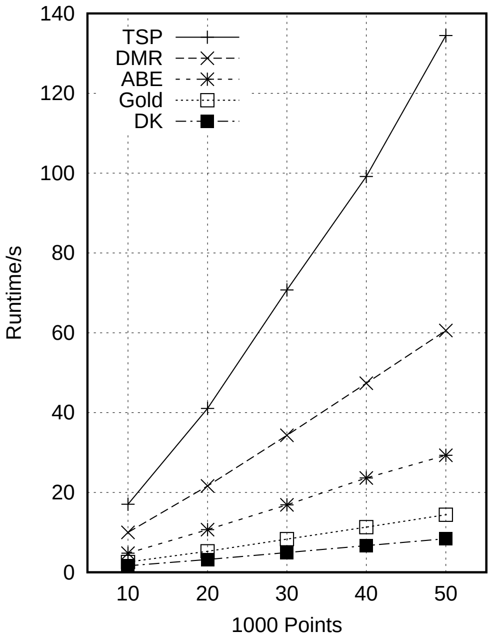

Figure 50: The running times for a curve with $\rho = 100$ points per period and varying number of periods. We used $n = i \cdot 10^4$ sampling points for $i = 1, 2, \ldots, 5$. The running times of the algorithms follow approximately the relation DK : Gold : ABE : DMR : TSP = 1 : 1.3 : 3 : 6 : 13.

Among the other algorithms, DMR does worst (this is to be expected, since it is designed to reconstruct open and closed curves and hence omits more edges), DK is best, followed by Gold and ABE.

6.3.2 Running Time

In order to compare efficiency we performed two experiments: One for varying sampling density and a fixed number of points, one for fixed sampling density and a varying number of points. The first experiment was described in the previous section. For the second experiment we used a curve with $\rho = 100$ points per period, varied the number n of points, and kept the number of periods at $p = n/100$. We used sampling densities for which all algorithms were able to solve the reconstruction problem. The results of the second experiment are shown in Figure 50.

All filtering algorithms start by computing the Delaunay triangulation of the sample set V. DK and Gold do linear additional work, DK selects the nearest of each points and then the nearest neighbor in the other halfspace for each point of degree one. Gold constructs the Voronoi diagram (= a linear time process, when the Delaunay diagram is available) and then performs a local test for the each pair of primal and dual edge. The running time of our Delaunay program [Dwy87] is $O(n)$ in the best case and $O(n \log n)$ in the worst case. For the instances in Figure 49 the running time improves as a function of the number of periods. This is due to the fact that the algorithm uses divide-and-conquer on the x-coordinates of the points and that the running time of the conquer step decreases as the number of periods increases (observe that the number of points contained in a vertical stripe decreases as the number of periods increases). In Figure 49 the running times of DK and Gold are shifts of the cost of the Delaunay computation.

The algorithms ABE and DMR use $O(n \log n)$ (DMR may use $O(n^2)$ additional time in the worst case, but we do not know of a worst-case example) additional time. ABE constructs the Delaunay triangulation of $V \cup V'$, where V' is the set of Voronoi vertices of V. The running time of the second Delaunay computation seems to be independent of the number of periods. DMR uses a data structure for proximity queries which has a running time of $\Theta(n \log n)$.

The running time of the TSP-algorithm also conforms to intuition. For fixed problem size the running time in addition to the time for computing the Delaunay triangulation decreases as a function of sampling density (= increases as a function of the number of periods), since the reconstruction problem becomes simpler as the sampling density increases. For sparse samples the number of phases of the cutting plane algorithm goes up. Moreover, the connected components heuristic is no longer able to find the required cuts. For fixed sampling density and varying n the running time grows linearly in n. The explanation is that a few (typically 5 or 6) iterations of the cutting plane algorithm suffice for sampling densities where the other algorithms are able to find solutions, that the linear time heuristic for finding cuts is sufficient, and that the expected running time of the simplex algorithm is linear in the number of constraints and variables.

6.4 Robustness

The DMR and ABE implementations suffer from robustness problems when run with the floating point kernel of LEDA. We discuss the issue for the ABE algorithm. The ABE first constructs the Voronoi diagram of V and then the Delaunay triangulation of $V \cup V'$, where V' is the set of Voronoi vertices of V. If V contains nearly co-circular points (which might be a frequent situation in the curve reconstruction scenario), V' contains clusters of points and this makes the construction of the Delaunay diagram of $V \cup V'$ a difficult task.

The robustness problems can be removed by switching to the rational kernel of LEDA which guarantees that all geometric predicates are evaluated correctly (and uses a floating point filter for speed). The switch to the rational kernel can cause a significant increase in running time. It is the greatest increase that we have ever experienced. The reason is that the construction of the Delaunay diagram of $V \cup V'$ requires many exact tests, if V contains many co-circular points.

We propose to overcome the robustness problem in the ABE algorithm by rounding the Voronoi vertices to a fine grid. This will identify Voronoi vertices that lie close together.

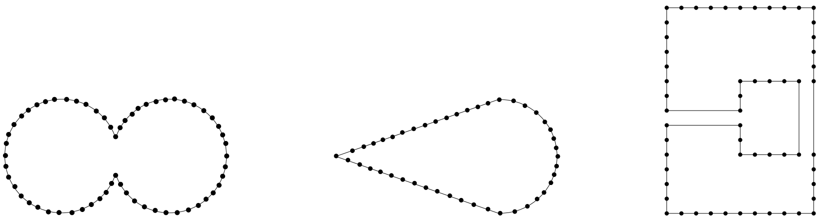

Figure 51: Part (a) shows an example where the heuristics NM, NA, and FA fail, part (b) an example where CH, NN, and MF fail. Part (c) shows a three optimal tour different from the traveling salesman tour.

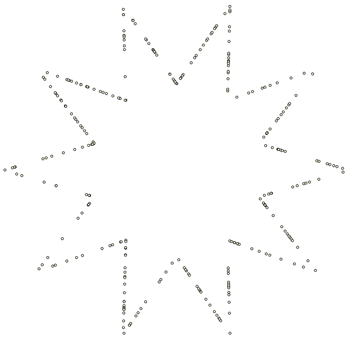

Figure 52: A star with 10 spikes and 300 sample points.

6.5 TSP Heuristics

We have implemented some TSP- and tour improvement heuristics. The TSP heuristics are an iterative version to compute the Held-Karp Bound, Christofides Algorithm (CH), the nearest merger algorithm (NM), the nearest neighbor heuristic (NN), the multiple fragment greedy heuristic (MF), the nearest addition algorithm (NA) and the farthest addition algorithm (FA). In addition we have implemented the *2-opt* and *3-opt* improvement heuristics. For details of these algorithms we refer to [Ben90].

No heuristic reconstructs the tulip of Figure 1 correctly. Furthermore no heuristic solves the reconstruction problem, i.e., for all TSP-heuristics mentioned above there is a closed curve and an arbitrary dense sample for which the algorithm does not find the optimal tour and hence fails to reconstruct. For the improvement heuristics an initial (bad) tour is also chosen. Figure 51 shows examples in which the heuristics fail.

The heuristics NM, NA and FA fail even on smooth curves. They do not reconstruct any of the curves considered in Section 6.3. The remaining heuristics (CH, NN, and MF) reconstruct uniformly sampled smooth curves (even the minimum spanning tree does) but fail by construction in sharp corners.

The improvement heuristics do quite well. We give experimental results for the curve

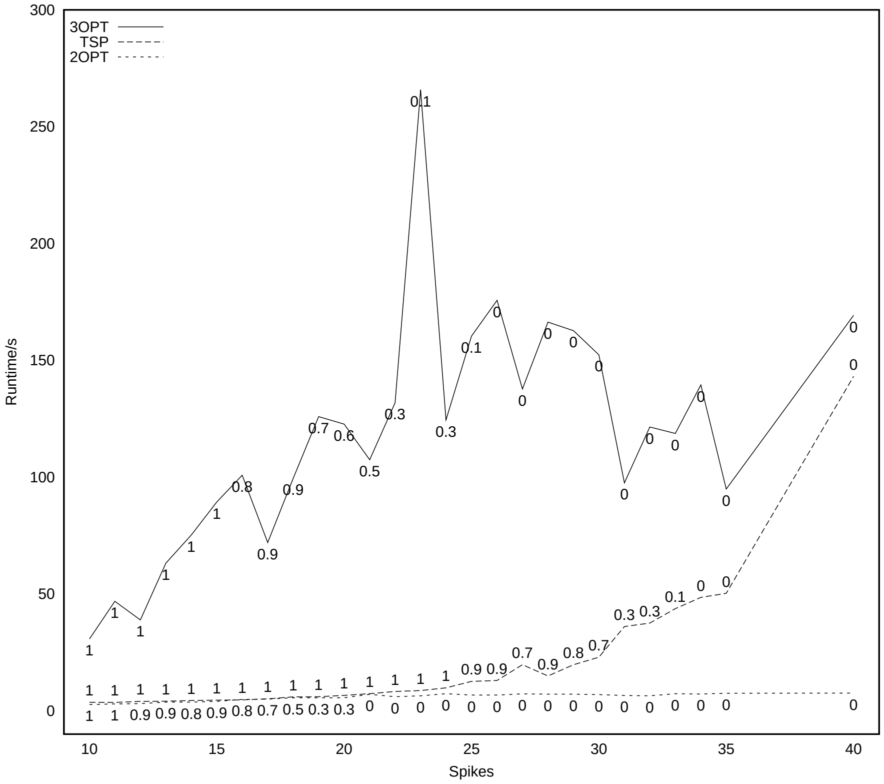

Figure 53: The abscissa shows the number of spikes of the star and the ordinate shows the running time (average over ten runs). The numbers on the curves indicate the fraction of problems solved. For all experiments we used $n = 3000$ points.

shown in Figure 52, a simple star with a varying number of spikes. The results are shown in Figure 53. For dense samples both heuristics succeed, 3-opt can cope with smaller sampling density than 2-opt and the TSP-algorithm can cope with even smaller sampling density. This was to be expected. For sampling densities where the heuristics worked, the TSP-algorithm is almost as fast as 2-opt and considerably faster than 3-opt. We find this surprising. We have to admit however, that our implementations of the improvement heuristics are fairly crude.

7 Discussion

7.1 Curve Reconstruction

Our algorithm solves the curve reconstruction problem for collections of closed curves. There are several possible extentions of the algorithm.

One desire could be, as in the case of smooth curves, to extend the set of curves to collections of semi–regular open and closed curves. Funke and Ramos [FR01] gave an algorithm that provably reconstructs such curves, but their sample condition is not monotone.

A further useful extention of the class of curves could concern curves with branching points. For example the drawing of implicit functions or the reconstruction of road or river networks requires the capability of handling branching points.

In the case of open curves the reconstruction is usually justified by a curve, i.e. , beside the polygonal reconstruction, the algorithm returns a curve for which the given sample is dense with respect to a sampling condition that is only slightly weaker than the original sampling condition (see [DMR00, FR01]). The graph returned by the TSP–algorithm is not justified by a curve.

7.2 Shape Reconstruction and the Minimization Principle

Although Giesen was the first to propose the minimal tour for curve reconstruction, there is earlier work that proposes minimization or maximization for other shape reconstruction problems.

For the surface reconstruction problem from contours, Keppel [Kep75] proposes to maximize the volume in 1975. In 1977, Fuchs et al. [FKU77] propose area minimization and presented an algorithm that can efficiently mimimize the area if both contours are single closed polygons. This algorithm was extended to the case where one contour consists of two closed polygons and the other contour consists of one closed polygon by Kaneda et al. [HSKK98] in 1998.

In the more general case of arbitrary point clouds, O'Rouke [O'R81] also proposed area minimization in 1977. He derives a heuristic for area mimimization similar to the merging algorithms (i.e. , start with a tour through three cities and iteratively add another city to the tour) for the traveling salesman problem.

7.3 Surface Reconstruction from Contours

As mentioned, area minimizing is a successful method for the problem of reconstructing surfaces from planar contours. Exact algorithms are only known if both contours are single closed polygons [FKU77], or if one contour consists of two single closed polygons and the other contour is one single closed polygon [HSKK98]. These algorithms could be extended to handle contours that consist of more polygons, but the running time would explode.

Another possiblity is to formulate the problem as an integer linear program. Similar to the TSP case, we introduce a binary decision variable for every triangle of the Delaunay triangulation of the union of the two point sets. Let F be the set of all triangles and a_f denote the area of the triangle f, for all $f \in F$. The restrictions are that we have to select exactly one triangle that is adjacent to a segment of a polygon and either zero or two triangles adjacent to any other edge. These conditions lead to the following integer linear program. Let V be the set of all sample points, E be the set of all edges, E^s be the set of all edges that are segments of a polygon and $E^r = E \setminus E$ the remaining edges.

$$\begin{array}{lrcll} \min & \sum_{f \in F} a_f x_f & & & \\ s.t. & \sum_{f \in F | e \subset f} x_f & = & 1 & \text{for all edges } e \in E^s \\ & \sum_{g \in F \setminus \{f\} | e \subset f} x_g & \geq & x_f & \text{for all triangles } f \text{ and edges } e \in E^r \text{ with } e \subset f. \\ & \sum_{f \in F | e \subset f} x_f & \leq & 2 & \text{for all edges } e \in E^r \end{array}$$

The set of equalites ensures that exactly one triangle adjacent to every segment of a contour is selected and the set of inequalities ensures that for all other edges either zero or two adjacent triangles are selected.

Together with Christian Fink, we have implemented a branch–&–cut algorithm for the above integer linear program. We used the inequalities above as initial constraint system and separated the following class of inequalities. Let u, v, w be three adjacent points in one polygon. Than the inequalities

$$\sum_{f \in F | v \in f, |S \cap f| = 1} x_f \geq 1 \text{ for all sets } S \text{ with } \{u\} \subset S \subset V \setminus \{v, w\}$$

are valid for the above polytope. They express the following. Look at the triangles in a valid surface that are adjacent to a node v. The edges of these triangles that do not contain v must form a path between u and w. We separate the inequalities by iterating over all points v and computing the minimal cut separating the two adjacent points of v say u and w in the graph over $G \setminus \{v\}$ and edge capacities $c_{xy} = x^*_{vxy}$, where x^*_{vxy} is the actual LP solution of triangle vxy.

The algorithm is quite efficient in practice. We are able to solve all reconstruction problems of the Organ Data Set of Dr. Gill Barequet [Bar]. The reconstruction quality seems to be very high. For an example see Figure 54.

There is no algorithm for the surface reconstruction problem from planar contours that comes with any guarantee. If the contours are close together and the samples in one contour are dense, the sample set would satisfy the sampling condition that is defined for the surface reconstruction problem from arbitrary point clouds. For this problem, an algorithm that is provably correct for smooth surfaces is known (see below).

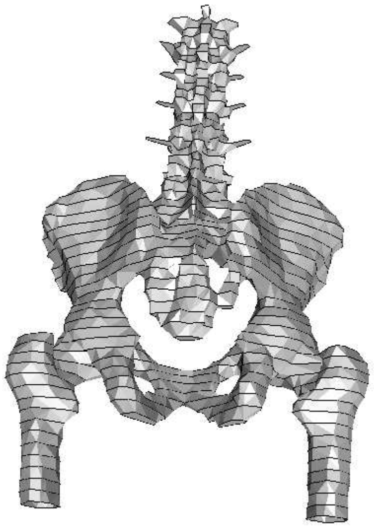

Figure 54: The reconstruction of the pelvis of the Organ Data Set.

7.4 Surface Reconstruction

We have also tried to apply the area minimization paradigm to the general surface reconstruction problem from point clouds. As in the above case, we introduce a binary decision variable for every triangle of the Delaunay Triangulation and formulate the inequality that requires every edge to have either zero or two adjacent triangles. Furthermore, we require that at least three adjecent triangles for every sample point are selected and that the reconstructed surface is connected. These conditions are encoded in the following integer linear program.

$$
\begin{array}{lrcll}
\min & \sum_{f \in F} a_f x_f & & & \\
s.t. & \sum_{g \in F \setminus \{f\} | e \subset g} x_g & \geq & x_f & \text{for all triangles } f \text{ and edges } e \subset f \\
 & \sum_{f \in F | e \subset f} x_f & \leq & 2 & \text{for all edges } e \\
 & \sum_{f \in F} |f \cap S||f \cap V \setminus S \setminus Z| x_f & \geq & 6 - 2|Z| & \text{for all } Z \subset V \text{ with } |Z| \leq 2 \text{ and} \\
 & & & & \emptyset \subset S \subset V \setminus Z
\end{array}
$$

The first two classes of inequalities state that for every edge there are either 0 or 2 adjacent triangles selected and the last class of inequalities expresses the three–connectivity of the graph of the induced edges, i.e. , the edges for which two adjacent triangles are selected.

We implented a branch and cut algorithm for the above integer linear program. Beside the inequalities above, we include some inequalities that force the triangle adjacent to one sample point v to form a cycle around v. Bauer [Bau97] describes some facets of the circuit polytope.

Branch and Cut does not yet work in practice, even if we restrict our attention to surfaces with genus 0 to obtain additional inequalities (the total number of triangles is $2|V|-4$ and in every subset there are at most $2|V|-5$ triangles). We can solve instances with up to 40 sample points. This number is much too small to describe interesting surfaces. We looked at some instances, where we know the triangulated surface with mimimal area and observed that the gap between the solution of the linear program and the integer program was very large, so that there is no hope that this integer linear program can be used unless we find a very effective class of valid inequalities.

Amamy, Giesen and John [AGJ00] propose using a combinatorial algorithm to detect most of the triangles of the surface and use integer linear programming to make some local repairs.

The heuristic of O'Rouke is not sufficient for computing an appropriate surface for complicated point cloulds.

Amenta and Bern [AB98] have developed an algorithm that provably reconstructs a smooth surface, if a sampling condition similar to the medial axis condition in the case of curves is satisfied. They define a triangle to be correct, iff the dual edge in the Voronoi diagram intersects the surface. This definition makes also sense if the surface is not smooth. In this case no algorithm with theoretical gurarantees is known. It also remains an open question whether the surface with mimimal area is the correct reconstruction even for smooth surfaces.

Summary

An instance of the curve reconstruction problem is a finite sample set V of an unknown collection of curves γ. The task is to construct a graph G on V so that two points in V are connected by an edge of G iff the points are adjacent on γ. The curve reconstruction problem and the related surface reconstruction problem have received a lot of attention in the graphics and the computational geometry community. We are interested in *reconstruction algorithms with guaranteed performance*, i.e., algorithms which provably solve the reconstruction problem under certain assumptions on γ and V.

Many curve reconstruction algorithms have been proposed in the past; we restrict our discussion to algorithms that provably solve the reconstruction problem for a certain class of curves and under certain assumptions on the sample set. The algorithms differ with respect to the following aspects:

- Whether a collection of curves or just a single curve can be handled.
- Whether (collections of) open and closed curves can be handled or only (collections of) closed curves.
- Whether the sampling must be uniform or not. Uniform sampling with density ϵ requires that the sample set V contain at least one point from every curve segment of length ϵ. In non-uniform sampling, the sampling frequency may depend on local properties of the curve, e.g., can be lower in parts of low curvature.
- Whether non-smooth curves can be handled or not. A smooth curve has a tangent everywhere.

For uniformly sampled collections of closed smooth curves, several methods are known to work, ranging over minimum spanning trees [FG94], α-shapes [BB97, EKS83], β-skeletons [KR85], and r-regular shapes [Att97]. A survey of these techniques appears in [Ede98]. The case of non-uniformly sampled collections of closed smooth curves was first successfully treated by Amenta, Bern and Eppstein [ABE98] and has been subsequently improved by [DK99, Gol99]. Non-uniformly sampled collections of open and closed smooth curves were treated in [DMR00]. All papers mentioned so far require the curves to be smooth.

Giesen [Gie99] recently obtained the first result for non-smooth curves. He considered the class of *benign semi-regular curves*. An (open or closed) curve is *semi-regular*, if a left and a right tangent exists in every point of the curve; the two tangents may however be different. A semi-regular curve is *benign*, if the turning angle at every point of the curve is less than π. Giesen showed that the Traveling Salesman tour of the sample set V solves the curve reconstruction problem for uniformly sampled benign closed semi-regular curves. More precisely, he showed that for every benign semi-regular closed curve γ, there exists a positive ϵ, so that the optimal Traveling Salesman tour of V is the polygonal reconstruction of γ provided that for every $x \in \gamma$ there is a $p \in V$ with $\|xp\| \leq \epsilon$, where $\|xy\|$ is the Euclidean distance of the two points x and y. Giesen's result is a prove of existence; he did not quantify ϵ in terms of properties of the curve γ. We extend Giesen's result in several directions:

- We relate ϵ to local properties of the curve γ and show that the optimal Traveling Salesman tour solves the reconstruction problem even if sampling is non-uniform. For smooth curves our sampling condition is similar to the one used in [ABE98, DK99, Gol99, DMR00].
- We show that the Traveling Salesman path is able to reconstruct open curves for a suitable sampling condition. We treat the case of paths with and without specified endpoints.
- We show that the optimal Traveling Salesman tour (path) can be constructed in polynomial time if our sampling condition is satisfied.
- We give a simplified proof that the Traveling Salesman tour (path) solves the curve reconstruction problem.
- We show that an extension of the Traveling Salesman tour algorithm is able to reconstruct non-uniformly sampled collections of closed non-smooth curves.

For computational purposes it is desirable to restrict the search for the reconstruction to a sparse graph defined on the sample set. We show that the edges of the polygonal reconstruction are in the Delaunay diagram for a slightly strengthened sampling condition. Furthermore, we relate our sampling condition to the sampling conditions used by most of the other papers, and we relate our result to so-called necklace tours. Necklace tours are a polynomially solvable case of the Traveling Salesman problem. Curve reconstruction problems are allowed to "invent" curves, if the input does not satisfy the required sampling condition. This issue is also discussed.

We have implemented a number of curve reconstruction algorithms. The JAVA-Applet `http://review.mpi-sb.mpg.de:81/Curve-Reconstruction/` makes our implementations available. Our experiments show that the TSP-based curve reconstruction is able to solve the reconstruction problem for surprisingly small sampling density and that its speed is comparable to Delaunay diagram-based reconstruction algorithms.

Zusammenfassung

Die Eingabe eines Kurvenrekonstruktionsproblems ist eine endliche Menge V von Samplepunkten auf einer unbekannten Kurve γ. Die Aufgabe besteht darin, einen Graphen $G = (V, E)$ zu konstruieren, in dem zwei Punkte genau dann durch eine Kante verbunden sind, wenn die Punkte in γ benachbart sind. Dem Kurvenrekonstruktionsproblem und dem verwandten Oberflächenrekonstruktionsproblem wurde viel Aufmerksamkeit von Seiten der Computergrafik und der Computergeometrie entgegengebracht. Wir sind an *Algorithmen mit garantierter Korrektheit* interessiert, d.h. an Algorithmen, die unter bestimmen Annahmen an γ und V garantiert die richtige Lösung zurückliefern.

In der Vergangenheit wurden viele Algorithmen für das Kurvenrekonstructionsproblem vorgeschlagen; wir beschränken unsere Aufmerksamkeit auf Algorithmen, die das Rekonstruktionsproblem beweisbar für eine bestimmte Klasse von Kurven unter bestimmten Annahmen an die Menge von Samplepunkten lösen. Die Algorithmen unterscheiden sich bezüglich folgender Aspekte:

- ob eine Menge von Kurven oder nur einzelne Kurven behandelt werden können,
- ob jeweils offene und geschlossene Kurven oder nur geschlossene Kurven behandelt werden können,
- ob die Menge der Samplepunkte gleichmäßig sein muß oder nicht. Ein gleichmäßiges Sample mit Dichte ϵ hat die Eigenschaft, daß V ein Samplepunkt auf jedem Kurvenstück der Länge ϵ enthält. Dahingegen hängt beim ungleichmäßigen Sample die Sampledichte von lokalen Eigenschaften der Kurve ab, d.h. die Sampledichte kann geringer sein, wenn die Kurve in einem Bereich weniger gekrümmt ist,
- ob nur glatte Kurven behandelt werden können oder nicht. Eine glatte Kurve besitzt in jedem Punkt eine Tangente.

Für gleichmäßige Samplemengen von geschlossenen, glatten Kurven gibt es mehrere beweisbare Methoden, wie z.B. minimale aufspannende Bäume [FG94], α-shapes [BB97, EKS83], β-skeletons [KR85], und r-regular shapes [Att97]. Ein Überblick über diese Methoden ist in [Ede98] zu finden. Der Fall von ungleichmäßigen Samplemengen wurde erstmals von Amenta, Bern und Eppstein [ABE98] erfolgreich behandelt und schrittweise durch Algorithmen wie z.B. [DK99, Gol99] verbessert. Ungleichmäßige Samplemengen von offenen und geschlossenen Kurven wurden in [DMR00] behandelt. Alle bisher erwähnten Artikel funktionieren nur für glatte Kurven.

Giesen [Gie99] hat kürzlich das erste Ergebnis für Kurven gezeigt, die nicht glatt sind. Er betrachtete *zuträglich einseitig glatte* Kurven. Wir nennen eine (offene oder geschlossene) Kurve einseitig glatt, wenn in jedem Punkt die linke und rechte Tangente existiert; allerdings dürfen die beiden Tangenten verschieden sein. Eine einseitig glatte Kurve heißt zuträglich, wenn in jedem Punkt der Winkel zwischen rechter und linker Tangente weniger als π beträgt. Giesen zeigte, daß die Traveling Salesman Tour durch die Punkte V das Kurvenrekonstruktionsproblem für gleichmäßige Samplemengen von einzelnen geschlossenen zuträglich einseitig glatten Kurven löst. Genau genommen hat er gezeigt, daß für jede zuträglich einseitig glatte geschlossene Kurve γ ein positives ϵ existiert, so daß die Traveling Salesman Tour durch die Punkte V die polygonale Rekonstruktion ist, falls für jeden Punkt $x \in \gamma$ ein Punkt $p \in V$ existiert, mit $\|xp\| \leq \epsilon$, wobei $\|xy\|$ den Euklidischen Abstand zwischen x und y bezeichnet. Giesen's Ergebnis ist eine Existenzaussage; er bestimmt ϵ nicht in Abhängigkeit von Eigenschaften der Kurve γ. Wir erweitern Giesen's Ergebnis in mehrere Richtungen:

- Wir beziehen ϵ auf lokale Eigenschaften der Kurve und zeigen, daß die Traveling Salesman Tour das Kurvenrekonstruktionsproblem auch für ungleichmäßige Samplemengen löst. Für glatte Kurven ist unsere Bedingung an das Sample ähnlich zu denen, die in [ABE98, DK99, Gol99, DMR00] benutzt wurden.

- Wir zeigen, daß der Traveling Salesman Weg das Rekonstruktionsproblem für offene Kurven lösen kann. Wir betraten die Fälle mit und ohne vorgegebene Endpunkte.

- Wir zeigen, daß die kürzeste Traveling Salesman Tour bzw. der kürzeste Traveling Salesman Weg in polynomieller Zeit berechnet werden kann, wenn unsere Bedingung an die Samplemenge erfüllt ist.

- Wir geben einen einfacheren Beweis, daß die Traveling Salesman Tour bzw. der Traveling Salesman Weg das Kurvenrekonstruktionsproblem löst.

- Wir zeigen eine Erweiterung von dem Traveling Salesman basierten Algorithmus, der Mengen von zuträglich einseitig glatten geschlossenen Kurven für ungleichmäßige Samplemengen rekonstruieren kann.

Aus Effizienzgründen ist es wünschenswert, die Suche nach der Rekonstruktion auf einen dünnen Graphen über der Samplemenge zu suchen. Wir zeigen, daß die Kanten der Rekonstruktion im Delaunay Diagramm enthalten sind, wenn man die Bedingung an die Samplemenge etwas verschärft. Außerdem setzen wir unsere Bedingung an die Samplemenge in Beziehung zu der Bedingung, die in den meisten anderen Artikeln benutzt wurde. Zusätzlich vergleichen wir unser Ergebnis zu sogenannten *Necklace Tours*. Necklace Tours sind Instanzen des Traveling Salesman Problems, die in Polynomzeit gelöst werden können. Kurvenrekonstruktionsalgorithmen dürfen Kurven "erfinden", wenn die Samplebedingung nicht erfüllt ist. Dieser Aspekt wird auch betrachtet.

Wir haben einige Kurvenrekonstruktionsalgorithmen implementiert. Das JAVA-Applet `http://review.mpi-sb.mpg.de:81/Curve-Reconstruction/` macht diese Implementierungen verfügbar. Unsere Experimente zeigen, daß Traveling-Salesman basierte Algorithmen das Kurvenrekonstruktionsproblem für überraschend dünne Samplemengen lösen kann und daß die Laufzeit mit der anderer Algorithmen vergleichbar ist.

References

[AB98] N. Amenta and M. Bern. Surface reconstruction by voronoi filtering. In *Proceedings of the 14th ACM Symposium on Computational Geometry (SCG'98)*, 1998.

[ABE98] N. Amenta, M. Bern, and D. Eppstein. The crust and the β-skeleton: Combinatorial curve reconstruction. *Graphical Models and Image Processing*, pages 125–135, 1998.

[AGJ00] U. Adamy, J. Giesen, and M. John. New techniques for topologically correct surface reconstruction. In *IEEE Visualization 2000*, 2000. Accepted for publication.

[Att97] D. Attali. r-regular shape reconstruction from unorganized points. In *Proceedings of the 13th Annual ACM Symposium on Computational Geometry (SCG'97)*, pages 248–253, 1997.

[Bar] Dr. Gill Barequet's home page. `http://www.cs.technion.ac.il/~barequet/`.

[Bau97] P. Bauer. The circuit polytope: Facets. *Mathematics of Operations Research*, 22:110–145, 1997.

[BB97] F. Bernardini and C.L. Bajaj. Sampling and reconstructing manifolds using α-shapes. In *Proceedings of the 9th Canadian Conference on Computational Geometry (CCCG'97)*, pages 193–198, 1997.

[BDvD$^+$98] R. E. Burkard, V. G. Deineko, R. van Dal, J. A. A. van der Veen, and G. J. Woeginger. Well-solvable special cases of the traveling salesman problem: A survey. *SIAM Review*, 40:496–546, 1998.

[Ben90] J. L. Bentley. Experiments on traveling salesman heuristics. In *Proceedings of the 1st Symp. on Discrete Algorithms*, 1990.

[Blu67] H. Blum. A transformation for extracting new descriptors of shape. In *Models for the Perception of Speech and Visual Form*, pages 362–380. MIT Press, 1967.

[CCPS98] W.J. Cook, W.H. Cunningham, W.R. Pulleyblank, and A. Schrijver. *Combinatorial Optimization.* John Wiley & Sons, Inc, 1998.

[Chv83] Vasek Chvátal. *Linear programming.* W.H. Freeman and Co., 1983.

[DK99] T.K. Dey and P. Kumar. A simple provable algorithm for curve reconstruction. In *Proceedings of the 10th ACM-SIAM Symposium on Discrete Algorithms (SODA'99)*, pages 893–894, 1999.

[DMR00] T.K. Dey, K. Mehlhorn, and E.A. Ramos. Curve reconstruction: Connecting dots with good reason. *Computational Geometry: Theory and Applications*, 15(4):229–244, 2000. `www.mpi-sb.mpg.de/~mehlhorn/ftp/cure.ps.gz`.

[Dwy87] R.A. Dwyer. A faster divide-and-conquer algorithm for constructing Delaunay triangulations. *Algorithmica*, 2:137–151, 1987.

[Ede98] H. Edelsbrunner. Shape reconstruction with the Delaunay complex. In *LATIN'98: Theoretical Informatics*, volume 1380 of *Lecture Notes in Computer Science*, pages 119–132, 1998.

[EKS83] H. Edelsbrunner, D.G. Kirkpatrick, and R. Seidel. On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, 29(4):71–78, 1983.

[FG94] L.H. Figueiredo and J.M. Gomes. Computational morphology of curves. *Visual Computer*, 11:105–112, 1994.

[FKU77] H. Fuchs, Z. M. Kedem, and S. P. Uselton. Optimal surface reconstruction from planar contours. *Graphics and Image Processing*, 20(10):693–702, 1977.

[FR01] S. Funke and E.A. Ramos. Reconstructing collections of curves with corners and endpoints. In *Proceedings of the 8th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'01)*, 2001. Accepted for publication.

[GGJ78] M. Garey, R. Graham, and D. Johnson. Some NP-complete geometric problems. In *In Proceedings STOC-78*, pages 10–21, 1978.

[Gie99] J. Giesen. Curve reconstruction, the TSP, and Menger's theorem on length. In *Proceedings of the 15th Annual ACM Symposium on Computational Geometry (SCG'99)*, pages 207–216, 1999.

[Gol99] C. Gold. Crust and anti-crust: A one-step boundary and skeleton extraction algorithm. In *Proceedings of the 15th Annual ACM Symposium on Computational Geometry (SCG'99)*, pages 189–196, 1999.

[HSKK98] M. Hong, T.W. Sederberg, K.S. Klimaszewski, and K. Kaneda. Triangulation of branching contours using area minimization. *International Journal of Computational Geometry & Applications*, 8(4):389–406, 1998.

[ILO99] ILOG. *ILOG CPLEX 6.5 : user's manual.* ILOG, Bad Homburg, march 1999 edition, 1999.

[JT92] J. Jaromczyk and G. Toussaint. Relative neighborhood graphs and their relatives. In *Proc. IEEE, 80(9):1502–1517, 1992.*, 1992.

[Kep75] E. Keppel. Approximating complex surfaces by triangulation of contour lines. *IBM J. of Research and Development*, 19:2–11, 1975.

[KR85] D.G. Kirkpatrick and J.D. Radke. A framework for computational morphology. *Computational Geometry*, pages 217–248, 1985.

[LED] LEDA (Library of Efficient Data Types and Algorithms). `www.mpi-sb.mpg.de/LEDA/leda.html`.

[LLRKS92] Eugene L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, and D.B. Shmoys. *The traveling salesman problem : a guided tour of combinatorial optimization.* Wiley-interscience series in discrete mathematics and optimization. Wiley, / reprinted with subject index edition, 1990/1992.

[Lov79] L. Lovász. *Combinatorial problems and exercises.* North-Holland Publishing Co., Amsterdam, 1979.

[Meh84] K. Mehlhorn. *Data Structures and Algorithms 2: Graph Algorithms and NP-Completeness.* Springer, 1984.

[MN99] K. Mehlhorn and S. Näher. *The LEDA Platform for Combinatorial and Geometric Computing.* Cambridge University Press, 1999.

[NW88] George L. Nemhauser and Laurence A. Wolsey. *Integer and Combinatorial Optimization.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, New York;Chichester;Brisbane, 1988.

[O'R81] J. O'Rourke. Polyhedra of minimal area as 3D objects models. In *Proceedings of the Seventh IJCAI*, pages 664–666, 1981.

[Pap77] C. Papadimitriou. Euclidean TSP is NP-complete. *Theoretical Computer Science*, 4:237–244, 1977.

[Sch86] Alexander Schrijver. *Theory of linear and integer programming.* Wiley, repr. 94 edition, 1986.

[USG] USGS EROS Data Center. `http://edcdaac.usgs.gov/`.

[VHP] Visual Human Project. `http://www.nlm.nih.gov/research/visible/visible_human.html`.

[Wol98] Laurence A. Wolsey. *Integer programming.* Wiley-interscience series in discrete mathematics and optimization. Wiley & Sons, New York, 1998.

[Wun97] R. Wunderling. Paralleler und Objektorientierter Simplex–Algorithmus. Technical report, Konrad-Zuse-Zentrum für Informationstechnik Berlin, 1997.