Geometry and the Travelling Salesman Problem

by

Sándor P. Fekete

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 1992

Abstract

The objective of this Ph.D. thesis is to explore some connections between problems arising from Combinatorial Optimization and problems from Computational Geometry.

It deals mainly with geometric problems that are in some way related to the Travelling Salesman Problem. The first two parts, ("Area Optimization" and "Angle-Restricted Tours") discuss aspects of two problems of combinatorial geometry. The other two parts ("Geometric TSP Optimality" and "Error Analysis for the TSP") are more closely related to particular aspects of the TSP.

The "Minimum Area Polygon" problem (MAP) asks for a simple polygon with a given set of vertices for which the enclosed area attains the minimum. Pick's theorem provides a relation between the area of a simple polygon and the number of grid points it meets; this yields a simple combinatorial lower bound for the area of a polygon. Considering this bound leads to the "Grid Avoiding Polygon" problem (GAP), which asks for a simple polygon with a given set of (grid) vertices that does not meet any additional grid points. We prove that GAP is NP-complete, implying NP-completeness of MAP. This result answers a question stated by Suri in 1989. We also show that the respective maximization problems are NP-complete and closely related to MAP and GAP. We give upper and lower bounds on approximation factors for the "Maximum Area Polygon" problem. Finally, we show that it is NP-hard to minimize the volume of the k-dimensional faces of a simple d-dimensional polyhedron with a given set of vertices. This answers a generalization of a question of O'Rourke.

For a given set A of angles, the problem "Angle-Restricted Tour" (ART) is

to decide whether a finite set of points in the Euclidean plane allows a closed tour where subsequent points are connected by straight lines and adjacent lines may only subtend angles from the set A. For the angle-set $A = [0, \pi]$, we get "pseudoconvex tours". We prove that every set of at least 5 points allows a pseudoconvex tour. We discuss the complexity of detecting angle-restricted tours for other angle sets, most notably for $A \subseteq \{-\frac{\pi}{2}, \frac{\pi}{2}, \pi\}$.

In the third part, we discuss generalized convexity and its relevance for TSP optimality. It turns out that a generalized convex tour allows a particularly elegant geometric proof of optimality, called a *moat packing*, as introduced by Jünger and Pulleyblank.

Finally, we consider backward error analysis for the TSP. Geometrically speaking, we examine the question: How far is a given set of vertices from one that forms a generalized convex arrangement?

Acknowledgements

My advisor Bill Pulleyblank has been a great role model for me in more than one respect. Even at a distance of 500 miles, he sensed when I was in need of advice or encouragement, or when I could use some constructive criticism. He presented me with a considerable number of interesting problems, gave me the freedom to come up with my own questions, and took an active part in pursuing them. At the same time, he managed to teach me about many important aspects of being a professional mathematician, on details (e.g. *how* to give a mathematical talk) as well as on overall aspects (e.g. *where* to give a mathematical talk).

Anna Lubiw has always had an open door and an interested ear when I came to bother her about questions of computational geometry. On more than one occasion when I needed a well-founded opinion, she has led me in the right direction.

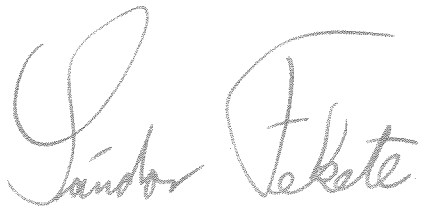
Bill Cunningham has supplied me with many useful and humorous comments and suggestions. Charlie Colbourn has pointed out some research and people from the field of Computer Science, which turned out to be very beneficial; his detailed advice on style and referencing will continue to be valuable. Like the other members of my examining committee, David Avis has carefully read my thesis and helped to improve it.

Gerhard Woeginger worked with me on the problem of Angle-Restricted Tours. Jit Bose directed my attention to Minimum Area Polygons. Joe O'Rourke asked me about the complexity of minimizing surface area in higher dimensions. Sven Schuierer and Derick Wood talked to me about aspects of generalized convexity and restricted orientations. Jon Rokne pointed out David Salesin's thesis. Heiko Knospe, Nora Sleumer, Naji Mouawad and Peter Yamamoto participated in several interesting discussions.

Thanks to Helga and Stefan for their patience during the intensive phase of my work and to Sharon, Nora and Maurizio for their hospitality during the end of my time at Waterloo.

I would like to thank everybody else in the Department of Combinatorics and Optimization for helping to make my stay successful and enjoyable. It's been fun!

I am grateful for financial support during the course of this research, provided by the German National Fellowship Foundation ("Studienstiftung des deutschen Volkes"), the Institute for Computer Research, the Department of Combinatorics and Optimization, and my advisor William R. Pulleyblank.



To my friends and my family

Contents

1 Introduction **1**

2 The Area of Simple Polygons **8**

2.1 Introduction 8

2.2 Pick's Theorem and Area Calculation 13

2.3 Special Cases 19

3 Minimal Area **23**

3.1 Basic Observations on Planar Cubic Digraphs 24

3.2 Embedding the Planar Cubic Digraph 26

3.3 Perturbing the Embedding 30

3.4 Replacing Nodes by Point Sets 31

3.5 A Hamiltonian Path Induces a Small Polygon 37

3.6 A Small Polygon Induces a Hamiltonian Path 40

3.6.1 Basic Observations on a Small Polygon 41

3.6.2 How the Hamiltonian Path Is Induced 42

3.6.3 Connections between Boxes 44

3.6.4 H is a Hamiltonian Path 49

3.7 Approximating MAP 50

4 Maximal Area **53**

4.1 GAXP and MAXP are NP-complete 53

4.2 A $\frac{1}{2}$-Approximation Algorithm for MAXP 55

4.3 A Negative Result on MAXP-Approximation 62

4.4 A Game-Theoretical Problem 65

5 Higher Dimensions **70**

5.1 Volume in Higher Dimensions 71

5.2 Minimizing Surface Area 74

5.3 Minimizing Face Volume 77

6 Pseudoconvex Tours **85**

6.1 Convexity and Pseudoconvexity 85

6.2 Setting up the Proof 88

6.3 Proving Theorem 6.1.2 102

6.4 Pseudoconvex Paths 114

7 Angle-Restrictions **116**

7.1 Introduction 116

7.2 Orthogonal Tours 119

7.3 Acute and Obtuse Angles 131

7.4 Other Questions 136

8 Geometric TSP Optimality **138**

8.1 Polyhedral Combinatorics and the Subtour Polytope 138

8.2 Geometric Interpretation and Moat Packings 141

8.3 Convexity 144

8.3.1 Definitions 145

8.3.2 Generalized Convexity 150

8.3.3 Tours and Curves 151

8.4 Convexity and Moat Packings 155

9 Error Analysis for the TSP **160**

9.1 Introduction 160

9.2 L_1 Distances 164

9.2.1 L_∞ Error Bounds 169

9.2.2 L_1 Error Bounds 173

9.3 Euclidean Distances 177

9.3.1 General Properties 178

9.3.2 Polygonal Error Bounds 185

9.3.3 Euclidean Error Bounds 193

A Rectilinear Planar Layouts **199**

Bibliography **204**

List of Figures

2.1 Pick's theorem 13

2.2 Covering a large square with parallelograms 14

2.3 A set of grid vertices with a Grid Avoiding Polygon and a Grid Approximating Polygon 18

2.4 A horizontally convex set of thickness 2 allows a Grid Avoiding Polygon 20

2.5 Two horizontally convex sets, with and without a Grid Avoiding Polygon 22

3.1 Making all edges mandatory or optional for both their vertices . . . 25

3.2 A planar cubic digraph D 27

3.3 How to connect the vertices to the edge segments 29

3.4 An embedding of D 30

3.5 Horizontal and vertical switch 32

3.6 The two types of link boxes 33

3.7 How to choose the boxes for an optional edge: Case 1 33

3.8 How to choose the boxes for an optional edge: Case 2 34

3.9 How to choose the boxes for an optional edge: Case 3 34

3.10 How to choose the boxes for an optional edge: Case 4 35

3.11 Odd 3-bends and 4-bends 35

3.12 Even 3-bends and 4-bends 36

3.13 The two terminal boxes 37

3.14 The point set $\overline{P}$ 38

3.15 A small polygon for $\overline{P}$ 39

3.16 How to represent an optional edge contained in H. 40

3.17 How to represent an optional edge not contained in H. 41

3.18 The circled points cannot be included in the polygon without enclosing additional grid points 43

3.19 The auxiliary graph C for the polygon of Figure 1.21 44

3.20 The Hamiltonian path H induced by C 45

3.21 A vertex of degree 3 in H and the corresponding situation in C . . 45

3.22 Three connections for a switch box 46

3.23 Adjacencies in a switch box with three connections 47

3.24 Two connections for a link box 48

3.25 Adjacencies in a link box with two connections 49

3.26 An example where "Greedy-Build" performs badly 52

4.1 GAXP solves GAP: Turning a small polygon inside out 54

4.2 The simple polygon P_1 57

4.3 The points $q_2^{(j+1)}$ and $q_{x_j-1}^{(j)}$ can see each other 58

4.4 The simple polygon P_2 59

4.5 An example for which the bound $\frac{1}{2}$ is tight 60

4.6 P_1 has area close to $\frac{1}{2}MAX(P)$ 61

4.7 P_2 has area close to $\frac{1}{2}MAX(P)$ 61

4.8 Finding a polygon with area at least $(\frac{2}{3}+\varepsilon)AR(conv(P))$ solves GAP 63

4.9 The circled point cannot see any whole edge of the surrounding polygon 67

4.10 $CAMP_k \leq \frac{k+1}{2k+1}$ 68

4.11 The strategy does not generalize if Q is not a triangle 69

5.1 Why Pick's theorem does not generalize easily to 3 dimensions . . . 72

5.2 Extending a planar tour into a cone in 3-space 75

6.1 A Euclidean TSP instance with an easy solution 86

6.2 The bad shape 87

6.3 Extending a pseudoconvex tour 88

6.4 A γ-segment 89

6.5 Using a γ-segment when $\varepsilon_1 \cup \varepsilon_2$ contains more than hull point . . . 91

6.6 Using a γ-segment when $\varepsilon_1 \cup \varepsilon_2$ contains only one hull point 92

6.7 How to solve Case 0 93

6.8 How to solve Case 1 94

6.9 How to solve Case 2 95

6.10 How to solve Case B.3 97

6.11 How to solve Case B.4 99

6.12 How to solve Case B.5+ 101

6.13 The setup for the induction step 102

6.14 Rearranging T in order to get a γ-segment 105

6.15 Potential γ-segments in T 106

6.16 A pseudoconvex tour for y right of $\overrightarrow{q_2q_4}$ 107

6.17 There must be a point $h_{i_{34}}$ left of $\overrightarrow{q_2q_3}$ and right of $\overrightarrow{q_4q_5}$ 108

6.18 $h_{i_{34}}$ lies left of $\overrightarrow{q_2q_4}$ and left of $\overrightarrow{q_3q_5}$ 109

6.19 $h_{i_{45}}$ lies right of $\overrightarrow{q_2q_4}$ 109

6.20 $h_{i_{45}}$ lies right of $\overrightarrow{q_1q_2}$ 110

6.21 $h_{i_{34}+1}$ lies right of $\overrightarrow{q_1q_2}$ 110

6.22 The final pseudoconvex tour 112

6.23 A pseudoconvex tour that encounters a vertex more than once . . . 113

6.24 Any point set has a pseudoconvex spanning path 115

7.1 Shifting vertical classes removes the dotted "crossovers" 120

7.2 Box type A: OUT1 and OUT2 leave on the same side. 123

7.3 Box type B: OUT1 and OUT2 leave on different sides. 124

7.4 A cubic digraph D 124

7.5 The representation for D 125

7.6 Finding strongly feasible A_3-tours 127

7.7 Merging subtours to find weakly feasible A_3-tours 129

7.8 A point set without obtuse spanning paths 132

7.9 Every finite set of points has an acute path. 133

7.10 Sets that do not allow acute tours 134

7.11 Some difficulties for acute tours 135

8.1 A moat 143

8.2 Every norm is defined by its unit ball 147

8.3 The flowershop metric 148

8.4 The Livermore metric 152

8.5 A d-convex tour requires more than just a d-convex support 154

8.6 Inserting the new moat and redirecting the old moats 156

8.7 The new moat packing is tight and feasible 158

9.1 An L_1-convex polygon 166

9.2 An L_1-convex point set does not necessarily have a unique shortest tour 168

9.3 An L_1-convex point set is equal to its surface 169

9.4 A point set with L_1-surface and L_∞ error bounds 170

9.5 Including an interior point p into P_{nw} 171

9.6 The lower bound is sufficient to achieve L_1-convexity 172

9.7 A polygonal metric L_P and a dual metric L_P 174

9.8 A point set with an L_P-convex polygon for the surface points and L_{P^*} error bounds 174

9.9 A point set with L_1 error bounds around the surface points 175

9.10 How far do the points in P have to be moved to be comparable? . . 176

9.11 Lemma 9.3.1 is not necessarily true if B is not convex. 180

9.12 The sets M_j for an arrangement of squares 182

9.13 Examining intersecting convex curves can be reduced to examining polygons 184

9.14 There are at most k nontrivial M_j 186

9.15 Picking vertices of P^* 187

9.16 The situation if there are no static vertices. 189

9.17 The situation if there are static vertices. 192

9.18 An arrangement of disks that forces a unique choice of intersection points 195

9.19 All touching points must be rational 196

9.20 Another arrangement with a unique solution 197

A.1 A planar digraph D^* 200

A.2 The dual of D^* with edge directions induced by the linear order O_F 201

A.3 A planar rectilinear layout for D^* 203

Chapter 1

Introduction

Mathematicians working in the area of Combinatorial Optimization are used to dealing with discrete objects like finite sets, permutations or graphs. This reflects the necessities of finding algorithmic solutions for optimization problems. For a problem that is not intrinsically discrete, it can be beneficial to transform it into a purely combinatorial task. Probably the best known example for this approach is the simplex method for Linear Programming: The problem of finding a point in a polyhedron that optimizes a linear function is reduced to considering a graph that is formed by vertices and edges of the polyhedron.

On the other hand, Geometry is the prototypical field dealing with continuous objects: subsets of the Euclidean plane, the motion of objects, manifolds. The classical methods of Euclidean Geometry are quite algorithmic, but the underlying philosophy is to treat lines or circles as continuous entities.

The study of geometry has also focussed on discrete problems. But until recent years, mathematicians interested in Discrete Geometry were usually more concerned

about nonalgorithmic extremal aspects than about questions of algorithmic optimization.

Since about the 1970s, the necessity to consider geometric questions in areas as computer graphics, image processing, robotics, pattern recognition, and architecture has led to increasing interest in discrete geometric algorithms. Design and analysis of geometric algorithms is the subject of the field of Computational Geometry.

The objective of this Ph.D. thesis is to explore some connections between questions arising from Combinatorial Optimization and problems from Computational Geometry.

One of the best known problems of combinatorial optimization is the Travelling Salesman Problem (TSP): For a given finite set of cities, find a closed tour that visits them all. We deal mainly with geometric problems that are in some way related to the Travelling Salesman Problem. Some of them ("Area Optimization", "Pseudo-convex Tours") are combinatorial problems in their own right, others ("Generalized Convexity, TSP Optimality, and Moat Packings", "Error Analysis for the TSP") are more closely related to particular aspects of the TSP.

Geometrically speaking, the Travelling Salesman Problem is to construct a (simple) polygon with a given set of vertices for which the *perimeter* is minimal. The "Minimum Area Polygon" problem asks for a simple polygon for which the enclosed *area* attains the minimum.

In Chapter 2, we discuss the calculation of the area of a simple polygon in the Euclidean plane. The relation between area of a simple polygon with integer

vertices and the set of enclosed grid points, is described by Pick's Theorem (1899):

$$\mathrm{AR}(\mathcal{P}) = \frac{1}{2}b(\mathcal{P}) + i(\mathcal{P}) - 1,$$

where $b(\mathcal{P})$ and $i(\mathcal{P})$ denote the number of grid points on the boundary and in the interior of $\mathcal{P}$.

When considering simple polygons with a given vertex set, Pick's Theorem provides a combinatorial bound on the maximum and minimum area simple polygons. We describe how we can optimize the enclosed area for some special cases of the given vertex set and point out some difficulties that arise in the case of arbitrary vertex sets.

It follows from Pick's Theorem that every simple polygon on a given set of n vertices with integer coordinates must have an area of at least $\frac{n}{2} - 1$. This lower bound is met if and only if we can find a *Grid Avoiding Polygon* that does not include any other grid points on its boundary or in its interior.

In Chapter 3, we prove that it is NP-complete to decide whether a given set of vertices allows a Grid Avoiding Polygon. (The reduction is from Hamiltonian Circuit in planar cubic directed graphs.) It follows that finding a Minimum Area Polygon is NP-hard. This answers a generalization of a question stated by Suri in 1989. Since it takes only a short time to calculate the area of a simple polygon and compare it to a given number, the problem is in NP, thus NP-complete. This differs from the situation for minimizing the perimeter, where it is still unknown whether the problem is in NP.

In Chapter 4, we extend our NP-completeness result for the maximization problem. We discuss questions of approximating a Maximum Area Polygon and show that we can find an approximation in time $O(n \log n)$ that comes within a factor

of $\frac{1}{2}$ of the optimum. We prove that coming within a constant factor of at least $\frac{2}{3} + \varepsilon$ of the area of the convex hull is NP-complete. We conclude the chapter by discussing how a vertex set P should be chosen inside of a convex region in order to minimize the maximal area of a simple polygon with vertex set P.

The question of minimizing perimeter or area of a simple polygon can be generalized to higher dimensions. In Chapter 5, we prove that it is NP-hard to find a simple d-dimensional polyhedron on a given set of vertices, for which the sum of volumes of its k-dimensional faces is minimal. This answers a generalization of a question of O'Rourke from 1980.

Any simple (not self-intersecting) Euclidean Travelling Salesman Tour that makes only right-hand turns is convex and therefore optimal. If we allow a tour to be self-intersecting while only making right-hand turns, we get the notion of a *Pseudoconvex Tour*. The example of three points surrounding a fourth point shows that not every finite set allows a pseudoconvex tour. In Chapter 6, we prove that every set of at least 5 points has a pseudoconvex tour.

The question of existence of pseudoconvex tours can be generalized as follows: For a given set A of angles, the problem "Angle-Restricted Tour" (ART) is to decide whether a finite set of points in the Euclidean plane allows a closed tour where subsequent points are connected by straight lines and adjacent lines may only subtend angles from the set A. This means that we are restricted in the way we can change direction relative to the orientation on a previous edge. As pointed out by Wood, our notion of restricted *relative* orientations is an interesting alternative to restricted orientations that are *fixed in space*.

In Chapter 7, we give complete solutions for all subcases of "Orthogonal Tours", where the feasible angle set is a subset of $\{+\frac{\pi}{2}, -\frac{\pi}{2}, \pi\}$: For the full set, the problem is NP-complete. (The reduction is from Hamiltonian Circuit in grid graphs.) For the subset $\{+\frac{\pi}{2}, -\frac{\pi}{2}\}$, the problem is polynomial with a complexity of $O(n^2 \log n)$. For the subset $\{+\frac{\pi}{2}, \pi\}$, the problem is NP-complete. (The reduction is from HC in cubic directed graphs.) We also present some results and open questions concerning "Acute Tours" (where all turns have to be at least as sharp as $\frac{\pi}{2}$) and "Obtuse Paths" (where all turns have to be not sharper than a specific lower bound).

Chapter 8 considers a geometric interpretation of the subtour constraints of the Travelling Salesman Problem and some related questions.

For a given set of vertices in the plane, a *moat* is a simply closed strip of constant width that separates two nonempty complementary subsets of the vertices. A *moat packing* is a collection of moats with pairwise disjoint interior. Any tour of the vertices has to cross any given moat at least twice, so twice the sum of moat widths in a moat packing is a lower bound for the length of a tour. This makes it interesting to find a *tight* moat packing for which this value is as large as possible. In some cases, we may be able to prove optimality of a tour by displaying a moat packing where the lower bound is met.

The question for an optimal moat packing can be formulated as the dual of the problem of optimizing over the subtour polytope. It can also be interpreted geometrically if the distances between points are given by shortest paths for some metric of $\mathbb{R}^2$.

The notion of a convex tour for the Euclidean TSP has been useful in several contexts. There are practical cases of geometric Travelling Salesman Problems

where the distances are not described by the Euclidean norm, but by some other metric d of $\mathbb{R}^2$. This makes it interesting to generalize convexity for these metrics. We can do this for a metric d by requiring any two points in a convex set to be connected by a d-shortest path that does not leave the set. In this sense, a d-convex tour allows us to connect any two tour vertices by a path that runs inside the region enclosed by the tour.

We conclude Chapter 8 by showing that every d-convex tour allows a tight packing of d-moats and is therefore optimal.

In numerical analysis, Forward Error Analysis deals with solutions of a problem that are obtained by imprecise calculations: How far is a calculated solution from the true solution for a given problem instance? Backward Error Analysis examines the situation where the imprecision in the input data is of greater concern: How far is a particular problem instance from an instance that has a known solution?

This idea of Backward Error Analysis can be applied to examine combinatorial optimization problems for which a true solution can only be found at great expense: Given an instance of a hard problem — how far is it from an instance for which we can find the true solution in short time?

In Chapter 9, we examine this notion of Backward Error Analysis for the TSP. The property that provides the optimum tour for a given set of tour vertices is generalized convexity. Geometrically speaking, we examine the question: How far is a given set of vertices from forming a convex arrangement?

There are two metrics involved in this question: The distance metric of the TSP defines the particular type of convexity that we have to consider. The second metric

describes the error bounds around the tour vertices, i.e. the distances between corresponding vertices in the given set and a set that forms a convex arrangement.

We consider several combinations of distance metrics and error metrics. We show that it is easy to solve the question for L_1 distances and L_∞ errors and also give a polynomial method for L_1 distances with L_1 errors. Both results can be generalized to polygonal norms. (A polygonal norm $L_\mathcal{P}$ has a centrally symmetric convex $2k$-gon as its unit ball.) For Euclidean distances, we show how it can be decided in polynomial time whether a given $L_\mathcal{P}$ error bound around each point is sufficient to transform a point set into a convex arrangement. This is closely related to the notion of *convex stabbing*: Does a given family of sets allow a convex curve that intersects them all?

The notion of convex stabbing was introduced by Tamir in 1987. Goodrich and Snoyeink have given a solution for the special case of a family of parallel line segments. Our result on $L_\mathcal{P}$ norms yields a polynomial solution for the case of a family of congruent convex polygons with a fixed number of vertices.

We conclude Chapter 9 by discussing difficulties arising from convex stabbing of disks.

Chapter 2

The Area of Simple Polygons

2.1 Introduction

The Travelling Salesman Problem is probably the best known of all optimization problems. It seems impossible to give an account of all the different aspects, variations and special cases that have been studied. (A good overview is provided by the book edited by Lawler, Lenstra, Rinooy Kan and Shmoys [54].)

From a geometrical point of view, the Euclidean Travelling Salesman Problem is to find a polygon with a given set of vertices that has shortest perimeter. Since it is not hard to see that an optimal polygon cannot be self-intersecting (an easy consequence of the triangle inequality), we may restrict our attention to simple polygons that have a given set of vertices.

It seems very natural to look for a simple polygon with a given set of vertices which minimizes another basic geometric measure: the enclosed area.

MINIMUM AREA POLYGON (MAP)

Given a finite set P of points in the Euclidean plane. Among all the simple polygons with vertex set P, find one with minimal enclosed area.

We write MAXP for the related problem of *maximizing* the enclosed area.

The problem MAP has gained some importance for questions related to pattern recognition, which are often concerned with simple polygons. At the first Canadian Conference on Computational Geometry in 1989, Suri asked for the complexity of MAP.

While there has been some research on extremal polygons on a given vertex set (see Boyce, Dobkin, Drysdale and Guibas [6] and Eppstein, Overmars, Rote and Woeginger [27]), the main attention has focussed on finding subpolygons with certain special properties, such as convexity. Typically, the results are fairly general and "area" is only one special case of a measure of simple polygons for which an algorithm works — "perimeter" usually being among the others.

In this chapter, we focus on the particularities of dealing with the area of polygons. We start with a number of basic observations which make it comparatively easy to calculate the area of a given simple polygon. It turns out that the area of a polygon with rational vertices can be calculated in linear time. This contrasts sharply to determining the perimeter of a given polygon (see Garey, Graham and Johnson [33]). Since there is no efficient way known comparing the sum of a finite number of square roots with a given rational, it is still an open problem whether the Euclidean Travelling Salesman Problem belongs to the class NP.

One particularly nice aspect of the area of a simple polygon is provided by *Pick's theorem.* It states that the area $AR(\mathcal{P})$ of any simple polygon $\mathcal{P}$ with integral vertices can be expressed by the number of grid points it encounters:

$$AR(\mathcal{P}) = b(\mathcal{P})/2 + i(\mathcal{P}) - 1,$$

where $b(\mathcal{P})$ is the number of grid points on the boundary of $\mathcal{P}$ and $i(\mathcal{P})$ is the number of grid points in the interior of $\mathcal{P}$. This means that we can think of optimizing the area of a simple polygon with a given set of vertices as a problem about grid points: How can we connect the given points by a simple polygon, such that we encounter and enclose as few or as many other grid points as possible?

If we are given n vertices, the best that we could do is avoid having *any* other additional grid point in the polygon. Since the given points have to be contained in the boundary of the polygon, we would get an area of $\frac{n}{2} - 1$ for such a GRID-AVOIDING POLYGON (GAP).

On the other hand, there are geometric aspects that make it seem harder to optimize the area enclosed by a polygon than to minimize its perimeter: We can think of a tour as a set of line segments and of a polygonal region as a set of triangles. But while the vertices of a short edge are necessarily at a close distance, the same need not be true for a triangle with a small area. The points in a triangle can be very far apart, while the enclosed area is arbitrarily small. These difficulties were mentioned by Boyce, Dobkin, Drysdale and Guibas in [6], where it was noted that "small perimeter implies that the vertices are well localized, but small area does not." Another difficulty described in [27] by Eppstein, Overmars, Rote and Woeginger was the problem of possible self-intersection: While it automatically does not occur in a polygon with minimal perimeter, it can be algorithmically

problematic to maintain simplicity when building up a polygon from triangles.

It may be surprising that until now, it has not been shown that it is NP-hard to find a minimum area (simple) polygon with a given set of vertices. This proof is a major subject of the following chapter. In fact, we prove that it is NP-complete to determine whether there exists a simple polygon with a given vertex set whose area equals the lower bound given by Pick's theorem that we described above.

From the described problems of dealing with the area of polygons, one has to expect that the technical details of such a proof require a lot of careful analysis, even when a construction with a simple underlying idea is found. Therefore it should not be too surprising that our NP-completeness proof for Minimum Area Polygons is long and technically involved.

We deal with area minimization and related questions in Chapters 2, 3, 4 and 5. The contents is organised as follows.

In this Chapter 2, we give basic definitions and a description of Pick's theorem. We show how the area of a simple polygon can be calculated and prove that MAP is in NP. We note that Pick's theorem yields very easy upper and lower bounds for the area of a simple polygon on a given vertex set. For the minimization problem, meeting this bound means finding a polygon that does not contain any other than the given grid points on its boundary or in its interior. For the maximization problem, it means finding a polygon that contains all grid points that lie in the convex hull of the given grid points. We describe some easy special cases where we can meet the lower bound and point out some difficulties for the general case.

Chapter 3 contains an NP-completeness proof for the problem GAP: Is there a

simple polygon that connects a given set of grid points and does not contain any other grid points on its boundary or in its interior? Since this question is a strong version of MAP, the result implies NP-completeness of MAP.

In Chapter 4, we shift our attention to the problem MAXP of *maximizing* the area of a simple polygon with given vertices. We show how a similar NP-completeness result of MAXP follows directly from the result for GAP. We give a very fast ($O(n \log n)$) heuristic for finding a simple polygon on a given set of points that encloses more than $\frac{1}{2}$ of the area of their convex hull, yielding a $\frac{1}{2}$-approximation method for MAXP. This result is best possible for any approach that uses the area of the convex hull as an upper bound for the optimal area. Since finding a simple polygon on n points takes $\Omega(n \log n)$ in the algebraic computation tree model, we cannot hope to find a faster method. (See Ben-Or [2] for an explanation of the model.) On the other hand, we show that it is NP-complete to decide whether there is a polygon of more than $\frac{2}{3}$ of the area of the convex hull. Finally, we discuss a game theoretic problem closely related to maximizing area of a polygon.

In Chapter 5, we consider related problems in higher dimensions. We discuss the calculation of the volume of a d-dimensional polyhedron and point at difficulties with generalizations of Pick's theorem. Then we generalize the geometric problems TSP and MAP for two dimensions to arbitrary dimensions: We show show that in any fixed dimensions k and d, finding a simple d-dimensional polyhedron with a given set of vertices that has minimal volume of its k-dimensional faces is NP-hard. This answers and generalizes a question stated by O'Rourke [66], [62].

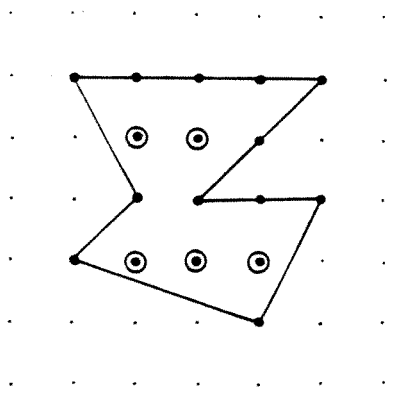


Figure 2.1: Pick's theorem

2.2 Pick's Theorem and Area Calculation

Let $\mathcal{P}$ be a polygon, given by a set of vertices and a set of edges. We call $\mathcal{P}$ *simple* if any vertex is only contained in two edges and nonadjacent edges do not intersect. Now consider a simple polygon $\mathcal{P}$ with grid points as vertices. What is its enclosed area $AR(\mathcal{P})$?

Perhaps one of the most surprising and elegant answers is provided by Pick's theorem:

Theorem 2.2.1 (Pick [71]) *Let $\mathcal{P}$ be a simple polygon with integer vertices; let $i(\mathcal{P})$ be the number of grid points contained in the interior of $\mathcal{P}$ and let $b(\mathcal{P})$ be the number of grid points on the boundary of $\mathcal{P}$. Then*

$$AR(\mathcal{P}) = \frac{1}{2}b(\mathcal{P}) + i(\mathcal{P}) - 1.$$

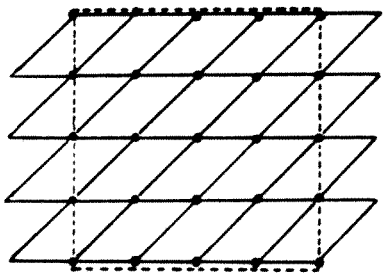


Figure 2.2: Covering a large square with parallelograms

An elegant proof can be found in Coxeter [16], pp. 208–209; we give a sketch of the idea:

Proof: It is straightforward to check that triangulating the polygon reduces the problem to showing the claim for triangles; the expression $\frac{1}{2}b(\mathcal{P}) + i(\mathcal{P}) - 1$ is additive for the two parts that are created by cutting the polygon along an inner chord.

Similarly, triangulating the set of grid points contained in the interior or boundary of a triangle reduces the problem to showing that every triangle with integer vertices that does not contain any other grid points has area $\frac{1}{2}$. We call such a triangle an *empty triangle.*

Consider an empty triangle with vertices $(0,0),(x_1,y_1),(x_2,y_2)$. Adding the vertex (x_1+x_2, y_1+y_2) yields a parallelogram. Because of symmetry, this has to be empty as well.

Now consider fitting together N such parallelograms, four at each vertex, to

cover a large square. (See Figure 2.2.) Every parallelogram is incident to four lattice points. Conversely, any lattice point that does not lie close to the periphery of the square is incident to four of the parallelograms. If the square is large enough, the peripheral error becomes negligible and the number of parallelograms needed is asymptotically close to the number of lattice points. Since this number is asymptotically close to the area of the region, the area of the parallelogram must be 1.

□

For a discussion of alternative proof methods see the article by Niven and Zuckermann [61]. There are numerous generalizations to other than the orthogonal grid, e.g. by Ren and Reay [82].

Pick's theorem can provide us with an easy way to see that the area of any simple polygon with rational vertices is rational: Simply multiply all of the coordinates by a suitably large integer I, such that all numbers become integer. This increases the area by a factor of I^2. Since the area of the enlarged polygon is rational by Pick's theorem, this must also be true for the original polygon. If the initial coordinates of the vertices are integer, Pick's theorem shows that the area must be a multiple of $\frac{1}{2}$.

Even though Pick's theorem provides a very elegant formula for calculating the area of a simple polygon, it is not very practical for actually calculating areas. An easy method is indicated by the following observation:

Lemma 2.2.2 *Let Δ be a triangle with vertices $(0,0),(x_1,y_1),(x_2,y_2)$.*
Then $AR(\Delta) = \frac{1}{2}(x_1y_2 - y_1x_2)$.

Proof: Adding the vertex $(x_1 + x_2, y_1 + y_2)$ yields a parallelogram with side vectors (x_1, y_1), (x_2, y_2). It is well known from Linear Algebra that the area of this parallelogram is equal to the determinant formed by the two vectors, resulting in a value of $x_1 y_2 - y_1 x_2$.

□

This makes it desirable to partition a polygon into a small number of triangles quickly. One easy way is to partition the polygon into triangles by a set of non-crossing interior chords between polygon vertices. This kind of partition is usually referred to as a *triangulation.* It is not very hard to see that a triangulation can always be found in polynomial time. A celebrated result by Chazelle [12] states that a triangulation of a simple polygon can even be found in linear time, so we can calculate the area of a simple polygon in time that is linear to the size of the input coordinates. As a corollary, it follows that a solution for MAP can be verified efficiently:

Corollary 2.2.3 MAP $\in$ NP.

Proof: As an input for MAP, consider a set of n vertices, each described by a pair of integers. A polygon can be described as a (cyclic) permutation of the vertices. Simplicity can be checked easily. (Chazelle's triangulation method yields a way to accomplish this in linear time.) Calculate the area of the polygon by determining the sum of triangle areas in a triangulation. The numbers involved cannot become too big, since we only have to consider a linear number of products of at most two coordinate numbers and a linear number of additions.

□

As already discussed in the introduction, Pick's theorem yields a combinatorial interpretation for minimizing or maximizing the area of a polygon. (See Figure 2.3.) Any grid point that is contained in the boundary contributes $\frac{1}{2}$ to the area of the polygon, any grid point in the interior contributes 1. The best we can do when minimizing the area is to avoid including any grid points other than the given n vertices, thus getting a polygon of area $\frac{n}{2} - 1$.

If we want to *maximize* the area, we have to include as many additional grid points as possible into the polygon, in a way that each of them contributes as much as possible. Since no grid point on the boundary of the convex hull of the given vertex set can be contained in the interior, they can at most contribute $\frac{1}{2}$. Any other grid point that is not given as a vertex contributes 1 when contained in the interior of the polygon.

We summarize these upper and lower bounds:

Theorem 2.2.4

Let P be a set of n points in the plane that all have integer coordinates. Let $h_i(P)$ denote the number of points of the integer grid that are not contained in P and strictly inside the convex hull, and $h_b(P)$ the number of grid points not in P that are on the boundary of the convex hull. Then for any simple polygon $\mathcal{P}$ on the vertex set P, we have

$$\frac{n}{2} - 1 \leq AR(\mathcal{P}) \leq \frac{n}{2} + \frac{h_b(P)}{2} + h_i(P) - 1.$$

□

These bounds suggest the following stronger versions of MAP and MAXP:

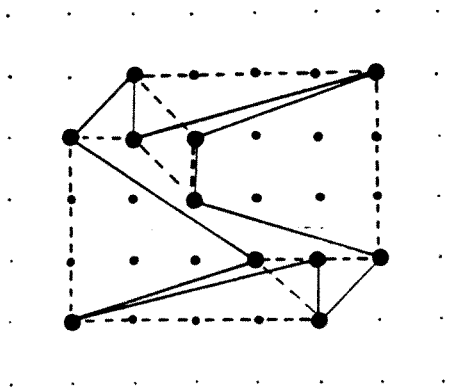


Figure 2.3: A set of grid vertices with a Grid Avoiding Polygon and a Grid Approximating Polygon

---

GRID AVOIDING POLYGON (GAP) *Given n grid points in the plane. Is there a simple polygon on this vertex set that does not contain any other grid points on its boundary or in its interior, and therefore has area* $\frac{n}{2} - 1$*?*

GRID APPROXIMATING POLYGON (GAXP) *Given n grid points in the plane. Is there a simple polygon on this vertex set that contains all grid points in the convex hull as well as possible: The grid points on the boundary of the convex hull lie on the boundary and all other grid points belonging to the convex hull lie in the interior? Such a polygon must have area* $\frac{n}{2} + \frac{h_b(P)}{2} + h_i(P) - 1$.

Even though these problems may seem combinatorially easier (all we have to do is avoid including integer points or making sure that they are all included in the polygon), we see in the following chapter that they are NP-complete, and hence as difficult as the original problem.

2.3 Special Cases

In this section, we discuss a few easy special cases for optimizing the area of a simple polygon and point out a source of difficulties in more general situations.

Corollary 2.3.1 *Given a* convex *set P of grid vertices, i.e. a set for which each grid point contained in the convex hull* $\text{conv}(P)$ *belongs to P. Then any simple polygon with vertex set P has area $\frac{n}{2} - 1$.*

In this situation, we do not have any choice about including additional grid points, so Pick's theorem guarantees an area of $\frac{n}{2} - 1$. A set P of grid vertices has $h_b(P) = h_i(P) = 0$ if and only if it forms a convex subset of the integer grid. Thus, Corollary 2.3.1 characterizes exactly the situation where lower and upper bound from Theorem 2.2.4 coincide.

Consider some "horizontal" direction, and the orthogonal "vertical" direction. (We do *not* require the "horizontal" direction to be parallel to an axis of the integer grid.) A set $M \subseteq \mathbb{R}^2$ is said to be *horizontally convex*, if any horizontal line intersects M in a connected set. We say that a set of grid points is horizontally convex, if no horizontal line segment between two points in the set contains a grid point that is not in the set. A horizontal line l *splits* a set of grid points P, if each of the open halfplanes bounded by l contains at least one point of P. We say that a horizontally convex set P of grid points has *thickness* t, if any line l that splits P and contains a point of the integer grid contains at least t points of P.

Theorem 2.3.2 *Given a set P of grid vertices. If P is horizontally convex and has thickness at least 2 for some horizontal orientation, then P has a Grid Avoiding Polygon.*

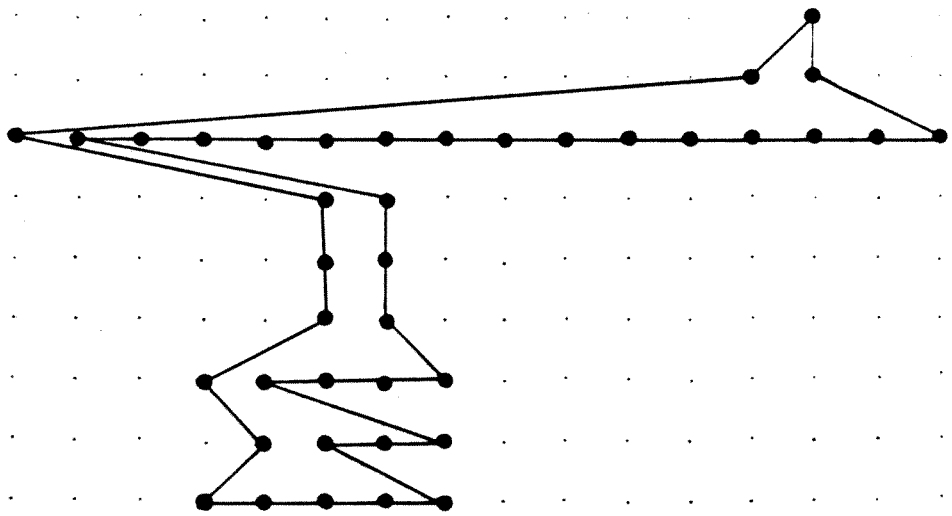


Figure 2.4: A horizontally convex set of thickness 2 allows a Grid Avoiding Polygon

---

Proof: See Figure 2.4. Consider the subsets of grid vertices that are contained in the same horizontal line. We can think of these subsets $S_1, \ldots, S_m$ as ordered by their vertical coordinates. The points in subset S_i can be ordered by their horizontal coordinates. Let p_{ij} denote the jth point in S_i and k_i the number of points in S_i. Then connect the points by the polygon

$$\langle p_{11}, \cdots, p_{m1}, \cdots, p_{mk_m}, p_{(m-1)2}, \cdots, p_{(m-1)k_{(m-1)}}, \cdots, p_{22}, \cdots, p_{2k_2}, p_{1m_1} \cdots, p_{11} \rangle.$$

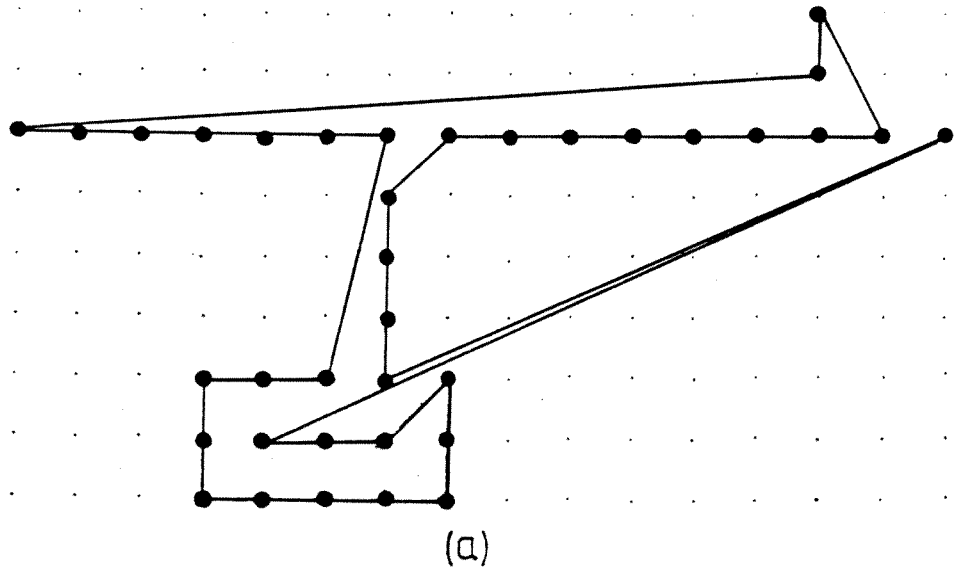
By construction, this polygon does not contain any other grid points.

□

The constructed polygon is monotone with respect to the vertical coordinates. (Conversely, the existence of a unimodal Grid Avoiding Polygon requires the vertex set to be horizontally convex.)

If we drop the thickness condition, matters can get difficult, as shown in Figure 2.5. Both point sets are horizontally convex. Careful analysis shows that P_2 does not allow a Grid Avoiding Polygon. The additional point in P_1 allows a narrow triangle of area $\frac{1}{2}$ that picks up a point far away, thereby "deblocking" the point set P_2 and allowing the polygon shown.

These "probes" make it very difficult to localize arrangements of points. We show in Chapter 3 how they can be controlled and used to show the NP-completeness of GAP.



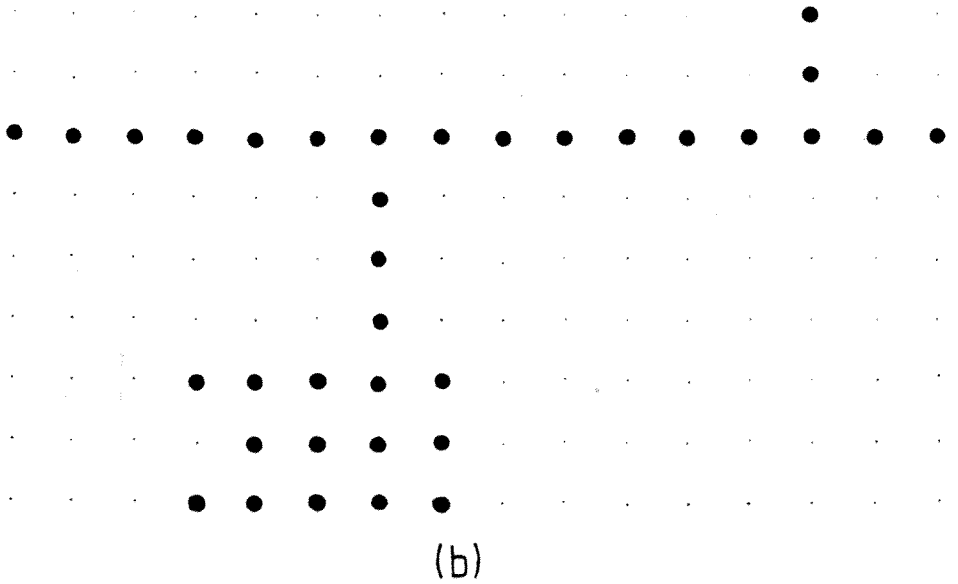


Figure 2.5: Two horizontally convex sets, with and without a Grid Avoiding Polygon

Chapter 3

Minimal Area

We have seen that calculating the enclosed area of a simple polygon is easier than calculating its perimeter — at least as long as nobody knows a simple way to compare sums of square roots. (See Garey, Graham and Johnson [33].) In this chapter we demonstrate, however, that determining a polygon with smallest area is as hard as determining the minimal perimeter:

Theorem 3.0.3 GAP *is NP-complete.*

We have already shown that the problem is contained in NP. To show that it is NP-hard, we give a reduction of HAMILTONIAN CYCLE IN PLANAR CUBIC DIRECTED GRAPHS. That is, for a given planar cubic digraph D, we construct a point set P_D in polynomial time, such that P_D admits a GRID AVOIDING POLYGON if and only if D has a Hamiltonian Cycle.

The idea for this construction is as follows:

After some minor rearrangements of the planar cubic directed graph, it is suitably embedded in the plane, such that all edges are rectilinear sets of line segments. Then the embedding is suitably scaled up and perturbed, to guarantee there are no collinearities between nonadjacent end points of line segments. Finally, these end points are replaced by suitable sets of grid points ("boxes"), such that a Hamiltonian path in the graph corresponds to a very narrow polygon that does not encounter any other grid points. Each set of points corresponding to an edge that is used in a Hamiltonian path is collected in one connected "branch" of the polygon, while the point sets corresponding to an edge that is not used by the path are split into two sets that are contained in two separate branches of the polygon. (This is somewhat similar to the idea contained in the NP-hardness proof for HAMILTONIAN CYCLE IN GRID GRAPHS described in Itai, Papadimitriou, Szwarcfiter [46], Johnson and Papadimitriou [47], and Papadimitriou and Vazirani [74].)

The layout of the points is chosen in a way that these branches can only be put together in a certain way without including any extra grid points.

As pointed out in the introduction, dealing with areas instead of distances makes it very hard to localize neighbours in a set of points. We achieve the desired localization by the perturbation mentioned above.

3.1 Basic Observations on Planar Cubic Digraphs

Consider any planar cubic digraph D. There are a few easy assumptions that we can make about the digraph D when we want to test it for Hamiltonicity: If D has a vertex with indegree 3 or outdegree 3, there can be no Hamiltonian Circuit, so all

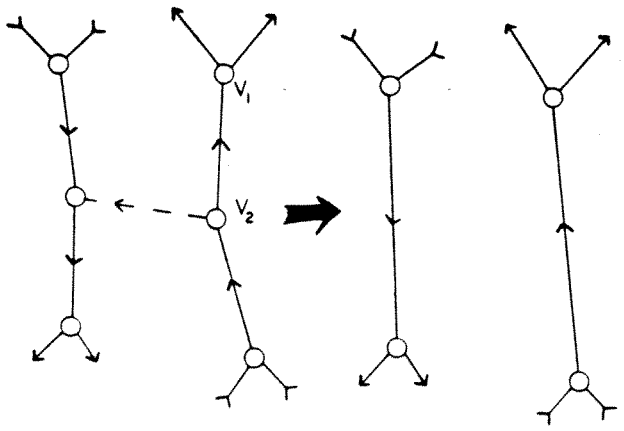


Figure 3.1: Making all edges mandatory or optional for both their vertices

vertices must have either indegree 2 or outdegree 2. Let the first type of vertices be called *in-vertices*, the second *out-vertices*. An edge is *mandatory* for a vertex, if and only if it is the only incoming or outgoing edge; otherwise it is *optional* for the vertex.

Assume there was an edge that was mandatory for one of its vertices v_1, but optional for its other vertex v_2. We could delete the other optional edge of v_2 without changing Hamiltonicity of the graph. The resulting vertices of degree 2 and their two adjacent edges can be replaced by a single edge. We can continue this process until all edges are either mandatory or optional for both their vertices. (See Figure 3.1.)

So we may assume that any edge is either mandatory or optional for both its vertices. This implies that D is bipartite, since any mandatory edge goes from an

in-vertex to an out-vertex and any optional edge goes from an out-vertex to an in-vertex. Furthermore, the optional edges of D form a set of vertex-disjoint cycles in the undirected graph G, obtained by replacing all arcs of D with edges. We denote by m the number of vertices of D. Then D has $\frac{3m}{2}$ edges.

3.2 Embedding the Planar Cubic Digraph

Now we show how to embed D in the plane, such that all edges are rectilinear paths, the optional edges have precisely one bend, the mandatory edges have three or five bends. Furthermore, all end points of line segments ("nodes") will have coordinates which are multiples of $\frac{1}{5}$ in the range between $\frac{3}{5}$ and m.

We start by identifying the cycles formed by the optional edges. As noted above, they are vertex-disjoint and even, so we can easily choose one (of two possible) perfect matchings in each cycle and contract all its edges, resulting in a 4-regular planar digraph $\overline{D}$ with $\frac{m}{2}$ vertices, m edges and $\frac{m}{2}+2$ faces. (See Figure 3.2; a graph $\overline{D}$ to D is shown in Figure A.1 in Appendix A.) As described in Appendix A, this digraph $\overline{D}$ can be embedded into the plane as a planar rectilinear layout with no two collinear line segments.

We modify the planar rectilinear layout to get an embedding of D where every optional edge has precisely one bend, every mandatory edge has three or five bends and all coordinates are multiples of $\frac{1}{5}$ in the range between $\frac{3}{5}$ and m.

Every vertex segment v has two edge segments adjacent to it that correspond to optional edges, one o_+ coming in and one o_- going out. Similarly, every vertex

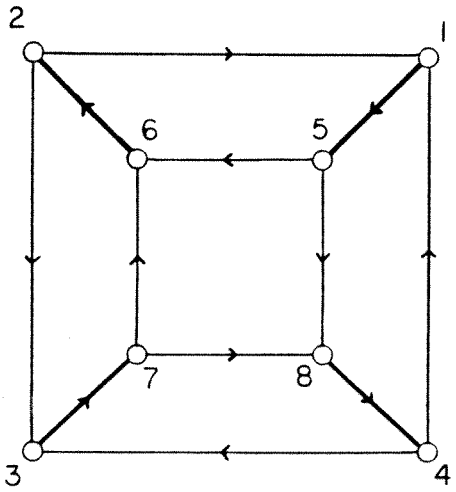


Figure 3.2: A planar cubic digraph D

segment is adjacent to one incoming mandatory vertex segment e_+ and one outgoing mandatory vertex segment e_-.

Let $\langle v_1, v_2 \rangle$ be the directed optional edge in D that was contracted to v. Because of planarity, the two edges o_+ and e_- adjacent to v_1 must follow each other when going clockwise around v; similarly, o_- and e_+ must follow each other.

Now place v_1 at $(\mathcal{O}_E(o_-) \pm \frac{1}{5}, \mathcal{O}_V(v))$. Place v_2 at $(\mathcal{O}_E(o_+), \mathcal{O}_V(v) \pm \frac{1}{5})$, such that v_1 lies between o_- and o_+ and v_2 lies on o_+.

The optional edge from v_1 to v_2 is led along the segments v and o_+, the optional edge adjacent to v_1 is connected to o_- along v and the optional edge adjacent to v_2 is led along o_+.

The mandatory edge coming into v_1 from e_+ is led from $(\mathcal{O}_E(e_+), \mathcal{O}_V(v) \pm \frac{1}{5})$

over $(\mathcal{O}_E(o_-) \pm \frac{1}{5}, \mathcal{O}_V(v) \pm \frac{1}{5})$ to $v_1 = (\mathcal{O}_E(o_-) \pm \frac{1}{5}, \mathcal{O}_V(v))$. ($e_+$ is lengthened or shortened appropriately.) Because of the aforementioned neighbourhood between e_+ and o_- when going clockwise around v, we do not get any interference with other line segments adjacent to v. Interference with other line segments is excluded by the way the new segments stay close to the original locations.

Finally, the mandatory edge going out from v_2 is led to e_- either with three bends:

$$\text{from } v_2 = \left(\mathcal{O}_E(o_+), \mathcal{O}_V(v) \pm \frac{1}{5}\right) \text{ over } \left(\mathcal{O}_E(o_+) \pm \frac{1}{5}, \mathcal{O}_V(v) \pm \frac{1}{5}\right)$$
$$\text{and } \left(\mathcal{O}_E(o_+) \pm \frac{1}{5}, \mathcal{O}_V(v) \mp \frac{2}{5}\right) \text{ to } \left(\mathcal{O}_E(e_-) \pm \frac{1}{5}, \mathcal{O}_V(v) \mp \frac{2}{5}\right).$$

or with one bend:

$$\text{from } v_2 = \left(\mathcal{O}_E(o_+), \mathcal{O}_V(v) \pm \frac{1}{5}\right) \text{ to } \left(\mathcal{O}_E(e_-) \pm \frac{1}{5}, \mathcal{O}_V(v) \pm \frac{1}{5}\right).$$

Again, e_- is lengthened or shortened appropriately and the neighbourhood of e_- and o_+ assures us that we do not get any interferences. For a display of some typical cases, see Figure 3.3.

This results in an embedding of D where every optional edge has precisely one bend, every mandatory edge has three or five bends and all coordinates are multiples of $\frac{1}{5}$ in the range between $\frac{3}{5}$ and m. (See Figure 3.4.)

We noted above that the mandatory edge $e_* = \langle v_a, v_z \rangle$ is led through the vertical line segment from $t := (m, t_y)$ for some t_y to $s := (m, s_y)$ for some s_y. If we replace the edge $\langle v_a, v_z \rangle$ by the two edges $\langle s, v_z \rangle$ and $\langle v_a, t \rangle$, we get a digraph D' that has a Hamiltonian path from s to t if and only if D has a Hamiltonian cycle. (For reasons that will become clear in Chapter 4 on Maximal Area, we make the position of t

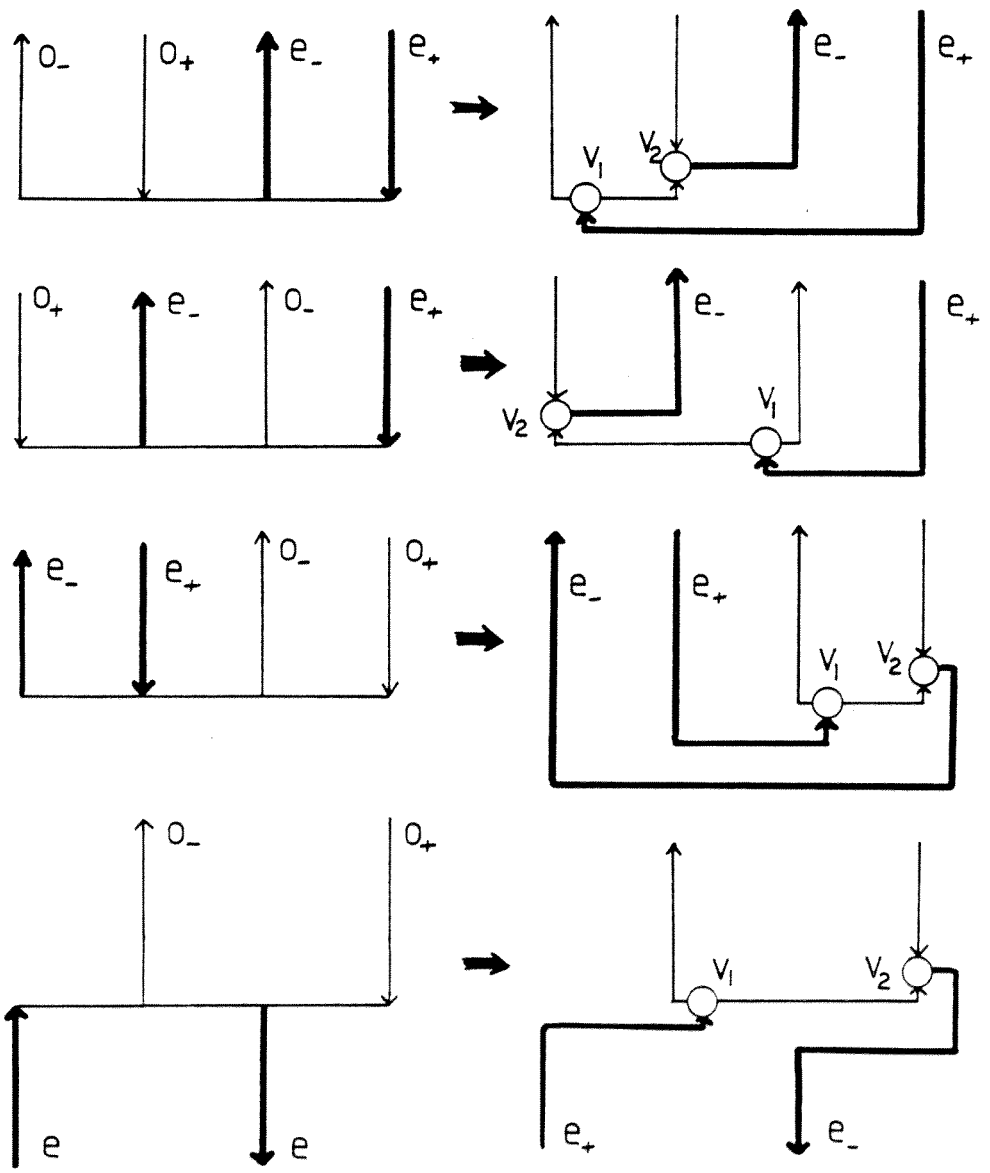


Figure 3.3: How to connect the vertices to the edge segments

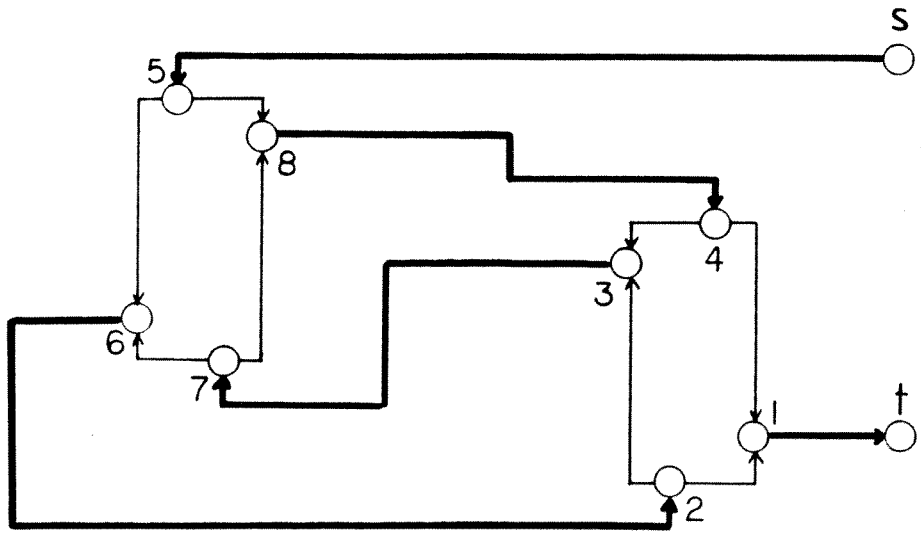


Figure 3.4: An embedding of D

extremal by moving t to $(m+1, t_y)$, thereby extending the adjacent line segment by 1.) We denote by G' the undirected version of D'.

3.3 Perturbing the Embedding

In the following, a *node* is a grid point in the embedding that separates two straight line segments. So a node either represents a vertex of D' or is a grid point where two line segments meet that belong to the representation of the same edge of D'. We let $\overline{m}$ denote the number of nodes. It is not hard to see that $\overline{m} \leq \frac{9m}{2}$.

Next, we perturb the embedding of D'. We start by multiplying all the distances in the embedding by a factor of $5N^3$, where $N := \overline{m}^{\overline{m}}$.

We think of x-coordinates as coordinates in the horizontal direction and y-coordinates as coordinates in the vertical direction. In the next step, partition the nodes into *vertical* classes. Two nodes belong to the same vertical class if they have identical y-coordinates and if they are connected by a path of at most two line segments of this y-coordinate. So each class consists of two or three nodes, where three nodes occur in the case of an out-vertex, with the two adjacent nodes representing the bends of the two adjacent optional edges. This implies that there are less than $\frac{\overline{m}}{2}$ vertical classes.

Shift all points in the ith verticalclass by the vector $(0, \overline{m}^i)$. Similarly, define horizontal classes and shift the points in the ith horizontal class by the vector $(N\overline{m}^i, 0)$.

3.4 Replacing Nodes by Point Sets

Finally, we replace each node by an appropriate set of points that form a connected subset of the integer grid — called a *box*.

Every degree 3 node is replaced by a *switch* box — see Figure 3.5. The circled point denotes the location of the node, the arrows indicate the edges adjacent to the node. Dotted locations correspond to points that may or may not be in the box, depending on the adjacent optional edges — we discuss this below. Out-vertices (represented by nodes with both optional edges leaving horizontally) are replaced by a horizontal switch; in-vertices (with optional edges entering vertically) by a vertical switch.

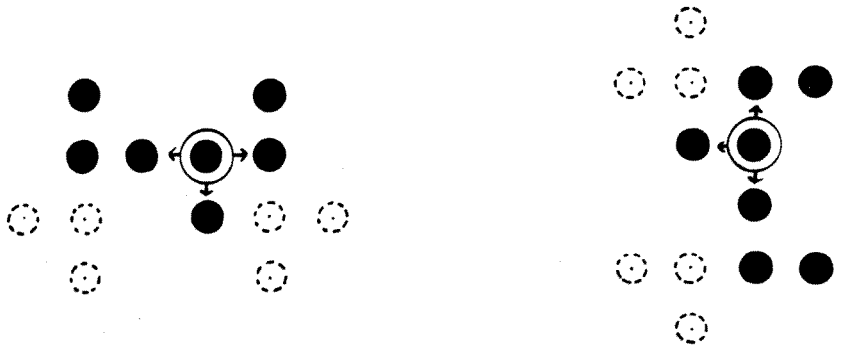


Figure 3.5: Horizontal and vertical switch

Every node representing a bend in an optional edge is replaced by one of the *link* boxes shown in Figure 3.6. The choice of link box depends on the configuration of the optional edge and the two adjacent mandatory edges. There are four different cases — shown in Figures 3.7- 3.10. Depending on the case, we also have to add three points to one of the adjacent boxes — again see Figures 3.7- 3.10.

We use *bend* boxes to represent the nodes at bends of the mandatory edges. Again, they are positioned such that the circled point is placed on the node and the two line segments for the mandatory edge run as indicated.

Running through any mandatory edge e from its start to its end vertex, we encounter a sequence of 3 or 5 nodes. The first is odd, the second is even, etc. The parity of the bend boxes is chosen accordingly. The type is chosen in the following way: If at its first bend, e makes a right-hand turn, we place a 3-bend; if it is a left-hand turn, we place a 4-bend. Again, the circled position p is placed on the node and the points p_1 and p_2 indicate the direction of the mandatory edge. On

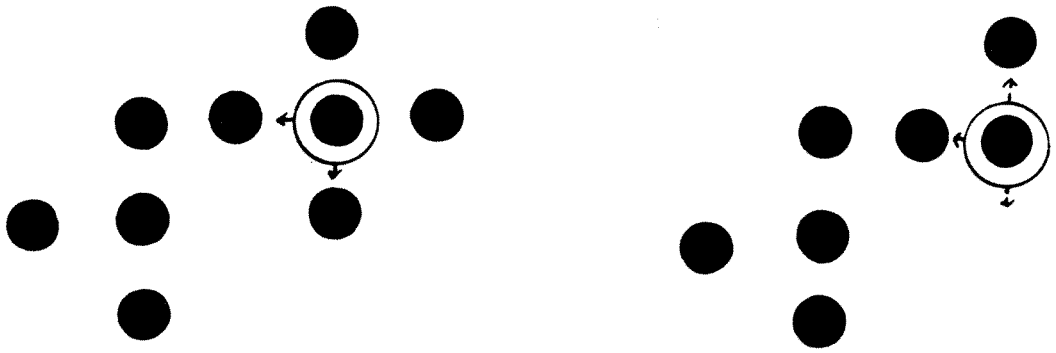


Figure 3.6: The two types of link boxes

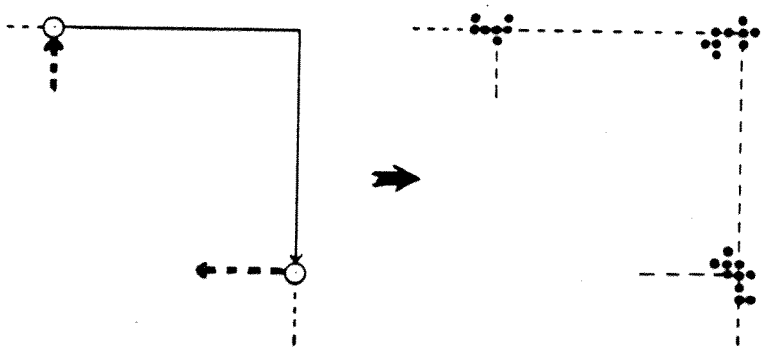


Figure 3.7: How to choose the boxes for an optional edge: Case 1

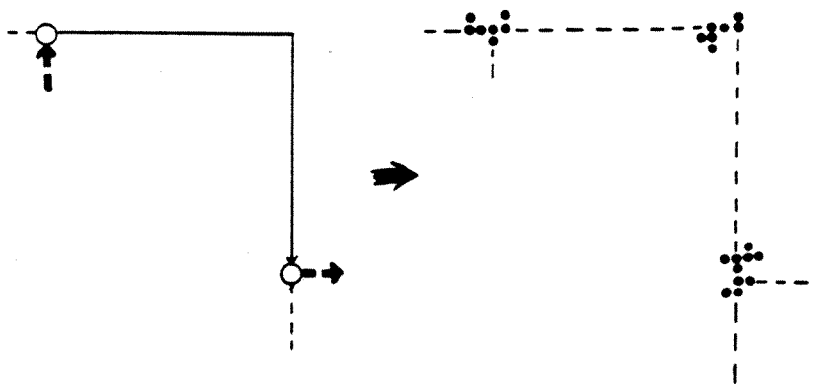


Figure 3.8: How to choose the boxes for an optional edge: Case 2

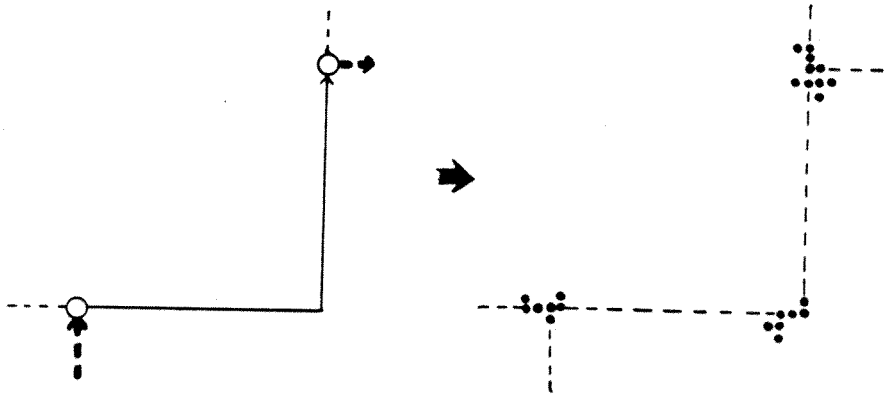


Figure 3.9: How to choose the boxes for an optional edge: Case 3

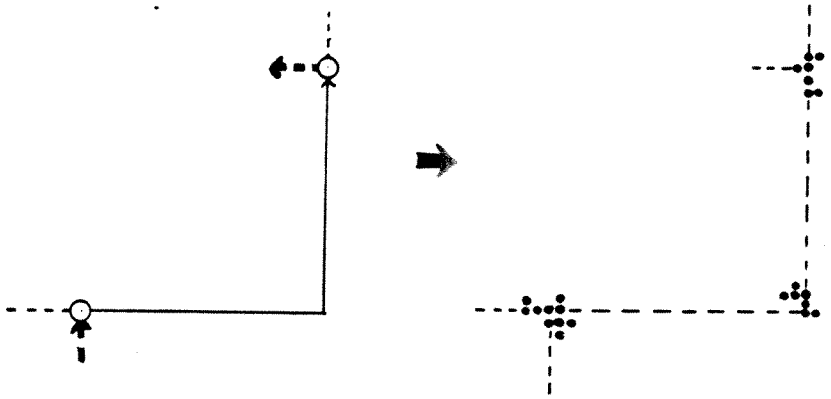


Figure 3.10: How to choose the boxes for an optional edge: Case 4

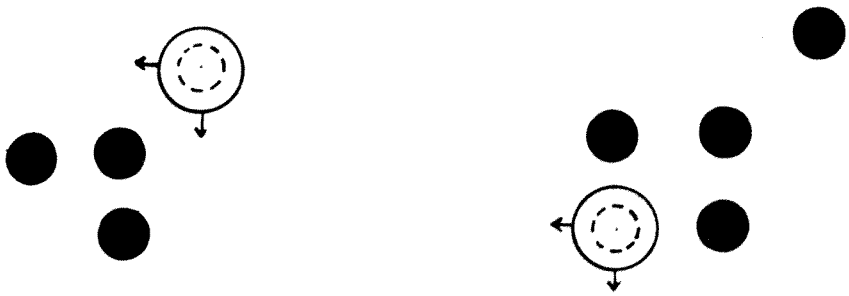


Figure 3.11: Odd 3-bends and 4-bends

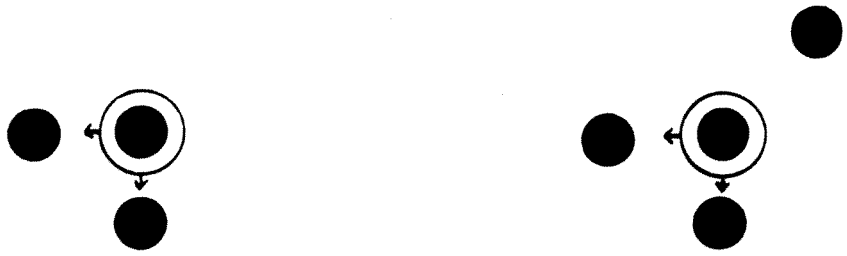


Figure 3.12: Even 3-bends and 4-bends

---

any following bend we use the opposite bend type if the edge makes the same type of turn and the same bend type if the edge makes the opposite type of turn. This implies that the points of an odd bend box are always to the right of the mandatory edge. (See Figure 3.14 for the overall situation.)

The nodes corresponding to s and t are replaced by *terminal boxes* as shown in Figure 3.13. (With respect to later usage in the following chapter on area maximization, we make sure that the points t_1 and t_2 of the box for t are extremal in horizontal direction. t can be moved horizontally as far to the right as necessary.) The grid points in all these boxes form a point set $\overline{P}$. A straightforward estimate yields $n := \mid \overline{P} \mid \leq 13\overline{m}$.

We claim: *The point set $\overline{P}$ is the vertex set of a simple polygon of area $\frac{n}{2} - 1$* if and only if *D has a Hamiltonian cycle.*

Figure 3.13: The two terminal boxes

3.5 A Hamiltonian Path Induces a Small Polygon

First we show that any Hamiltonian path H in D' induces a polygon of area $\frac{n}{2} - 1$:

Any mandatory edge in D' is represented by a partial polygon that connects all its bends and the two switches representing its end vertices.

Any optional edge in D' that is contained in the Hamiltonian path is represented by one partial polygon that connects its link and the two switches representing its end vertices — see Figure 3.16. (The cases 3 and 4 are similar to case 2.)

Any optional edge in D' that is *not* contained in the Hamiltonian path is represented by two partial polygons that each connect a part of the link to one of the two switches representing its end vertices — see Figure 3.17.

Finally, the two end boxes for s and t are connected to the adjacent bends in the obvious manner. Clearly, this results in a simple polygon that does not contain any other grid points, so the claim follows from Pick's theorem.

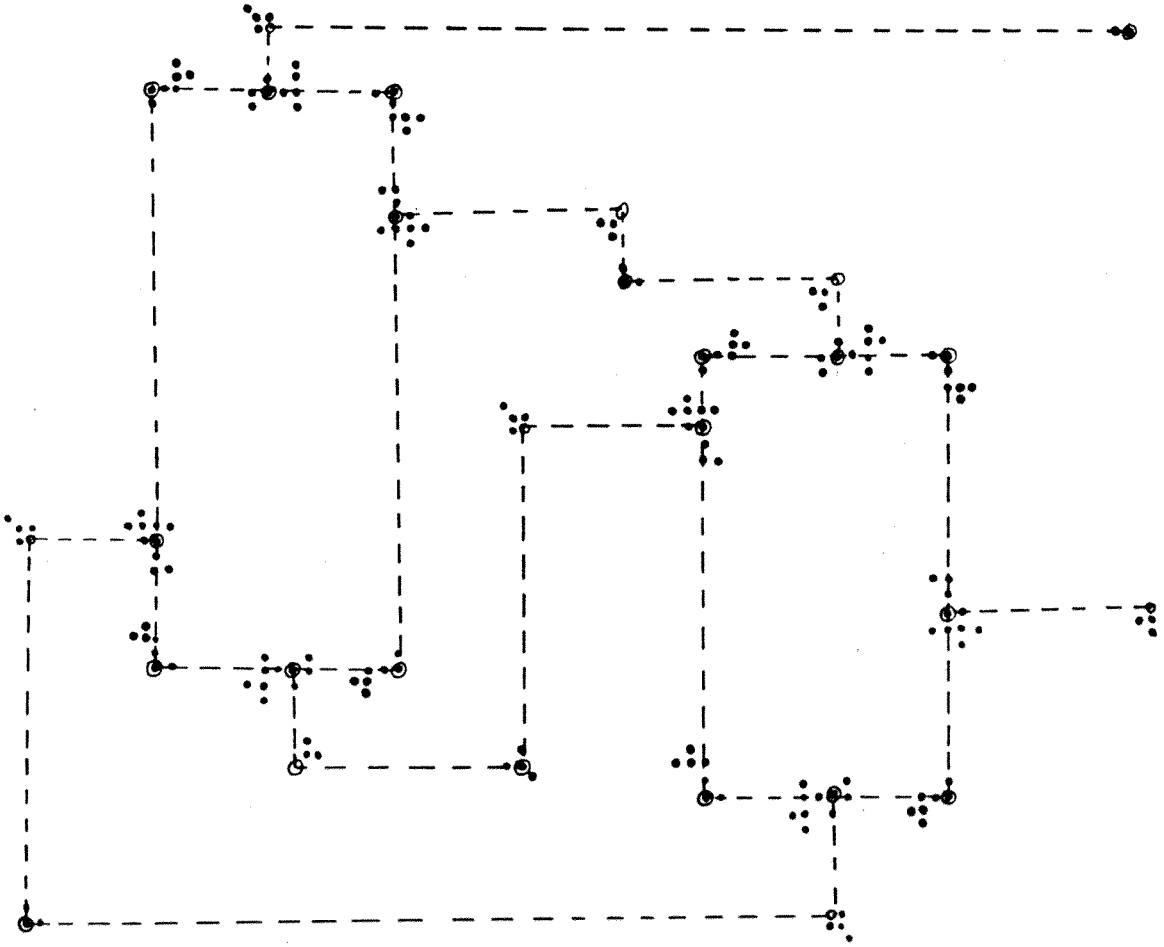


Figure 3.14: The point set $\overline{P}$

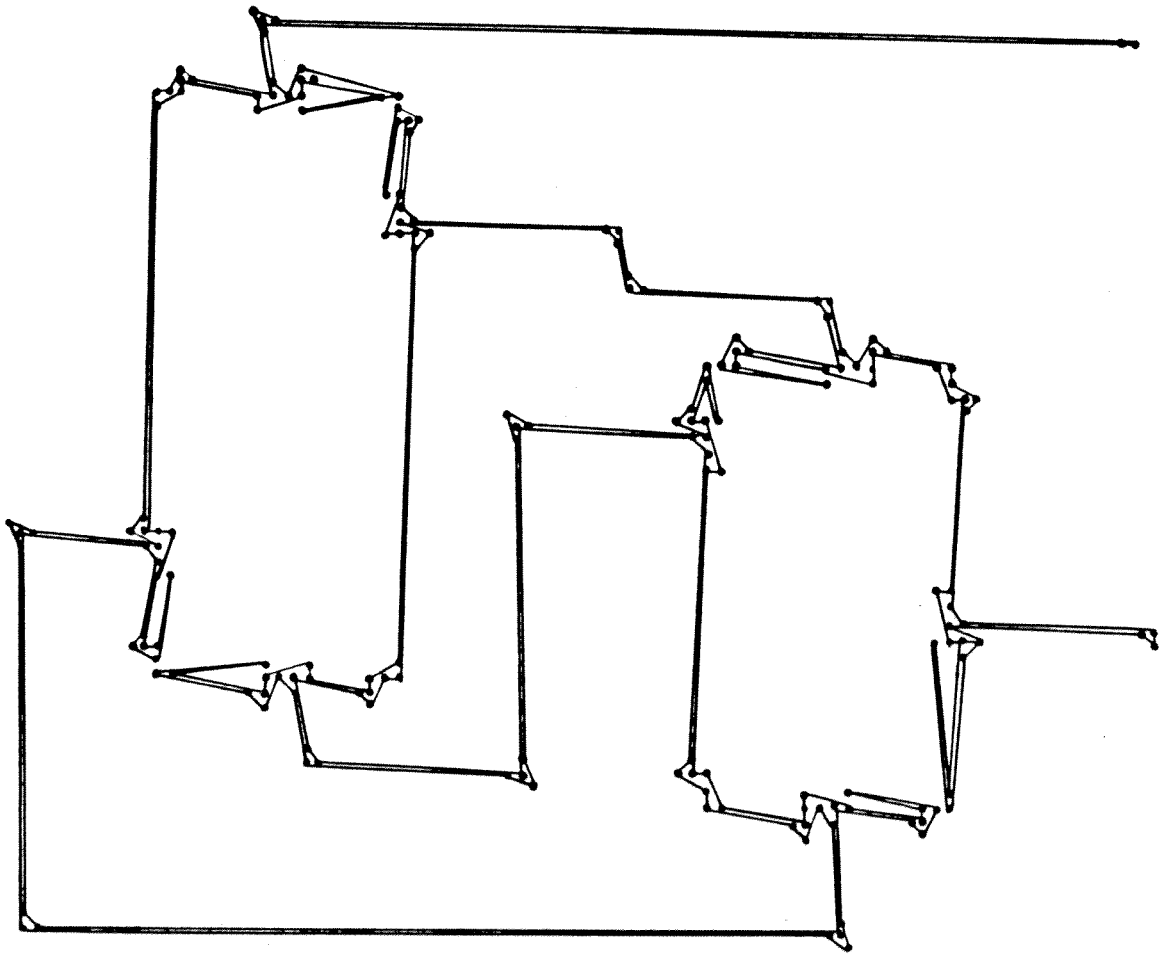


Figure 3.15: A small polygon for $\overline{P}$

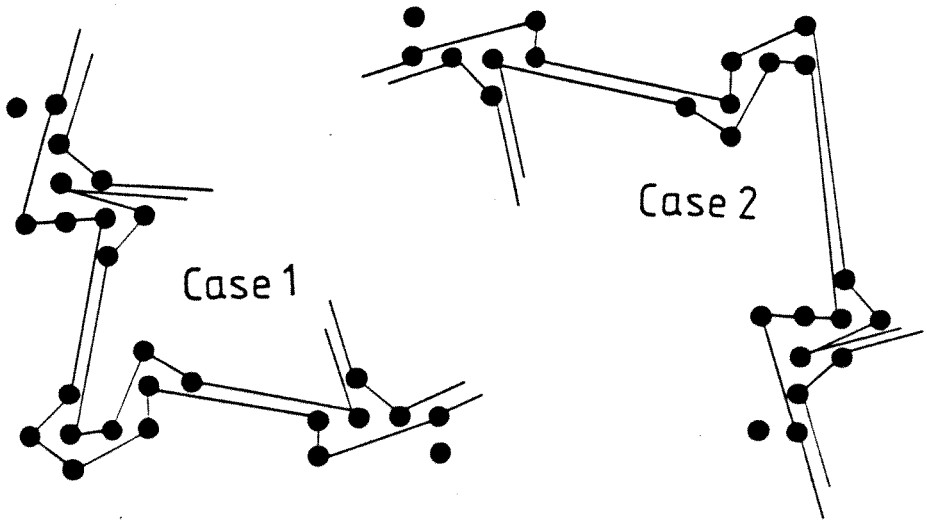


Figure 3.16: How to represent an optional edge contained in H.

3.6 A Small Polygon Induces a Hamiltonian Path

We show that any simple polygon on $\overline{P}$ of area $\frac{n}{2} - 1$ induces a Hamiltonian path in D'. Considering a triangulation of such a polygon $\overline{\mathcal{P}}$, it turns out that the perturbation performed on the points restricts us to triangles that connect points from adjacent boxes. Careful analysis of possible connections by triangles in the polygon ultimately shows that they form a particular structure that can be used to construct a Hamiltonian Path.

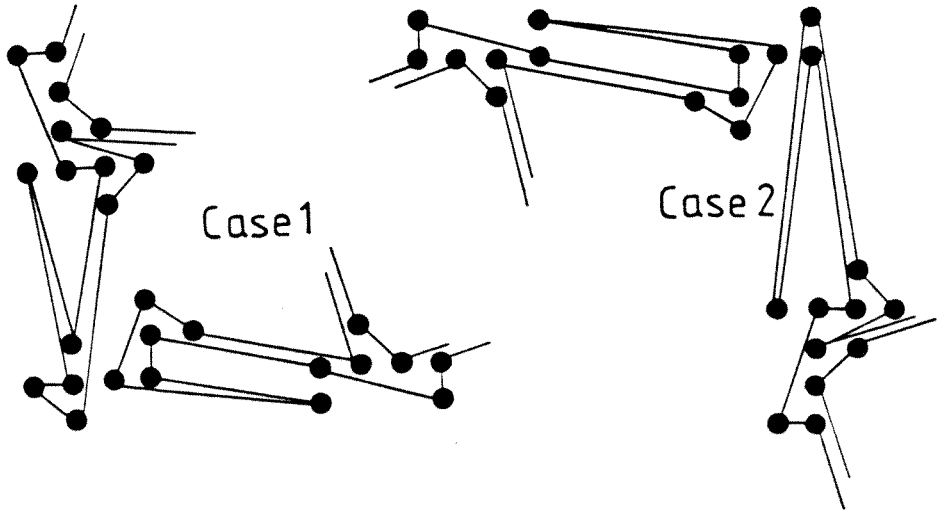


Figure 3.17: How to represent an optional edge not contained in H.

3.6.1 Basic Observations on a Small Polygon

Assume that we have a small polygon $\overline{\mathcal{P}}$ for the n vertices in the set $\overline{P}$ such that $AR(\overline{\mathcal{P}}) = \frac{n}{2} - 1$. A triangulation partitions $\overline{\mathcal{P}}$ into $n-2$ triangles, each having area at least $\frac{1}{2}$, so each of them must have area exactly $\frac{1}{2}$. For the rest of this section, let Δ_i be a triangle in some triangulation of $\overline{\mathcal{P}}$.

Because of the perturbation that we performed, these Δ_i are quite restricted:

Lemma 3.6.1 *Any triangle $\Delta_i = (p, q, r)$ with $AR(\Delta) = \frac{1}{2}$ must have p, q and r from the same horizontal or vertical class. Furthermore, two of the points must be from the same box.*

Proof: By construction, we may assume that

$$p = (0,0) + (Nn^{i_1}, n^{i_2}),$$

$$q = (a_1N^3, a_2N^3) + (Nn^{j_1}, n^{j_2}),$$

$$r = (b_1N^3, b_2N^3) + (Nn^{k_1}, n^{k_2}).$$

Now

$$AR(\Delta) = \frac{1}{2}\left((q_1 - p_1)(r_2 - p_2) - (r_1 - p_1)(q_2 - p_2)\right).$$

Note that $-n < a_1, a_2, b_1, b_2 < n$ and $1 \leq i_1, i_2, j_1, j_2, k_1, k_2 < \frac{n}{2}$. If neither $i_1 = j_1 = k_1$ nor $i_2 = j_2 = k_2$, the assumption $AR(\Delta_i) = \frac{1}{2}$ easily leads to a contradiction.

Now it is not hard to see that p, q and r cannot be from three different boxes.

□

We say that two boxes are *connected*, if $\overline{\mathcal{P}}$ contains a quadrangle Q of area 1, such that the points in each box contain exactly two of the vertices of Q. Because of Lemma 3.6.1, two boxes can only be connected if they are in the same horizontal or vertical class. Moreover, connecting two *links* of the same class leaves a set of points of the *switch* between them that cannot be included in the polygon without enclosing additional grid points — see Figure 3.19. This implies that only switches can be connected to more than two other boxes.

3.6.2 How the Hamiltonian Path Is Induced

We define the (undirected) auxiliary graph C:

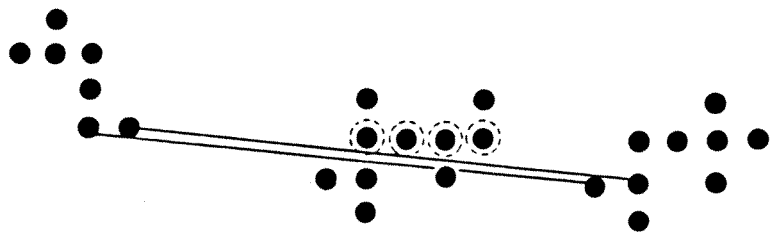


Figure 3.18: The circled points cannot be included in the polygon without enclosing additional grid points

---

Definition 3.6.2 *Every switch box B of $\overline{P}$ with neighbour boxes B_1, B_2 and B_3 is represented by three vertices $v(B, B_1)$, $v(B, B_2)$ and $v(B, B_3)$. Similarly, every link or bend box B with neighbour boxes B_1 and B_2 is represented by two vertices $v(B, B_1)$ and $v(B, B_2)$. Finally, every terminal box B with neighbour box B_1 is represented by one vertex $v(B, B_1)$.*

There is an edge between two corresponding vertices $v(B, B_i)$ and $v(B_i, B)$ of different boxes B and B_i if and only if B and B_i are connected. There is an edge between two vertices $v(B, B_1)$ and $v(B, B_2)$ of the same box B if and only if B is connected to both B_1 and B_2 and there is a path of edges or interior chords of the polygon $\overline{P}$ from B_1 to B_2 that visits only points of B.

For all non-terminal boxes B, delete any vertex $v(B, B_j)$ that is not connected to any of the other vertices $v(B, B_i)$ representing the same box B. The resulting subgraph of C is C^*. C^* defines a subgraph H of D' as follows:

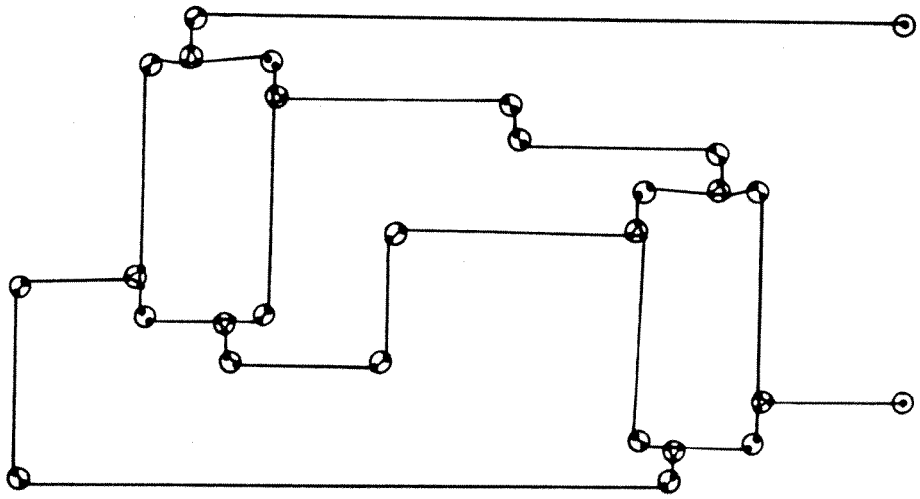


Figure 3.19: The auxiliary graph C for the polygon of Figure 1.21

An edge between two adjacent vertices v_1 and v_2 in D' is included in H, if and only if there is a path in C^* between vertices of the switch boxes for v_1 and v_2 that contains only vertices of non-switch boxes.

The remainder of this section shows that H is a Hamiltonian path in D'.

3.6.3 Connections between Boxes

Assume H has a vertex of degree 3.

This implies that there is a switch B that is connected to all three adjacent boxes B_1, B_2 and B_3 such that for each of the adjacent B_i, there is another $B_j \neq B_i$,

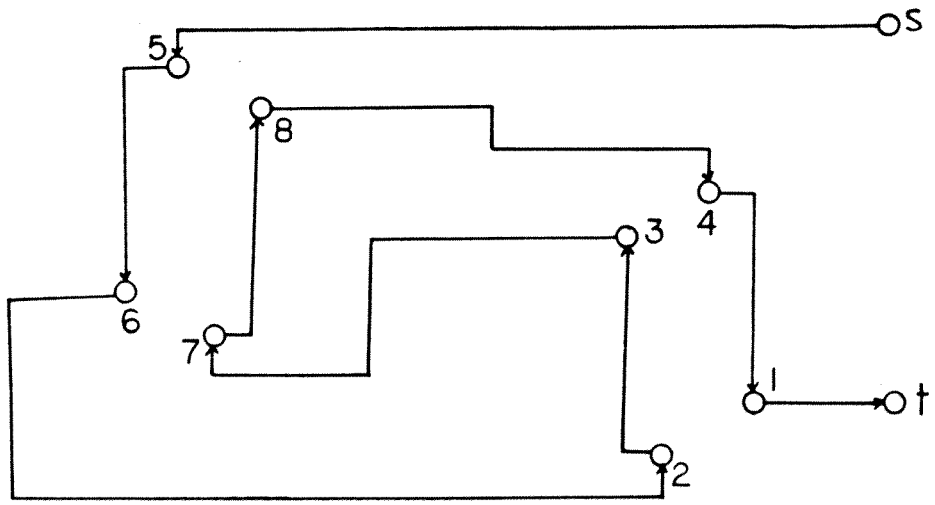


Figure 3.20: The Hamiltonian path H induced by C

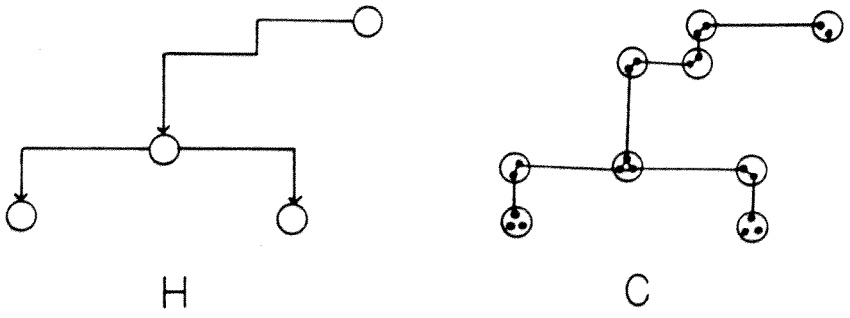


Figure 3.21: A vertex of degree 3 in H and the corresponding situation in C

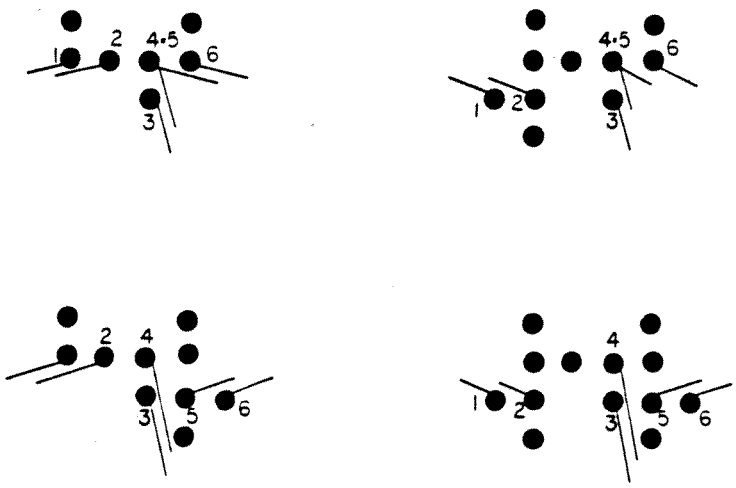


Figure 3.22: Three connections for a switch box

---

such that there is a path of edges or interior chords of $\overline{\mathcal{P}}$ connecting B_i and B_j that visits only vertices of B. Furthermore, for each of the adjacent links B_i, both vertices $v(B_i, B)$ and $v(B_i, \overline{B})$ in C must have degree 2. This means that there is a path of edges or interior chords of $\overline{\mathcal{P}}$ connecting the switches B and $\overline{B}$ that visits only vertices of B_i.

First consider the adjacencies in the switch. One easily sees that there must be edges that essentially look like those in Figure 3.22. (The four cases correspond to the different situations arising from the addition of points as in Figures 3.7, 3.8, 3.9 and 3.10: A switch can have in addition (1) no extra points, (2) three extra points on the left, (3) three extra points on the right, or (4) three extra points on the left and three extra points on the right.)

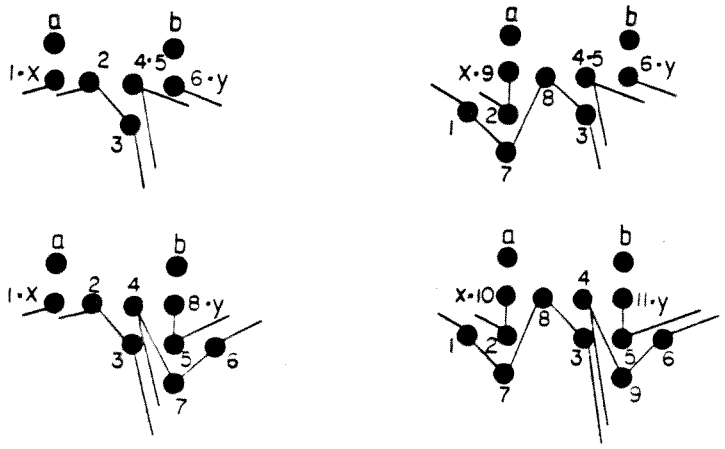


Figure 3.23: Adjacencies in a switch box with three connections

---

Representatively, we discuss case (4). (The other cases are similar — see Figure 3.23). All conclusions result from the fact that no grid points other than those of boxes may be contained in the polygon and the condition that the vertices 1 and 2, 3 and 4, 5 and 6 may not be connected by a set of edges within the box. The point 1 can not be connected to 2, so it must be connected to 7. The only other possible connection for 7 is 8, leaving only 10 for 2. 3 must be connected to 8. The only connection for 6 is 9, then 9 must be connected to 4. It follows that 5 is adjacent to 11.

In all cases, we are left with the points x, y, a and b. It is not hard to see that at least one of the *pivot points* a and b has to be adjacent to two points from a neighbouring link.

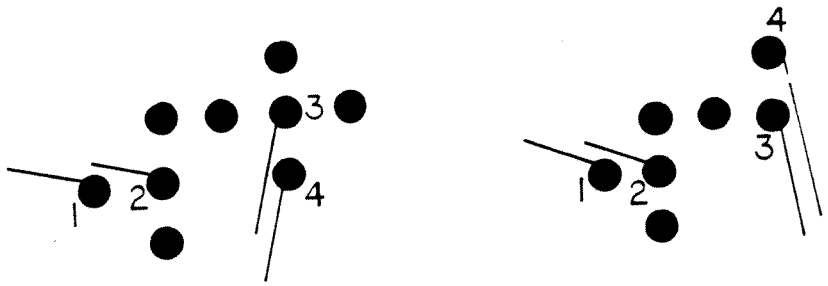


Figure 3.24: Two connections for a link box

Now consider the adjacencies in the links next to the switch. Representatively, we discuss the more complicated case: First we remark that the connection from below cannot be adjacent to 9 instead of 3 or 4. The point 1 can not be connected to 2, so it must be connected to 5. The only other possible connection for 5 is 6, leaving only 7 for 2. 4 must be connected to 8 and 8 to 9. Finally, 3 must be connected to 6 and 7 to 9.

At no stage would it be possible to have adjacencies to points outside of the box without including extra grid points in $\overline{\mathcal{P}}$. So none of the pivot points can be included in a connected link with two connections. This contradicts our assumption of having a degree 3 node in H.

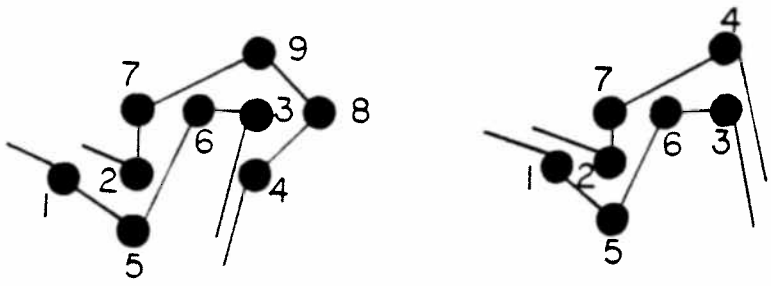


Figure 3.25: Adjacencies in a link box with two connections

3.6.4 H is a Hamiltonian Path

We have seen that H has maximum degree 2. Furthermore, every switch node has at most one adjacent link node, so every vertex in H is adjacent to at most one optional edge.

We show that the undirected version of H is connected.

Let v_a and v_z be two vertices in G' and B_a and B_z the corresponding switch or terminal boxes. One easily checks that for both B_a and B_z, there must be triangles Δ_a and Δ_z with two vertices from the respective boxes. Since $\overline{\mathcal{P}}$ is connected, there is a sequence of triangles $\Delta_1, \ldots, \Delta_k$ with the following properties:

1. $\Delta_a = \Delta_1$.

2. Δ_i and Δ_{i+1} share two vertices.

3. $\Delta_z = \Delta_k$.

By Lemma 3.6.1, each of the Δ_i has two vertices from the same box. This implies that the sequence of Δ_i induces an (undirected) path between v_1 and v_2 in H.

We conclude that the undirected version of H is a Hamiltonian path in G'; since every vertex in H has at most indegree 1 and at most outdegree 1, H must be a Hamiltonian path in D'.

□

3.7 Approximating MAP

Since MAP is NP-complete, it is highly unlikely that we could ever find an exact polynomial algorithm for finding a minimum area polygon. This would make it interesting to get good approximation methods. A common method for measuring the quality of an approximation is to consider the *relative error*: For a problem instance x with true optimum $opt(x)$, the relative error of an approximate solution $app(x)$ is

$$\left| \frac{app(x) - opt(x)}{opt(x)} \right|.$$

One particularly strong tool for finding approximative solutions to a combinatorial optimization problem is a *fully polynomial approximation scheme* (see Papadimitriou and Steglitz [73]). This is a method which, for any positive number ε, determines an approximate solution having relative error less than ε in time that is polynomial in both input size n and $\frac{1}{\varepsilon}$.

Theorem 3.0.3 implies that approximating MAP is hard:

Corollary 3.7.1 *If there is a fully polynomial approximation scheme for* MAP, *we can solve GAP in polynomial time.*

Proof:

For a set of n vertices, there can be no simple polygon of area less than $\frac{n}{2} - \frac{1}{2}$ if there is none of area $\frac{n}{2} - 1$. Choosing $\varepsilon := \frac{1}{2n}$ would therefore yield a polynomial method for finding a polygon with area no greater than

$$(1+\varepsilon)\left(\frac{n}{2}-1\right) = \frac{n}{2} + \frac{1}{4} - 1 - \frac{1}{2n} < \frac{n}{2} - \frac{1}{2},$$

i.e. a grid avoiding polygon.

□

But even a method that can guarantee a solution not exceeding a reasonable constant multiple of the optimal solution is not known. This contrasts to the TRAVELLING SALESMAN PROBLEM that allows the well-known approximation method by Christofides (see [14] or [73]) that uses a shortest spanning tree to find a solution within a factor of $\frac{3}{2}$ of the optimum.

The main reason for this difficulty lies in the more complicated geometric relation between the boundary of a simple polygon and its enclosed area. A short tour can be constructed of easy pieces like shortest spanning trees, while a similar approach for small area (building up a simple polygon by greedily adding triangles) tends to run into difficulties. We briefly discuss difficulties for a natural approach:

"Greedy-Build"

1. Start with the smallest nondegenerate triangle in the vertex set.

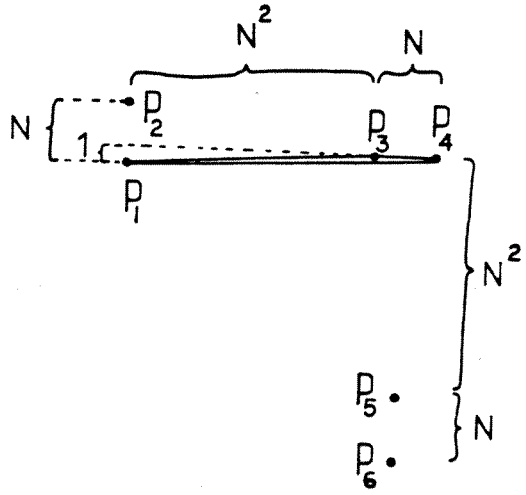


Figure 3.26: An example where "Greedy-Build" performs badly

---

2. As long as there are vertices not contained in the polygon, choose the smallest nondegenerate triangle formed by an edge of the current polygon and a vertex outside of the polygon that can see the edge.

3. Add the triangle to the polygon and continue at 2.

It may very well be that a vertex outside a simple polygon cannot see any complete edge. Even if this difficulty does not occur, the approach can yield arbitrarily bad results:

Example 3.7.2 See Figure 3.26.
"Greedy-Build" constructs the polygon $\langle p_1, p_5, p_6, p_4, p_2, p_3, p_1 \rangle$ that has an area of more than $\frac{N^4}{2}$. The optimum area is $N^3 + \frac{N^2}{2}$ and is assumed e.g. by the polygon $\langle p_1, p_3, p_6, p_5, p_4, p_2, p_1 \rangle$.

Chapter 4

Maximal Area

For some practical purposes, one might be more interested in simple polygons with a large enclosed area than in those with a small area. As we will see in this chapter, the relation between the problems MAXP and MAP is very close.

4.1 GAXP and MAXP are NP-complete

Using our results on GAP, we show that GRID APPROXIMATING POLYGON is NP-complete.

Theorem 4.1.1 *GAXP is NP-hard.*

Proof: See Figure 4.1. Consider the point set P in the NP-hardness proof of GAP. In any simple polygon with vertex set P of area $\frac{n}{2} - 1$, the points $t_1 :=$

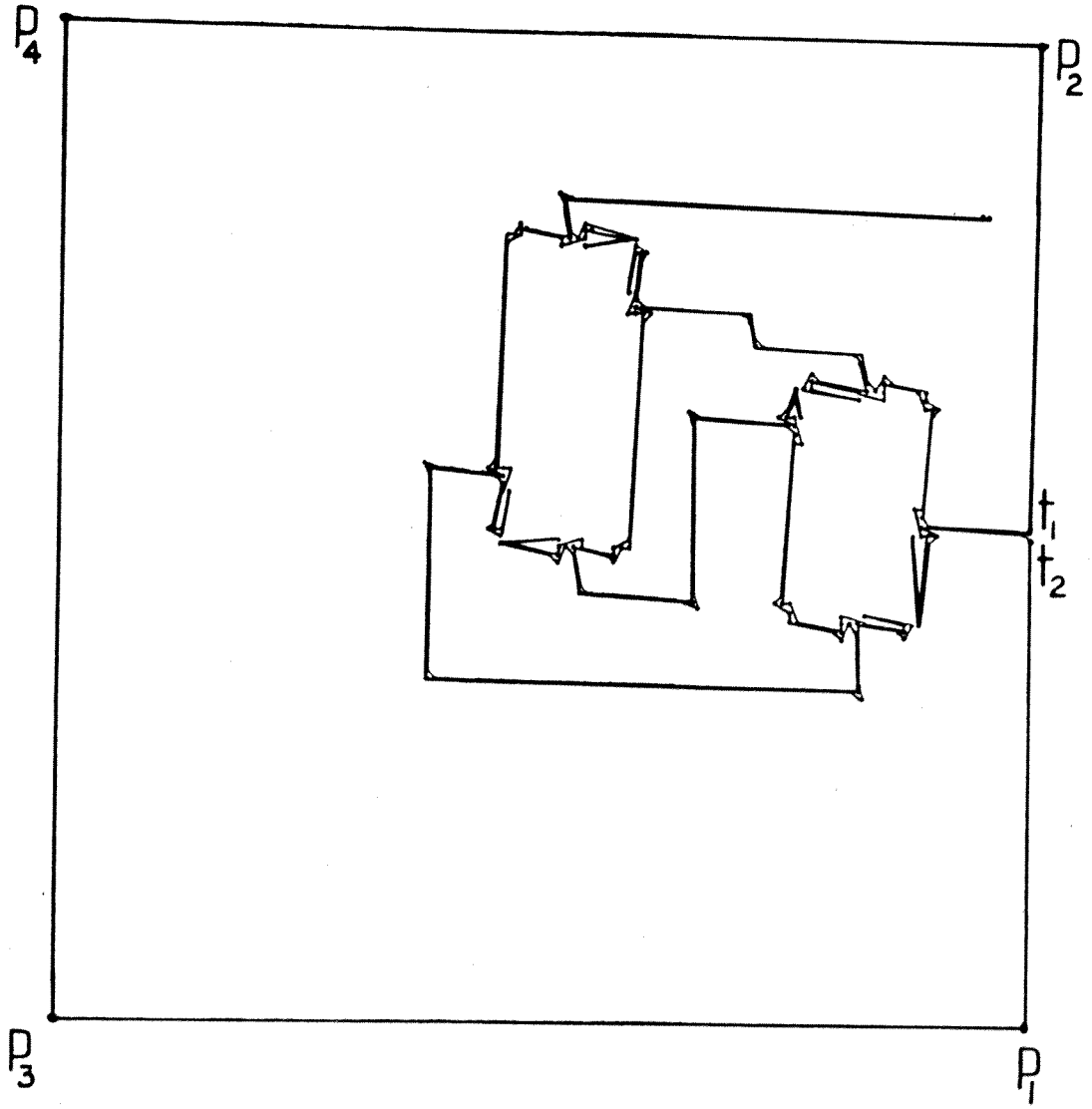


Figure 4.1: GAXP solves GAP: Turning a small polygon inside out

$(t_x, t_y - 1)$ and $t_2 := t_1 - (0,1)$ in the terminal box for t are connected to each other. By construction, all other grid points lie to the left of the vertical line through t_1 and t_2 (We made sure of this in the previous section in order to guarantee this property.) Then add the points

$$\begin{aligned} p_1 &:= t_2 \quad -(0, N^4) \\ p_2 &:= t_1 \quad +(0, N^4) \\ p_3 &:= p_1 \quad -(2N^4 + 1, 0) \\ p_4 &:= p_2 \quad -(2N^4 + 1, 0) \end{aligned}$$

to P to get the set $\overline{P}$. It is straightforward to see that there is a simple polygon $\mathcal{P}$ on the vertices P that satisfies GAP if and only if there is a simple polygon $\overline{\mathcal{P}}$ on the vertices $\overline{P}$ that satisfies PICK'S BOUND FOR MAXP. (The polygon $\overline{\mathcal{P}}$ is simply the complement of $\mathcal{P}$ in the square with vertices p_1, p_2, p_3, p_4.)

□

Similarly as for GAP and MAP, we get the following corollary from the NP-completeness of GAXP.

Corollary 4.1.2 *MAXP is NP-complete.*

4.2 A $\frac{1}{2}$-Approximation Algorithm for MAXP

A lower time bound for any approximation algorithm for MAP is given from the difficulty of just finding a simple polygon on a given set of n vertices. Lower time bounds are usually based on the model of *algebraic computation trees.* In this model, we assume that the evaluation of any algebraic function or any comparison

of two numbers can be carried out in one step; see Ben-Or [2]. It is well known that finding a simple polygon with a given set of vertices has a lower bound of lower bound of $\Omega(n \log n)$ in the model of algebraic computation trees. (See Exercise 3.6.1, p. 142 in the book by Preparata and Shamos [73].)

Drysdale and Jaromczyk [23] have established a lower bound of $\Omega(n \log n)$ for finding a polygon with k vertices chosen from a set of n points in the plane, where k is a given parameter.

In the following, we describe an $O(n \log n)$ method to obtain a simple polygon on a given set of points P whose area is bigger than half the area of the convex hull, $conv(P)$, of P. Since the area of the convex hull is an upper bound for any solution of MAXP, this yields a fast approximation method for MAXP.

Theorem 4.2.1 *Let P be a set of $n+1$ points in the plane. We can determine a simple polygon $\mathcal{P}$ on P that has area larger than $\frac{1}{2}AR(\text{conv}(P))$. This can be done in time $O(n \log n)$.*

Proof: Let p_0 be a point on the convex hull of P. In time $O(n \log n)$, sort the points p_i of P by the slope of the lines $l(p_0, p_i)$, such that the neighbours of p_0 on the convex hull are the first and the last point, respectively. If there is a set of points for which the slope is the same, break the tie by ordering them in increasing distance from p_0, except when those points have the smallest of all slopes, in which case we take them in order of decreasing distance from p_0. (The latter corresponds to the line through p_0 and p_n.) Connecting the points p_i in this order yields a simple polygon $\mathcal{P}_1$ on P — see Figure 4.2.

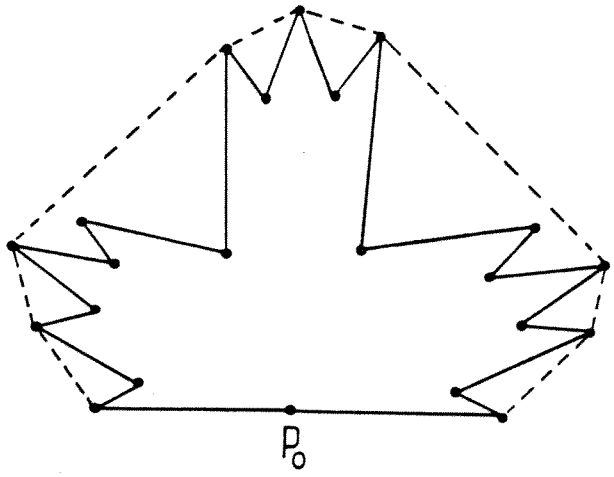


Figure 4.2: The simple polygon $\mathcal{P}_1$

If $AR(\mathcal{P}_1) > \frac{1}{2}AR(conv(P))$, we are done. Suppose this is not the case. Then the set $\mathcal{Q} := conv(P) \setminus \mathcal{P}_1$ has area at least $\frac{1}{2}AR(conv(P))$. Let $\overline{P}$ be the set of points that lie on the boundary of $conv(P)$. (There are well-known methods to determine the convex hull of a simple polygon in time $O(n)$, e.g. see McCallum and Avis [56]. Since $\mathcal{P}_1$ is star-shaped from p_0, we can achieve the same objective in a very straightforward and much simpler fashion.)

In the following, we write $\overline{q}_i$ for a point $q_i \in P$ if and only if $q_i \in \overline{P}$. $\mathcal{Q}$ consists of $h \geq 1$ polygons $Q_j := \langle \overline{q}_1^{(j)}, q_2^{(j)}, \ldots, q_{z_j-1}^{(j)}, \overline{q}_{z_j}^{(j)}, \overline{q}_1^{(j)} \rangle$, where $\overline{q}_1^{(j)}$ and $\overline{q}_{z_j}^{(j)}$ are the only vertices of Q_j that belong to $\overline{P}$.

Now for any j, the points $q_2^{(j+1)}$ and $q_{z_j-1}^{(j)}$ can see each other in the polygon $\mathcal{P}_1$ — see Figure 4.3. By construction, the angle $\varphi := \angle(q_2^{(j)}\overline{p}_0 q_{z_j-1}^{(j)}$ satisfies $0 < \varphi < \pi$. Let C_φ be the set of all points of the plane inside this angle φ. Since all vertices of

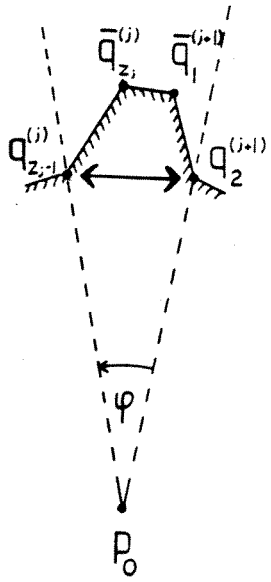


Figure 4.3: The points $q_2^{(j+1)}$ and $q_{z_j-1}^{(j)}$ can see each other

---

$\mathcal{P}_1$ between $\overline{q}_{z_j}^{(j)}$ and $\overline{q}_1^{(j+1)}$ belong to $\overline{P}$, the set $C_\varphi \cap \mathcal{P}_1$ is convex and the line from $q_2^{(j+1)}$ to $q_{z_j-1}^{(j)}$ runs through the inside of $\mathcal{P}_1$.

With the same argument, it is not hard to see that the edge from $\overline{p}_n$ to $q_{z_h-1}^{(h)}$ lies in $\mathcal{P}_1$. This implies that the polygon

$$\mathcal{P}_2 = \langle \overline{p}_0, q_2^{(1)}, \ldots, q_{z_1-1}^{(1)}, q_2^{(2)}, \ldots, q_{z_j-1}^{(j)}, q_2^{(j+1)}, \ldots, q_2^{(h)}, \ldots, q_{z_h-1}^{(h)}, \overline{p}_n, \ldots, \overline{p}_1, \overline{p}_0 \rangle$$

is simple: First travel all the points inside the hull, then go back along the hull — see Figure 4.4.

Since $\mathcal{P}_2$ contains $\mathcal{Q}$ as a proper subset, we get $AR(\mathcal{P}_2) > \frac{1}{2} AR(conv(P))$, concluding the proof.

□

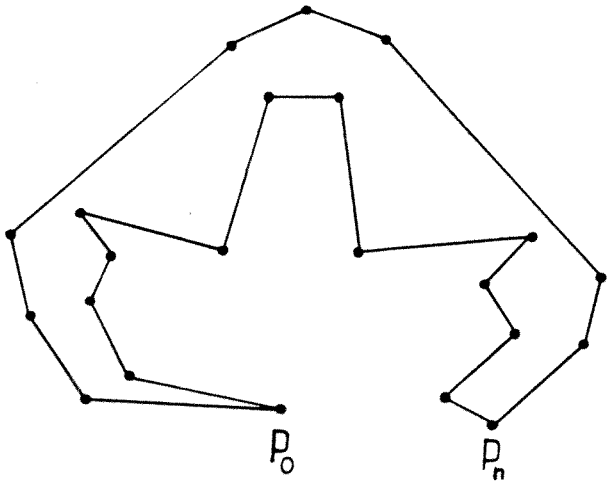


Figure 4.4: The simple polygon $\mathcal{P}_2$

All the estimates involved are tight:

Theorem 4.2.2 *Let $MAX(P)$ denote the largest area enclosed by a simple polygon on P and $\mathcal{P}^*$ be a solution obtained by the described heuristic.*

There are classes of point sets for which the inequalities

1. $MAX(P) \leq AR((\text{conv}(P))$
2. $MAX(P) \geq \frac{1}{2} AR(\text{conv}(P))$
3. $AR(\mathcal{P}^*) \geq \frac{1}{2} MAX(P)$

are tight. Moreover, the estimate for the above heuristic is still tight if $\mathcal{P}^*$ is obtained by trying all *possible starting points p_0 on the convex hull.*

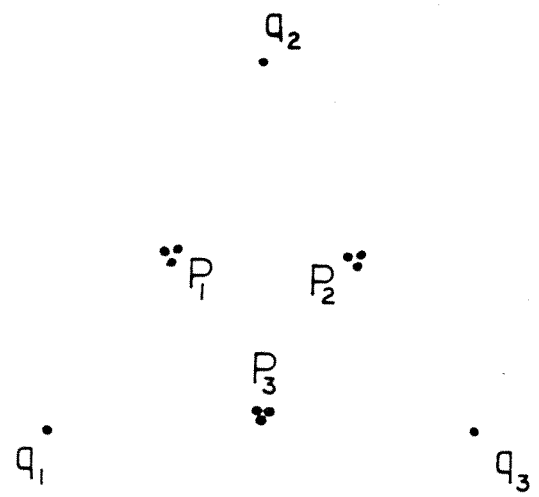


Figure 4.5: An example for which the bound $\frac{1}{2}$ is tight

---

Proof:

1. Any simple polygon is contained in the convex hull.

2. Consider a set of $(n+1)^2$ points that form an $(n+1)*(n+1)$-grid. By Pick's theorem, the area of *any* simple polygon on these points must be $\frac{n^2+2n-1}{2}$. For large n, the quotient with $AR(conv(P)) = n^2$ gets arbitrarily close to $\frac{1}{2}$.

3. Consider the situation shown in Figure 4.5: The points q_1, q_2, q_3 form a suitable approximation of an equilateral triangle spanning the convex hull. The sets P_1, P_2, P_3 all consist of points that are close to the midpoints of the sides of the triangle. It is easy to see that the best solution has an area close to that of the convex hull. Now choose q_1 as the point p_0 in our heuristic. Sorting the slopes from q_1, we get $\mathcal{P}_1$ as shown in Figure 4.6. The area of $\mathcal{P}_1$ is close to

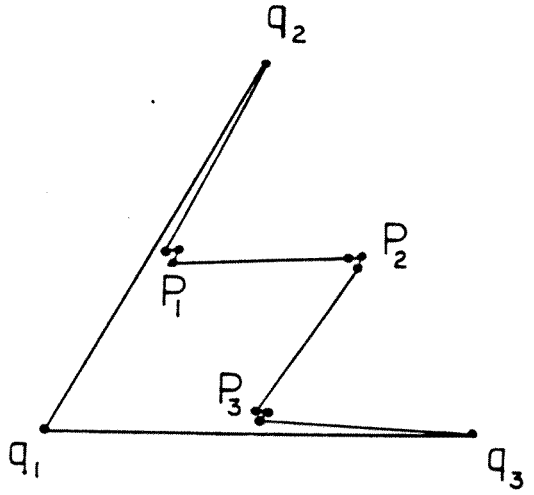


Figure 4.6: $\mathcal{P}_1$ has area close to $\frac{1}{2}MAX(P)$

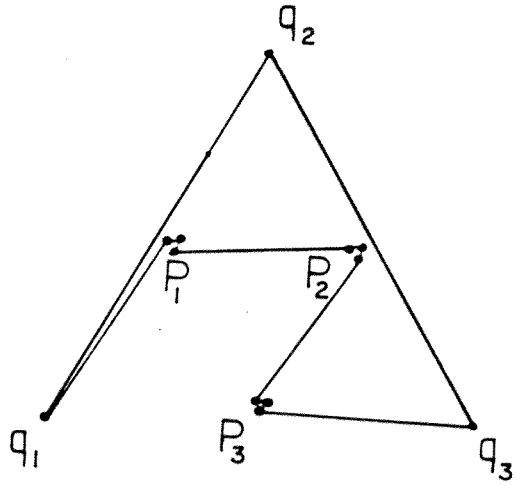


Figure 4.7: $\mathcal{P}_2$ has area close to $\frac{1}{2}MAX(P)$

$\frac{1}{2}AR(conv(P))$. The same holds for the area of $\mathcal{P}_2$ — see Figure 4.7. Since the situation is symmetric, we get no different result by choosing q_2 or q_3 instead of q_1 as the point p_0 for constructing $\mathcal{P}_1$.

□

4.3 A Negative Result on MAXP-Approximation

We can use a similar idea as in Theorem 4.1.1 to show that there is little hope of finding a simple polygon that encloses more than $\frac{2}{3}$ of the area of the convex hull:

Theorem 4.3.1 *Let $0 < \varepsilon < \frac{1}{3}$. If there is a polynomial algorithm that decides for any vertex set P whether there is a simple polygon of area at least $(\frac{2}{3} + \varepsilon)AR(\text{conv}(P))$, then we can solve GAP in polynomial time.*

Proof: The idea is similar to the proof of Theorem 4.1.1. For any instance of GAP, we can add a suitable set of points outside, such that deciding the existence of a polygon with $\frac{2}{3} + \varepsilon$ of the area of the convex hull solves the GAP instances constructed in the NP-completeness proof of GAP.

Given a set of n grid points that form an instance for GAP as constructed in the NP-completeness proof of GAP. As in the proof of Theorem 3.0.3, we can assume that they lie inside the box B with vertices $(\frac{N^4}{2}, \frac{N^4}{2})$, $(\frac{N^4}{2}, -\frac{N^4}{2})$, $(\frac{3N^4}{2}, \frac{N^4}{2})$, $(\frac{3N^4}{2}, -\frac{N^4}{2})$, where $N = n^n$. The terminal box for t is moved to the right so that the points t_1 and t_2 have the coordinates $(2N^4, 1)$ and $(2N^4, 0)$. Finally, add the

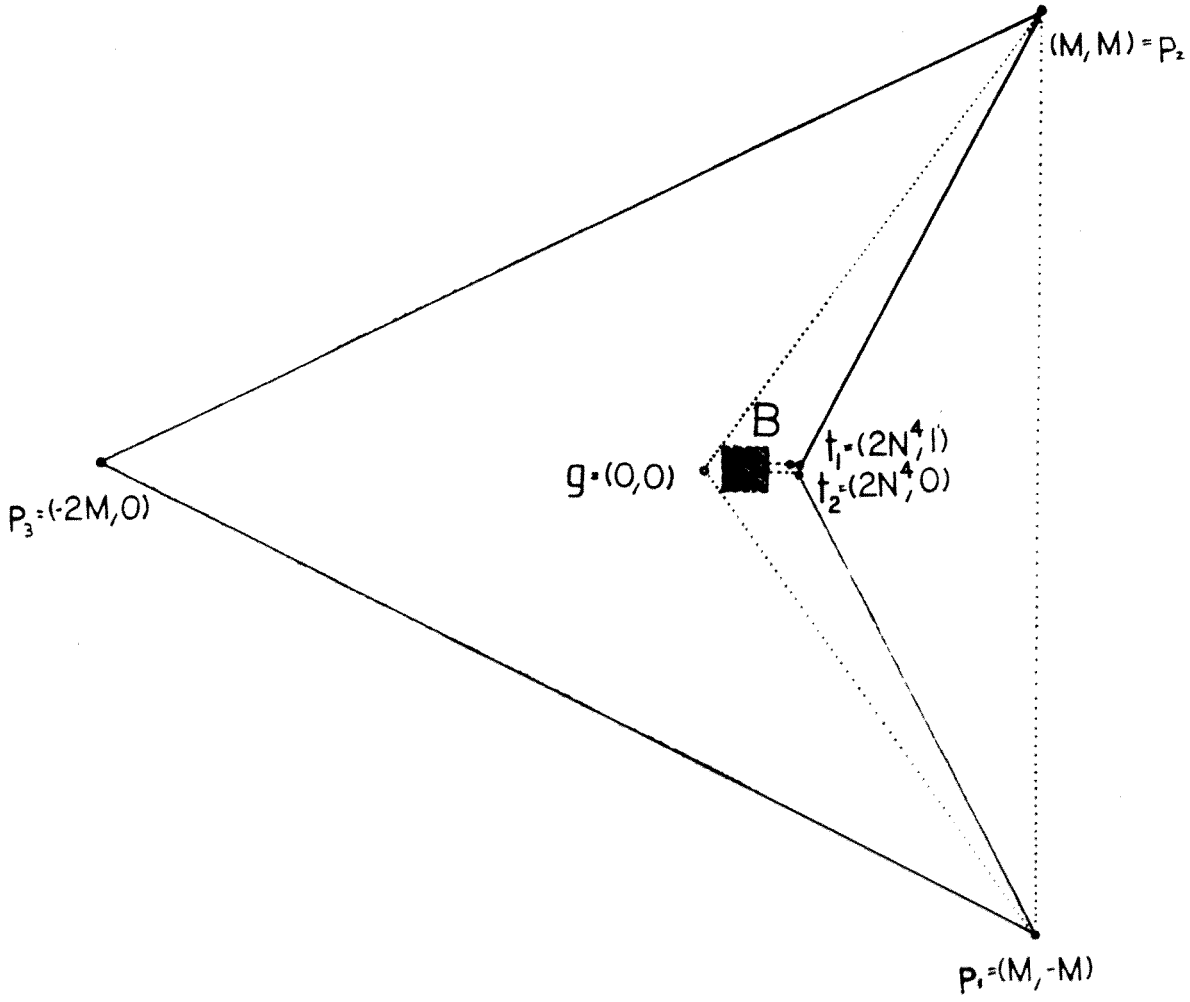


Figure 4.8: Finding a polygon with area at least $(\frac{2}{3} + \varepsilon)AR(P)$ solves GAP

points $p_1 = (M, -M)$, $p_2 = (M, M)$ and $p_3 = (-2M, 0)$, where

$$M = \frac{4N^4 - 1}{6\varepsilon} + \frac{2N^4 - n + 1}{4N^4 - 1}.$$

(See Figure 4.8.) If desired, we can make M integral by scaling the whole arrangement with a suitable integer. Assuming that N is large enough, it follows from the definition of M that $M > 2N^4$ and the convex hull is formed by p_1, p_2 and p_3 and has area $AR(conv(P)) = 3M^2$.

Consider a simple polygon $\mathcal{P}$ in which p_1 is connected to t_1 and p_2 to t_2. If $GAP(P)$ denotes the minimum area solution of the GAP instance, we get

$$AR(\mathcal{P}) = 2M^2 + 2N^4M - \frac{M - 2N^4}{2} - GAP(P).$$

Now we note that

$$\frac{1}{2} > \frac{1}{4}\left(\frac{2N^4 - n + 1}{2N^4 - 2}\right)^2 > 0.$$

Adding

$$\left(\frac{4N^4 - 1}{6\varepsilon}\right)^2 + \left(N^4 - \frac{n}{2} + \frac{1}{2}\right)$$

to these inequalities and using the definition of M yields

$$\left(2N^4 - \frac{1}{2}\right)M + N^4 - \frac{n}{2} + 1 > 3\varepsilon M^2 > \left(2N^4 - \frac{1}{2}\right)M + N^4 - \frac{n}{2} + \frac{1}{2}.$$

Since either $GAP(P) = \frac{n}{2} - 1$ or $GAP(P) \geq \frac{n}{2} - \frac{1}{2}$, this implies that $AR(\mathcal{P}) \geq (\frac{2}{3} + \varepsilon)AR(conv(P))$, if and only if $GAP(P) = \frac{n}{2} - 1$.

p_3 cannot be connected to any other points than p_1 and p_2: We would have a triangle outside of $\mathcal{P}$ that is formed by p_3, one of the points in B and (w.l.o.g.) p_1. Such a triangle has area at least $M^2 - MN^4$, which leaves an area of at most $2M^2 + N^4M$ for $\mathcal{P}$. Since this is smaller than the value $2M^2 + 2N^4M - \frac{M-2N^4}{2} - \frac{n}{2} + \frac{1}{2}$

that does not meet the required bound, the area of $\mathcal{P}$ is too small, regardless of other connections.

This concludes the proof.

□

4.4 A Game-Theoretical Problem

The preceding discussion dealt with the question of how to find a polygon that covers a large portion of the convex hull for a given set of points. From a game-theoretical point of view, it is interesting to examine the converse question:

CHOOSING AREA-MINIMIZING POINTS (CAMP)

Given a convex polygon Q with vertex set Q. How should we choose a set P of vertices in the polygon, such that for the resulting point set $\overline{P} := Q \cup P$,

$$\frac{MAX(\overline{P})}{AR(Q)}$$

becomes as small as possible?

(We can think of this problem as setting fence posts on a convex piece of land. An opponent then gets to draw a fence that has to use all of the posts and takes everything inside the fence. We want to keep his share as small as possible.)

It follows from Theorem 4.2.1 that for any choice of P, there is a simple polygon

$\mathcal{P}$ on $\overline{P} := Q \cup P$ such that

$$\frac{AR(\mathcal{P})}{ARconv((P))} > \frac{1}{2}.$$

With the help of Pick's theorem, we can show that this bound is asymptotically tight:

Theorem 4.4.1 *For any convex polygon $\mathcal{Q}$ with vertex set Q, we can add a sufficiently large set of points P inside $\mathcal{Q}$, such that for the resulting point set $\overline{P} := Q \cup P$, the value $MAX(\overline{P})$ is arbitrarily close to $\frac{1}{2}AR(Q)$.*

Proof:

We multiply the coordinates by a suitably large integer factor. Consider the set G of all resulting grid points contained in $\mathcal{Q}$ and let $G_i \subseteq G$ be the set of all grid points contained in the interior of $\mathcal{Q}$; similarly, let $G_b \subseteq G$ be the set of all grid points on the boundary of $\mathcal{Q}$. If we choose the scaling factor high enough, the ratio $\frac{|G_i|}{|G|}$ gets arbitrarily close to 1. Furthermore, Pick's theorem assures us that the ratio $\frac{|G_b|/2+|G_i|-1}{\mathrm{AR}(\mathcal{Q})}$ gets arbitrarily close to 1. (The ratio is equal to 1 for rational vertices. If the vertices of $\mathcal{P}$ are not all rational, we can approximate the area by a polygon contained in $\mathcal{P}$ that has rational vertices.) On the other hand, Pick's theorem tells us that the area of *any* polygon on the vertex set G is $\frac{|G|}{2} - 1$.

Choosing P such that $\overline{P} = G$ results in $\frac{\mathrm{MAX}(\overline{P})}{AR_{(\mathcal{Q})}}$ getting arbitrarily close to $\frac{|G|/2-1}{|G_b|/2+|G_i|-1}$, thus arbitrarily close to $\frac{|G_i|/2}{|G_i|} = \frac{1}{2}$.

□

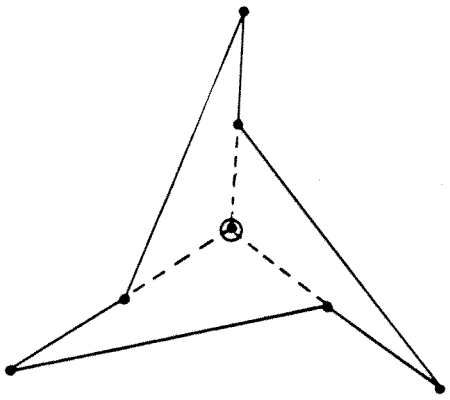


Figure 4.9: The circled point cannot see any whole edge of the surrounding polygon

---

Farming on the resulting pieces of land may not be too efficient: In the limit, the fence is a spacefilling curve.

We have assumed that the original polygon Q is convex. If we drop this condition, there may not be a simple polygon on $\overline{P}$ that does not leave Q:

Example 4.4.2 *See the polygon in Figure 4.9. (In Computational Geometry, it is known as "GFP" for "Godfried's Favourite Polygon", since Godfried Toussaint has used it as a counterexample for many conjectures.) By placing one additional point as shown, it becomes impossible to find a* simple *polygon on all the points that does not leave the original area.*

It may be an interesting question to study the extremal values of

$$CAMP_k = \min_{|P|=k} \left(\frac{\text{MAXP}(Q \cup P)}{AR(Q)} \right)$$

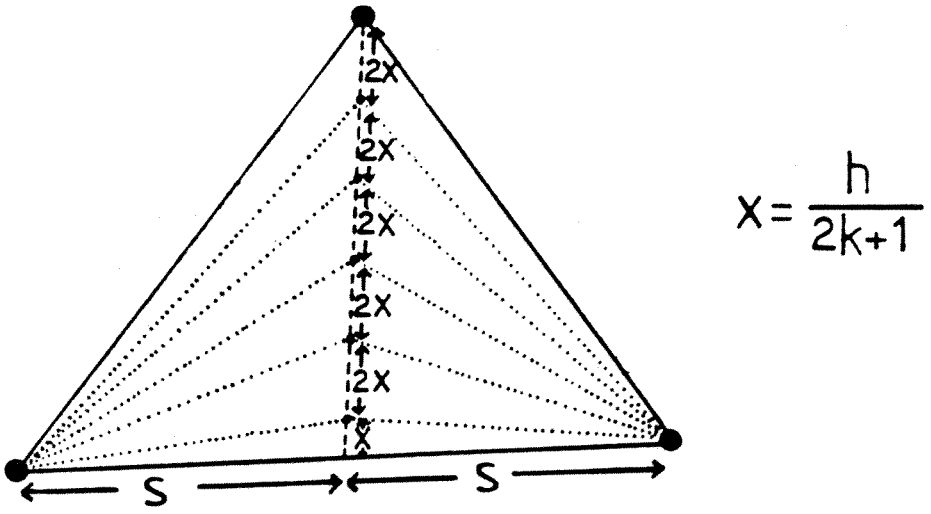


Figure 4.10: $CAMP_k \leq \frac{k+1}{2k+1}$

for a fixed k and convex polygons described by their vertex sets Q. It seems that $CAMP_k$ is extremal if Q is a triangle; the example in Figure 4.10 proves that $CAMP_k \leq \frac{k+1}{2k+1}$. Some case analysis shows that this bound is tight for $k = 1, 2, 3$. Note, however, that this strategy (choosing P on a line through one of the vertices that bisects the area) does in general *not* provide an area close to $\frac{1}{2}AR(Q)$ if Q is not a triangle — see Figure 4.11.

Determining $CAMP_k$ for a given Q, i.e. optimizing the choice of the additional points does not seem to be easy. For the special case $k = 1$, assume that Q has the edges e_i. it is not hard to show that for an optimal choice of $P = \{p\}$ for $k = 1$, there must be three distinct edges e_j, e_k, e_l such that $AR(\Delta_j) = AR(\Delta_k) = AR(\Delta_l)$, where Δ_j, Δ_k, Δ_l are the triangles spanned by p and e_j, e_k, e_l, respectively. This yields a straight-forward $O(n^3)$ algorithm for finding the optimal position of p. It should not be too hard to improve on this complexity.

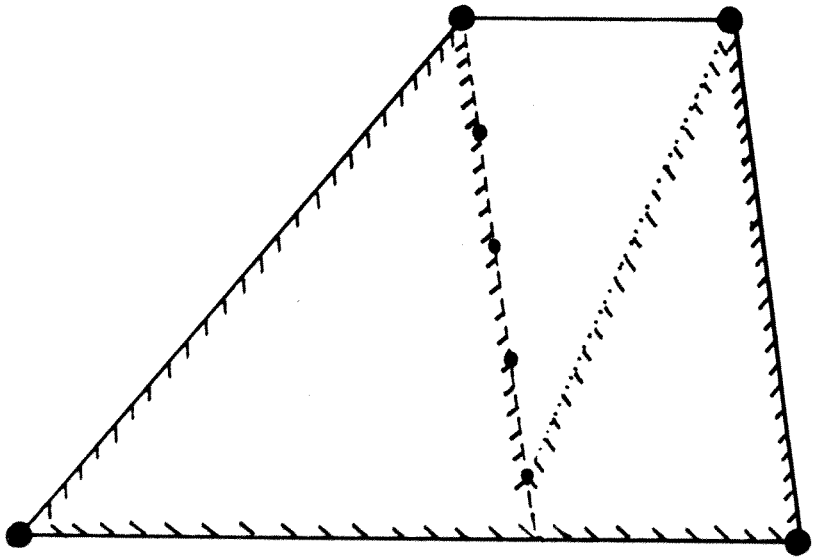


Figure 4.11: The strategy does not generalize if Q is not a triangle

Chapter 5

Higher Dimensions

In this chapter we study several higher dimensional generalizations. After discussing the question of calculating volume, we show that in any fixed dimensions k and d, finding a simple d-dimensional polyhedron with a given set of vertices that has minimal volume of its k-dimensional faces is NP-hard. This answers and generalizes a question stated by O'Rourke [66], [62].

Before we discuss how we can generalize the calculation and optimization of area to higher-dimensional polyhedra, we define what kind of polyhedra we want to consider:

Definition 5.0.3 *A d-dimensional polyhedron $\mathcal{P}$ is called* simple, *if it is homotopic (topologically equivalent) to a d-dimensional sphere.*

It is feasible *for a given vertex set P, if every vertex of $\mathcal{P}$ belongs to P and every point in P is contained in at least $d-1$ different faces of $\mathcal{P}$.*

The generalization of simplicity from the two-dimensional situation is clear. The reason for considering feasibility in the stated form is the following: We want all points in the given set P to carry some significance for the polyhedron. In the two-dimensional case, however, a "vertex" of a simple polygon is usually not required to be locally extreme; it can very well be the common end point of two adjacent collinear edges. We account for this situation with the above definition. Any point in P is at least contained in an edge of $\mathcal{P}$.

5.1 Volume in Higher Dimensions

Pick's Theorem has been generalized to three-dimensional space, although there are easy examples showing that counting grid points alone is not enough:

Example 5.1.1 (Reeve [81])

Consider the point set $(0,0,0),(1,0,0),(0,1,0),(1,1,r)$ *for any positive integer* r. *The tetrahedron with this vertex set contains no more than those* 4 *grid points, but the enclosed volume can become arbitrarily large.*

Reeve [81] has shown that this difficulty can be overcome by counting points with *half*-integer coordinates. One special case of his fairly general results yields the following formula for three-dimensional volume:

Theorem 5.1.2 *Let* $\mathcal{P}$ *be a convex three-dimensional polyhedron with integer vertices. Denote by* $i(\mathcal{P})$ *the number of points with integer coordinates inside* $\mathcal{P}$, $i_2(\mathcal{P})$ *the number of points with half-integer coordinates inside* $\mathcal{P}$, $b(\mathcal{P})$ *the number of*

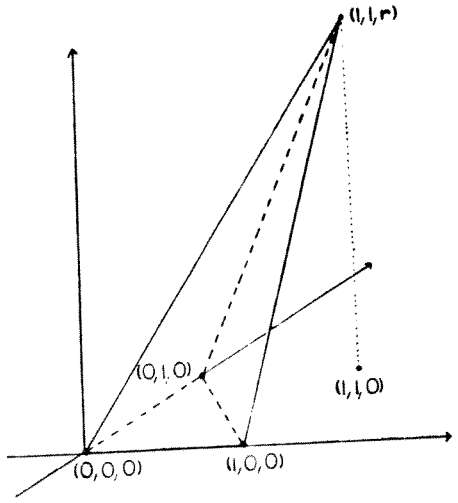


Figure 5.1: Why Pick's theorem does not generalize easily to 3 dimensions

---

integer points on the boundary of $\mathcal{P}$, $b_2(\mathcal{P})$ *the number of half-integer points on the boundary of* $\mathcal{P}$. *Then*

$$6\,VOL(\mathcal{P}) = \frac{1}{2}\,(2b(\mathcal{P}) - b_2(\mathcal{P})) + i_2(\mathcal{P}) - 2i(\mathcal{P}).$$

In addition,

$$4b(\mathcal{P}) - b_2(\mathcal{P}) = 6.$$

For calculating the volume of a higher-dimensional simple polyhedron, a similar approach as for calculating the area of a simple polygon seems appropriate: Subdivide the polyhedron into simpler parts and calculate their individual volumes. Lemma 2.2.2 can be generalized easily:

Lemma 5.1.3 *Let S be a d-dimensional simplex with vertices $(0,\ldots,0)$, $(x_{11},\ldots,x_{1d}),\ldots,(x_{d1},\ldots,x_{dd})$. Let $X = (x_{ij})$. Then*

$$VOL_d(S) = \frac{1}{d!}|\det(X)|.$$

Proof: We can proceed straightforwardly by induction over d, see Benson [3].

□

But while triangulating a (two-dimensional) polygon is a comparatively easy problem, the analogous task is a lot trickier in three dimensions. A detailed discussion can be found in a very recent article by Rupert and Seidel [86]. We just mention some of the main difficulties.

It is not true that any "tetrahedralization" of a polyhedron without any additional vertices (Steiner points) has the same number of tetrahedra — a fact that was already known to Schönhardt [89]. Moreover, it is NP-complete to decide whether a polyhedron can be decomposed at all without using any Steiner points — see Rupert and Seidel [86].

However, these difficulties need not be of great concern for us: All we need for calculating the volume of a d-dimensional polyhedron is a polynomial method to partition it into a polynomial number of tetrahedra. Von Hohenbalken [42] has solved the problem of subdividing a convex d-dimensional polyhedron into a small number of d-dimensional simplices. Therefore it is sufficient to partition into convex pieces. In three dimensions, this has been done by Chazelle in [10], who gave a method to decompose into $O(n^2)$ pieces and also showed that $\Omega(n^2)$ convex pieces may be necessary. (His method is based on the exact RAM model of computation; the behaviour of the coordinates of the resulting Steiner points

is not clear from his exposition.) For higher dimensions, a brute-force method for partitioning a simple polyhedron into $O(n^d)$ convex pieces proceeds by cutting it with axis-parallel hyperplanes through its vertices.

5.2 Minimizing Surface Area

The following problem has been proposed by O'Rourke [66], [62]:

MINIMAL SURFACE POLYHEDRON (SURF):

Given a finite set P of points in 3-dimensional Euclidean space. Among all simple polyhedra that are feasible for a vertex set P, find the one with the smallest surface area.

This is a special case of the following problem:

MINIMAL FACE POLYHEDRON (FACE):

Let $2 \leq d$ and $1 \leq k \leq d$. Given a finite set P of points in d-dimensional Euclidean space. Among all simple polyhedra that are feasible for vertex set P, find the one with the smallest volume of its k-dimensional faces.

It turns out that FACE is NP-hard for any choice of d and k. We first show that the special case SURF is NP-hard. The concluding section describes how the reduction can be generalized to prove NP-hardness of FACE.

Theorem 5.2.1 SURF *is NP-hard.*

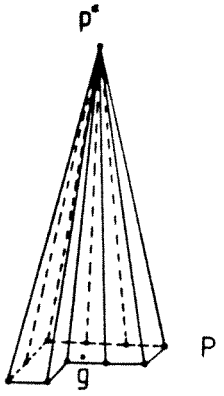


Figure 5.2: Extending a planar tour into a cone in 3-space

Proof: We describe a reduction of HAMILTONIAN CYCLE IN GRID GRAPHS. Take any instance of HCGG, i.e. a grid graph G with n vertices. This grid graph G can be canonically represented by a set P_G of n points in the plane $E_2 = \{(x, y, 0) \mid x, y \in \mathbb{R}\}$. Let $g = (x_g, y_g, 0)$ be the center of mass of the points in P_G and $p^* = (x_g, y_g, H)$, where $H = 2n^3$. (See Figure 5.2.)

We show that for the set $P := P_G \cup \{p^*\}$, there is a polyhedron with the required properties of surface area $n^4 + n^2 + \frac{1}{8}$ or less *if and only if* there is a (Euclidean) tour of length n or less in P, i.e. a Hamiltonian cycle in the grid graph.

If there is a tour of length n, it induces a polyhedron with a surface area of

$$A + \sum_{i=1}^{n} \frac{\sqrt{H^2 + d_i^2}}{2},$$

where A is the area enclosed by the tour and d_i is the distance of $(x_g, y_g, 0)$ from the ith tour edge. Clearly, we have $A \leq n^2$ and $d_i < n$. Thus we get the surface

area to be at most

$$n^2 + \sum_{i=1}^{n} \frac{\sqrt{H^2 + n^2}}{2}.$$

Now since $H = 2n^3$ and

$$H^2 + n^2 < \left(H + \frac{n^2}{2H}\right)^2,$$

we have

$$\sqrt{H^2 + n^2} < H + \frac{1}{4n},$$

so the surface area is at most

$$n^2 + n\left(\frac{H}{2} + \frac{1}{8n}\right) = n^4 + n^2 + \frac{1}{8}.$$

To see the converse, assume that there is no tour of length n. First note that for every simple polyhedron $\mathcal{P}$ on P, the domain $\mathcal{P}_2 = \mathcal{P} \cap E_2$ must be a simple two-dimensional polygon. (Each vertex in P_G must be close to interior points of $\mathcal{P}$; furthermore, the set of inner points of $\mathcal{P}$ close to p^* has to be simply connected and nonempty.) This implies that the edges of $\mathcal{P}_2$ must form a tour. Let A be the enclosed area of this tour. Discounting this face of $\mathcal{P}$, we can partition the surface of $\mathcal{P}$ into a set of triangles by connecting p^* to all other vertices in P.

If there is no tour of length n, the length of the shortest tour must be at least $n + \sqrt{2} - 1 > n + \frac{1}{3}$ and the surface area

$$A + \sum_{i=1}^{n} \frac{s_i\sqrt{H^2 + d_i^2}}{2}$$

(where s_i is the length of the ith edge) is at least

$$0 + \frac{H(n + \frac{1}{3})}{2} = n^4 + \frac{n^3}{3}.$$

This is bigger than $n^4 + n^2 + \frac{1}{8}$ as soon as n exceeds 3.

□

5.3 Minimizing Face Volume

We conclude this chapter by proving the NP-hardness of FACE:

Theorem 5.3.1 *FACE is NP-hard.*

In the proof of Theorem 5.3.1, we use the following lemma:

Lemma 5.3.2 *Let $1 \leq k$ and $P = \{e_1, \ldots, e_{k+1}\}$, where e_i denotes the ith unit vector. Then the k-dimensional regular simplex S_k spanned by P has k-dimensional volume*

$$VOL_k(S_k) = \frac{\sqrt{k+1}}{k!}.$$

Proof:

We proceed by induction over k. The claim holds for $k = 1$. Assume it was true for $k-1$. Then the distance of e_{k+1} from S_{k-1} is $\sqrt{1 + k\left(\frac{1}{k}\right)^2}$ and we get

$$VOL_k(S_k) = \frac{1}{k}\sqrt{\frac{k+1}{k}}\, VOL_{k-1}(S_{k-1}) = \frac{\sqrt{k+1}}{k!}.$$

□

We now proceed to prove Theorem 5.3.1. The idea is to add a set of $(d-2)$ points at a large distance from an instance of a 2-dimensional problem to transform it into a d-dimensional one.

We distinguish the following three cases:

(A) $k = d$

(B) $k = 1$

(C) $1 \leq k < d$.

For all these cases we use the following:

Let P_G be a set of n points in the plane $E_2 = \{(x_1, x_2, 0, \ldots, 0) \mid x_1, x_2 \in \mathbb{R}\}$ that either represents an instance of MAP (case (A)) or a grid graph G (cases (B) and (C)). Let g be the center of mass of the points in P_G and $p_i = g + H \cdot e_i$ for $i = 3, \ldots, d$, where $H = 9d^2n^3$ and e_i denotes the ith unit vector. We write $P_j := \bigcup_{i=3}^{j}\{p_i\} \cup P_G$ and $E_j := \{(x_1, \ldots, x_d) \in \mathbb{R}^d \mid x_{j+1}, \ldots, x_d = 0\}$. For any d-dimensional simple polyhedron $\mathcal{P}_d$ feasible for the vertex set P_d, the corresponding j-dimensional subpolyhedron induced on P_j is denoted by $\mathcal{P}_j := \mathcal{P}_d \cap E_j$.

As in the previous proof, we may assume that $\mathcal{P}_2 = \mathcal{P} \cap E_2$ is a simple polygon. Furthermore $\mathcal{P}_j$ is simple if and only if $\mathcal{P}_{j-1}$ is simple. Finally, each of the points p_i is connected to any other vertex by an edge of $\mathcal{P}_d$.

Now consider case (A).

By Theorem 3.0.3, it is NP-complete to decide whether a given set of n grid points allows a simple polygon of area $\frac{n}{2} - 1$. With respect to Corollary 5.3.3, we may assume that if there is no such polygon, there is none of area less than $\frac{n-1}{2}$.

We claim that there is a simple polyhedron feasible for the vertices P_d of volume at most

$$VOL_d = \frac{2}{d!}H^{d-2}(\frac{n}{2} - 1)$$

if and only if there is a simple polygon on the vertices P_G with area at most $\frac{n}{2} - 1$.

This claim follows by induction over j: Consider the (j-dimensional) volume $VOL_j(\mathcal{P}_j)$ and note that

$$VOL_j(\mathcal{P}_j) := VOL_{j-1}(\mathcal{P}_{j-1})\frac{H}{j}, \text{ so}$$

$$VOL_j(\mathcal{P}_j) = \frac{2}{j!}H^{j-2}(\frac{n}{2} - 1).$$

Next consider case (B).

We show that there is a simple polyhedron $\mathcal{P}_d$ that is feasible for the vertex set P_d with sum of edge lengths $LEN(\mathcal{P}_d)$ not exceeding

$$n + \frac{(d-2)(d-3)}{2}H\sqrt{2} + (d-2)nH + \frac{n^3(d-2)}{2H},$$

if and only if there is a (Euclidean) tour of length n or less in P_G, i.e. a Hamiltonian cycle in the grid graph.

If there is a tour of length n, it induces a polyhedron $\mathcal{P}$, such that the sum of edge lengths in $\mathcal{P}_{d-k}$ satisfies

$$LEN(\mathcal{P}_j) = LEN(\mathcal{P}_{j-1}) + \sum_{i=3}^{j-1} dist(p_j, p_i) + \sum_{q \in P_G} dist(p_j, q).$$

Now

$$dist(p_j, p_i) = H\sqrt{2}$$

and

$$dist(p_j, q) = \sqrt{H^2 + s_q^2},$$

where $s_q = dist(q, g) \leq n$, and

$$\left(H + \frac{n^2}{2H}\right)^2 > H^2 + n^2,$$

thus

$$\sum_{q \in P_G} dist(p_j, q) < nH + \frac{n^3}{2H}.$$

Using these relations, we get

$$LEN(\mathcal{P}_d) < n + \frac{(d-2)(d-3)}{2} H\sqrt{2} + (d-2)nH + \frac{n^3(d-2)}{2H}.$$

Now assume that there is no tour of length n in P_G. Then it follows from the previous estimates that

$$LEN(\mathcal{P}_d) \geq n + \sqrt{2} - 1 + \sum_{i=3}^{d} ((i-2)H + nH), \text{ so}$$

$$LEN(\mathcal{P}_d) \geq n + \sqrt{2} - 1 + \frac{(d-2)(d-3)}{2} H\sqrt{2} + (d-2)nH.$$

Since $\sqrt{2} - 1 > \frac{(d-2)n^3}{2H}$, the claim holds.

Finally, consider Case (C).

For a simple polygon $\mathcal{P}_d$ on P_d, let $FACE_k(\mathcal{P}_d)$ denote the sum of volumes of its k-dimensional faces. We show that there is a simple polyhedron $\mathcal{P}_d$ feasible for the vertex set P_d with

$$\begin{aligned} FACE_k(\mathcal{P}_d) \;<\; & \binom{d-2}{k+1} \frac{\sqrt{k+1}}{k!} H^k \\ & + \binom{d-2}{k} \left(nH + \frac{n^3}{2H}\right) \frac{\sqrt{k}}{(k-1)!} H^{k-1} \\ & + \binom{d-2}{k-1} n(1 + \frac{n}{2H^2}) \frac{\sqrt{k-1}}{k!} H^{k-1} \\ & + \binom{d-2}{k-2} \frac{2}{(k-2)!} H^{k-2} n^2. \end{aligned}$$

if and only if there is a (Euclidean) tour of length n or less in P_G, i.e. a Hamiltonian cycle in the grid graph.

Consider the k-dimensional faces f of $\mathcal{P}_d$. There are four cases:

1. f is determined by a set of $k+1$ of the $d-2$ points p_i. In this case f is a regular k-dimensional simplex and we get

$$VOL_k(f) = \frac{\sqrt{k+1}}{k!}H^k.$$

There are $\begin{pmatrix} d-2 \\ k+1 \end{pmatrix}$ faces of this form.

2. f is determined by a set of k of the $d-2$ points p_i and one of the points in P_G. In this case f consists of a regular $(k-1)$-dimensional simplex and a single point at distance at most

$$\sqrt{n^2+H^2} < H + \frac{n^2}{2H}.$$

We get

$$\frac{\sqrt{k}}{(k-1)!}H^k \leq VOL_k(f) < \left(1 + \frac{n^2}{2H^2}\right)\frac{\sqrt{k}}{(k-1)!}H^k.$$

There are $n\begin{pmatrix} d-2 \\ k \end{pmatrix}$ faces of this form.

3. f is determined by a set of $k-1$ of the $d-2$ points p_i and an edge in $\mathcal{P}_2$.

Let (q_1, q_2) be the edge of length s_i and $l_1 = dist(q_1, g)$. The volume of the $k-1$ dimensional simplex $S_{k-2}^{(q_1)}$ formed by q_1 and the $k-1$ points p_i is

$$VOL_{k-1}(S_{k-2}^{(q_1)}) = \frac{\sqrt{l_1^2+H^2}}{k-1}VOL_{k-2}(S_{k-2}).$$

The distance of q_2 from $S_{k-2}^{(q_1)}$ lies between $s_i(1-\frac{k}{H})$ and s_i. Since

$$H \leq \sqrt{l_1^2 + H^2} < H + \frac{n^2}{2H},$$

we get

$$s_i(1-\frac{k}{H})\frac{\sqrt{k-1}}{k!}H^{k-1} \leq VOL_k(f) < s_i(1+\frac{n^2}{2H^2})\frac{\sqrt{k-1}}{k!}H^{k-1}.$$

There are $\begin{pmatrix} d-2 \\ k-1 \end{pmatrix}$ faces of this form for each edge of $\mathcal{P}_2$.

If we sum this over all (tour) edges for a fixed simplex S_{k-2}, we get a lower bound of

$$s(T)(1-\frac{k}{H})\frac{\sqrt{k-1}}{k!}H^{k-1}$$

and an upper bound of

$$s(T)(1+\frac{n^2}{2H^2})\frac{\sqrt{k-1}}{k!}H^{k-1}$$

where $s(T)$ is the length of the tour.

4. f is determined by a set of $k-2$ of the $d-2$ points p_i and the simple polygon $\mathcal{P}_2$. In this case, we get

$$VOL_k(f) = VOL_k(\mathcal{P}_k)$$

and therefore (cf. **Case (A)**)

$$VOL_j(f) = \frac{2}{(k-2)!}H^{k-2}A,$$

where A is the area enclosed by $\mathcal{P}_2$; since $0 < A < n^2$, we get

$$0 < VOL_k(f) < \frac{2}{(k-2)!}H^{k-2}n^2.$$

There are $\begin{pmatrix} d-2 \\ k-2 \end{pmatrix}$ faces of this form.

Now assume there is a tour of length n. Using the above estimates, we see that this tour induces a polyhedron $\mathcal{P}_d$, such that

$$\begin{aligned} FACE_k(\mathcal{P}_d) \;<\; & \binom{d-2}{k+1}\frac{\sqrt{k+1}}{k!}H^k \\ & + \binom{d-2}{k}\left(nH+\frac{n^3}{2H}\right)\frac{\sqrt{k}}{(k-1)!}H^{k-1} \\ & + \binom{d-2}{k-1}(n+\frac{n^3}{2H^2})\frac{\sqrt{k-1}}{k!}H^{k-1} \\ & + \binom{d-2}{k-2}\frac{2}{(k-2)!}H^{k-2}n^2. \end{aligned}$$

For the converse assume that there is no tour of length n. Then $s(T) \geq n+\sqrt{2}-1$ and

$$\begin{aligned} FACE_k(\mathcal{P}_d) \;>\; & \binom{d-2}{k+1}\frac{\sqrt{k+1}}{k!}H^k \\ & + \binom{d-2}{k}nH\frac{\sqrt{k}}{(k-1)!}H^{k-1} \\ & + \binom{d-2}{k-1}(n+\sqrt{2}-1)\left(1-\frac{k}{H}\right)\frac{\sqrt{k-1}}{k!}H^{k-1}. \end{aligned}$$

Since

$$\begin{aligned} &(d-2)\frac{\sqrt{k}}{\sqrt{k-1}}\frac{n^3}{2}+\frac{n^3}{2H}+(n+\sqrt{2}-1)k+\frac{2(k-1)^2kn^2}{(d-2)\sqrt{k-1}} \\ &\qquad < 9(\sqrt{2}-1)d^2n^3 = (\sqrt{2}-1)H, \end{aligned}$$

we conclude that the inequality

$$\binom{d-2}{k}\left(\frac{n^3}{2}\right)\frac{\sqrt{k}}{(k-1)!}H^{k-2}$$

$$
\begin{aligned}
&+ \left(\begin{array}{c} d-2 \\ k-1 \end{array} \right) \left(\frac{n^3}{2} \right) \frac{\sqrt{k-1}}{k!} H^{k-3} \\
&+ \left(\begin{array}{c} d-2 \\ k-1 \end{array} \right) (n+\sqrt{2}-1)k \frac{\sqrt{k-1}}{k!} H^{k-2} \\
&+ \left(\begin{array}{c} d-2 \\ k-2 \end{array} \right) \frac{2}{(k-2)!} H^{k-2} n^2 \\
&< \left(\sqrt{2}-1\right) \left(\begin{array}{c} d-2 \\ k-1 \end{array} \right) \frac{\sqrt{k-1}}{k!} H^{k-1}
\end{aligned}
$$

holds. Comparing the estimate for $FACE_k(\mathcal{P}_d)$ to the given bound, this implies that the value for $FACE_k(\mathcal{P}_d)$ does not meet the given bound.

This concludes the proof.

□

Corollary 5.3.3 *If there is a fully polynomial approximation scheme for FACE, then we can solve FACE in polynomial time.*

Proof:

By our above proof, we can guarantee that the next best solution to an instance is at least some positive number r larger than the critical value that we have to check. Since this value is always bounded above by some polynomial $b(n)$, choosing $\varepsilon < \frac{r}{b(n)}$ guarantees the optimum.

□

Chapter 6

Pseudoconvex Tours

The research for this chapter was done in cooperation with Gerhard Woeginger.

6.1 Convexity and Pseudoconvexity

Consider the Euclidean TSP instance shown in Figure 6.1. Intuitively, the solution shown looks optimal. Why is this true?

It is not hard to see that the given tour is the only noncrossing one, and hence the only 2-optimal tour. (A tour is 2-optimal if it cannot be changed into a shorter tour by replacing any 2 tour edges by 2 non-tour edges. It is a simple consequence of the triangle inequality that a self-intersecting tour is not 2-optimal.)

More geometrically, we observe that all the vertices of the tour lie on the boundary of their convex hull, i.e. they form a *convex* arrangement. The associated *convex*

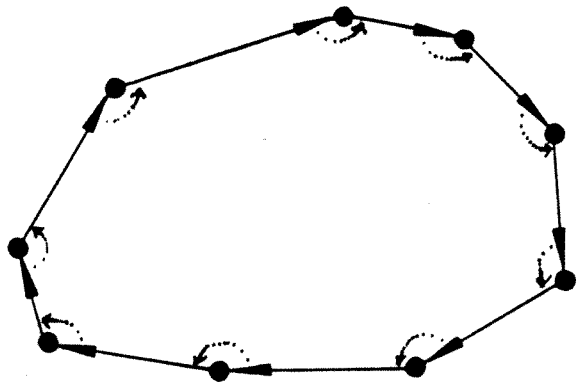


Figure 6.1: A Euclidean TSP instance with an easy solution

---

tour is the boundary of the convex hull. We discuss generalizations of convex tours for other than the Euclidean distance function in Chapter 8 and Chapter 9.

If we think of a tour as being a closed polygon consisting of directed line segments between consecutive vertices, we can characterize a (non-degenerate) convex tour by the following two conditions:

1. It is *simple*, i.e. noncrossing.
2. For any three consecutive vertices p_{i-1}, p_i, p_{i+1}, the angle $\angle(p_{i-1}, p_i, p_{i+1})$ lies in the nonnegative interval $[0, \pi]$.

All angles are considered as contained in the interval $(-\pi, \pi]$. We follow the usual convention that for three points x, y, z, $\angle(x, y, z)$ describes the angle by which the ray $\overrightarrow{yx}$ has to be rotated in counterclockwise fashion around y to place it over $\overrightarrow{yz}$.

We can think of 2. as a local condition that requires us to make a "right-hand turn" when following the line segments from p_{i-1} to p_i and then from p_i to p_{i+1}.

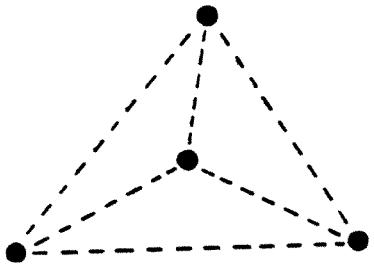


Figure 6.2: The bad shape

If we relax the conditions on convexity by dropping the global condition 1., we get so-called *pseudoconvex tours.*

Observation 6.1.1 *Not every finite set of points has a pseudoconvex tour — see the four points in Figure 6.2. We call this arrangement the* bad shape.

Surprisingly enough, this is the only counterexample:

Theorem 6.1.2 *Let P be a set of $n \geq 5$ points in the Euclidean plane. Then P has a pseudoconvex tour.*

The main objective of this chapter is to prove Theorem 6.1.2.

The basic idea is the following — see Figure 6.3: Consider the convex hull of the point set P. Assuming that we have a pseudoconvex tour of the points I_P inside the hull, we will try to find an edge of this inner tour that we can replace by a sequence of all the points on the hull H_P. It turns out that a particular arrangement of directed edges (a "γ-segment") in the inner tour guarantees the existence of such a

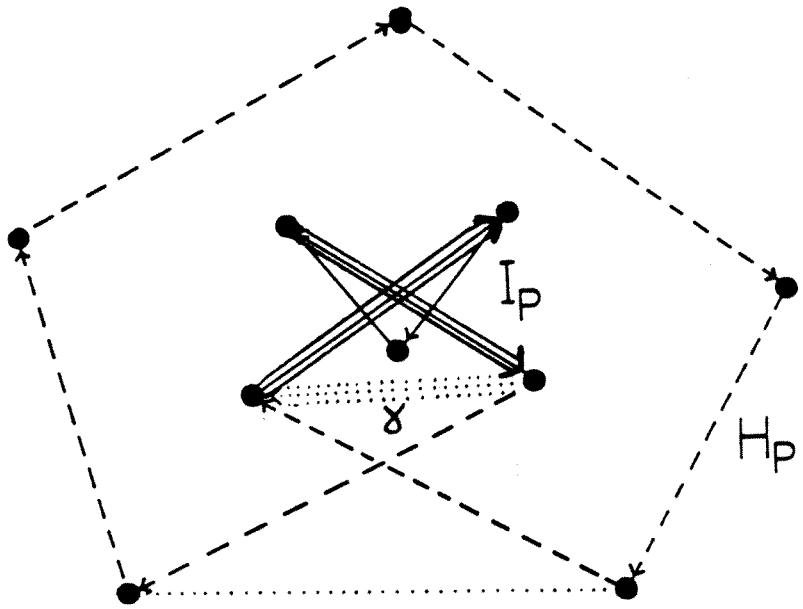


Figure 6.3: Extending a pseudoconvex tour

---

replaceable edge. The proof proceeds by induction over the number of convex layers in the so-called *onion-decomposition* of P. Most of the work consists of identifying γ-segments.

6.2 Setting up the Proof

Before we start the proof by induction, we introduce some basic definitions, state and prove a key lemma and take care of some starting cases that arise when the set of inner points is too small to allow a pseudoconvex tour. For easier description, we assume that the points are in general position. We discuss the situation for nongeneral position at the very end.

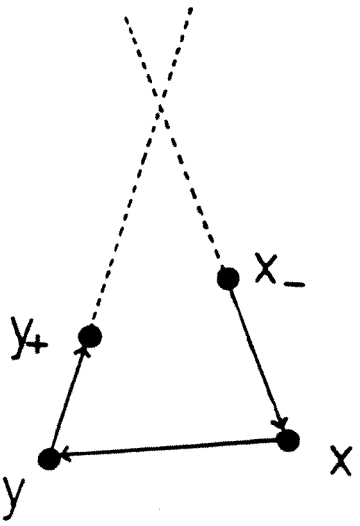


Figure 6.4: A γ-segment

Definition 6.2.1 *For any point set P, let H_P be the set of points on the boundary of the convex hull. We call H_P the set of* hull points. *We write $I_P := P \setminus H_P$ for the set of its interior points.*

For two distinct points x and y, $\overline{xy}$ denotes the line that they determine. $\overrightarrow{xy}$ is the directed edge from x to y. A point z lies right of $\overrightarrow{xy}$, *if the angle $\angle(x,y,z)$ is positive, i.e. contained in the interval $(0,\pi)$.*

Let I be a point set with a pseudoconvex tour $\mathcal{T}$. Assume there is a consecutive sequence x_-, x, y, y_+ in $\mathcal{T}$, such that the lines $\overline{x_-x}$ and $\overline{yy_+}$ do not intersect left of $\overrightarrow{xy}$. (See Figure 6.4.) In this case we say that $\mathcal{T}$ contains a γ-segment $\overrightarrow{xy}$.

The main significance for γ-segments arises from the following Lemma 6.2.2, which shows that we can include the next layer of points in a pseudoconvex tour of

the inner layers by using a γ-segment. Lemma 6.2.2 provides the main tool in the induction step.

Lemma 6.2.2 *Let I be a finite set of points with a pseudoconvex tour $\mathcal{T}$ and a γ-segment $\overrightarrow{xy}$. Let H be a finite set distinct from I such that H is the set of points on the boundary of the convex hull of $H \cup I$. Then there is a pseudoconvex tour in $H \cup I$ in which the elements of H follow each other consecutively in the order in which they appear when going clockwise around the hull.*

Proof: Assume that in $\mathcal{T}$, x_- is the predecessor of x, and y_+ is the successor of y. Consider the halfplane ε_1 lying to the right of $\overrightarrow{x_-x}$ and the halfplane ε_2 to the right of $\overrightarrow{yy_+}$. As the bounding lines cut the convex hull of P, both halfplanes ε_1 and ε_2 must each contain at least one point of H.

If the set $\varepsilon_1 \cup \varepsilon_2$ contains at least two points of H, then we are done — see Figure 6.5: In this case, we can find two points $h_0 \in \varepsilon_1$ and $h_1 \in \varepsilon_2$ that are neighbors in the clockwise ordering $h_0, h_1, \ldots, h_{m-1}$ of H. By construction, both the angles $\angle(x_- x h_1)$ and $\angle(h_0 y y_+)$ are positive, so we get the pseudoconvex tour $\langle x, h_1, \ldots, h_{m-1}, h_0, y, y_+, \mathcal{T}', x\rangle$, where $\mathcal{T}'$ denotes the part of $\mathcal{T}$ leading from y to x. (Whenever we use $\mathcal{T}'$ in a figure, it denotes an appropriate part of a pseudoconvex tour.)

If the set $\varepsilon_1 \cup \varepsilon_2$ contains only a single point h_1 of H, then $h_1 \in \varepsilon_1 \cap \varepsilon_2$. Consider the segment $\overrightarrow{x_-x}$ instead of $\overrightarrow{xy}$. Let $x_=$ be the predecessor of x_- in $\mathcal{T}$ and the halfplanes ε_3 and ε_4 lie to the right of $\overrightarrow{x_= x_-}$ and $\overrightarrow{xy}$, respectively. By definition, $\overline{x_-x}$ and $\overline{yy_+}$ do not intersect to the left of $\overrightarrow{xy}$. This implies that the set of all points left of $\overrightarrow{yy_+}$ and left of $\overrightarrow{x_-x}$ lies to the right of $\overrightarrow{xy}$. (See Figure 6.6.) This

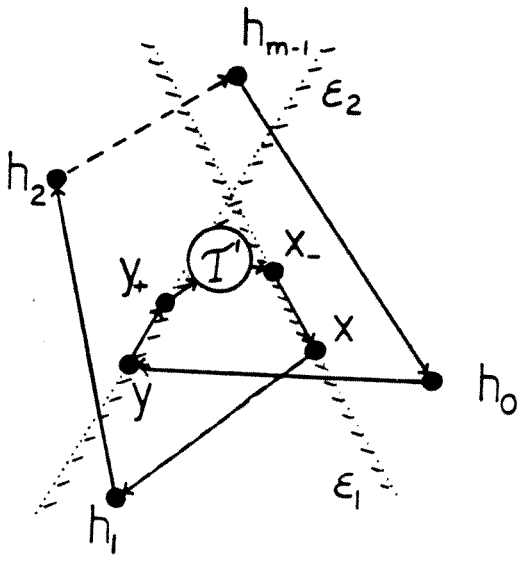


Figure 6.5: Using a γ-segment when $\varepsilon_1 \cup \varepsilon_2$ contains more than hull point

---

means that ε_4 contains all the points in $H \setminus \{h_1\}$. As in the previous case, we can find two consecutive points h_i and h_{i+1}, such that $h_i \in \varepsilon_4$ and $h_{i+1} \in \varepsilon_3$.

By construction, the angles $\angle(x_= x_- h_{i+1})$, $\angle(x_- h_{i+1} h_{i+2})$, $\angle(h_{i-1} h_i x)$, $\angle(h_0 x y)$ are all positive, so $\langle x_=, x_-, h_{i+1}, \ldots, h_{i-1}, h_i, x, y, y_+, \mathcal{T}', x_=, x_- \rangle$ is a pseudoconvex tour.

□

The following Lemma 6.2.3 covers all the starting cases for the induction. The reader should be warned that the number of cases that need to be considered makes the proof quite tedious.

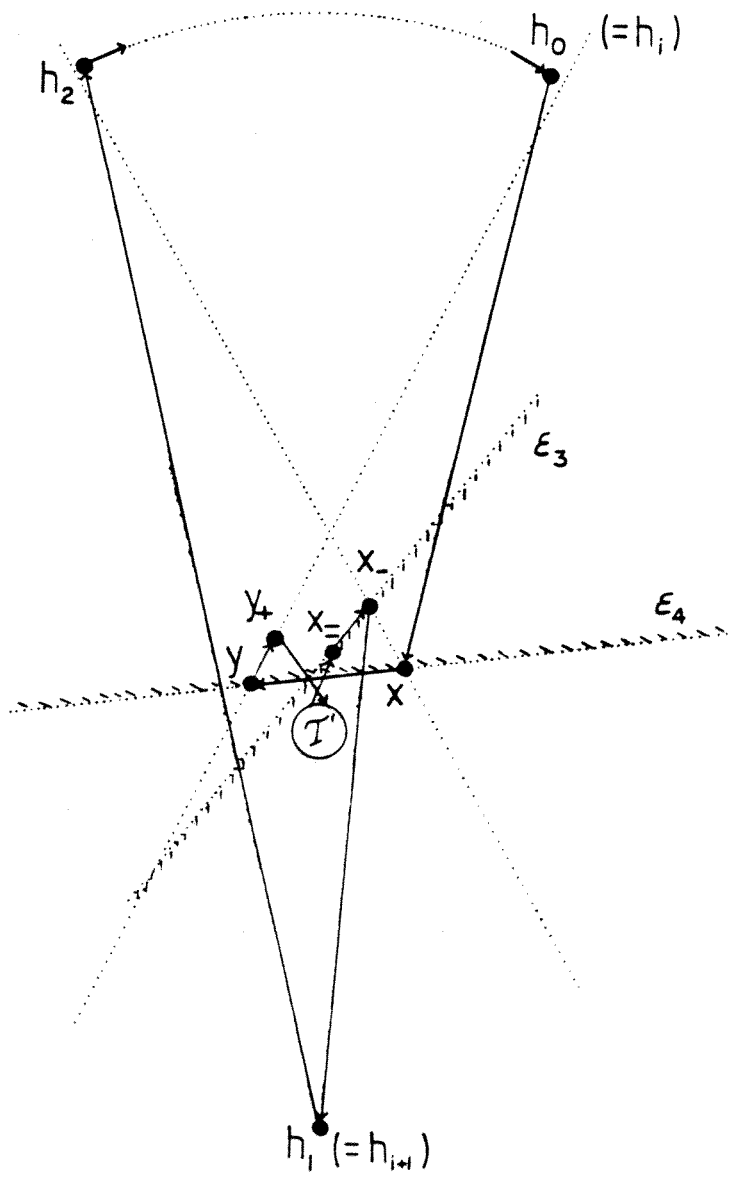


Figure 6.6: Using a γ-segment when $\varepsilon_1 \cup \varepsilon_2$ contains only one hull point

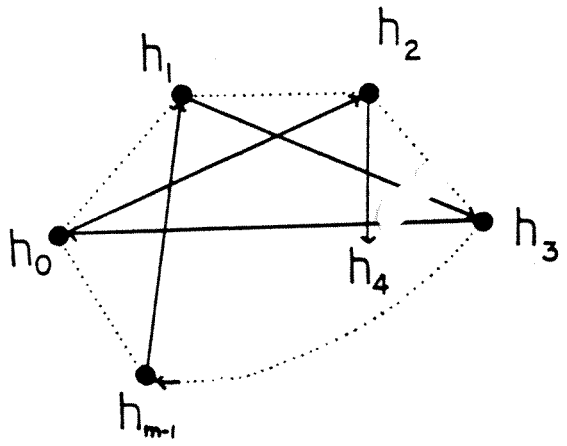


Figure 6.7: How to solve Case 0

Lemma 6.2.3 *Let P be a set of points in the Euclidean plane that is in general position and not a bad shape. Let I_P be the set of its interior points. If $|I_P| < 5$ and does not allow a pseudoconvex tour, then P has a pseudoconvex tour containing a γ-segment.*

Proof: Let H_P with $|H_P| = m$ be the hull points of P. Let the points of H_P be ordered as $h_0, \ldots, h_{m-1}$ when going clockwise around the hull.

We distinguish cases according to the form of I_P.

(CASE 0) $|I_P| = 0$:

See Figure 6.7. For $m = 3, 4$, the tour $\langle h_0, \ldots, h_{m-1}, h_0 \rangle$ is convex, thus pseudoconvex. Since there must be two consecutive angles whose sum does not exceed π, the tour contains a γ-segment. For $m \geq 5$, we can rearrange the order of h_{m-1},

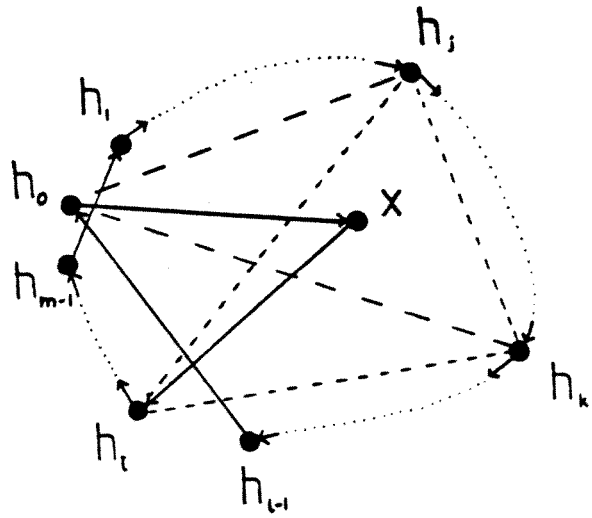


Figure 6.8: How to solve Case 1

---

h_0, h_1, h_2, h_3 in the convex tour $\langle h_0, \ldots, h_{m-1}, h_0 \rangle$ in order to get a pseudoconvex tour with a γ-segment: Clearly, the tour $\langle h_{m-1}, h_1, h_3, h_0, h_2, h_4, \ldots, h_{m-1} \rangle$ is pseudoconvex and $\overrightarrow{h_3 h_0}$ is a γ-segment.

(CASE 1) $|I_P| = 1$:

See Figure 6.8. Let x denote the only point in I_P; since $|P| \geq 5$, we have $|H_P| \geq 4$. There exists a triangle $\Delta h_j h_k h_l$ of points $h_j, h_k, h_l \in H_P$ containing x in its interior. Without loss of generality assume $0 < j < k < l$ and x is contained in the triangle $\Delta h_0 h_j h_k$. Then $\langle h_0, x, h_l, \ldots, h_{m-1}, h_1, \ldots, h_{l-1}, h_0 \rangle$ is a pseudoconvex tour and $\overrightarrow{x h_l}$ a γ-segment in it.

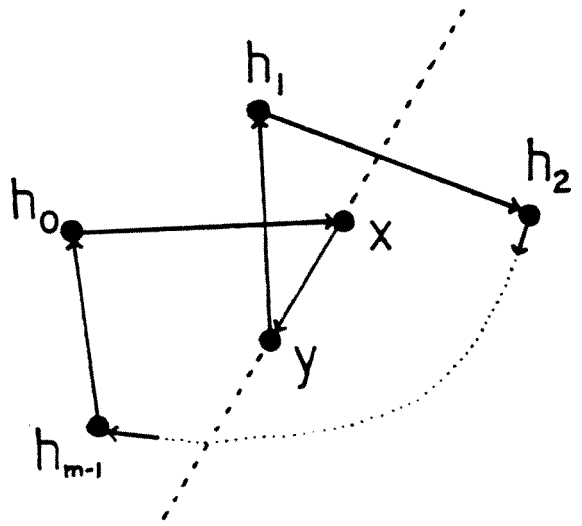


Figure 6.9: How to solve Case 2

(CASE 2) $|I_P| = 2$:

See Figure 6.9. Let x and y denote the two points in I. The line through x and y partitions the set H_P into two subsets. Without loss of generality assume that h_0 and h_1 lie on the right of the directed line $\overrightarrow{xy}$.

Then $\langle h_0, x, y, h_1, h_2, \ldots, h_{m-1}, h_0 \rangle$ is a pseudoconvex tour and $\overrightarrow{xy}$ a γ-segment in it.

(CASE B) I_P is a bad shape.

Let $I_P := \{q_0, q_1, q_2, q_3\}$, where q_0 is the point contained in the convex hull of the other three, which may have clockwise order q_1, q_2, q_3.

According to the number m of hull points, we further distinguish:

(CASE B.3) $m = 3$.

The lines $\overline{h_0q_0}, \overline{h_1q_0}$ and $\overline{h_2q_0}$ partition the triangle $\Delta h_0h_1h_2$ into the six triangles

$$\Delta_1 := \Delta h_0q_0m_2,$$

$$\Delta_2 := \Delta h_1m_2q_0,$$

$$\Delta_3 := \Delta h_1q_0m_0,$$

$$\Delta_4 := \Delta h_2m_0q_0,$$

$$\Delta_5 := \Delta h_2q_0m_1,$$

$$\Delta_6 := \Delta h_0m_1q_0,$$

where m_0, m_1, m_2 are the intersection points of $\overline{h_0q_0}$ and $\overline{h_1h_2}$, $\overline{h_1q_0}$ and $\overline{h_0h_2}$, $\overline{h_2q_0}$ and $\overline{h_0h_1}$, respectively.

Since q_0 is contained in the convex hull of q_1, q_2 and q_3, there must be one of these three points such that it lies in triangle Δ_i and none of the two other points lie in Δ_{i-1}, Δ_i or Δ_{i+1}. By symmetry, we may assume that q_1 lies in Δ_1 and neither q_2 nor q_3 lies in Δ_6, Δ_1 or Δ_2. This implies that either $q_3 \in \Delta_4$ or $q_3 \in \Delta_5$.

Now consider the line $\overline{q_1q_2}$, which has to intersect $\overrightarrow{h_1h_2}$ and either $\overrightarrow{h_0h_1}$ or $\overrightarrow{h_0h_2}$: If $\overline{q_1q_2}$ intersects $\overrightarrow{h_0h_1}$, then the angle $\angle(h_0q_1q_2)$ is positive and one easily checks that all the angles in the tour $\langle h_0, q_1, q_2, q_3, q_0, h_1, h_2, h_0\rangle$ are positive — see Figure 6.10(a): $\angle(q_1q_2q_3)$, since q_1, q_2, q_3 follow each other clockwise on the convex hull of I_P; $\angle(q_2q_3q_0)$, since q_2, q_3 follow each other clockwise on the convex hull of I_P and q_0 lies inside the convex hull of I_P; $\angle(q_3q_0h_1)$, since q_3 lies in $\Delta_3 \cup \Delta_4 \cup \Delta_5$; $\angle(q_0h_1h_2)$, since h_1, h_2 follow each other clockwise on the convex hull of H_P and q_0 lies inside the convex hull of H_P; $\angle(h_1h_2h_0)$, since h_1, h_2, h_0 follow each other clockwise on the

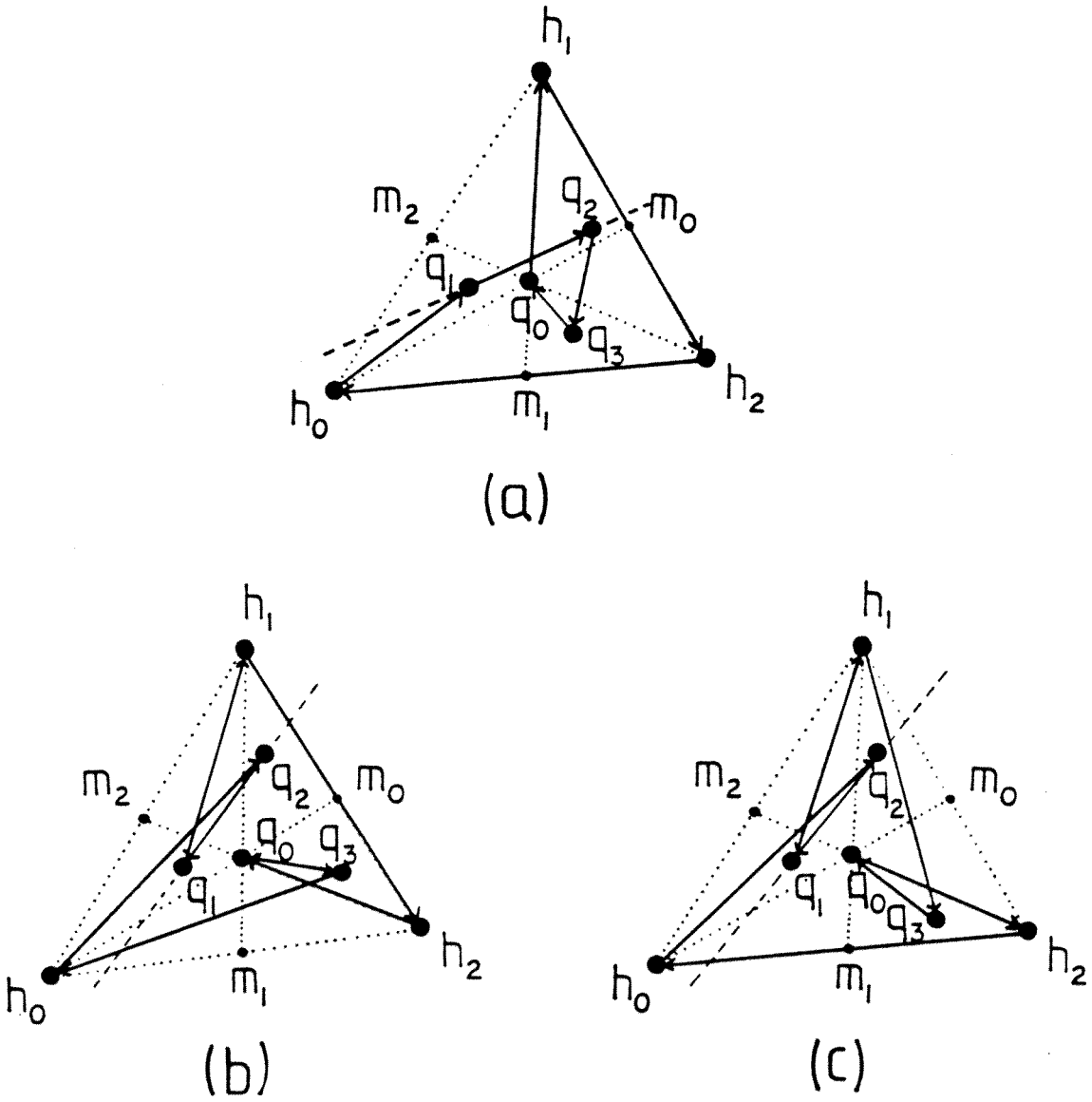


Figure 6.10: How to solve Case B.3

convex hull of H_P; $\angle(h_2h_0q_1)$, since h_2, h_0 follow each other clockwise on the convex hull of H_P and q_1 lies inside the convex hull of H_P. Clearly, $\overrightarrow{h_2h_0}$ is a γ-segment.

If $\overline{q_1q_2}$ intersects $\overrightarrow{h_0h_2}$, then the angles $\angle(h_0q_2q_1)$ and $\angle(q_2q_1h_1)$ are positive and q_2 lies in Δ_3. As stated above, either $q_3 \in \Delta_4$ or $q_3 \in \Delta_5$.

For $q_3 \in \Delta_4$, the angle $\angle(h_2q_0q_3)$ is positive and one easily checks that all the angles in the tour $\langle h_0, q_2, q_1, h_1, h_2, q_0, q_3, h_0\rangle$ are positive — see Figure 6.10(b): $\angle(q_1h_1h_2)$, since h_1, h_2 follow each other clockwise on the convex hull of H_P and q_1 lies inside the convex hull of H_P; $\angle(h_1h_2q_0)$, since h_1, h_2 follow each other clockwise on the convex hull of H_P and q_0 lies inside the convex hull of H_P; $\angle(q_0q_3h_0)$, since q_3 lies in Δ_4; $\angle(q_3h_0q_1)$, since q_2 lies in Δ_3. Clearly, $\overrightarrow{q_2q_1}$ is a γ-segment.

For $q_3 \in \Delta_5$, the angle $\angle(h_2q_3q_0)$ is positive and one easily checks that all the angles in the tour $\langle h_0, q_2, q_1, h_1, q_3, q_0, h_2, h_0\rangle$ are positive — see Figure 6.10(c): $\angle(q_1h_1q_3)$, $\angle(h_1q_3q_0)$, $\angle(q_3q_0h_2)$, since q_3 lies in Δ_5; $\angle(q_0h_2h_0)$, since h_2, h_0 follow each other clockwise on the convex hull of H_P and q_0 lies inside the convex hull of H_P; $\angle(h_2h_0q_2)$, since h_2, h_0 follow each other clockwise on the convex hull of H_P and q_2 lies inside the convex hull of H_P. Again, $\overrightarrow{q_2q_1}$ is a γ-segment.

(CASE B.4) $m = 4$.

Let s be the intersection point of the diagonals $\overline{h_0h_2}$ and $\overline{h_1h_3}$. Without loss of generality, let q_0 be contained in the triangle Δ with vertices s, h_3, h_2. Furthermore, we may assume that we have numbered q_1, q_2 and q_3 in a clockwise fashion such that q_1 lies left of $\overrightarrow{h_0q_0}$ and q_2 right of $\overrightarrow{h_0q_0}$. Both angles $\angle(h_0h_2h_3)$ and $\angle(h_2h_3h_1)$ are positive because of the clockwise order of the h_i.

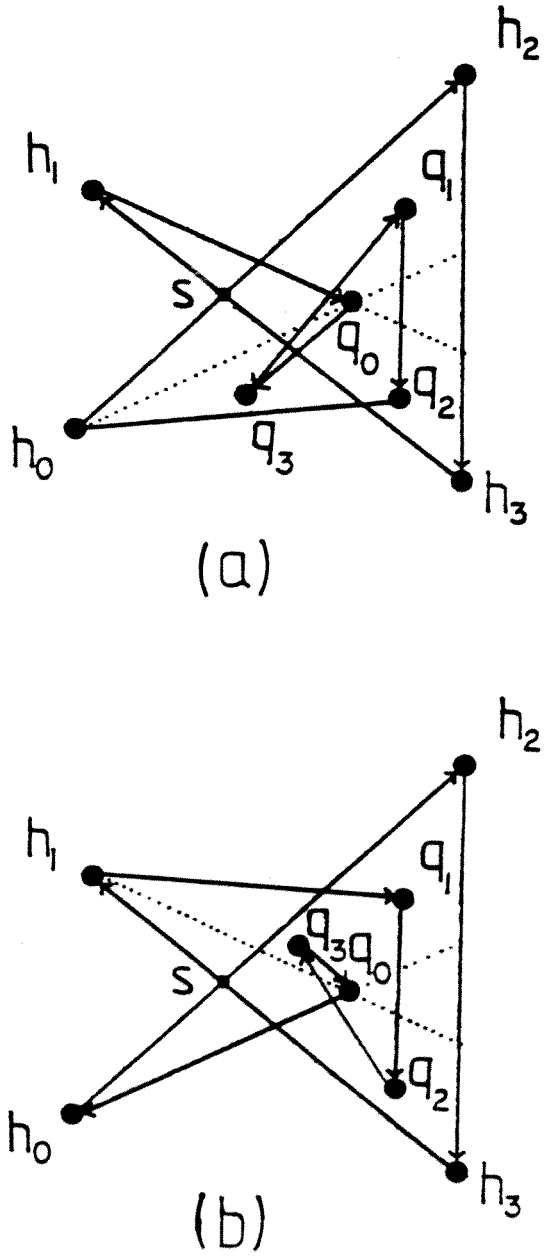


Figure 6.11: How to solve Case B.4

If q_3 lies to the right of $\overrightarrow{h_1q_0}$, the angle $\angle(h_1q_0q_3)$ is positive. It is not hard to check that all other angles in the tour $\langle h_0, h_2, h_3, h_1, q_0, q_3, q_1, q_2, h_0\rangle$ are positive — see Figure 6.11(a):

$\angle(h_3h_1q_0)$, since q_0 lies in Δ; $\angle(q_0q_3q_1)$, since q_3, q_1 follow each other clockwise on the convex hull of I_P and q_0 lies inside the convex hull of I_P; $\angle(q_3q_1q_2)$, by the way we numbered the q_i; $\angle(q_1q_2h_0)$, since $\angle(q_2h_0q_1)$ is positive by numbering of the q_i; $\angle(q_2h_0h_2)$, since q_0 lies in Δ and q_2 right of the line $\overrightarrow{h_0q_0}$.

If q_3 lies to the left of $\overrightarrow{h_1q_0}$, it follows that q_3 must lie to the left of $\overrightarrow{h_0q_0}$, otherwise $\langle q_1, q_2, q_3, q_1\rangle$ cannot run clockwise around q_0. By the same argument, q_2 must lie to the right of $\overrightarrow{h_1q_0}$, and q_1 to the left of $\overrightarrow{h_1q_0}$.

We easily check that all other angles in the tour $\langle h_0, h_2, h_3, h_1, q_1, q_2, q_3, q_0, h_0\rangle$ are positive - see Figure 6.11(b): $\angle(h_3h_1q_1)$, since q_0 lies in Δ and q_1 right of $\overrightarrow{h_1q_0}$; $\angle(h_1q_1q_2)$, since $\angle(q_2h_1q_1)$ is positive as q_1 and q_2 lie left and right of $\overrightarrow{h_1q_0}$; $\angle(q_1q_2q_3)$, by the way we numbered the q_i; $\angle(q_2q_3q_0)$, since q_2, q_3 follow each other clockwise on the convex hull of I_P and q_0 lies inside the convex hull of I_P; $\angle(q_3q_0h_0)$, since q_3 lies left of $\overrightarrow{h_0q_0}$; $\angle(q_0h_0h_2)$, since q_0 lies in Δ.

In both cases, $\overrightarrow{h_2h_3}$ is a γ-segment.

(CASE B.5+) $m \geq 5$.

See Figure 6.12. Consider the lines $l_1 := \overline{q_0q_1}$ and $l_2 := \overline{q_2q_3}$. The idea is to find two appropriate pairs of points $\langle h_{i_1}, h_{i_1+1}\rangle$ and $\langle h_{j_1}, h_{j_1+1}\rangle$ for which we can insert $\langle q_0, q_1\rangle$ and $\langle q_2, q_3\rangle$ into the convex tour $h_{i_1}, \ldots, h_{j_1}, \ldots, h_{i_1}$. To get a pseudoconvex

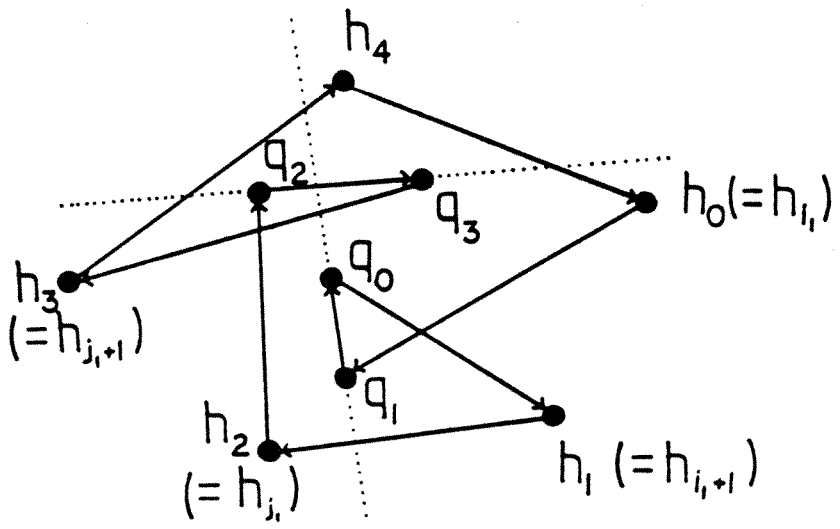


Figure 6.12: How to solve Case B.5+

---

tour, we have to make sure that these two pairs are disjoint. For easier notation, we write e_i for a pair $\{h_i, h_{i+1}\}$ of consecutive points of the hull H_P.

There are at least 3 different pairs e_{i_1}, e_{i_2} and e_{i_3} that are not separated by l_1; similarly, there are at least 3 pairs e_{j_1}, e_{j_2} and e_{j_3} that are not separated by l_2. Assume that $e_{i_2} \cap e_{j_k} \neq \emptyset$ for $k = 1, 2, 3$, i.e. the edge e_{i_2} intersects all three edges e_{j_1}, e_{j_2}, e_{j_3}. It is not hard to see that this can only happen if e_{j_1}, e_{j_2}, e_{j_3} are adjacent edges and (for appropriate numbering), $e_{i_2} = e_{j_2}$. It follows that there must be a j_k for which $e_{i_1} \cap e_{j_k} = \emptyset$; let this be the case for j_1.

Without loss of generality, let the angles $\angle(h_{i_1} q_0 q_1)$ and $\angle(h_{j_1} q_2 q_3)$ be positive. (Otherwise exchange the order of q_0 and q_1 or q_2 and q_3, respectively, in the following tour to make it pseudoconvex.) Then we can change the convex tour of H_P into $\langle h_{i_1}, q_0, q_1, h_{i_1+1}, \ldots, h_{j_1}, q_2, q_3, h_{j+1}, \ldots, h_{i_1} \rangle$. By choice of i_1 and j_1, the angles

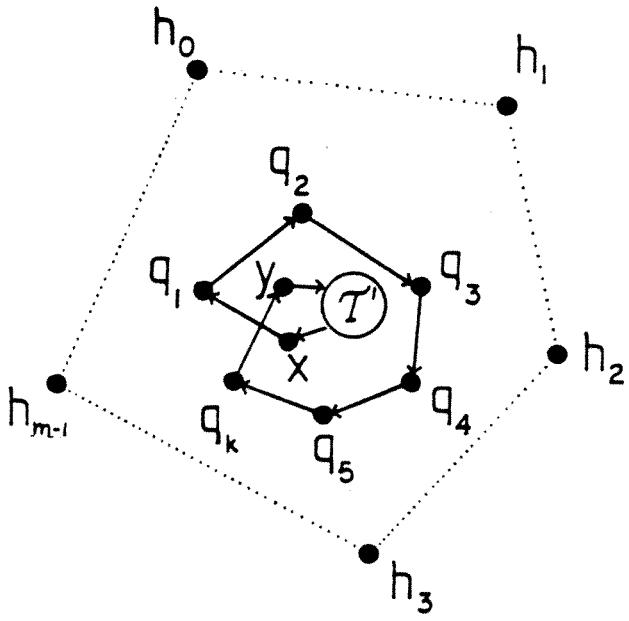


Figure 6.13: The setup for the induction step

$\angle(q_0 q_1 h_{i_1+1})$ and $\angle(q_2 q_3 h_{j_1+1})$ are both positive and the tour pseudoconvex. By construction, $\overrightarrow{q_0 q_1}$ is a γ-segment.

This concludes the proof.

□

6.3 Proving Theorem 6.1.2

With the help of Lemma 6.2.2 and Lemma 6.2.3, we can prove the main Theorem 6.1.2. As stated before, we use induction on the number of convex layers; Lemma 6.2.3 is used to start the induction, while Lemma 6.2.2 is the main tool for establishing the induction step.

As mentioned before, the main idea is to include the points of one layer at a time in their clockwise order into a tour of the inner points. (See Figure 6.13.) Since we use γ-segments for this purpose, we may have to rearrange the resulting tour in order to get a γ-segment that we can use for including the *next* layer. Since there are numerous cases, we use indirect argumentation to establish the validity of the induction step.

Let H_P with $|H_P| = m$ be the extremal points of P. Let the points of H_P be ordered as $h_0, \ldots, h_{m-1}$ when going clockwise around the hull H_P and let $I_P = P \setminus H_P$ denote the interior points of P. Let H_I with $|H_I| = k$ be the hull points of I_P and I_I the interior points of I_P.

If I_P satisfies the assumptions of Lemma 6.2.3, we are done: I_P allows a pseudoconvex tour with a γ-segment, so by Lemma 6.2.2, this tour can be extended into a pseudoconvex tour of P. So we only have to consider point sets P for which this is not the case.

Assume that $\overline{P}$ is the smallest of these sets that does not have *a pseudoconvex tour where the points of $H_{\overline{P}}$ follow each other in the order in which they appear around the hull.*

We observe that the set $I_{\overline{P}}$ with its interior points $I_{I_{\overline{P}}}$ cannot satisfy the assumptions of Lemma 6.2.3: One of the constructions in Lemma 6.2.3 would create a pseudoconvex tour of $I_{\overline{P}}$ with a γ-segment, therefore (by Lemma 6.2.2) we would get a pseudoconvex tour for $\overline{P}$ in which the elements of $H_{\overline{P}}$ follow each other in order. Using our minimality assumption on $\overline{P}$, we conclude that $I_{\overline{P}}$ allows a pseudoconvex tour $\mathcal{T}$, where the elements of $H_{I_{\overline{P}}}$ follow each other in order as $q_1, \ldots, q_k$.

In the following, we either locate a γ-segment in $\mathcal{T}$, find a way to rearrange $\mathcal{T}$ in order to create a γ-segment, or show that $\mathcal{T}$ has a structure that allows us to extend $\mathcal{T}$ into a tour of $\overline{P}$ without using γ-segments.

We distinguish three cases:

(CASE 1) $k \leq 4$:

Since the sum of angles in the polygon with vertices $q_1, \ldots q_k$ is $\frac{\pi}{2}$ or π, it is not hard to see that one of the segments $\overrightarrow{q_1q_2}$, $\overrightarrow{q_2q_3}$, $\overrightarrow{q_{k-1}q_k}$ is a γ-segment.

(CASE 2) $k \geq 6$:

If x lies left of $\overrightarrow{q_1q_3}$, the segment $\overrightarrow{q_1q_2}$ is a γ-segment, contradicting our assumption on $\overline{P}$. Similarly, if y lies left of $\overrightarrow{q_{k-2}q_k}$, the segment $\overrightarrow{q_{k-1}q_k}$ is a γ-segment. If neither is the case, we can rearrange the sequence of the q_i in $\mathcal{T}$ to get

$$\langle x, q_1, q_3, q_{k-1}, q_2, q_4, \ldots, q_{k-2}, q_k, y, \mathcal{T}, x \rangle,$$

thereby again creating a pseudoconvex tour with $\overrightarrow{q_3q_5}$ being a γ-segment. (See Figure 6.14)

(CASE 3) $k = 5$:

Let x be the predecessor of q_1 in $\mathcal{T}$ and y be the successor of q_5 in $\mathcal{T}$. Let x_- be the predecessor of x and y_+ the successor of y. Now see Figure 6.15. We note that:

- If the point x lies to the left of $\overrightarrow{q_1q_3}$, $\overrightarrow{q_1q_2}$ is a γ-segment in $\mathcal{T}$.
- If the lines $\overline{q_1q_2}$ and $\overline{q_3q_4}$ do not intersect to the left of $\overrightarrow{q_2q_3}$, $\overrightarrow{q_2q_3}$ is a γ-segment in $\mathcal{T}$.

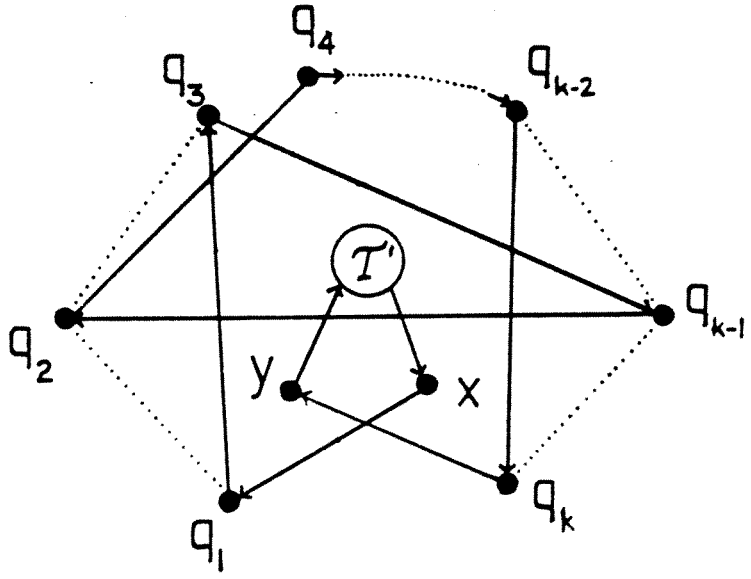


Figure 6.14: Rearranging $\mathcal{T}$ in order to get a γ-segment

---

- If the lines $\overline{q_2q_3}$ and $\overline{q_4q_5}$ do not intersect to the left of $\overrightarrow{q_3q_4}$, $\overrightarrow{q_3q_4}$ is a γ-segment in $\mathcal{T}$.
- If the lines $\overline{q_3q_4}$ and $\overline{q_5y}$ do not intersect to the left of $\overrightarrow{q_4q_5}$, $\overrightarrow{q_4q_5}$ is a γ-segment in $\mathcal{T}$.
- If the line from y through y_+ separates q_4 and q_5, $\overrightarrow{q_5y}$ is a γ-segment in $\mathcal{T}$.

Therefore assume for the rest of this proof that

- x lies right of $\overrightarrow{q_1q_3}$.
- The line $\overline{yy_+}$ does not separate q_4 and q_5.
- $\overline{q_1q_2}$ and $\overline{q_3q_4}$ intersect in a point s_{23} that lies to the left of $\overrightarrow{q_2q_3}$. q_2 lies between q_1 and s_{23}, and q_3 between s_{23} and q_4.

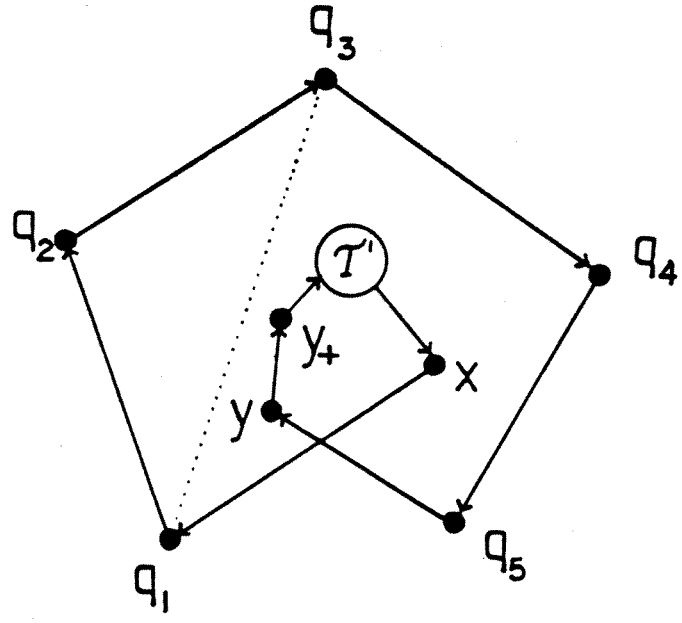


Figure 6.15: Potential γ-segments in $\mathcal{T}$

- $\overline{q_2q_3}$ and $\overline{q_4q_5}$ intersect in a point s_{34} that lies to the left of $\overrightarrow{q_3q_4}$. q_3 lies between q_2 and s_{34}, and q_4 between s_{23} and q_5.

- $\overline{q_3q_4}$ and $\overline{q_5y}$ intersect in a point s_{45} that lies to the left of $\overrightarrow{q_4q_5}$. q_4 lies between q_3 and s_{45}, and q_5 between s_{45} and y.

With the first two relations, it follows that the tour

$$\langle x_-, x, q_1, q_3, q_5, q_2, q_4, y, y_+, \ldots, x_- \rangle$$

is pseudoconvex with $\overrightarrow{q_3q_5}$ being a γ-segment, if y lies to the right of $\overrightarrow{q_2q_4}$. (See Figure 6.16.)

Therefore assume that y lies left of $\overrightarrow{q_2q_4}$. Let t_1 be the intersection of $\overline{q_2q_4}$ and $\overline{q_5y}$. Clearly, t_1 lies between q_2 and q_4.

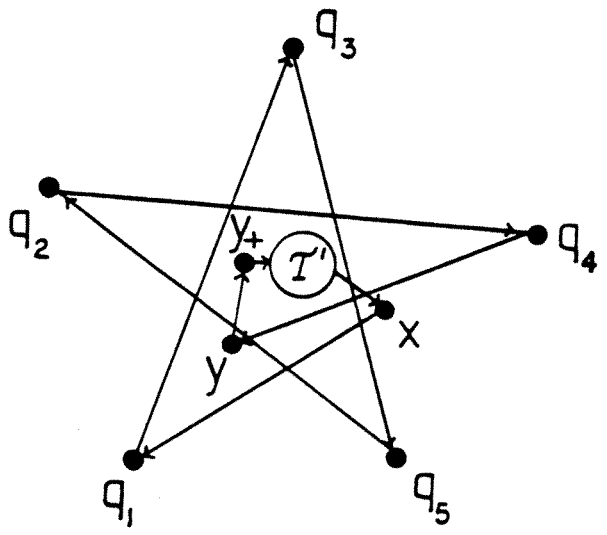


Figure 6.16: A pseudoconvex tour for y right of $\overrightarrow{q_2q_4}$

If none of the points $h_i \in H$ lies left of $\overrightarrow{q_2q_3}$ and left of $\overrightarrow{q_4q_5}$, the edge $\overrightarrow{q_3q_4}$ can be used to extend $\mathcal{T}$ to a pseudoconvex tour of $\overline{P}$ by inserting the points h_i in the way they appear clockwise around H: Assuming that h_0 is the first of the h_i that lies left of $\overrightarrow{q_4q_5}$, h_0 lies right of $\overrightarrow{q_2q_3}$ and h_{m-1} right of $\overrightarrow{q_4q_5}$.

This means that the tour

$$\langle x, q_1, q_2, q_3, h_0, \ldots, h_{m-1}, q_4, q_5, y, y_+, \ldots, x_- \rangle$$

is pseudoconvex, contradicting our assumption on $\overline{P}$. (See Figure 6.17.) We conclude that there is a point $h_{i_{34}}$ left of $\overrightarrow{q_2q_3}$ and left of $\overrightarrow{q_4q_5}$; without loss of generality assume that $h_{i_{34}+1}$ lies to the right of $\overrightarrow{q_2q_3}$.

Considering the position of $h_{i_{34}}$ relative to the points q_2, q_3, q_4, q_5 and s_{34} (see Figure 6.18), we see that $h_{i_{34}}$ lies left of $\overrightarrow{q_2q_4}$ and left of $\overrightarrow{q_3q_5}$.

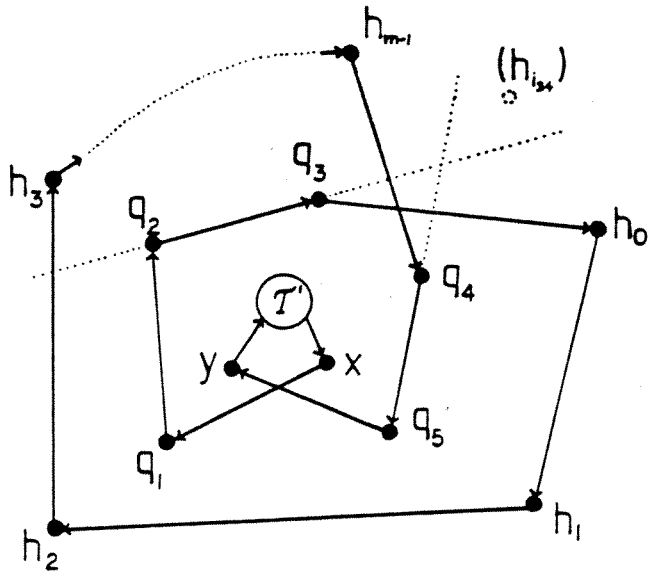


Figure 6.17: There must be a point $h_{i_{34}}$ left of $\overrightarrow{q_2q_3}$ and right of $\overrightarrow{q_4q_5}$

Considering the position of $h_{i_{45}}$ relative to q_3, q_4, q_5, t_1 and s_{45}, we see that we have a similar situation as for $h_{i_{34}}$ relative to q_2, q_3, s_{34}, q_4 and q_5 — see Figure 6.19. Thus, $h_{i_{45}}$ lies right of $\overrightarrow{t_1q_4}$, i.e. right of $\overrightarrow{q_2q_4}$. This implies $h_{i_{34}} \neq h_{i_{45}}$.

Next consider the intersection of $\overline{q_5y}$ with the boundary of the triangle $\Delta = (q_4s_{23}q_2)$, as shown in Figure 6.20. Since $\{s_{45}\} = \overline{q_3q_4} \cap \overline{q_5y}$ lies right of $\overrightarrow{q_4q_5}$, the point s_{45} cannot belong to the triangle; so $\overline{q_5y}$ must intersect $\overline{q_2s_3}$ in a point t_2 between q_2 and s_3. This means that $h_{i_{45}}$ lies right of $\overrightarrow{t_2s_3}$, i.e. right of $\overrightarrow{q_1q_2}$.

Finally, $h_{i_{34}+1}$ must lie right of $\overrightarrow{q_1q_2}$ (see Figure 6.21): By construction, $h_{i_{34}+1}$ must lie between $h_{i_{34}}$ and $h_{i_{45}}$ when going clockwise around the hull H; if we do cross $\overline{q_1q_2}$ on this way, this cannot happen after we cross $\overline{q_2q_3}$. This means that $h_{i_{34}+1}$ lies on the same side of $\overrightarrow{q_1q_2}$ as $h_{i_{45}}$, i.e. on the right.

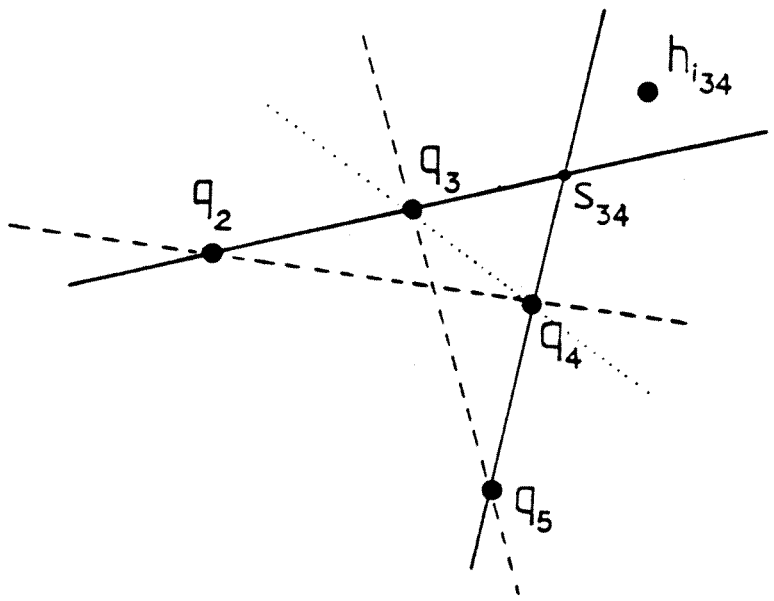


Figure 6.18: $h_{i_{34}}$ lies left of $\overrightarrow{q_2q_4}$ and left of $\overrightarrow{q_3q_5}$

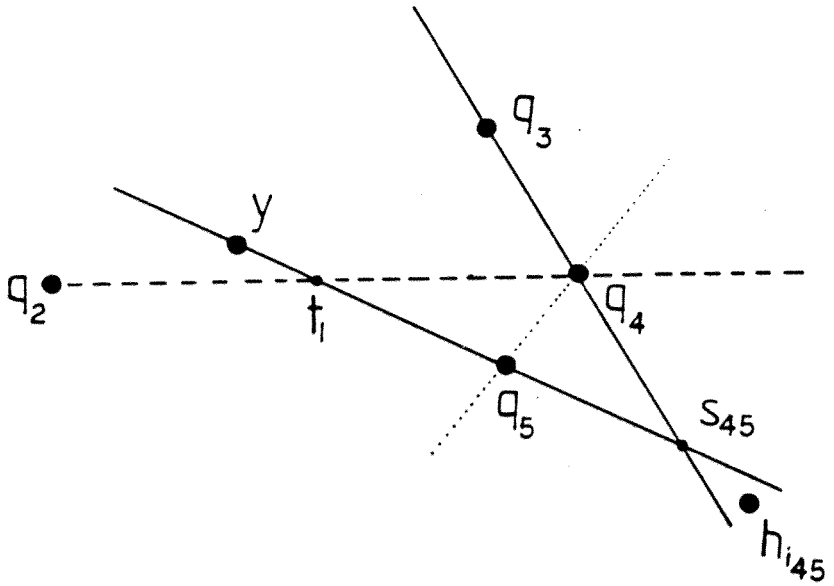


Figure 6.19: $h_{i_{45}}$ lies right of $\overrightarrow{q_2q_4}$

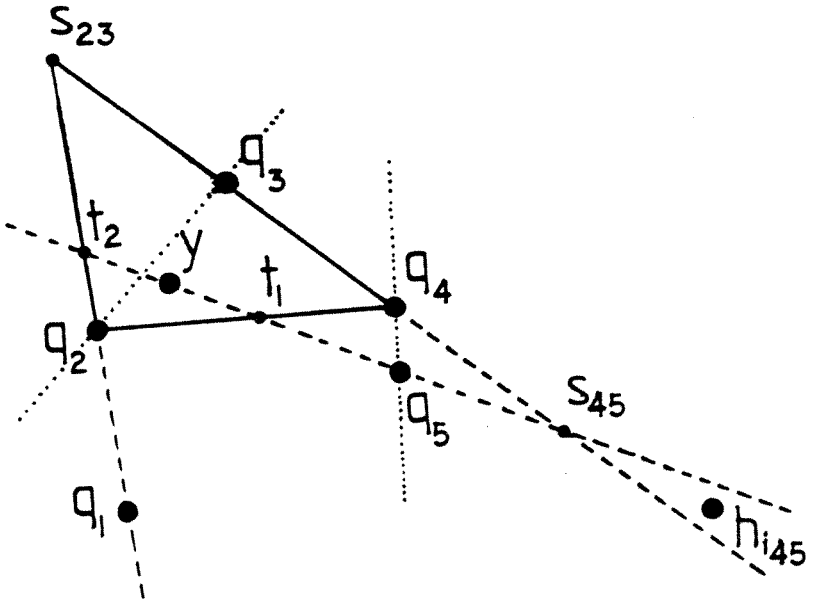


Figure 6.20: $h_{i_{45}}$ lies right of $\overrightarrow{q_1 q_2}$

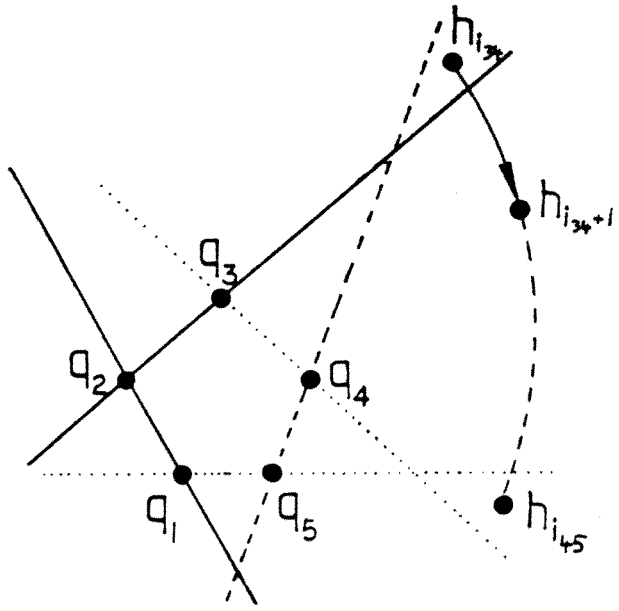


Figure 6.21: $h_{i_{34}+1}$ lies right of $\overrightarrow{q_1 q_2}$

But now all the angles in the tour

$$\langle x, q_1, q_2, h_{i_{34}+1}, \ldots, h_{i_{34}}, q_5, q_3, q_4, y, y_+ \ldots, x_-, x \rangle$$

are positive — see Figure 6.22:

The angle $\angle(q_1 q_2 h_{i_{34}+1})$, as stated; $\angle(q_2 h_{i_{34}+1} h_{i_{34}+2})$, since $h_{i_{34}+1}, h_{i_{34}+2}$ follow each other clockwise on the convex hull of $\overline{P}$ and q_2 lies inside the convex hull of $\overline{P}$; $\angle(h_{i_{34}-1} h_{i_{34}} q_5)$, since $h_{i_{34}-1}, h_{i_{34}}$ follow each other clockwise on the convex hull of $\overline{P}$ and q_5 lies inside the convex hull of $\overline{P}$; $\angle(h_{i_{34}} q_5 q_3)$, as stated; $\angle(q_5 q_3 q_4)$, since $\angle(q_3 q_4 q_5)$ is positive; $\angle(q_3 q_4 y)$, since q_3, q_4 follow each other clockwise on the convex hull of $I_{\overline{P}}$ and y lies inside the convex hull of $I_{\overline{P}}$; $\angle(q_4 y y_+)$, since $\angle(q_5 y y_+)$ is positive and $\overline{y y_+}$ does not separate q_4 and q_5.

We conclude that $\overline{P}$ *does* have a pseudoconvex tour with the elements of the hull $H_{\overline{P}}$ following each other in order.

This completes the proof.

□

The time complexity $O(n^2)$ claimed above follows by constructing the so-called onion decomposition of a planar point set into convex layers in $O(n \log n)$ time (see [11, 24]). Starting from the inside, at each of the $O(n)$ stages, we include another layer of the decomposition into the tour; detecting γ-segments (or the constellation used at the end of the preceding proof) can be done in $O(n)$ time, resulting in the above overall complexity.

If the points in P are not in general position, there may be no pseudoconvex tour that encounters each of the vertices exactly once. Easy examples for this situation

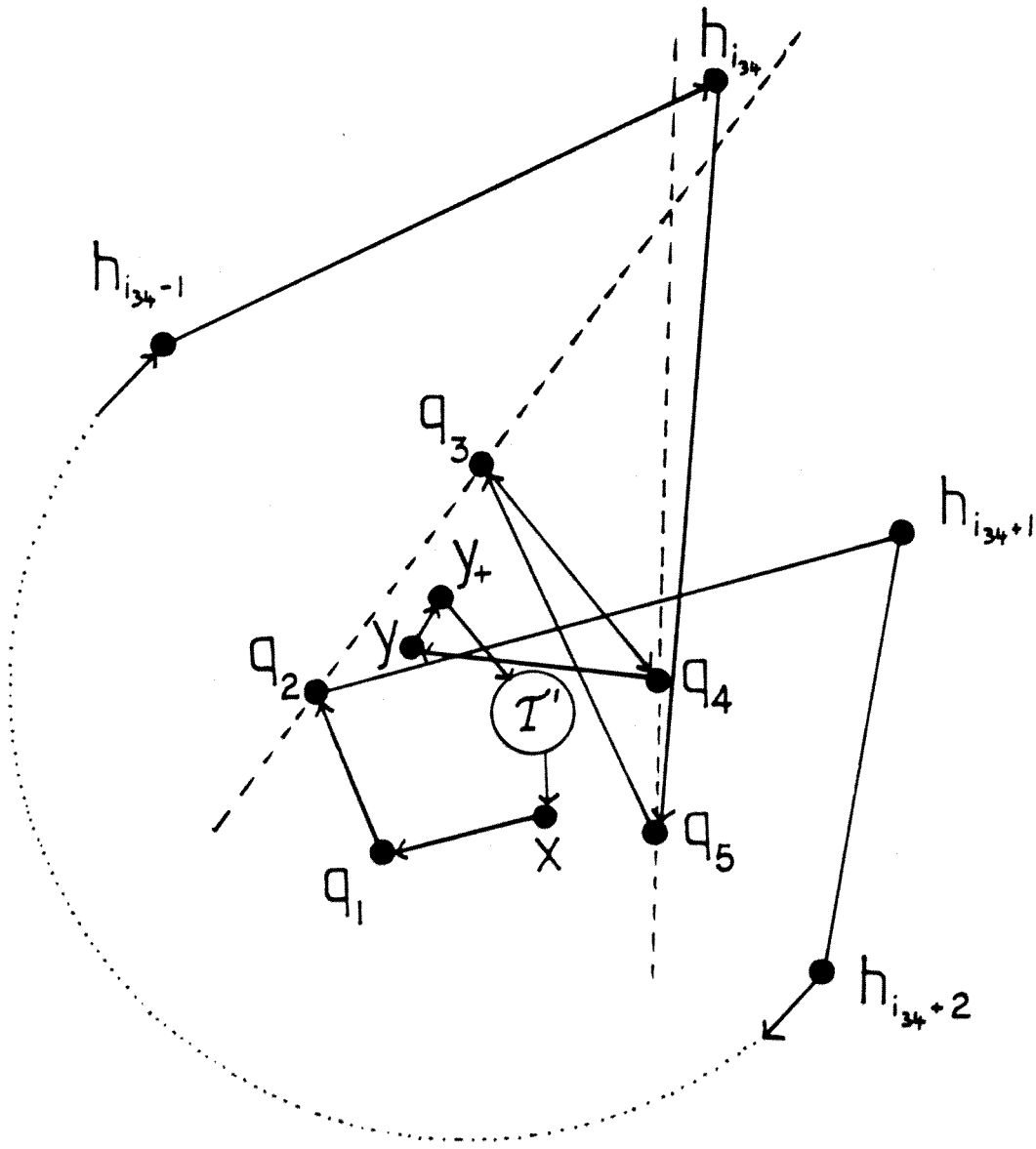


Figure 6.22: The final pseudoconvex tour

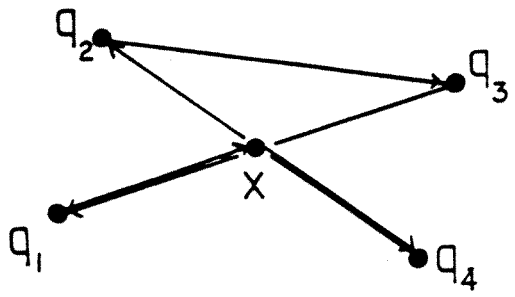


Figure 6.23: A pseudoconvex tour that encounters a vertex more than once

arise if P is a set of collinear points or if P consists of the four vertices of a rectangle together with a fifth point at the intersection of the diagonals.

We may choose to permit a vertex to be "run over" by other edges — see Figure 6.23, where we have the pseudoconvex tour $\langle x, q_4, q_2, q_3, q_1, x \rangle$. A tour that satisfies the angle-restriction but has vertices that are contained in the interior of some edge $\overline{p_i p_{i+1}}$ is called *weakly feasible.* (In a weakly feasible tour every vertex may only be used once as an end point of an edge.) A tour that satisfies the angle-restriction and encounters each vertex exactly once is called *strongly feasible.*

If we are content with weakly feasible tours, it is not hard to check that the above steps of the proof remain valid even if the points are not in general position. (All regions in the argumentations include their boundaries if we allow the use of the angles 0 and π.)

Corollary 6.3.1 *Let P be any set in the plane with $|P| \geq 5$ in general position. Then P has a strictly feasible pseudoconvex tour and all angles lie in the open interval $(0, \pi)$.*

6.4 Pseudoconvex Paths

Any nondegenerate convex tour is a pseudoconvex tour that is simple. While the existence of a convex tour is a very special property of a point set, we have seen that any set of at least 5 points has a pseudoconvex tour. If we relax the question for a *Hamiltonian cycle* with nonnegative angles to the question for a *Hamiltonian path* with nonnegative angles, the situation becomes considerably easier as the simple proof for the following Theorem 6.4.1 shows.

Theorem 6.4.1 *Any set P of n points allows a nonintersecting pseudoconvex spanning path.*

Proof: See Figure 6.24. Choose any point p_0 on the convex hull of P to be the starting point and remove p_0 from P. Find some line through p_0 that does not intersect the convex hull of P and turn it clockwise until it hits the first point in P. This point is the next point in our path, we remove it from P and repeat the procedure. The resulting path is always be intersection-free. (It is not hard to see that we follow the convex layers of P in an inward spiral.)

□

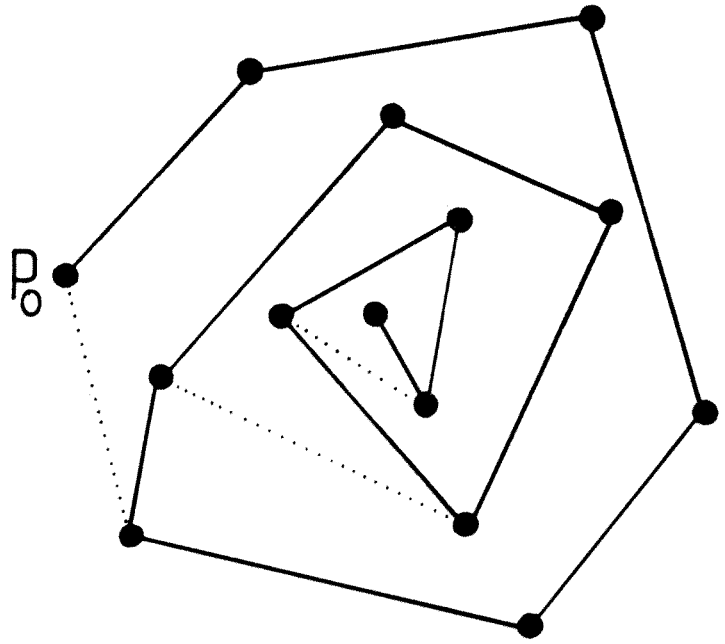


Figure 6.24: Any point set has a pseudoconvex spanning path

Chapter 7

Angle-Restrictions

The research for this chapter was done in cooperation with Gerhard Woeginger.

7.1 Introduction

In this chapter, we generalize the question for pseudoconvex tours to other angle sets. Let A be a set of feasible angles, $A \subseteq (-\pi; +\pi]$. For a set P of $n \geq 3$ points in the Euclidean plane, consider the tours $\langle p_1, p_2, \ldots, p_n, p_1 \rangle$. We call a tour *weakly feasible with respect to* A, or an *A-tour* for short, if all the angles $\angle(p_i p_{i+1} p_{i+2})$, $0 \leq i \leq n-1$, are elements of the angle set A. As in the case of pseudoconvex tours (p. 112), we call a feasible A-tour *strongly feasible*, if none of the segments $\overline{p_i p_{i+1}}$ contains another point of P, $0 \leq i \leq n-1$.

This yields the following problem:

ANGLE-RESTRICTED TOUR (ART) *Given a set $A \subseteq (-\pi; +\pi]$ of angles. The problem "Angle-Restricted Tour" (ART) is to decide whether a set P of n points in the Euclidean plane allows an A-tour, i.e. a closed directed tour consisting of straight line segments, such that all angles between consecutive line segments are from the set A.*

We can think of this as the question of deciding whether a machine with restricted mobility is able to make a roundtrip through a given set of points. For this motivation, Culberson and Rawlins [17] have discussed the problem of finding a simple polygon with a given sequence of angles.

Restricted orientations have also been examined because of their relevance for computer graphics and VLSI design — see Rawlins and Wood [77], [78], [79], [80], Schuierer [90] and Widmayer, Wu and Wong [93]. We discuss a different aspect of restricted orientations and convexity in Chapter 8.

While these orientations are fixed relative to the surrounding space, an angle-restricted tour has to deal with restricted *relative* orientations. Thus, angle-restricted tours can also be considered as an interesting variation on restricted orientations (Wood [92]).

As in the case of pseudoconvex paths, a (weakly) feasible spanning path with respect to some angle set A is defined like a (weakly) feasible spanning tour with the only difference that a path is not closed.

A trivial necessary condition on some angle set A to allow point sets with A-tours is the following: There exist angles $\alpha_1, \ldots, \alpha_k \in A$, k nonnegative integers

$c_1, \ldots, c_k$ that are not all zero and an arbitrary integer c_{k+1}, such that

$$c_1\alpha_1 + c_2\alpha_2 + \ldots + c_k\alpha_k = c_{k+1} \cdot 2\pi$$

holds. Otherwise, a tour could never close. Clearly, this condition is not necessary for the existence of A-paths.

One interesting question is whether there is any simple connection between the size of A and the complexity of ART. It is not hard to see that ART is easy when $|A| = 1$: Assume $A = \{\alpha_1\}$. Then the existence of a strongly feasible A-tour for some point set P, $|P| = n$, can be checked in the following way. It is easy to see that the starting segment of an A-tour completely determines all the following segments in the tour. Hence, we simply check all $n - 1$ possible starting segments emanating from some fixed point in P.

The problem becomes considerably harder if $|A| = 2$. As shown in the following Section 7.2, there exist angle sets with two elements for which determining the existence of A-tours is NP-complete and there exist two-element angle sets for which determining the existence of A-tours can be done in polynomial time.

In the following sections, we discuss the situation for the set of obtuse angles $A^{\text{obtu}} = \{\alpha \mid \alpha < -\pi/2 \text{ or } \pi/2 < \alpha\}$; the set of acute angles $A^{\text{acut}} = \{\alpha \mid -\pi/2 < \alpha < \pi/2\}$; the set of orthogonal angles $A^{\text{orth}} = \{-\pi/2, +\pi/2, \pi\}$. Tours or paths that are weakly feasible with respect to A^{acut}, A^{obtu}, or A^{orth} are called *orthogonal, acute,* or *obtuse* tours or paths, respectively.

7.2 Orthogonal Tours

In this section we treat the problem of detecting strongly feasible orthogonal tours. Let $A_1 = \{-\pi/2, +\pi/2, \pi\}$, $A_2 = \{\pi/2, \pi\}$ and $A_3 = \{-\pi/2, +\pi/2\}$. We prove that detecting strongly feasible A_1- and A_2-tours is NP-complete while finding A_3-tours can be done in polynomial time.

We start with the NP-completeness result on A_1-tours. A related result was derived by Rappaport [75, 76] who gave an NP-completeness proof for the case of A_1-tours that are *not allowed to be self-intersecting*.

Theorem 7.2.1 *Given a set P of n points in the plane, deciding whether P allows a strongly feasible A_1-tour is NP-complete.*

Proof: We show that the NP-complete problem HAMILTONIAN CYCLE IN GRID GRAPHS (Itai, Papadimitriou, Szwarcfiter [46]) can be reduced to detecting A_1-tours. A grid graph $G = (V, E)$ has a set V of n integer grid points. There is an edge between two vertices if and only if the corresponding points are at distance 1.

In our reduction, we first partition the points in V into horizontal classes. Two points belong to the same horizontal class if they have identical y-coordinates and if they are connected by a path that uses only edges of the graph with this y-coordinate. The points in the i-th horizontal class are shifted by the vector $(0, 1/4^i)$. (The reason for this choice is to give all horizontal classes distinct y-coordinates.) In an analogous way, we define vertical classes and shift them by vectors $(1/4^i, 0)$. Clearly, all points adjacent in the grid graph maintain their horizontal (vertical) connections. All points not adjacent in the grid graph either keep the distinctness

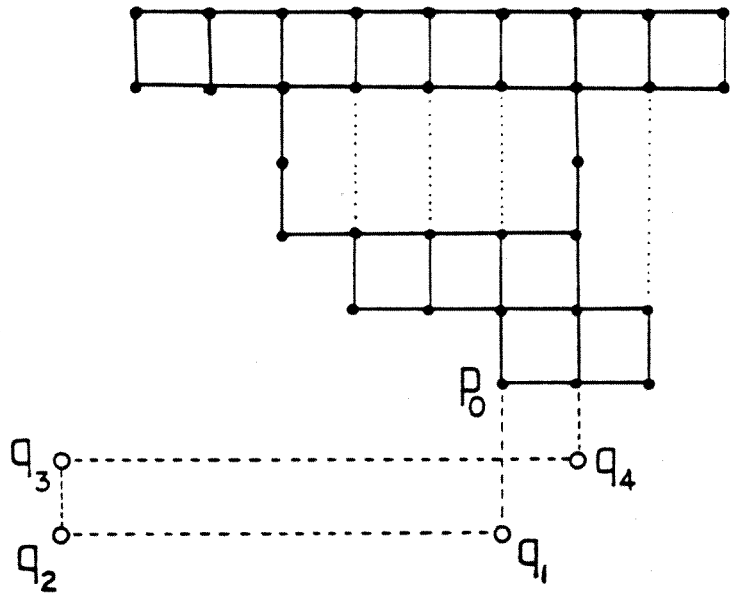


Figure 7.1: Shifting vertical classes removes the dotted "crossovers"

---

of their vertical (horizontal) coordinates, or these coordinates are made distinct by the shifts.

Finally, we consider the leftmost point p_0 with smallest y-coordinate. Let p_1 be its horizontal neighbour. As shown in Figure 7.1, we add the four points q_1, q_2, q_3, q_4 to the point set, such that q_1 has only neighbours p_0 and q_2, q_2 has only neighbours q_1 and q_3, q_3 has only neighbours q_2 and q_4, q_4 has only neighbours q_3 and p_1, and q_2 becomes lowermost and leftmost. Clearly, there can be no orthogonal tour of the point set where q_2 is not adjacent to both q_1 and q_3, so this arrangement forces every strongly feasible A_1-tour to be axes-parallel. Now it is easy to see that a Hamiltonian Cycle in the graph exactly corresponds to a strongly feasible A_1-tour in the shifted point set.

□

An alternative to adding the points q_1, q_2, q_3, q_4 is provided by Theorem 7.2.7: Assuming that the grid graph is connected, the only orientation for which we can connect all vertices in *any* manner is axes-parallel.

Next, we describe the NP-completeness proof for detecting A_2-tours.

Theorem 7.2.2 *Given a set P of n points in the plane, deciding whether P allows a strongly feasible A_2-tour is NP-complete.*

Proof: We show that the problem of detecting an *axes-parallel* A_2-tour is NP-complete. The claim then follows by adding some extra points as in the proof of the above theorem. We show the NP-completeness by reducing Hamiltonian Cycle in cubic directed graphs to it. (See Plesńik [72] or Garey and Johnson [34], p.199.) So let $D = (V, E)$ be some cubic, directed graph with $|V| = n$. As in Chapter 3 (see section 3.1, pp. 23ff.), we may assume the following properties: All vertices have either indegree or outdegree two, partitioning V into *in-vertices* and *out-vertices.* This partition induces a bipartition of the graph and the edges are either *mandatory* (being the only edge leaving or entering for both its end points) or *optional* (being one of two edges leaving or entering for both end points.) Furthermore, the optional edges form a set of disjoint (undirected) cycles in D. Let $C_1, \ldots, C_z$ denote the set of these cycles.

We construct a point set P_D that allows an axes-parallel A_2-tour if and only if G has a Hamiltonian Cycle.

In a first step, we choose n disjoint parallel boxes in the plane as shown in Figure 7.5. Let v_1 be some out-vertex in C_1 and $c_1 = |C_1|$. Let $v_2, \ldots, v_{c_1}$ be the

other vertices of C_1 in an order in which they appear when running through C_1. Then assign the first c_1 boxes to $v_1, \ldots, v_{c_1}$. In the same manner, assign the other boxes to the vertices in $C_2 \ldots, C_z$.

Next, the optional edges are represented by horizontal edges connecting the appropriate boxes, such that all edges get different vertical coordinates. Clearly, the boxes for v_1 and v_{c_1} have their two edges lying on the same side ("type A"), while any other box has one edge on each side ("type B"). For a box of type B, we refer to the appropriate adjacent optional edges as the "right" and the "left" optional edge. After placing the optional edges adjacent to v_1 at any vertical level, we run through the boxes for the vertices $v_2, \ldots, v_{c_1-1}$ of type B to place the other optional edges of C_1. The right edge for v_i is placed *below* the left edge whenever i is even — i.e. v_i an in-vertex. If i is odd, i.e. v_i an out-vertex, we place the right optional edge *above* the left optional edge.

After following this procedure for all C_i, we represent each mandatory edge by a rectilinear path consisting of two vertical and one horizontal line segment, such that the appropriate boxes are connected. Again, we make sure that all vertical coordinates are distinct.

Now consider some fixed out-vertex v with, and let Bo_v be its corresponding box. Depending on whether v is of type A or B, we distinguish:

(A) If both outgoing edges leave the box on the same side, we place a stretched copy of the set A as depicted in Figure 7.2 into the box Bo_v. This is done in such a way that the heights of OUT1 and OUT2 coincide with the heights of the line segments corresponding to the two outgoing edges. If (u, v) is the

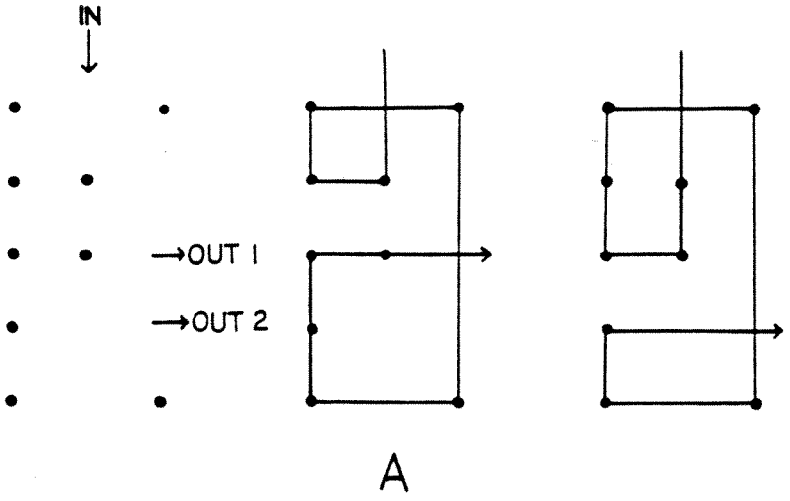


Figure 7.2: Box type A: OUT1 and OUT2 leave on the same side.

---

mandatory edge adjacent to v, reroute it into IN by placing one point at each right turn of (u, v) and three points at each left turn of (u, v).

(B) If both outgoing edges leave the box on different sides we use the point set B depicted in Figure 7.3 in a similar manner as A in (A).

Finally, if the vertex v is an in-vertex, we simply use reflected versions of the two point sets in (A) (Figure 7.2) and (B) (Figure 7.3). Inputs become outputs and vice versa. We indicate the resulting point sets by $\overline{A}$ and $\overline{B}$ in Figure 7.5.

To give some intuition on our constructions, we state the following observations. For an example, see Figure 7.5.

(a) Groups of points in different boxes have different coordinates with the exception of the three in- and outputs. Consequently, a strongly feasible tour enters

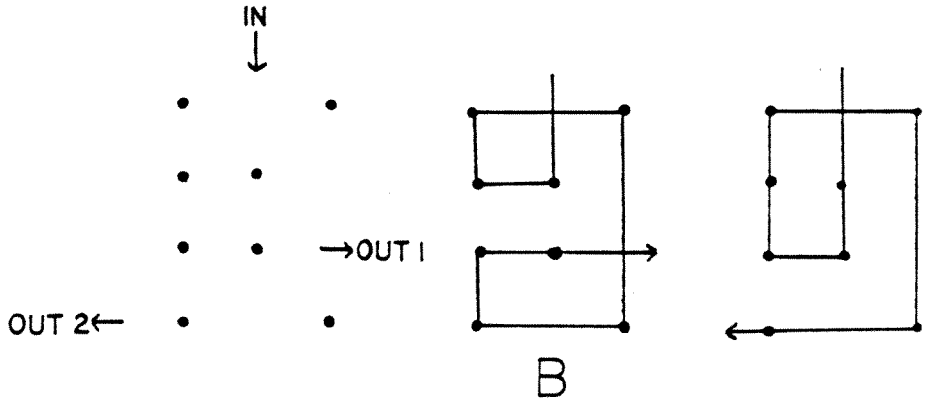


Figure 7.3: Box type B: OUT1 and OUT2 leave on different sides.

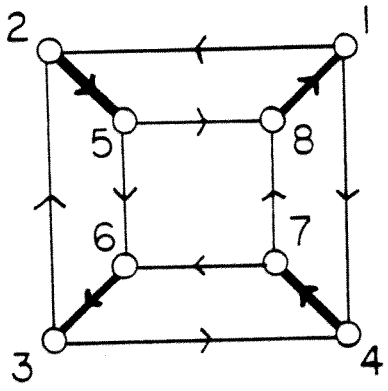


Figure 7.4: A cubic digraph D

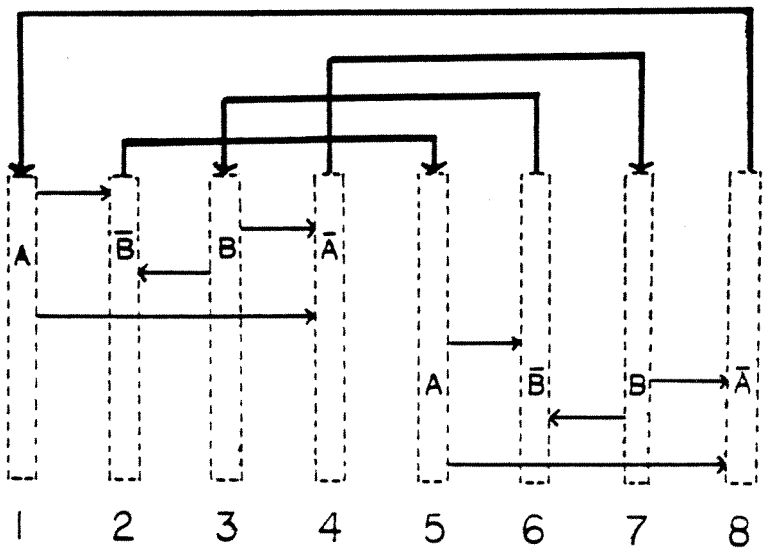


Figure 7.5: The representation for D

a box, run through all points, leave the box, and it never enters it again (otherwise it would have to use at least two inputs and at least two outputs).

(b) The point set used to reroute a mandatory edge *must* be used in each strongly feasible tour, since there is no possibility for a strongly feasible tour to enter through OUT1 and to leave through OUT2 (or to enter through OUT2 and to leave through OUT1) and to visit all points inside the box at the same time.

(c) If we reflect an A_3-path from a to b, it becomes a strongly feasible A_3-path from b to a. Hence, for in-vertices, the paths in the figures are strongly feasible paths from OUT to IN.

Now assume we are given a Hamiltonian Cycle for G. We want to construct a strongly feasible A_2-tour visiting all points in the point set P_G. First, we draw

in all line segments corresponding to edges in G that belong to the Hamiltonian Cycle. Next we consider a box corresponding to some out-vertex. We know at which positions the box is entered and at which positions it is exited again. We include the point set that was used to reroute the input from IN to the height of the ingoing edge. Then we use the corresponding segment sets shown in Figures 7.2 and 7.3 to connect IN to one of the outgoing segments. (In case of an in-vertex we use the reflected versions). Clearly, we end up with an A_2-tour.

Finally, assume that the constructed point set allows an A_2-tour $\mathcal{T}$ and consider some out-vertex v. The only possibilities for $\mathcal{T}$ to enter and to leave the box corresponding to v is via the line segments OUT1, OUT2 and IN corresponding to the three incident edges (all other points have distinct x- and y-coordinates). It is easily checked that the tour cannot visit all points inside the box in a valid way if it enters through OUT1 or OUT2 or if it leaves through IN. Hence, the tour must enter all boxes at places corresponding to ingoing edges and it must leave all boxes at places corresponding to outgoing edges. Contracting the subpaths of the tour in every single box yields a strongly feasible Hamiltonian Cycle in G. Our proof is complete.

□

The polynomial time result on A_3-tours is a direct implication of the following proposition due to O'Rourke [65].

Proposition 7.2.3 (O'Rourke [65]) *For a set P of n points in the Euclidean plane, it can be checked in $O(n \log n)$ time whether P allows a strongly feasible A_3-tour such that all segments in the tour are* axes-parallel *to the x- or to the y-axis.*

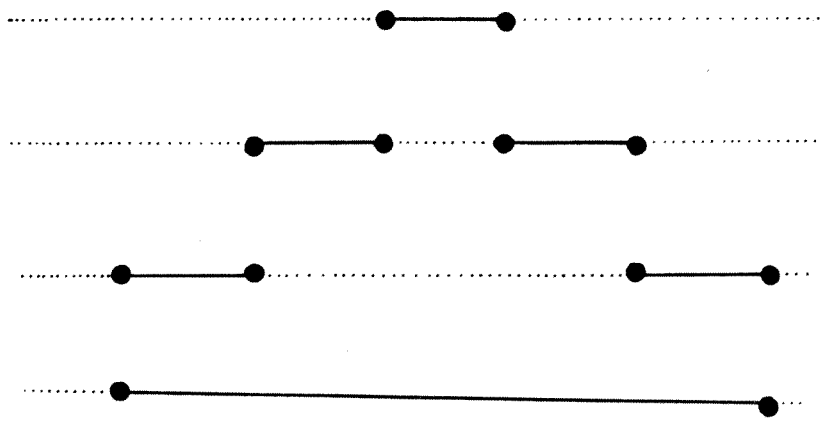


Figure 7.6: Finding strongly feasible A_3-tours

Furthermore, there can be at most one strongly feasible A_3-tour for any orientation of the axes.

This yields the following Corollary 7.2.4. For the sake of completeness, we describe the proof idea.

Corollary 7.2.4 *For a set P of n points in the Euclidean plane, it can be checked in $O(n^2 \log n)$ time whether P allows a strongly feasible A_3-tour.*

Proof: See Figure 7.6. Consider some fixed line segment s in some valid A_3-tour $\mathcal{T}$. Then all other line segments in $\mathcal{T}$ are either parallel or normal to s. That means we may assume an underlying coordinate system such that all segments in $\mathcal{T}$ are parallel to one of the axes.

As a first step we determine a set of $O(n)$ possible orientations for such a coordinate system: In $O(n^2 \log n)$ time, sort the $O(n^2)$ line segments into equivalence classes, where two segments belong to the same class if and only if they are parallel or orthogonal. We only have to check the orientations corresponding to classes with at least n segments, of which there are $O(n)$. In the rest of the proof, we show how to test the existence of an A_3-tour in $O(n \log n)$ time, if the coordinate system has been fixed. This gives us an overall time of $O(n^2 \log n)$, and our proof is complete.

Thus, we assume that the coordinate system has been fixed and we consider some horizontal line that contains at least one point of P. Let $p_1, p_2, \ldots, p_k$ be the points of P on this line sorted from left to right. Assume there exists an A_3-tour $\mathcal{T}$ that visits all points of P. In this case the point p_1 has two neighbors on the polygonal line $\mathcal{T}$, one of them vertical, the other horizontal. The only possible horizontal neighbor is the point p_2. Hence, we may connect p_1 to p_2. Similarly, the only possible horizontal neighbor of p_3 is p_4, the only possible horizontal neighbor of p_5 is p_6 and so on. This yields $k/2$ segments that must be part of $\mathcal{T}$. We repeat this process for every horizontal and every vertical line that contains at least one point of P. If we meet some line that contains an *odd* number of points, we immediately stop, as a tour cannot exist for the given orientation.

Therefore we end up with a set of n line segments, such that every point of P is incident to exactly two segments. If the segments form a single closed polygonal line we have constructed a (strongly feasible) A_3-tour, otherwise none exists for the orientation considered.

□

We can extend the result to not necessarily strongly feasible A_3-tours:

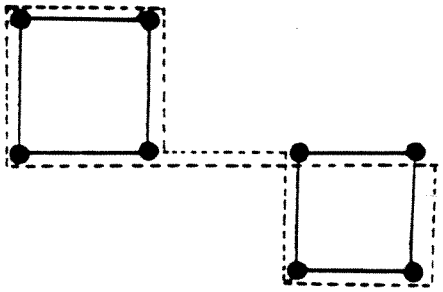


Figure 7.7: Merging subtours to find weakly feasible A_3-tours

---

Theorem 7.2.5 *Let P be a set of n points in the Euclidean plane. Then we can decide in $O(n^2 \log n)$ time whether P allows a weakly feasible A_3-tour.*

Proof: If a solution exists for some fixed coordinate system, the method described in Corollary 7.2.4 either detects a strictly feasibly tour (and we are done) or it finds a partition of P into disjoint sets $P_1, \ldots, P_k$ such that each P_i allows a weakly feasible A_3-tour.

Assume there exist two points $p_i \in P_i$ and $p_j \in P_j$ such that p_i and p_j have identical x-or y-coordinate. In this case, there exists an axes-parallel line containing a segment of the tour for P_i and a segment of the tour for P_j, and it is easy to see that the A_3-tours for P_i and P_j can be merged into a weakly feasible A_3-tour for $P_i \cup P_j$. (See Figure 7.7.)

We may repeat this merging procedure, so the problem reduces to determining whether there exists a coordinate system such that P is connected with respect to axes-parallel connections. This is easily done in $O(n^2 \log n)$ time by checking $n-1$ coordinate systems and sorting the corresponding x- and y-coordinates.

□

In [65], O'Rourke has stressed the aspect of uniqueness of strongly feasible A_3-tours. He motivates the question as a pattern recognition problem: We are trying to find a simple rectilinear polygon for a given set of vertices. For his uniqueness result, he assumes that the orientation of the polygon edges is axes-parallel in some given coordinate system.

If we have to find the underlying rectilinear polygonal shape of a given a set of vertices in the plane, we cannot necessarily assume knowledge of an underlying coordinate system. Interestingly enough, we can show that O'Rourke's uniqueness result remains valid even if no orientations are known in advance. This nicely illustrates our point about the relation between relative and fixed restricted orientations.

Definition 7.2.6 *Given a subset V of the Euclidean plane $\mathbb{R}^2$. For any orientation O of a Cartesian coordinate system, we define the* O-orthogonality graph $G_O(V) = (V, E)$ *as follows: Two vertices v_1 and v_2 are connected by an edge, if and only if they have the same x- or y-coordinate.*

Theorem 7.2.7 *Given a set P of n vertices in the Euclidean plane that all have rational coordinates. Then there is at most one orientation O for which $G_O(P)$ is connected.*

Proof: Suppose that we have two such orientations O_1 and O_2. Then we can write the direction of the x-axis in O_2 as (a, b) with $a, b \neq 0$ in O_1-coordinates. Since all the coordinates are rational, we may assume that a and b are relatively prime integers.

Furthermore we can assume that all O_1-coordinates of the given points are relatively prime integers. This implies that there are two points of the same y-coordinate that have a positive distance in x-direction that is not a multiple of $a^2 + b^2 > 1$, say k.

Since these two points can be connected by an axis-parallel path with orientation (a, b) and all endpoints of segments being grid points, the vector $(k, 0)$ must be an integer combination of the vectors (a, b) and $(-b, a)$, i.e. the system

$$\begin{aligned} ax - by &= k \\ bx + ay &= 0 \end{aligned}$$

must have an integer solution (x, y) — which is equivalent to

$$\begin{aligned} (a^2 + b^2)x &= ka \\ (a^2 + b^2)y &= kb. \end{aligned}$$

Since a and b are relatively prime, $a^2 + b^2$ and a are, so a must divide x. This implies that k is a multiple of $(a^2 + b^2)$ — a contradiction.

□

7.3 Acute and Obtuse Angles

In this section, we discuss tours and paths for acute or obtuse angles. The set A^{acut} can be interpreted as being required to make "sharp" turns at every vertex; conversely, A^{obtu} corresponds to the situation where we have to avoid too sudden changes of direction. Wolkowicz [95] pointed out the possibility that problems of

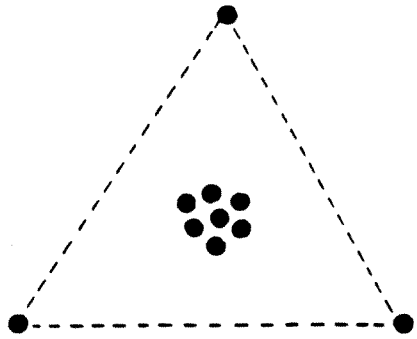


Figure 7.8: A point set without obtuse spanning paths

---

the latter kind might be interesting for the planning of aerial surveys, where a plane has to fly over a certain number of check points.

Mathematically, the feasible angle sets for pseudoconvex, acute and obtuse tours can all be represented as one half of the unit circle $S^1 = \mathbf{R} \bmod 2\pi$. The following statements, however show that the situation for each of the three angle sets is quite different from that for the others.

Example 7.3.1 *Obtuse spanning paths do not exist for all point sets. To see this, consider the three vertices of an equilateral triangle together with $n-3$ points in the interior of the triangle. (See Figure 7.3.1.) Any spanning path must contain one of the corner points somewhere in its middle. But then the corner point together with its two neighbors in the path form an acute angle.*

□

Theorem 7.3.2 *Every finite set of points has an acute path.*

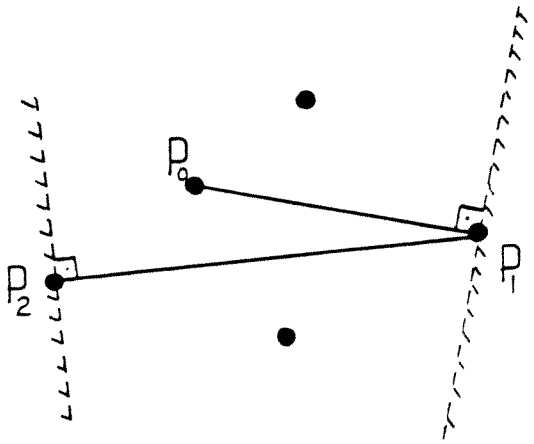


Figure 7.9: Every finite set of points has an acute path.

Proof: See Figure 7.9. Choose any point p_0 to be the starting point and remove p_0 from P. p_1 is a point in P that is farthest from p_0, and we remove p_1 from P. p_2 is a point in P that is farthest from p_1 and so on. This procedure can never produce an angle $\angle(p_i p_{i+1} p_{i+2})$ that is not acute, as in this case p_{i+2} is further from p_i than p_{i+1}, and we would have chosen p_{i+2} to be the successor of p_i.

□

The situation for acute tours is more complicated:

Example 7.3.3 *Let P_1 be a point set containing an odd number of points on the x-axis (perturbated a little bit to get a point set in general position). Suppose P_1 has an acute tour. We call a segment a western (eastern) segment, if the acute tour traverses it in direction of negative (positive) x. The tour must consist of an alternating sequence of eastern and western segments: If the tour traverses some western line segment $\overline{xy}$, it must traverse the following segment $\overline{yz}$ in eastern di-*

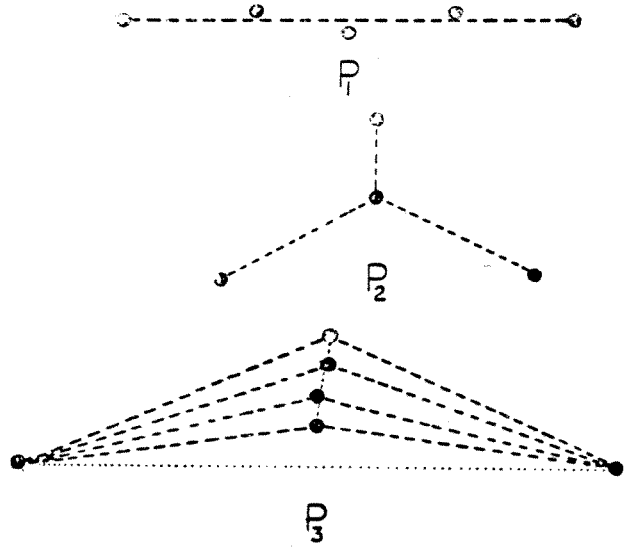


Figure 7.10: Sets that do not allow acute tours

---

rection to avoid an obtuse angle at the point y. This forces the tour to use an even number of line segments and, consequently, an even number of points.

An even number of points still does not guarantee an acute tour; see Figure 7.10.

Example 7.3.4 *The point set P_2 containing the four points $(0,-10)$, $(0,10)$, $(5,0)$, $(10,0)$ does not allow an acute tour. The point set P_3 containing the six points $(0,-10)$, $(0,10)$, $(1,0)$, $(2,0)$, $(3,\varepsilon)$ and $(4,\varepsilon)$ (where $\varepsilon < 10^{-10}$ is a small real) is in general position and does not allow an acute tour.*

Surprisingly enough, Example 7.3.4 cannot be generalized to any even number of points. In fact, we have the following conjecture:

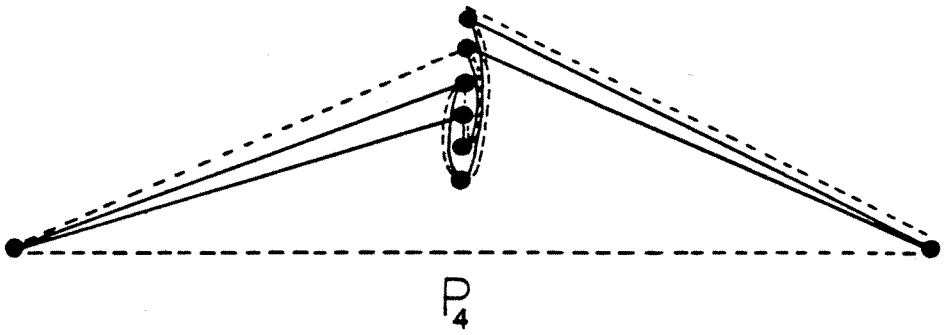


Figure 7.11: Some difficulties for acute tours

Conjecture 7.3.5 *Every set P of $2k \geq 8$ points admits an acute tour.*

After the construction of acute spanning paths in Theorem 7.3.2, one might be tempted to conjecture some kind of relation between acute paths, acute tours and tours of maximal overall length. However, the following Example 7.3.6 points out a few difficulties.

Example 7.3.6 *For a positive integer m, the point set P_4 contains the $2m$ points $(m^3, 0)$, $(-m^3, 0)$, $(0, 1), \ldots, (0, 2m-2)$ as shown in Figure 7.11. One easily checks that for $m \geq 4$, the following holds.*

- *P_4 allows an acute tour.*
- *No acute tour contains the diameter $\overline{(0, m^3)(0, -m^3)}$.*
- *It is impossible to get any acute tour by using the approach for constructing acute spanning paths as described in Theorem 7.3.2. (Regardless of the*

starting point, the diameter would be included in the second step at the latest.)

- *Any longest tour has length* $2m^3+\sqrt{m^6+(2m-2)^2}+\sqrt{m^6+(2m-3)^2}+4m-3$ *and contains the diameter (indicated by the broken lines in Figure 7.11), while a longest acute tour (solid lines) has only length* $\sqrt{m^6+(2m-2)^2}+\sqrt{m^6+(2m-3)^2}+\sqrt{m^6+(2m-4)^2}+\sqrt{m^6+(2m-5)^2}+4m-4$. *(Note that* $2m^3+1>\sqrt{m^6+(2m-4)^2}+\sqrt{m^6+(2m-5)^2}$ *for* $m \geq 4$.*)*

  □

Finding spanning paths that minimize the maximum angle or maximize the minimum angle seems to be difficult. By the above algorithm for finding acute spanning paths, we can always guarantee a maximum angle of absolute size at most $\pi/2$. A set of $n-1$ collinear points with an nth point far away from the line show that this guarantee is best possible.

As for a lower bound on the minimum angle, we have the following conjecture:

Conjecture 7.3.7
There always exists a spanning path with minimum angle at least $\pi/6$.

The equilateral triangle together with its centerpoint shows that no angle greater than $\pi/6$ can be guaranteed.

7.4 Other Questions

We conclude this chapter with some other open questions.

Conjecture 7.4.1 *It is polynomial to detect* $\{-2\pi/3, +2\pi/3\}$*-tours.*

Conjecture 7.4.2 *It is NP-complete to detect* $\{-\pi/3, +\pi/3, \pi\}$*-tours.*

To our knowledge, it has not been shown that detecting Hamiltonian Cycles in vertex-induced subgraphs of the hexagonal grid is NP-complete. A proof of this would certainly be helpful.

Problem 7.4.3 *What is the complexity of detecting acute/obtuse tours?*

For this problem as well as for Conjecture 7.3.7, some of the results on triangulations that minimize the maximal angle or maximize the minimal angle may be useful. It may appear as a good choice to consider Hamiltonian paths in a Delaunay triangulation, with any two adjacent edges in the path belonging to the same triangle. However, Delaunay triangulations or other minimum weight triangulations are in general *not* Hamiltonian. (See Dillencourt [18], [19].)

Chapter 8

Geometric TSP Optimality

8.1 Polyhedral Combinatorics and the Subtour Polytope

Consider the symmetric travelling salesman problem (TSP): For a complete graph K_n on n vertices, we are are given a weight c_e for each edge e. Each Hamilton cycle can be represented by its 0-1 incidence vector. We let T_n denote the set of these vectors. Finding a shortest tour is equivalent to the problem

$$\min\{cx \mid x \in T_n\}.$$

The polyhedral approach for this problem involves studying the associated polytope $Q^n \subset \mathbb{R}^n$ that is the convex hull of all 0-1 vectors representing tours of K_n:

$$\min\{cx \mid x \in Q^n\}.$$

Optimizing over Q^n is NP-hard, so it is unlikely that we will ever get a complete description of this polyhedron (see Karp and Papadimitriou [50]). For solving many

real world problems, however, it is not necessary to have complete knowledge of all the facet-defining inequalities of Q^n, as was demonstrated by the cutting plane approach by Dantzig, Fulkerson and Johnson [20], Grötschel [37], and Padberg and Hong [67]. A comprehensive survey of the polyhedral aspects of the TSP can be found in Grötschel and Padberg ([39] and [40]).

Some of the known inequalities describing Q^n have a particularly easy interpretation: The *nonnegativity constraints*

$$x_e \geq 0 \text{ for all } e \in E,$$

are valid, since no $x \in Q^n$ can have negative entries. Grötschel and Padberg [38] showed that these inequalities are facet-inducing for $n \geq 5$.

Furthermore, any $x \in Q^n$ must satisfy

$$x_e \leq 1 \text{ for all } e \in E.$$

For $S \subseteq V$ and $J \subseteq E$, we write $\delta(S) := \{[u,v] \in E \mid u \in S, v \in V\}$ and $x(J) = \sum_{e \in J} x_e$.

Clearly, every vertex must be adjacent to precisely two edges of a tour, so every $x \in Q^n$ must satisfy

$$x(\delta(v)) = 2 \text{ for all } v \in V.$$

These *degree constraints* form a maximal independent set of equations whose solution set contains Q^n — see Grötschel and Padberg [38].

For $S \subseteq V$, we write $\gamma(S) := \{(u,v) \in E \mid u, v \in S\}$.

No tour can contain a cycle of less than n vertices, so every set of less than n vertices must have at least 2 edges leaving it. Since the degree constraints assure this for single vertices, we can write this condition as

$$x(\delta(S)) \geq 2 \text{ for all } S \subseteq V \text{ and } 2 \leq |S| \leq n-2.$$

These *subtour elimination constraints* were first introduced by Dantzig, Fulkerson and Johnson [20]. Grötschel and Padberg showed that they are facet-inducing for $n \geq 4$.

The above inequalities define a polytope S^n that contains Q^n — the so-called *subtour polytope*. If $x \in S^n$ and x integer, x must be a tour. Thus the linear program

$$\min\{cx \mid x \in S^n\}$$

is a relaxation of the TSP.

Since $Q^n \subseteq S^n$, any optimal solution to minimizing over S^n is a lower bound for the optimal value of TSP. It was proved by Wolsey [96] that for any cost function c satisfying the triangle inequality, this bound can be at worst 2/3 of the optimum. It is an open conjecture that it is never worse than 3/4.

Grötschel, Lovász and Schrijver [41], and Karp and Papadimitriou [50], showed that a polynomial method for solving the separation problem for a polytope yields a polynomial method for optimization by means of the ellipsoid method. Padberg and Hong [67] (see also Padberg and Wolsey [70], and Padberg and Rao [68]) demonstrated how to solve separation for the subtour polytope in polynomial time by using the method of Gomory and Hu [35] for finding the minimum cost cut in a graph. Thus we know that optimization over the subtour polytope is possible

in polynomial time by means of the ellipsoid algorithm (see Grötschel, Lovász and Schrijver [41].) Since the ellipsoid method is, however, of purely theoretical value, it would be desirable to find a combinatorial method for this purpose. No such method is known at this time.

For a comprehensive study of optimizing over the subtour polytope, see Boyd [7] and also Boyd and Pulleyblank [8].

8.2 Geometric Interpretation and Moat Packings

If we think of the vertices V as locations and the cost c_e of an edge $e = [u, v]$ as describing the distance $d(u, v)$, we can write the optimization problem

$$\min\{cx \mid x \in S^n\}$$

as

$$\min \sum_{[u,v]\in E} d(u,v)x_{[u,v]}$$

subject to the constraints

$$\begin{array}{rcll} x_{[u,v]} & \geq & 0 \quad \text{for all} & [u,v] \in E, \\ x_{[u,v]} & \leq & 1 \quad \text{for all} & [u,v] \in E, \\ \sum_{[u,v]\in E} x_{[u,v]} & = & 2 \quad \text{for all} & v \in V, \\ \sum_{[u,v]\in \delta(S)} x_{[u,v]} & \geq & 2 \quad \text{for all} & \{S, \overline{S}\} \in \mathcal{M}. \end{array}$$

(We write $\mathcal{M}$ for the set of all partitions of V into two nonempty parts S and $\overline{S}$. $\mathcal{M}(u, v)$ denotes the set of all partitions for which u and v belong to different parts.)

The dual of this linear program is

$$\max \left(2 \sum_{S,\overline{S} \in \mathcal{M}} w_{\{S\overline{S}\}} + 2 \sum_{v \in V} r_v \right)$$

subject to the constraints

$$\begin{array}{rcll} w_{\{S\overline{S}\}} & \geq & 0 & \text{for all} \quad \{S, \overline{S}\} \in \mathcal{M}, \\ \sum_{\{S,\overline{S}\} \in \mathcal{M}(u,v)} w_{\{S\overline{S}\}} + r_u + r_v & \leq & d(u,v) & \text{for all} \quad u, v \in V. \end{array}$$

This linear program has exponentially many variables, but for any instance, there is an optimal solution for which the sets having $w_{\{S\overline{S}\}} > 0$ have the special structure of a *nested family*:

Definition 8.2.1 *A family of partitions* $(S_1, \overline{S_1}), \ldots, (S_k, \overline{S_k})$ *is called* nested, *if for any two partitions* $(S_i, \overline{S_i})$ *and* $(S_j, \overline{S_j})$, *we have* $S_i \cap S_j = \emptyset$ *or* $S_i \subseteq S_j$ *or* $S_j \subseteq S_i$.

For details on nested families, see Pulleyblank [71]; a proof for the above claim can be found in Cornuéjols, Fonlupt and Naddef [15].

If the locations of the $v \in V$ are embedded in the plane $\mathbb{R}^2$, and the distances are induced by some metric d (e.g. the Euclidean norm $\| \cdot \|_2$, also written as L_2), we can give a very elegant geometric interpretation of this problem.

Draw a d-ball of size r_v around each vertex v. Separate the two sets of vertices of a partition $\{S\overline{S}\}$ by a *moat* of width $w_{\{S\overline{S}\}}$ — i.e. a compact set such that any path between two vertices $v_1 \in S$ and $v_2 \in \overline{S}$ must have a connected section of at least length $w_{\{S\overline{S}\}}$ with respect to d.

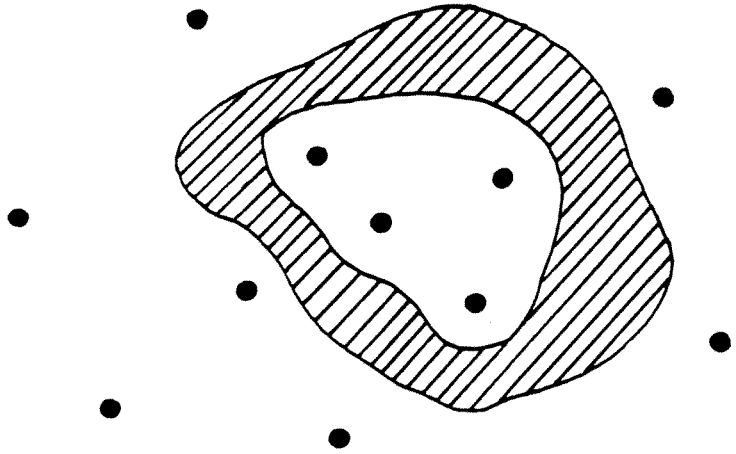


Figure 8.1: A moat

Now assume we can arrange these balls and moats in a way that they have pairwise nonintersecting interior. We call such an arrangement a *moat packing.* Clearly, the moat widths in a moat packing must satisfy the constraints of the linear program. Furthermore, any tour must cross any given ball or moat at least twice — so the objective function is a geometrically appealing lower bound for the length of a travelling salesman tour. If we have a moat packing for which this lower bound is met by a tour, we say that the moat packing is *tight.* If there is a tight moat packing for a given tour, it yields a fast proof for the optimality of the tour.

Jünger and Pulleyblank have shown in [48] that any nested solution of the linear program can be represented by a moat packing. (Because of the laminar structure, any solution can be transformed into nonnegative solution.) Their proof is valid regardless of the metric.

Geometric interpretations of combinatorial optimization problems and their use for heuristics have also been given for other problems than the TSP: Minimum spanning trees and minimum matching, along with TSP and Steiner trees are discussed in the survey paper by Jünger and Pulleyblank [49]; a discussion for the matching problem can be found in Jünger and Pulleyblank [48].

We will return to the subject of moat packings in section 8.4, after elaborating some topological aspects of realizing travelling salesman tours as curves embedded in a surface.

8.3 Convexity

The basic notion of a convex set is a set for which any two elements can be joined by a shortest path that stays within the set. These sets have been abstracted and generalized in various ways and their properties used in all different kinds of settings. (For a survey see Danzer, Grünbaum and Klee [21] and the shorter article by Klee [51].) Abstract and axiomatic generalizations were developed, most notably by Menger [57].

Recent developments in Computational Geometry and Computer Science have led to new interest in this topic. The study of restricted orientations, visibility and Voronoi diagrams (see Rawlins and Wood [77], [78], [79], [80], Schuierer [90] and Widmayer, Wu and Wong [93], as well as Klein [52], and Espie [24]) has led to the rediscovery of several different notions of convexity.

In this section, we introduce some basic concepts and definitions and show how they are useful for dealing with geometric realizations of the TSP.

8.3.1 Definitions

We follow the exposition of Klein [52].

Let M be a set. A *metric* on M is a function $d(.,.)$ that maps every pair (x, y) of elements of M into the nonnegative reals and satisfies the following three axioms:

$$\begin{aligned} M1: &\quad d(x,y) = 0 \text{ if and only if } x = y && \text{(identity)} \\ M2: &\quad d(x,y) = d(y,x) && \text{(symmetry)} \\ M3: &\quad d(x,z) \leq d(x,y) + d(y,z) && \text{(triangle inequality).} \end{aligned}$$

A *d-ball* is the set

$$B_d(x, \varepsilon) = \{y \in M \mid d(x,y) \leq \varepsilon\}$$

of all points not further than ε from the center x. A metric induces a *topology* on M: A set is *open*, if it contains a (sufficiently small) d-ball around each of its elements. Two different metrics d_1 and d_2 can induce the same topology for M — this is the case if every small d_1-ball contains a d_2-ball and vice versa. (Intuitively, this means that "closeness" of points is similar under the two metrics.)

Let M be a vector space. For simplicity, we restrict our attention to the vector space $\mathbb{R}^2$. (Most of the following properties and definitions hold for other spaces as well.) A *norm* of $\mathbb{R}^2$ is a function $\|\cdot\|$ that maps R^2 into the nonnegative reals and satisfies the following three axioms:

$$N1: \quad \|x\| = 0 \text{ if and only if } x = 0 \qquad \text{(identity)}$$
$$N2: \quad \|\lambda x\| = |\lambda|\|x\| \qquad \text{(scalar multiplicativity)}$$
$$N3: \quad \|x + y\| \leq \|x\| + \|y\| \qquad \text{(triangle inequality).}$$

Well-known examples are the Euclidean norm $\|(x_1, x_2)\|_2 = \sqrt{x_1^2 + x_2^2}$, the "Manhattan" norm $\|(x_1, x_2)\|_1 = |x_1| + |x_2|$, and the maximum norm $\|(x_1, x_2)\|_\infty = \max\{|x_1, |x_2|\}$. (We also write L_2, L_1 and L_∞ for these norms.) For a norm $\|\cdot\|$, $\overline{d}(x, y) := \|x - y\|$ defines a metric. Clearly, the metric induced by a norm is invariant under translations, i.e. the distance between x and y depends only on their relative positions. Because of $N2$, all balls are scaled copies of the unit ball $B := B_{\|\cdot\|}(0, 1)$. From its definition, it follows that the unit ball is compact, symmetric with respect to the origin, and convex.

Conversely, we can define a norm by its unit ball B: To calculate the distance of two points x and y, center B at x and consider the unique intersection point z of the ray from x through y with the boundary of B. Then define

$$d(x, y) := \frac{\|x - y\|_2}{\|x - z\|_2},$$

see Figure 8.2.

This definition is not dependent on the Euclidean norm for calculating the distances of x, y and z — we only need axiom $N2$. It follows that a norm is uniquely defined by its unit ball.

We note (see Rinow [84]):

Proposition 8.3.1 *All norm-induced metrics of $\mathbb{R}^2$ induce the same (Euclidean) topology.*

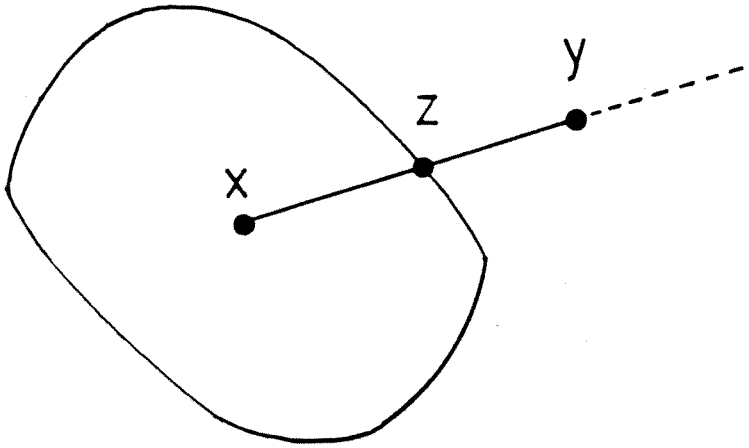


Figure 8.2: Every norm is defined by its unit ball

Norm-induced metrics can be characterized in the following way (see Lemma 1.2.1 in Klein [52]:)

Proposition 8.3.2 *A metric d on $\mathbf{R}^2$ is induced by a norm if and only if it has the following properties:*

1. *d induces the Euclidean topology.*
2. *The d-distance is invariant under translations.*
3. *The d-distance is additive on straight lines (that is, if x, y and z are points on a straight line with y between x and z, then $d(x,y) + d(y,z) = d(x,z)$ holds.)*

A *curve* is given by a continuous injective mapping c of the interval $[0,1]$ into $\mathbf{R}^2$. A curve is *closed*, if $c(0) = c(1)$. It is *simple*, if $c(s) \neq c(t)$ for all distinct

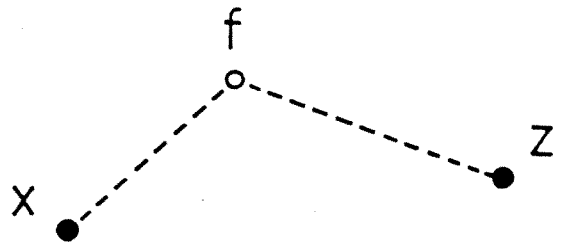


Figure 8.3: The flowershop metric

$s, t \in (0,1)$. The mapping induces an order of the points of the curve; in the following, we will refer to appropriate orderings of this type when we speak of the order of points. Given a metric d for $\mathbf{R}^2$, a curve is called *d-straight*, if for every set of three points x, y and z, such that y falls between x and z, the equality $d(x,y) + d(y,z) = d(x,z)$ holds. As noted above, the straight line (in the ordinary sense) between two points x and z is d-straight for any norm-induced metric d. Clearly, it is the only d-straight curve for the Euclidean distance.

It is not hard to see that d-straight curves between two given points are in general not unique — an example is given by the L_1-norm. It may be less obvious that there need not be any d-straight curves:

Example 8.3.3 (Klein [52]) *Let f be a fixed point in the plane. Then define*

$$d(x,z) = \begin{cases} \|x-f\|_2 + \|f-z\|_2 & \text{if } x \neq z \\ 0 & \text{if } x = z. \end{cases}$$

This "flowershop" metric can be interpreted as follows: When travelling from x to z, we have to go by f (the flowershop) to buy some flowers for the host at z. (See Figure 8.3.) Clearly, we have $d(x,z) < d(x,y)+d(y,z)$ for distinct x, y and z, unless $y = f$. So there can be no d-straight curve between x and f, hence none between x and z. Also note that this metric does not induce the Euclidean topology: For $x \neq f$ and $\varepsilon < \|x-f\|_2$, $B_d(x,\varepsilon) = \{x\}$.

The existence of d-straight curves is the subject of Menger's powerful "Verbindbarkeitssatz". A metric space (M,d) is complete if every d-Cauchy sequence in M has a limit in M.

Theorem 8.3.4 (Menger 1928, [57]) *Let (M,d) be a complete metric space and assume that to any two different points x and y, a point $z \notin \{x,y\}$ exists such that $d(x,z) = d(x,y) + d(y,z)$ holds. Then two points in M can always be connected by a d-straight curve.*

Proposition 8.3.5 (Jordan Curve Theorem) *A simple closed curve c subdivides the plane in exactly two connected open sets. Exactly one of those sets, called the interior of c, is bounded.*

See Blumenthal and Menger, [4].

Let c be a closed curve. Then the *support* of c is the smallest simply connected set that contains c. The *contour* of c is the boundary of the support.

8.3.2 Generalized Convexity

In [52], Klein has elaborated the theory of planar Voronoi diagrams for general metrics. In order to ensure suitable properties of the boundaries of Voronoi regions, he and Wood defined so-called "nice" metrics:

Definition 8.3.6 (Klein, Wood)

Let d be a metric of $\mathbb{R}^2$. For points $p, q \in \mathbb{R}^2$, we write

$$C(p,q) = \{z \in \mathbb{R}^2 \mid d(p,z) \leq d(q,z)\}$$
$$D(p,q) = \{z \in \mathbb{R}^2 \mid d(p,z) < d(q,z)\}.$$

A metric d on $\mathbb{R}^2$ is called nice *if it enjoys the following properties:*

1. *d induces the Euclidean topology.*
2. *The d-balls are bounded with respect to the Euclidean metric.*
3. *If $x, z \in \mathbb{R}^2$ and $x \neq z$ then there exists a point $y \notin \{x, z\}$ such that $d(x,z) = d(x,y) + d(y,z)$ holds.*
4. *For $p, q \in \mathbb{R}^2$, $J(p,q) := C(p,q) \cap \overline{D(q,p)}$ is a curve homeomorphic to $(0,1)$. (Here $\overline{D(q,p)}$ denotes the topological closure of $D(q,p)$.) The intersection of two such curves, $J(p,q)$ and $J(u,v)$, consists of finitely many connected components.*

Properties 1, 2 and 3 are satisfied by any norm-induced metric. Klein and Wood introduce the technical property 4 to make sure that the boundaries of Voronoi regions have a simple structure. For our purposes, we only need properties 1 and 3; we say a metric is a *Menger-metric* if it satisfies these two conditions.

Definition 8.3.7 *Let M be a subset of $\mathbb{R}^2$ and d be any metric of $\mathbb{R}^2$. M is* d-convex *if for any two points in the set, there is a d-straight curve connecting them that does not leave M.*

We get the usual convexity if d is the Euclidean metric. For $d = L_1$, we get *orthoconvexity*: A set is orthoconvex, if the intersection with any axis-parallel line is connected. Generalizing orthoconvexity has motivated the notion of *restricted orientation convexity*: For a given set of orientations $\mathcal{O}$, a set M is $\mathcal{O}$-convex, if any $\mathcal{O}$-parallel line has a connected intersection with M. Examining restricted orientation convexity has been one of the main objectives of Rawlins [77] and Schuierer [90] in their Ph.D. theses.

$\mathcal{O}$-convexity is a special case of d-convexity: For a given set of orientations $\mathcal{O}$, let B be a (centrally symmetric) unit ball defining some norm-induced metric d, such that $\mathcal{O}$ is the set of directions of the extreme points of B, as seen from the origin. It is not hard to see that for any such d, d-convexity implies $\mathcal{O}$-convexity. (In [90], Schuierer gave one particular way for defining d and proved the stated implication. We note that there are infinitely many different norms that define the same set of orientations.)

8.3.3 Tours and Curves

Consider a geometric Travelling Salesman Problem where the cities are given by vertices in the plane. For a considerable number of practical purposes, the distances are not described by the Euclidean norm. Examples of other distance functions include the L_1 norm, which plays a role in VLSI design, and the L_∞ norm, which describes the motion of a plotter. The following distance function originates from

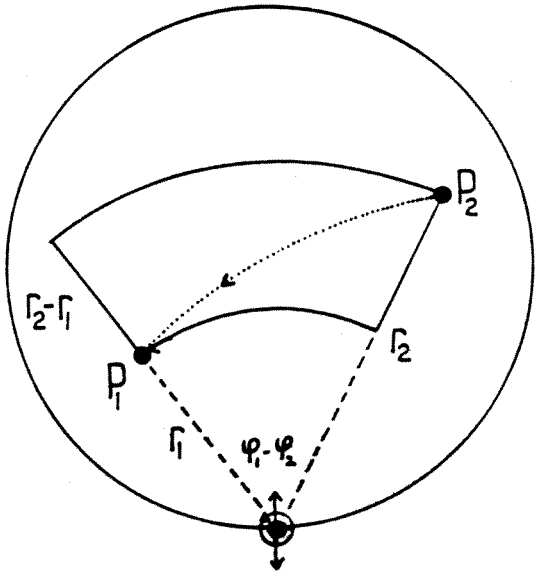


Figure 8.4: The Livermore metric

---

the testing of microchip wafers at the Livermore Laboratories. It shows that it is very easy to encounter problems where the distances are not described by a norm. See Figure 8.4.

Example 8.3.8 *Consider a set of vertices ("test points") on a disk ("wafer") and a distinguished ("attachment") point on the boundary of the disk. We have to repeatedly move the disk such that each of the vertices gets positioned at the origin ("quality sensor"). The disk can be moved by rotating it around the attachment point and moving the attachment point parallel to the x-axis. These two motions can be performed independently, so for two points $p_1 = (r_1, \varphi_1)$ and $p_2 = (r_2, \varphi_2)$ in polar coordinates, the distance is*

$$d(p_1, p_2) = \max\left(\frac{|r_1 - r_2|}{v_r}, \min\left(\frac{|\varphi_1 - \varphi_2|}{v_\varphi}, \frac{2\pi - |\varphi_1 - \varphi_2|}{v_\varphi}\right)\right).$$

We have to solve a Travelling Salesman Problem for this metric.

Despite all their differences, a considerable number of practical metrics have one thing in common: Reflecting the fact that we move from one tour vertex to the next, the distances between points correspond to shortest paths. Furthermore, two points are usually "close" if their Euclidean distance is small. This means that we are dealing with Menger-metrics.

We have already pointed out at the beginning of Chapter 6 how a convex arrangement of tour vertices leads to easy solvability:

Theorem 8.3.9 *Given a set of n vertices in the Euclidean plane, such that they lie on the boundary of their convex hull, i.e. the contour. Then the optimal solution of the Euclidean Travelling Salesman Problem for these vertices is given by the contour.*

Proof: The only way to get a noncrossing tour is given by the contour. It is an easy consequence of the triangle inequality that any optimal tour is noncrossing.

□

Convexity is a particularly easy optimality criterion for the TSP. We will see in the following chapter how it can be used for the purpose of error analysis.

What is it that makes convexity so useful for Euclidean distances? Can we generalize this property to other metrics?

We have already discussed notions of generalized convexity. Since we are dealing with Menger-metrics, we should concentrate on properties of shortest paths. This is done with the definition of d-convex sets. For tours, we can formulate the geometric properties as follows:

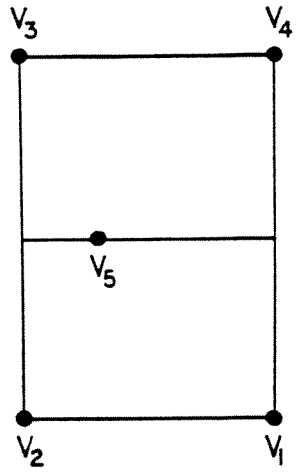


Figure 8.5: A d-convex tour requires more than just a d-convex support

---

Definition 8.3.10 *Consider the TSP for some Menger-metric of* $\mathbf{R}^2$. *A tour of* n *vertices in the plane is* d-convex, *if it can be represented by a curve* c *that consists of* n *d-straight curves connecting successive vertices, such that the support of* c *is d-convex and all vertices lie on the contour of* c.

We will see in the following section that any d-convex tour is indeed optimal. This need not be true if there are vertices that do not lie on the contour:

Example 8.3.11 *See Figure 8.5. Suppose the cost for travelling along the shown edges is equal to the Euclidean distance. The cost for any path outside those edges is 100 times the Euclidean distance. Then the tour* $\langle v_1, v_2, v_3, v_4, v_5, v_1 \rangle$ *has a d-convex support, but is longer than the tour* $\langle v_1, v_2, v_5, v_3, v_4, v_1 \rangle$.

8.4 Convexity and Moat Packings

We now return to moat packings.

We have seen that we can define unit balls, distances and convexity in a much more general sense than just for the ordinary Euclidean metric. We proceed to show that these notions can be incorporated in our geometric interpretation of balls and moats.

Theorem 8.4.1 *Let d be a Menger-metric, describing the distance function for a planar TSP. Consider a set of n vertices that form a d-convex arrangement, i.e. for which there is a d-convex tour. Then there is a tight moat packing for this tour.*

Proof:

We proceed by induction over n. All the "balls" will be d-balls of appropriate radius around vertices. (The reader may find it useful to consider the situation for the Euclidean metric, where the d-balls are disks. Our argumentation is based on weaker topological properties, since general d-balls do not have to be convex.) Each moat will separate two components of the plane that are at some d-distance w. A moat of width w will consist of the union of appropriate pieces of w-strips around some of the vertices it surrounds. (We remind that a *strip* of width w around some vertex v is given by $B_d(v,R) \setminus B_d(v,r)$, where $R - r = w$ as the difference between two d-balls.) We say a moat packing is tight, if it covers all the edges. A moat packing is feasible, if it is tight and nonoverlapping.

Let the tour be given as $T = \langle v_1, \ldots, v_{n-1}, v_n, v_1 \rangle$ and realised by an appropriate curve c, consisting of d-straight curves between successive vertices. Let c_1 and c_n

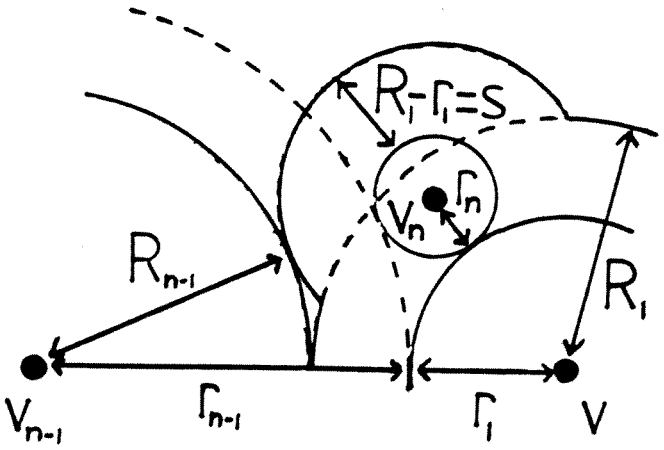


Figure 8.6: Inserting the new moat and redirecting the old moats

---

be the d-straight curves between v_1 and v_n, and v_n and v_{n-1}.

Let

$$r_n := \frac{1}{2}\left(d(v_1, v_n) + d(v_{n-1}, v_n) - d(v_{n-1}, v_1)\right),$$

$$r_1 := \frac{1}{2}\left(d(v_1, v_n) + d(v_{n-1}, v_1) - d(v_{n-1}, v_n)\right),$$

$$r_{n-1} := \frac{1}{2}\left(d(v_1, v_{n-1}) + d(v_{n-1}, v_n) - d(v_n, v_1)\right).$$

By triangle inequality, all these quantities are nonnegative. For $n = 3$, define a feasible moat packing by the three balls $B_d(v_1, r_1)$, $B_d(v_2, r_2)$, $B_d(v_3, r_3)$.

Suppose we have a tight moat packing for the tour $T' = \langle v_1, \ldots, v_{n-1}, v_1 \rangle$ on the $(n-1)$ vertices $v_1, \ldots, v_{n-1}$. For some d-straight curve c_{n-1} between v_1 and v_{n-1} that does not leave the support of c, the tour T' is also d-convex. If $r_n = 0$, we are done — the moat packing for T' is a moat packing for T.

The set of all moats that cross c_{n-1} can be partitioned into those that run in strips around v_1 and those that run in strips around v_{n-1}. Let R_1 be the sum of the strip widths around v_1 and R_{n-1} the sum of the strip widths around v_{n-1}. By definition, $R_1 + R_{n-1} = d(v_1, v_{n-1}) = r_1 + r_{n-1}$.

For $R_1 = r_1$, add only a ball of radius r_n to the set of moats and balls for T' to get a tight moat packing for T.

For $R_1 \neq r_1$, assume without loss of generality $R_1 > r_1$ and see Figure 8.6. The strip of width $s = R_1 - r_1 = r_{n-1} - R_{n-1}$ around v_1 is replaced by

$$B_d(v_1, R_1) \cup B_d(v_n, r_n + s) \setminus (B_d(v_1, r_1) \cup B_d(v_n, r_n)),$$

which means redirecting the strip so that it runs around the vertex v_n.

By construction, we get a tight moat packing. We have to show that it is nonoverlapping. Clearly, no intersection can occur between the moats separating v_1, v_{n-1}, and v_n. By induction hypothesis, the only overlapping of other moats can occur if a moat around v_n intersects a moat around some $v_i \neq v_1, v_{n-1}$.

So suppose $r_i + r_n + s > d(v_i, v_n)$, where r_i denotes the radius of an appropriate ball around v_i. Since T is convex, there must be a d-straight curve c_i between v_n and v_i that does not leave the support of c and therefore has to intersect c_{n-1} in some point p.

Now either $d(v_i, p) < r_i$ or $d(v_n, p) < r_n + s$. The former case implies that there is a point close to p on c_{n-1} that lies in the interior of $B_d(v_i, r_i)$ as well as in the interior of $B_d(v_1, R_1)$ or $B_d(v_{n-1}, R_{n-1})$ — a contradiction to the feasibility of the moat packing for T'. (See Figure 8.7(a).)

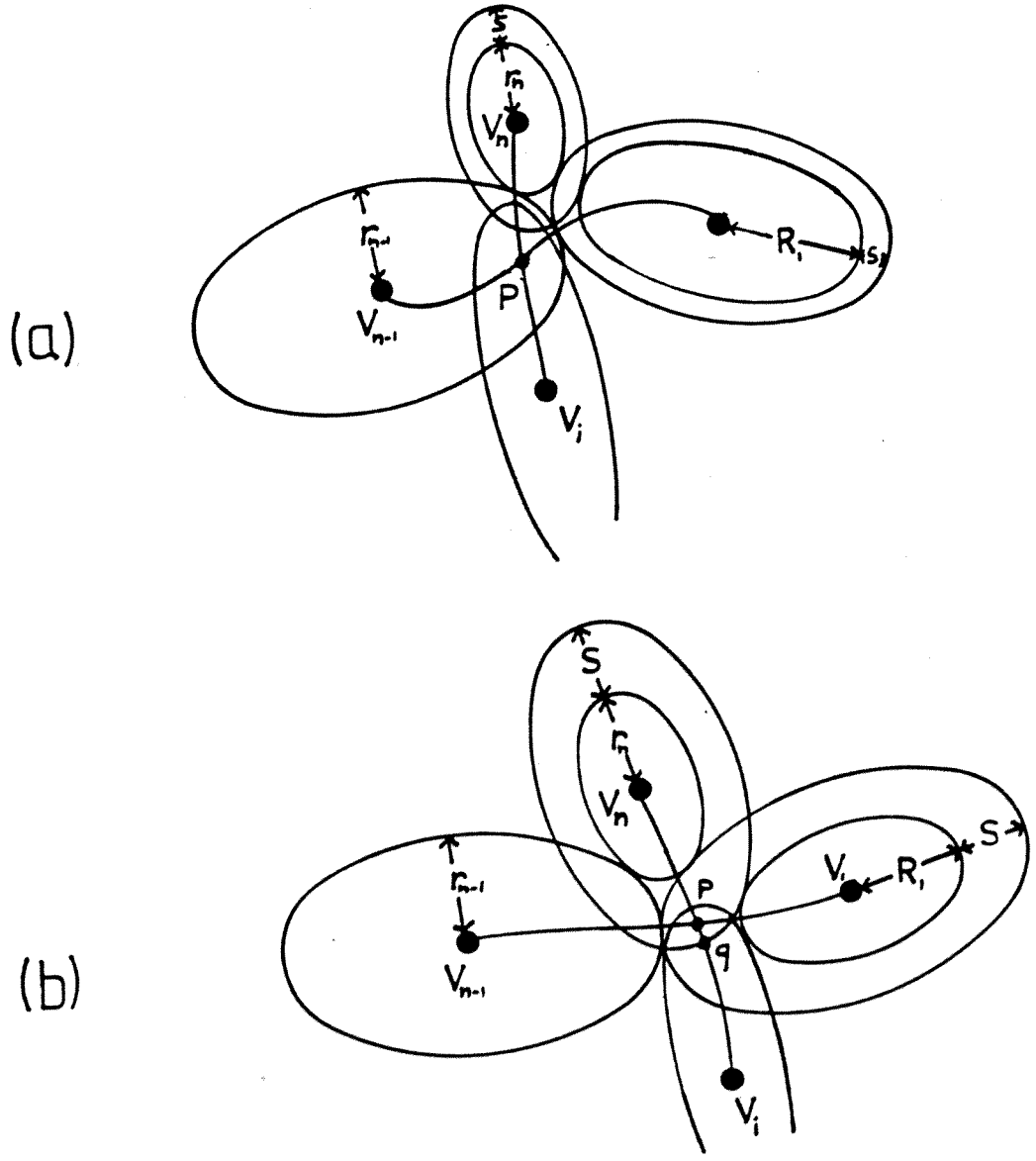


Figure 8.7: The new moat packing is tight and feasible

Therefore assume $d(v_n, p) < r_n + s$. Since

$$d(v_n, v_{n-1}) = r_n + s + R_{n-1},$$

$$d(v_1, v_{n-1}) = r_1 + s + R_{n-1},$$

$$d(v_1, v_{n-1}) = d(v_1, p) + d(p, v_{n-1}), \text{ and}$$

$$d(v_n, v_{n-1}) \leq d(v_n, p) + d(p, v_{n-1}),$$

we have

$$r_n + s - d(v_n, p) \leq r_1 + s - d(v_1, p).$$

This implies $d(v_n, v_i) \geq r_n + s$ — otherwise v_i would belong to the interior of $B_d(v_1, R_1)$. Then let q be the point on c_i with $d(v_n, q) = r_n + s$. (See Figure 8.7(b).) It follows that q belongs to $B_d(v_1, R_1)$ and to the interior of $B_d(v_i, r_i)$ — a contradiction to the assumption that $B_d(v_1, R_1)$ and $B_d(v_i, r_i)$ are nonoverlapping.

We conclude that the resulting moat packing for T is nonoverlapping, completing the proof.

□

Chapter 9

Error Analysis for the TSP

9.1 Introduction

Let us consider an instance I of a mathematical problem, given by a set of numerical input data. Our objective is to find a solution x. Because of limited accuracy in our calculations, we may only be able to find an approximate solution $\overline{x}$, even if we are given an algorithm that theoretically computes an exact solution. Under these circumstances, we would like to have some estimates on the error $\|x - \overline{x}\|$ that has been accumulated during the calculations. This type of error analysis has been studied extensively in the field of numerical analysis. It is usually referred to as *forward error analysis.*

For most practical purposes, we are in a situation where the *input data* is already inaccurate. Young and Gregory [97] state:

"If we take the point of view that any computed solution to a problem is the

exact solution to a slightly perturbed problem, then the details of what caused the computed solution to differ from the exact solution are not too important. This is the motivation for what is called *backward error analysis* as opposed to the usual *forward error analysis* [...]"

In other words: If we are given a suggested solution x, what is the error $\|I - \overline{I}\|'$ that separates the input data I from an instance $\overline{I}$ that has exact solution $\overline{x}$?

Backward error analysis was used by Givens [37] and Lanczos [53], but its greatest advocate has been Wilkinson [94].

When trying to solve a combinatorial optimization problem, we may have a third source of error besides imprecise calculations and inaccurate data: The difficulty of finding an exact solution within reasonable time may require us to accept an approximate solution. The bounds that we can give for the error on a solution that was achieved by a heuristic may be considerable: The well-known approximation method of Christofides for the symmetric TSP with triangle inequality yields a solution that can be 50% greater than the optimum. Still, currently there is no (polynomial) method known that can guarantee a smaller relative error.

Similar to the case of numerical problems, these combinatorial difficulties can make it interesting to consider the notion of backward error analysis. Since we may either have to pay a high price for an exact solution x of a given instance I or accept a high error $\|\overline{x} - x\|$ between an approximate solution $\overline{x}$ and x, we may prefer to solve an instance $\overline{I}$ that has a distance $\|\overline{I} - I\|'$ that is not too large. This means that we may rephrase the question of backward error analysis in the context of optimization:

How far is the instance I from an instance $\overline{I}$ that we can solve satisfactorily?

There are various special cases for which the TSP can be solved accurately and quickly. For the purposes of backward error analysis, we concentrate on an easy geometric criterion that we have already discussed in previous chapters: Generalized Convexity. This gives rise to a geometric question:

Given a finite set of points in the plane. How far is it from being convex?

There are two metrics involved in this question:

- The metric that defines the convexity of the point set. Since it also describes the distances between the points, we refer to this metric as the *distance metric* or the *distances* for short.
- The metric that measures the amounts by which each point in I must be moved to yield $\overline{I}$. We refer to this metric as the *error metric.*

Error analysis has also been examined in a purely geometric setting. But while the study of errors in arithmetic calculations has been extensively studied in the field of numerical analysis, the propagation of errors in geometric algorithms has only very recently become the subject of attention — see Hoffmann [43], Milenkovic [58], [59], [60], Li and Milenkovic [55], Salesin [87]. Since inaccuracy of geometric algorithms may cause inconsistencies in the combinatorial results, which may in turn cause severe problems in later stages of an algorithm, the question of robustness of geometric algorithms can become very important (Hoffman, Hopcroft and Karasick [44]).

Some of the interest from the geometric side has focussed on the question of (Euclidean) convexity. Milenkovic [58], [59], [60], and Li and Milenkovic [55] have studied the notion of *strong convexity.* They call a polygon *ε-strongly convex*, if any perturbation of all vertices by at most ε leaves it convex. The obverse of strong convexity is *weak convexity*: A polygon is *ε-weakly convex*, if it can be transformed into a convex polygon by moving each point by at most ε. In this sense, we are interested in weak convexity.

In his recent Ph.D. thesis [87], Salesin examines questions of strong and weak convexity. (He uses the terms *ε-convex* for ε-weakly convex and $-\varepsilon$-convex for ε-strongly convex.) Since he is interested in robust implementations of geometrical algorithms, his results are mostly of the "uncertainty interval" type:

Given a logical predicate with a parameter ε like "the polygon $\mathcal{P}$ is ε-convex", an uncertainty interval $[p.lo, p.hi]$ reflects the situation where it is known that $\mathcal{P}$ is *not* ε-convex for $\varepsilon < p.lo$ and $\mathcal{P}$ is ε-convex for $\varepsilon > p.hi$, but not known what the answer is for $p.lo \leq \varepsilon \leq p.hi$.

We have already discussed how we can successfully generalize the notion of convexity for other distance metrics than the Euclidean metric. We discuss ε-convexity for several combinations of distance metrics and error metrics.

Finally, we stress one important difference from Salesin's results: While he examines ε-convexity of *polygons*, we consider *point sets.* This reflects our interest in connecting the given vertices by a convex polygon without prescribing the optimal solution in advance. Not knowing the order of the points in the final convex solution may make matters considerably harder. (See Section 9.2.2 on L_1 distances and L_1 error bounds for an example.)

9.2 L_1 Distances

There are some Travelling Salesman Problems for which the distance between two points (x_1, y_2) and (x_2, y_2) is given by the L_1 metric as

$$\|(x_1, y_2) - (x_2, y_2)\|_1 = |x_1 - x_2| + |y_1 - y_2|.$$

Typical examples arise from VLSI routing, where the wire connections between adjacent pins have to consist of horizontal and vertical segments. Since these rectilinear paths resemble the route between two points in New York, the L_1 metric is also called the *Manhattan metric.*

In other applications, the distance is described by the L_∞ metric

$$\|(x_1, y_2) - (x_2, y_2)\|_\infty = \max\{|x_1 - x_2|, |y_1 - y_2|\}.$$

This situation occurs for the motion of a plotter, where vertical and horizontal motion are performed by independent motors, so the time required to connect two points is determined by the direction in which the greater motion occurs.

As described in the previous chapter, any norm-induced metric is characterized by its unit ball. L_1 and L_∞ metric are special cases of so-called *generalized polygonal metrics*: the unit ball is a convex $2k$-gon that is symmetric with respect to the origin. In the following, we write $L_\mathcal{P}$ if we refer to any generalized polygonal norm. Both L_1 the L_∞ metric have a square as the unit ball — with axes-parallel edges for L_∞, rotated by $\frac{\pi}{4}$ for L_1. This property makes the two metrics very similar for many purposes.

We mentioned in the previous chapter that for any norm-induced metric d, d-convexity can be expressed as *restricted orientation* or $\mathcal{O}$*-convexity*: For a set $\mathcal{O}$ of

orientations, a set $S \subseteq \mathbb{R}^2$ is convex, if any intersection with a line parallel to one of the orientations in $\mathcal{O}$ is given by the directions of the extreme points on the unit ball. The set of directions is given by the directions of the extreme points of the unit ball. This means that for any polygonal norm $L_{\mathcal{P}}$ that is defined by a unit ball with $2k$ vertices, $L_{\mathcal{P}}$-convexity is the same as $\mathcal{O}$ for a set of k orientations.

Rawlins and Wood [77], [78], [79], [80], Schuierer [90] have given a detailed description of the properties of $\mathcal{O}$-convex sets in general and $\mathcal{O}$-convex polygons in particular. Since it would take some fairly technical elaborations that would not provide any further insight for our purposes, we concentrate mostly on L_1-distances and hence L_1-convexity. In each case we briefly mention how to generalize the results to arbitrary $L_{\mathcal{P}}$.

Proposition 9.2.1 *A simple polygon Q is L_1-convex if and only if we can partition its edge set into at most 4 subsets, such that each subset forms a connected path that is monotonic with respect to x- and y-direction.*

Any path is an L_1-shortest path between its two end points, if and only if it is monotonic in the coordinate directions. An *axes-parallel* L_1-shortest (i.e. monotonic) path between two points is usually referred to as a *staircase.*

For other $L_{\mathcal{P}}$, we split the edge set of Q into $2k$ convex paths, each of which is monotonic with respect to two appropriate (adjacent) directions in $\mathcal{O}$. Staircases can be defined analogously. For details, see Schuierer [90]. Along these lines, all the following constructions can be carried out for $L_{\mathcal{P}}$ instead of L_1.

Monotonicity with respect to x- and y-directions is closely related to the following partial orders. We think of the positive and negative y-direction as "north" and

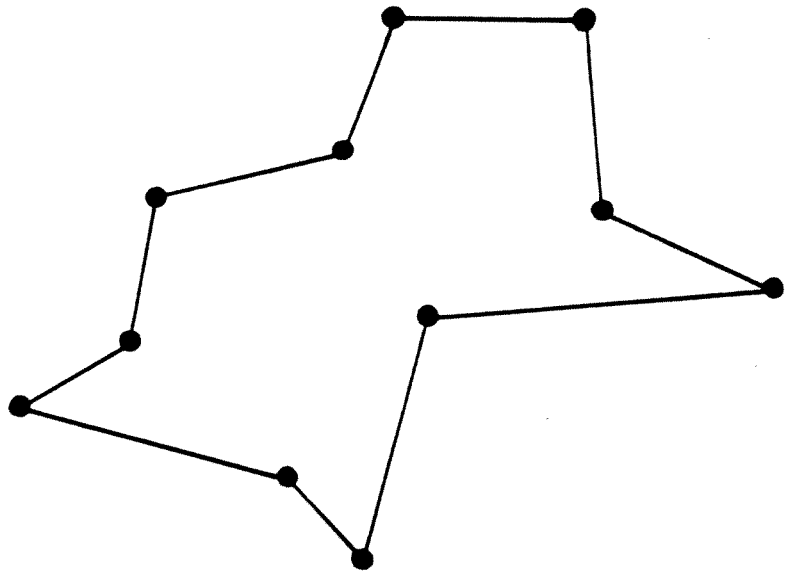


Figure 9.1: An L_1-convex polygon

"south", while the positive and negative x-direction are "east" and "west". These directions are indicated by the letters n, s, e, w in the following notation.

Definition 9.2.2 *For any two points $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$, we define*

$$p_1 \overset{ne}{\preceq} p_2 \text{ if and only if } x_1 \leq x_2 \text{ and } y_1 \leq y_2.$$

$$p_1 \overset{se}{\preceq} p_2 \text{ if and only if } x_1 \leq x_2 \text{ and } y_1 \geq y_2$$

$$p_1 \overset{nw}{\preceq} p_2 \text{ if and only if } x_1 \geq x_2 \text{ and } y_1 \leq y_2$$

$$p_1 \overset{sw}{\preceq} p_2 \text{ if and only if } x_1 \geq x_2 \text{ and } y_1 \geq y_2.$$

Similarly, we write

$$p_1 \overset{ne}{\prec} p_2 \text{ if and only if } x_1 < x_2 \text{ and } y_1 < y_2.$$

$$p_1 \overset{se}{\prec} p_2 \text{ if and only if } x_1 < x_2 \text{ and } y_1 > y_2$$

$$p_1 \overset{nw}{\prec} p_2 \text{ if and only if } x_1 > x_2 \text{ and } y_1 < y_2$$
$$p_1 \overset{sw}{\prec} p_2 \text{ if and only if } x_1 > x_2 \text{ and } y_1 > y_2.$$

Note that $p_1 \overset{ne}{\prec} p_2$ is *not* the same as $p_1 \overset{ne}{\preceq} p_2$ and $p_1 \neq p_2$.

If we are only given a point set without the additional information provided by a polygon, we would like to find ways to determine the existence of a convex tour. For standard (Euclidean) convexity, any finite point set P has a maximal set that can be connected by a convex tour that encloses all points in P. (This is the set of points on the boundary of the convex hull.) We call this set the *surface* of P. Moreover, the surface has a unique convex tour that follows the boundary of the convex hull.

For L_1-convexity, matters are slightly more complicated, as can be seen in Figure 9.2. We can, however, still define the notion of "surface" for L_1 convexity:

Definition 9.2.3 *Let P be a finite set of points. Then we have the following four sets:*

$$\begin{aligned}
P_{\text{ne}} &:= \{p \in P \mid \textit{there is no } q \in P : p \overset{ne}{\prec} q\},\\
P_{\text{se}} &:= \{p \in P \mid \textit{there is no } q \in P : p \overset{se}{\prec} q\},\\
P_{\text{nw}} &:= \{p \in P \mid \textit{there is no } q \in P : p \overset{nw}{\prec} q\},\\
P_{\text{sw}} &:= \{p \in P \mid \textit{there is no } q \in P : p \overset{sw}{\prec} q\}.
\end{aligned}$$

The surface *of P is* $\text{sur}(P) = P_{\text{ne}} \cup P_{\text{se}} \cup P_{\text{sw}} \cup P_{\text{nw}}$. *The* interior *of P is* $\text{int}(P) = P \setminus \text{sur}(P)$.

The following shows that the surface provides valuable information about convexity.

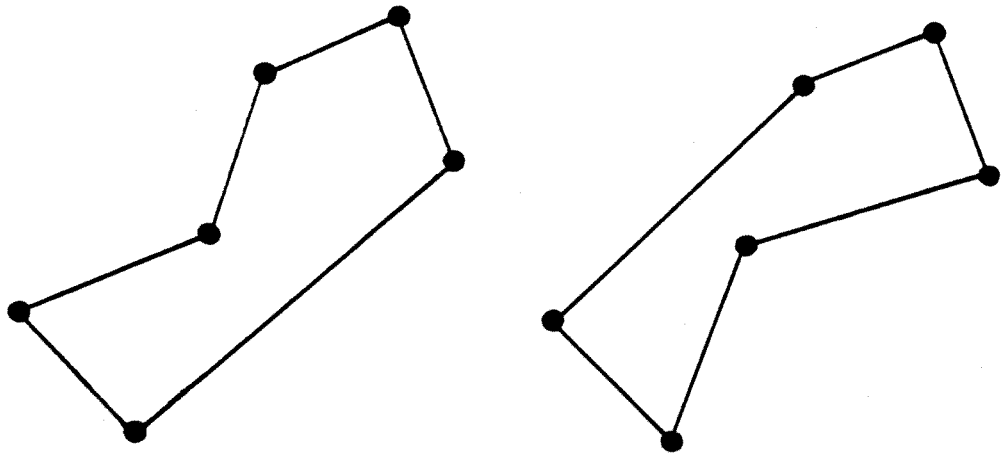


Figure 9.2: An L_1-convex point set does not necessarily have a unique shortest tour

---

Lemma 9.2.4 *A point set P allows an L_1-convex tour if and only if $P = \mathrm{sur}(P)$. In this case we say that P is convex.*

Proof: If P allows an L_1-convex tour, the claim is a direct consequence of Proposition 9.2.1. Conversely, suppose $P = sur(P)$. It follows from its definition that P_{ne} is totally ordered with respect to $\overset{se}{\preceq}$. Similarly, P_{se}, P_{sw}, P_{nw}, are ordered with respect to $\overset{sw}{\preceq}$, $\overset{nw}{\preceq}$, $\overset{ne}{\preceq}$. If there are $q \in P_{\mathrm{ne}} \cap P_{\mathrm{sw}}$ or $q \in P_{\mathrm{ne}} \cap P_{\mathrm{sw}}$ (as in Figure 9.2), eliminate them arbitrarily from one of the two intersecting sets. The points in P_{ne}, P_{se}, P_{sw}, P_{nw} can be connected by staircases in southeastern, southwestern, northwestern, northeastern direction. The union of these staircases forms a simple polygon that is L_1-convex by construction. See Figure 9.3.

□

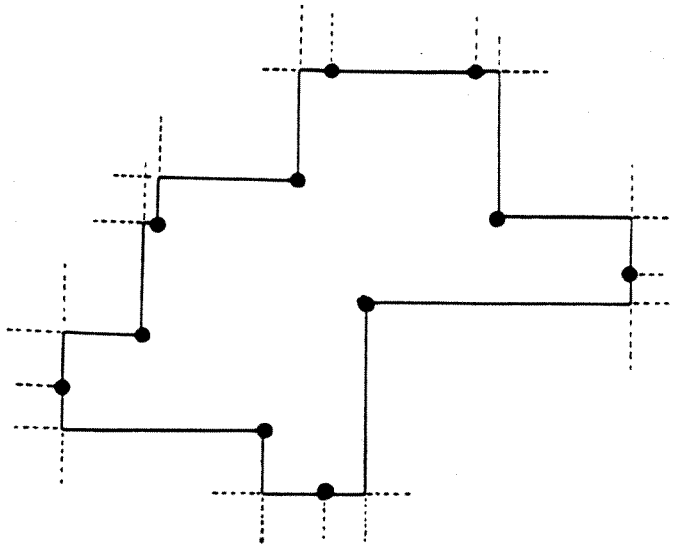


Figure 9.3: An L_1-convex point set is equal to its surface

If a set P is not convex, we consider its surface and interior. Our objective will be to move each point $p \in P$ by as little as possible (with respect to some specified error metric) to some new location p'. The p' form a set P' of *destination points.* We will try to move interior and surface points in a way that all interior points become part of the surface. The following two sections will show that coordination of the movement of all the points can be relatively easy or quite complicated, depending on the error metric that is used.

9.2.1 L_∞ Error Bounds

A typical situation for L_∞ error bounds arises when we describe the positions of the points by Cartesian coordinates with a limited number of digits. The result are error bounds in both coordinates that are independent of each other.

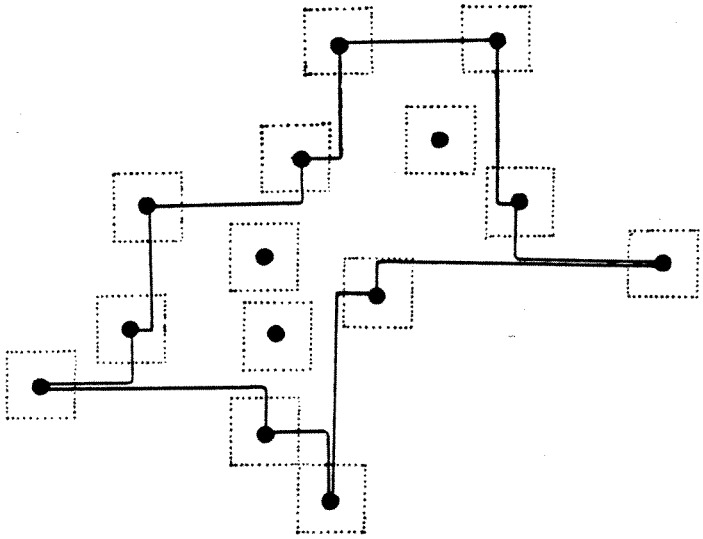


Figure 9.4: A point set with L_1-surface and L_∞ error bounds

---

The situation for a point set with L_∞ error bounds is shown in Figure 9.4. We can connect the surface points by four staircases in a way that the region inside the resulting polygon becomes as small as possible.

As mentioned before, we try to move interior and surface points such that all interior points become part of the surface. We can think of this process as moving the surface points until a staircase polygon through them intersects all error squares around the interior points. (See Figure 9.6.) Since this is fairly straightforward, we do not elaborate all of the technical details and instead sketch the way the necessary L_∞ distances can be calculated.

If p is any point of the interior, we can try including it into P_{ne}, P_{se}, P_{sw}, or P_{nw}. Suppose we try to include $p = (p_1, p_2)$ into $P_{\text{nw}} = \{q_i = (x_i, y_i) \mid i \in I\}$. (See Figure 9.5.) Since $p \notin P_{\text{nw}}$, we have to consider all q_i for which $p \overset{nw}{\prec} q_i$, i.e. for

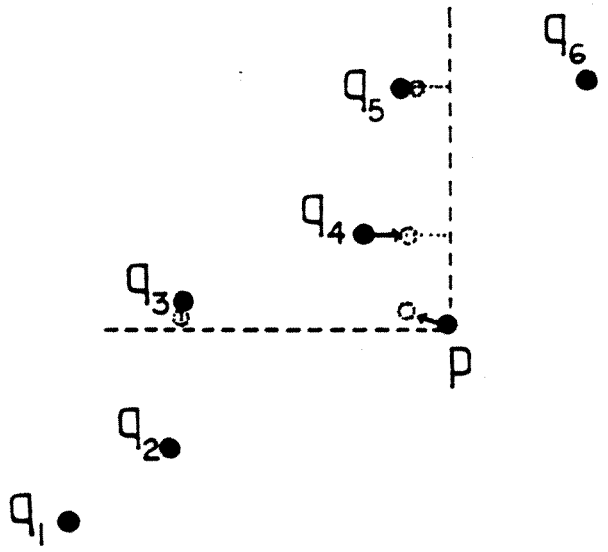


Figure 9.5: Including an interior point p into P_{nw}

---

which p and q_i are not comparable by $\overset{ne}{\preceq}$. Let J_{nw} be the index set of these q_i. To make p comparable with all $q_j, j \in J$, we either have to move each of p and q_j by at least $\frac{1}{2}(p_1 - x_j)$ in x-direction towards each other, or each by at least $\frac{1}{2}(y_i - p_2)$ in y-direction. We write $d_{\text{nw}}(p, q_j) := \min((p_1 - x_j), (y_j - p_2))$. Clearly, we need to move all points by at least $\frac{1}{2}\max\{d_{\text{nw}}(p, q_j) \mid j \in J\}$.

Analogously, we can consider including p into P_{ne}, P_{se}, P_{sw}. Then we need at least $\frac{1}{2}\max\{d_{\text{ne}}(p, q_j) \mid j \in J_{\text{ne}}\}$, $\frac{1}{2}\max\{d_{\text{se}}(p, q_j) \mid j \in J_{\text{se}}\}$, $\frac{1}{2}\max\{d_{\text{sw}}(p, q_j) \mid j \in J_{\text{sw}}\}$, respectively. Let $d_{\text{sur}}(p)$ be the smallest of the resulting 4 numbers; clearly, we have to move all the points in $sur(P) \cup \{p\}$ by at least this amount in order to include p into the surface. Overall, we get a lower bound of $\max\{d_{\text{sur}}(p) \mid p \in int(P)\}$.

Considering an L_1-convex polygon through the surface points that consists of 4 staircases, it is not hard to see that this lower bound is also sufficient to achieve

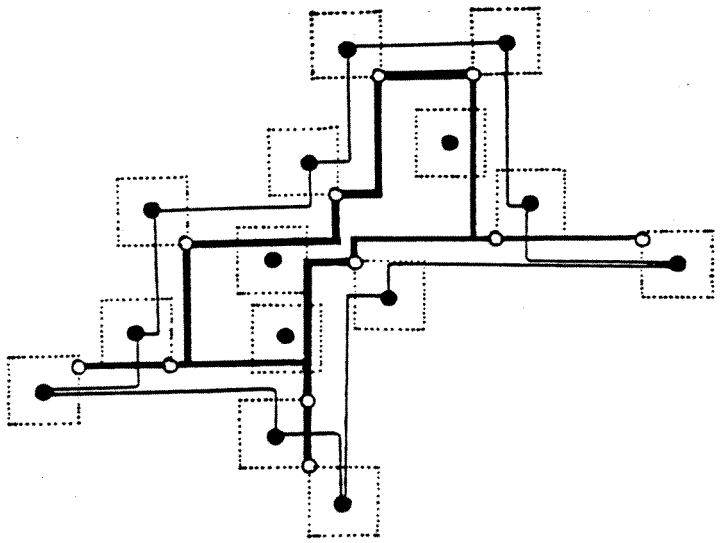


Figure 9.6: The lower bound is sufficient to achieve L_1-convexity

---

L_1-convexity of the whole point set. (See Figure 9.6.) For the sake of completeness, we describe the computational evaluation.

From the way we include an interior point p into the surface, it follows that a surface point in $P_{\text{nw}} \cap P_{\text{se}}$ does not need to be moved at all. (It can neither be contained in J_{ne} nor J_{sw}.) The same holds for surface points in $P_{\text{ne}} \cap P_{\text{sw}}$. Any surface point that is only contained in one of the sets P_{nw}, P_{ne}, P_{se}, P_{sw} is only moved in southeastern (southwestern, northwestern, northeastern) direction if it is contained in P_{nw} (P_{ne}, P_{se}, P_{sw}). Finally, consider a point $q = (x_q, y_q)$ that is contained in $P_{\text{ne}} \cap P_{\text{nw}}$. If

$$\max\{y_i \mid (x_i, y_i) \in sur(P) \text{ and } x_i < x_q\} < \max\{y_i \mid (x_i, y_i) \in sur(P) \text{ and } x_i > x_q\},$$

then move q in southeastern direction, otherwise in southwestern direction. Analogously, we can treat the points q in $P_{\text{ne}} \cap P_{\text{se}}$, $P_{\text{sw}} \cap P_{\text{se}}$, $P_{\text{sw}} \cap P_{\text{nw}}$.

We summarize:

Theorem 9.2.5 *For a given set $P = \{q_1, \ldots, q_n\}$, the minimal ε for which there is an L_1-convex set $P' = (q'_1, \ldots, q'_n)$ with $\|q_1 - q'_1\|_\infty \leq \varepsilon\}$ can be calculated in $O(n^2)$ time.*

It is straightforward to calculate the sets P_{nw}, P_{ne}, P_{se}, P_{sw} in time $O(n^2)$ (or in $O(n \log n)$ using a plane sweep). For any of the $O(n)$ interior points p, the distance $d_{\text{sur}}(p)$ can be calculated in linear time, resulting in an overall complexity of $O(n^2)$.

The simplicity of the described method hinges on the fact that for every point q on the surface, we can very easily determine an extreme point of the ε-balls around q that becomes the new location of q. (See Figure 9.6 again.) If we are guaranteed a similar situation by other metrics, we get a comparatively easy method.

In particular, we can get an $O(n^2)$ method for any combination of a polygonal (distance) metric $L_{\mathcal{P}}$ with another polygonal (error) metric $L_{\mathcal{P}^*}$ if $L_{\mathcal{P}^*}$ is dual to $L_{\mathcal{P}}$ in the following sense: The unit ball of $L_{\mathcal{P}^*}$ has edges that are orthogonal to the directions of the extreme points of $L_{\mathcal{P}}$. See Figure 9.7 and Figure 9.8.

9.2.2 L_1 Error Bounds

If we replace the L_∞ error metric of the previous section by the L_1 metric, the problem of making a point set L_1-convex becomes considerably harder. The main difficulty can be seen in Figure 9.9:

The choice of the new location p' of each of the surface points is no longer easy.

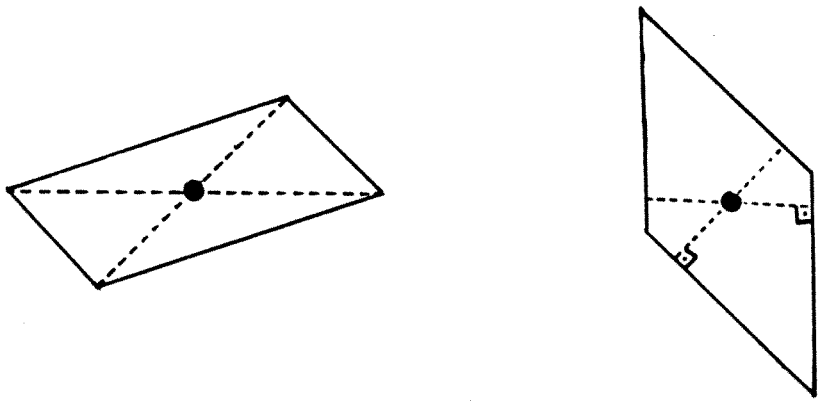


Figure 9.7: A polygonal metric $L_{\mathcal{P}}$ and a dual metric $L_{\mathcal{P}}$

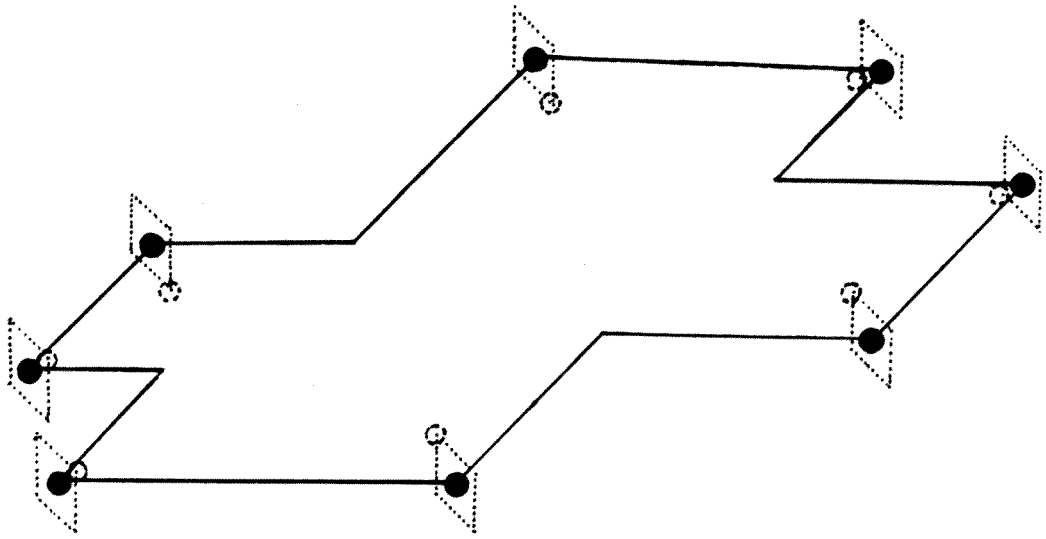


Figure 9.8: A point set with an $L_{\mathcal{P}}$-convex polygon for the surface points and $L_{\mathcal{P}}$-error bounds

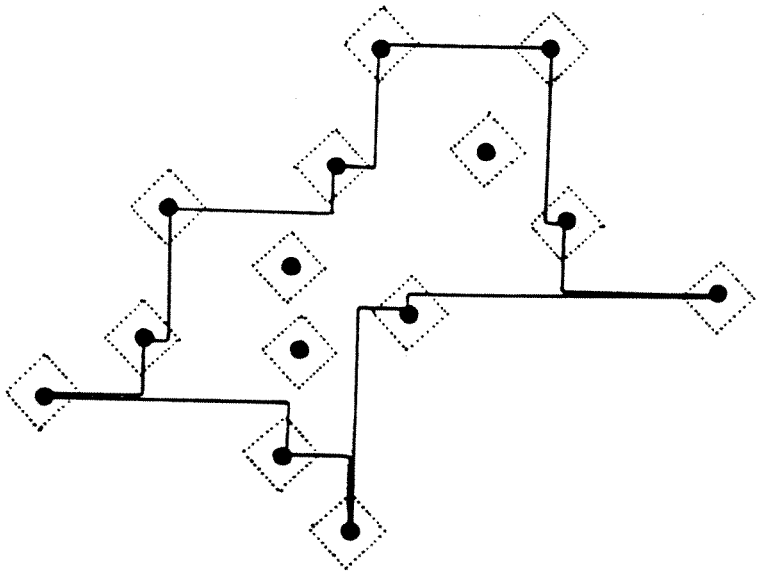


Figure 9.9: A point set with L_1 error bounds around the surface points

Instead, we have to "distribute" the allowed movement of a surface point between x- and y-direction.

We believe that a promising approach for solving the problem is to consider the following auxiliary problem first:

Problem 9.2.6 *Given a set $P = \{p_1, \ldots, p_n\}$. What is the minimal ε, such that there is a set $P' = \{p'_1, \ldots, p'_n\}$ with $\|p_i - p'_i\|_1 \leq \varepsilon$ that is totally ordered with respect to $\overset{ne}{\preceq}$?*

Once we have a method for Problem 9.2.6, we can consider partitioning P into four subsets that are transformed into the sets P'_{ne}, P'_{se}, P'_{sw}, P'_{nw}, of P'. We do not know how to find the right partition. This illustrates the difference to transforming a given *polygon* instead of a *point set*, where the right partition follows from the order of the vertices of the polygon.

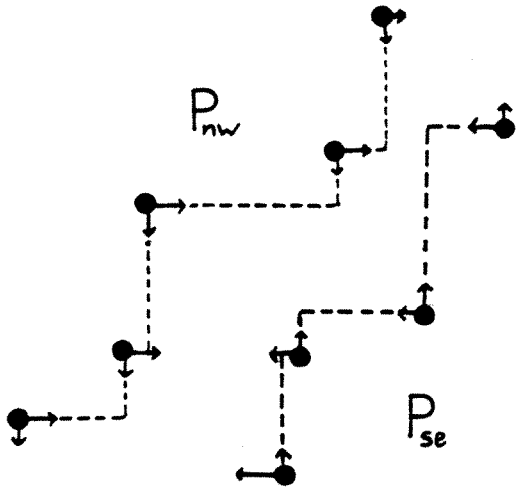


Figure 9.10: How far do the points in P have to be moved to be comparable?

When we consider ordering a set of points in northeastern direction, the points in P_{nw} and P_{se} play a special role, as they lie extremal among the incomparable points. It can be shown that we only need to consider the error bounds for the points in $P_{\text{nw}} \cup P_{\text{se}}$ when we are trying to order P. Furthermore, we can prove that we may assume that the final order of the destination points P' preserves the existing order of P with respect to $\overset{ne}{\preceq}$, so we only need to consider ordering incomparable pairs in P. Furthermore, we can show that we may assume that the points in P_{nw} are only moved in eastern and southern direction, while the points in P_{se} need only to be moved in northern and western direction.

After the described simplifications, Problem 9.2.6 is equivalent to the geometric problem of how to "merge" the two ordered chains P_{nw} and P_{se} in the most efficient manner. It may be possible to solve this problem by dynamic programming.

9.3 Euclidean Distances

In this section, we discuss the situation for Euclidean distances, corresponding to "normal" convexity. Whenever we speak just of "convexity", we refer to Euclidean convexity.

For some error metric d and a given ε, consider the set $B_i = \{p \in \mathbb{R}^2 \mid d(p, p_i) \leq \varepsilon\}$ for each point p_i. We will see that picking one point from each of the sets B_i in order to get a convex arrangement is equivalent to finding a convex curve that intersects all B_i. The latter is a variation on a problem known from computational geometry as *stabbing*. (See Edelsbrunner, Maurer, Preparata, Rosenberg, Welzl and Wood [25] about the problem of stabbing a family of sets with line segments.) We can state the question of finding a convex curve that intersects a family of sets as the problem of finding a *convex stabber*. As such it was posed by Tamir [91] in 1987 for the situation where the B_i are arbitrary compact sets. Goodrich and Snoyeink [36] have given an $O(n \log n)$ algorithm for finding a convex stabber for n parallel line segments.

We show that the problem is polynomial if the B_i are induced by a generalized polygonal error norm, i.e. there is a convex polygon B with $2k$ vertices such that for any p_i, we have $B_i = p_i + B$. On the other hand, we discuss some of the difficulties that arise when the error metric is given by the Euclidean norm and all the B_i are discs of radius ε.

9.3.1 General Properties

Consider some error norm $\|\cdot\|$, an $\varepsilon > 0$ and a set $P = \{p_1, \ldots, p_n\}$. Let B denote the set of all points that have $\|\cdot\|$-distance from the origin no greater than ε. For each point p_i, we let $B_i = \{p \in \mathbb{R}^2 \mid d(p, p_i) \leq \varepsilon\} = p_i + B$. Let $\mathcal{B}$ denote the family of all B_i. Deciding whether a given ε is large enough to get a convex arrangement is equivalent to the following problem:

Can we pick one point from each B_i such that the resulting set P' forms a convex arrangement?

We can ask this question regardless of whether the set B arises from a norm. For this section, we do not require B to be anything more than compact and convex. The sets B_i are given as $p_i + B$.

For Euclidean convexity, we can define the *surface*, $sur(P)$, of P as the set of all p_i that are extreme points of the convex hull of P. (This only includes points that are proper vertices of the convex hull.) For a point $p_i \in sur(P)$, we say that the related B_i is a *surface set*. Again, the interior $int(P)$ of P is $P \setminus sur(P)$. We say that a point $p \in \mathbb{R}^2$ lies *inside* a closed convex curve $c \subseteq \mathbb{R}^2$, if it lies in the convex hull of c, i.e. $p \in conv(c)$. A point p lies *strictly inside* c, if it lies inside c and not on c. Analogously, we speak of points lying *outside* and *strictly outside* c. For some subset of the plane D, we denote the boundary of D by ∂D. We write $\mathcal{C}$ for the set of all closed convex curves that intersect all sets of $\mathcal{B}$. (We admit degenerate cases, where "closed convex curve" stands for a line segment or a point.)

We have the following lemma:

Lemma 9.3.1 *Let c be a closed convex curve in the plane intersecting all sets of $\mathcal{B}$. Then every closed convex curve $\bar{c} \in \mathbb{R}^2$ lying inside of c intersects all sets of $\mathcal{B}$ if and only if it intersects all surface sets.*

Proof: We only prove the sufficiency. For each $p_j \in sur(P)$, there is a $v_j \in B$ such that $y_j := (p_j + v_j) \in B_j \cap \bar{c}$. For each $p_i \notin sur(P)$, there are $\lambda_j^{(i)}$ such that

$$p_i = \sum_{p_j \in sur(P)} \lambda_j^{(i)} p_j, \text{ where } \sum \lambda_j^{(i)} = 1.$$

Since $\bar{c}$ is convex, the point

$$y_i := \sum_{p_j \in \mathcal{S}} \lambda_j^{(i)} (p_j + v_j) = p_i + \sum_{p_j \in \mathcal{S}} \lambda_j^{(i)} v_j$$

lies inside of $\bar{c}$. Since B is convex, we have

$$v_i := \sum_{p_j \in \mathcal{S}} \lambda_j^{(i)} v_j \in B,$$

so $y_i \in B_i$.

Because c intersects B_i, there is a point $\overline{y_i} \in B_i$ outside of c, therefore outside of $\bar{c}$. So the line from y_i to $\overline{y_i}$ intersects $\bar{c}$ in a point $y_i^* \in B_i$.

□

Lemma 9.3.1 is not necessarily true if B is not convex, see Figure 9.11. With $q_1 := (0,0)$, $q_2 := (0,1)$, $q_3 := (1,-1)$, $q_4 := (-1,-1)$, let $B := \overline{q_1q_2} \cup \overline{q_1q_3} \cup \overline{q_1q_4}$ and $P := \{(0,0),(2,0),(1,-2),(1.25,-1)\}$. Then $sur(P) = \{(0,0),(2,0),(1,-2)\}$ and $c := \overline{q_2q_3} \cup \overline{q_3q_4} \cup \overline{q_4q_2}$ intersects all sets of $\mathcal{B}$. Furthermore, we can choose

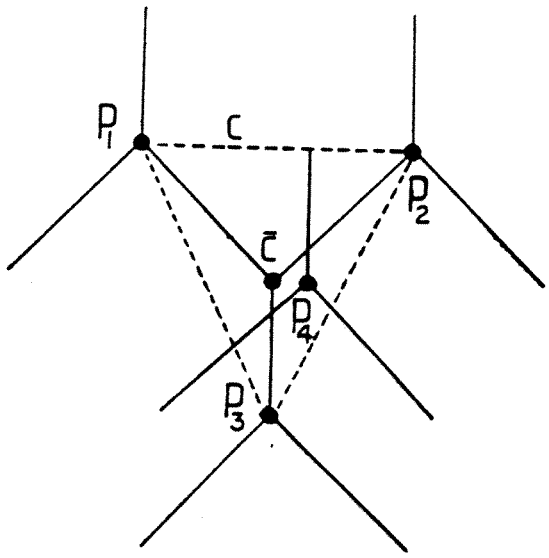


Figure 9.11: Lemma 9.3.1 is not necessarily true if B is not convex.

$$\bar{c} := \{(1,-1)\} = \bigcap_{p_i \in sun(P)} B_i,$$

which intersects all surface sets and lies inside of c, but does *not* intersect B_4.

We will show that it is sufficient to consider only polygons of very special properties instead of arbitrary convex curves. For this purpose, we need some topological tools.

The distance between two bounded convex sets $G, H \subseteq \mathbf{R}^2$ can be described by the *Hausdorff-distance* $\eta(G,H)$: For any positive $\xi \in \mathbf{R}$, we define the *ξ-widening* G_ξ of G by $G_\xi := \{x \in \mathbf{R}^2 | d(x,G) \leq \xi\}$, where $d(x,G)$ is the Euclidean distance between x and G. Then $\eta(G,H)$ between G and H is the smallest number $\eta \geq 0$ such that $G \subseteq H_\eta$ *and* $H \subseteq G_\eta$. It is straightforward to prove that the Hausdorff-distance is a metric. We say that a sequence $G_1, G_2, \ldots$ of compact convex regions

converges towards the region G, if the sequence $\eta(G, G_1), \eta(G, G_2), \ldots$ tends to 0. With this convention, we can consider converging sequences of regions. A detailed proof of the following lemma, can be found in Fejes-Tóth [30], pp.29–31:

Lemma 9.3.2 (Blaschke) *Every sequence of convex regions, which all lie within a bounded subset of the plane, has a convergent subsequence.*

With the help of Lemma 9.3.2 we can prove the following lemma:

Lemma 9.3.3 *Let $\mathcal{C}$ be the set of all convex curves that intersect all sets of $\mathcal{B}$, and assume that $\mathcal{C}$ is nonempty. Then there is a (bounded) curve $c^* \in \mathcal{C}$ with minimal enclosed area $AR(c^*)$.*

Proof: We consider mapping each $c \in \mathcal{C}$ to the size of its enclosed area $AR(c)$. Using the Hausdorff-metric on $\mathcal{C}$, we see that this mapping is continuous. Consider the set

$$conv(\mathcal{B}) := conv\left(\bigcup_{B_i \in \mathcal{B}} B_i\right).$$

For every $c \in \mathcal{C}$, we can consider a convex curve that does not leave $conv(\mathcal{B})$:

$$c|_{\mathcal{B}} := \partial\left(conv(c) \cap conv(\mathcal{B})\right).$$

Then $c|_{\mathcal{B}} \in \mathcal{C}$ and $AR(c|_{\mathcal{B}}) \leq AR(c)$. Therefore, it is sufficient to consider the enclosed area of curves from the set

$$\mathcal{C}|_{\mathcal{B}} := \{c|_{\mathcal{B}} \mid c \in \mathcal{C}\}.$$

By Lemma 9.3.2, every sequence $(c_k)_{k \in N}$ of intersecting convex curves has a subsequence c_{k_j} such that $conv\left(c_{k_j}\right)$ converges towards a closed convex set D.

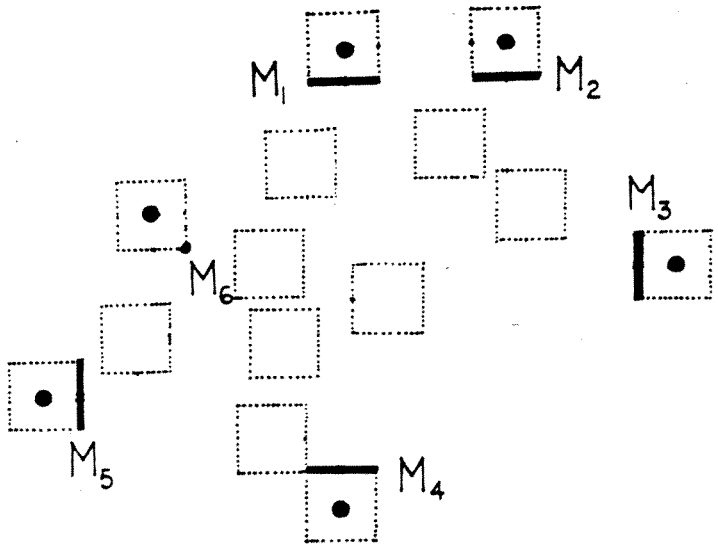


Figure 9.12: The sets M_j for an arrangement of squares

Now consider the boundary ∂D of D: Since all c_{k_j} have (Euclidean) distance 0 from all sets $B_i \in \mathcal{B}$, ∂D must have distance 0 from all B_i. As all the B_i are compact, ∂D has to intersect all B_i, thus $\partial D \in \mathcal{C}$.

So $\mathcal{C}|_{\mathcal{B}}$ is a compact set and the continuous function AR assumes a minimum in an element $c^* \in \mathcal{C}|_{\mathcal{B}}$.

□

Our next objective is to restrict the choice of p_i' from the B_i. In order to use Lemma 9.3.3, we need the following definitions:

For each $p_j \in \mathcal{S}$, define l_j as the set of all nontrivial linear functions which are maximized over P by p_j.

A *candidate set* is a set

$$M_j := \{q \in B_j \mid \exists l \in l_j : l(q) \leq l(r) \text{ for all } r \in B_j\},$$

i.e. the set of all points in B_j which *minimize* one of the linear functions in l_j over B_j. For $p_j \notin sur(P)$, we write $M_j = \emptyset$. We can think of these sets M_j as the sets of all points on a surface set that lie "towards the interior" of the arrangement formed by all the sets B_j. (See Figure 9.12.) Intuitively, it seems sufficient to intersect the surface sets at these points. We establish this property in the following Theorem 9.3.4.

Theorem 9.3.4 *Let there be a closed convex curve c intersecting all sets of $\mathcal{B}$. Then there exists a convex polygon $\mathcal{P}^*$ with the following properties:*

1. *$\mathcal{P}^*$ intersects all sets of $\mathcal{B}$.*
2. *For each vertex v_j of $\mathcal{P}^*$, there is a surface set B_{i_j} such that $\mathcal{P}^* \cap B_{i_j} = \{v_j\}$.*
3. *For any vertex v_j of $\mathcal{P}^*$, we have $v_j \in M_{i_j}$ if $(v_j) = \mathcal{P}^* \cap B_{i_j}$.*

Proof: Consider a closed convex curve $c^* \in \mathcal{C}$ of minimal enclosed area. The existence of c^* was proven in Lemma 9.3.3. In case of $AR(c^*) = 0$, i.e. c^* being a line segment or a point, assume that c^* has minimal length among all $c \in \mathcal{C}$ with $AR(c) = 0$.

For each surface set B_i, choose a point $y_i \in c^* \cap B_i$, yielding the finite set of points Y. $\partial conv(Y)$ lies inside of c^*, so by Lemma 9.3.1, it intersects all sets of $\mathcal{B}$. By minimality of c^*, we get $\partial conv(Y) = c^*$ for all Y chosen in the described manner. Therefore, c^* is a convex polygon $\mathcal{P}^*$.

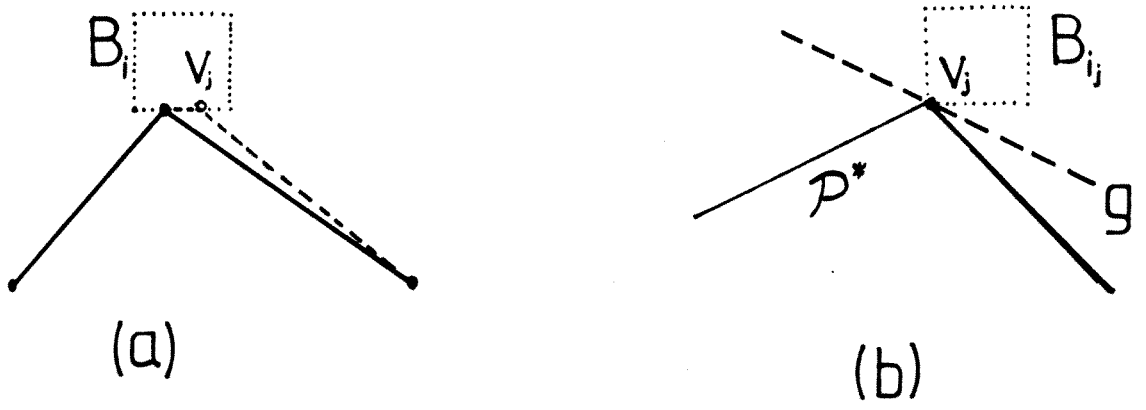


Figure 9.13: Examining intersecting convex curves can be reduced to examining polygons

---

Assume there was a vertex v_j of $\mathcal{P}^*$ with $|\mathcal{P}^* \cap B_i| \geq 2$ for all surface sets B_i with $v_j \in B_i$. Then we can choose a set Y that does *not* contain v_j, a contradiction. (See Figure 9.13(a).)

To see that *3.* must hold, note that there must be a line g separating the convex sets B_{i_j} and $\mathcal{P}^*$, such that $g \cap \mathcal{P}^* = \{v_j\}$. (See Figure 9.13(b).) Since all $B_k \neq B_{i_j}$ must have one point y_i in the same halfplane as $\mathcal{P}^*$, we see that for an appropriate linear function l induced by g, we get the two conditions

$$l(v_j) \geq l(y_k)$$

and

$$l(p_{i_j} - v_j) \geq l(p_k - y_k),$$

so consequently

$$l(p_{i_j}) \geq l(p_k).$$

Since the minimum of l over B_{i_j} is attained at v_j, we have $v_j \in M_{i_j}$.

□

9.3.2 Polygonal Error Bounds

As stated before, the situation where the error metric is given by a generalized polygonal norm corresponds to a B that is a centrally symmetric polygon with $2k$ vertices. In generalizing these B, we consider the case that *B is a convex polygon with k vertices.* To investigate the existence of a convex curve that intersects all sets B_i, we use Theorem 9.3.4, which reduces the problem to examining certain sets of points that constitute the vertices of the convex polygon $\mathcal{P}^*$ of Theorem 9.3.4 that intersects all sets B_j. When we talk of "picking" a vertex v from a set B_j, we imply that B_j is a set that satisfies $B_j \cap \mathcal{P}^* = \{v\}$.

As we have seen in the previous section, we can restrict ourselves to picking vertices of $\mathcal{P}^*$ from the candidate sets M_j. Our next steps are to show that we can apply further constraints, eventually leaving us with no more than $O(n^{3k})$ candidates for $\mathcal{P}^*$. Checking whether the boundary of a given convex polygon with $O(n)$ vertices intersects all n sets of $\mathcal{B}$ can be done in time $O(n^2)$, so we get an overall complexity of $O(n^{3k+2})$.

First we make the following observation:

Lemma 9.3.5 *There are at most $2k$ nontrivial M_j, i.e. sets that satisfy $|M_j| > 1$. We can restrict those to at most k nontrivial $\overline{M_j}$.*

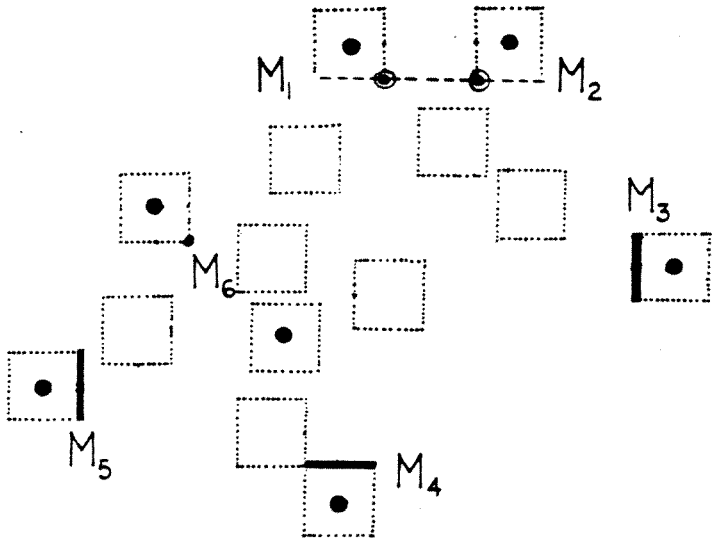


Figure 9.14: There are at most k nontrivial M_j

Proof: Each M_j consists of a connected set of edges of B_j. An edge e_i of B can can correspond to an edge in at most two of the M_j. Moreover, we see that e_i corresponds to edges in two different M_j (let it be in B_1 and B_2 by $b_1 := x_1 + e_i$ and $b_2 := x_2 + e_i$) only if b_1 and b_2 are collinear.

If $b_1 \cap b_2 = \emptyset$, let $x_1 < x_2$, i.e. x_1 be "left" of x_2. Using Theorem 9.3.4, we see that any vertex of $\mathcal{P}^*$ picked from b_1 has to be the right endpoint z_1 of b_1, any vertex of $\mathcal{P}^*$ contained in b_2 must be the left endpoint z_2 of b_2. So we can delete $b_1 \backslash \{z_1\}$ from M_1 (thereby getting $\overline{M_1}$) and $b_2 \backslash \{z_2\}$ from M_2, yielding $\overline{M_2}$.

If $b_1 \cap b_2 \neq \emptyset$, then similarly we can set $\overline{M_1} := b_1 \cap b_2$ and $\overline{M_2} := \emptyset$. By the same reasoning we can remove all other edges that might be contained in M_1 or M_2.

□

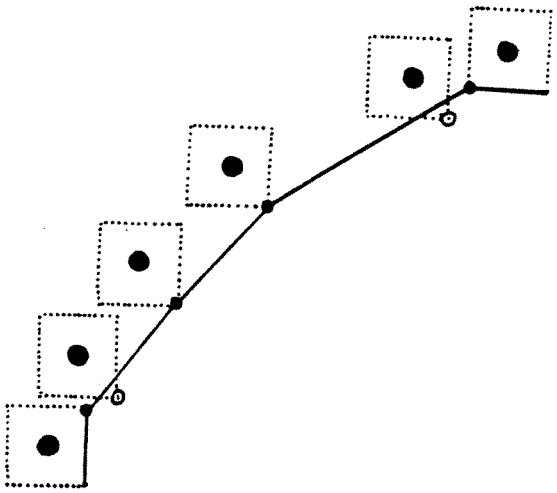


Figure 9.15: Picking vertices of $\mathcal{P}^*$

We observe that even a nontrivial $\overline{M_j}$ does *not* necessarily imply that the set B_j contains a vertex of $\mathcal{P}^*$ — see Figure 9.15. It is, however, not hard to see that for the vertices of $\mathcal{P}^*$ that correspond to the same vertex of B on their respective B_i, the following holds:

For all $\overline{M_i}$ that contain the same vertex v of B as $v + p_i \in \overline{M_i}$, the respective p_i form one consecutive sequence (say $p_1, \ldots, p_j$) as we go around the convex hull of P. Furthermore, $|\overline{M_2}| = \ldots = |\overline{M_{j-1}}| = 1$ and if $v + p_i$ and $v + p_{i+2}$ are vertices of $\mathcal{P}^*$, so is $v + p_{i+1}$.

We can use this to enumerate all possible choices of vertices from the $\overline{M_i}$ with $|\overline{M_i}| = 1$: From the sequence $p_1, \ldots, p_j$ for which $\overline{M_1}, \ldots, \overline{M_j}$ contain $v + p_i$, choose a smallest index a and a largest index z, such that precisely $v + p_a, \ldots, v + p_z$ are picked as vertices of $\mathcal{P}^*$. (Possibly $a = z$, i.e. only one of the $v + p_i$ is picked, or

$a > z$, i.e. none of the $v + p_i$ is picked.) Since there are only k vertices v of B, we get $O(n^{2k})$ different combinations.

Now consider the situation for any of these $O(n^{2k})$ choices, each consisting of $0 \leq s \leq kn$ vertices of known location. In the following, we speak of vertices with known location as *static* vertices of $\mathcal{P}^*$. The other $m \leq k$ vertices have to be chosen one each from the m nontrivial $\overline{M}_i$ and are called *mobile* vertices. For each $\overline{M}_i$, a mobile vertex can be chosen from one of k_i edges. We get $\prod_{i=1}^{m} k_i$ different combinations of edges to choose, where $\sum_{i=1}^{m} k_i = k$, resulting in $O(3^{\frac{k}{3}})$ combinations. In the following, consider one of these combinations at a time, so choose each mobile vertex v_i from an edge e_i. In determining the location of the mobile vertices, we make use of the fact that we can change the location of mobile vertices within their respective e_i until we are restricted by additional points.

Theorem 9.3.6 *Assume there is a convex polygon that intersects all sets B_i. Then there must be a polygon $\mathcal{P}^*$ that intersects all sets B_i and has the following properties:*

- *$\mathcal{P}^*$ has precisely one mobile vertex v_i in m of the edges e_i. The other vertices $v_{m+1}, \ldots, v_{m+s}$ are static.*
- *There are m anchor points $a_1, \ldots, a_m$ contained in specified edges of $\mathcal{P}^*$, such that each a_i is a vertex of some B_j.*
- *The static vertices $v_{m+1}, \ldots, v_{m+s}$ and the anchor points $a_1, \ldots, a_m$ determine $\mathcal{P}^*$.*

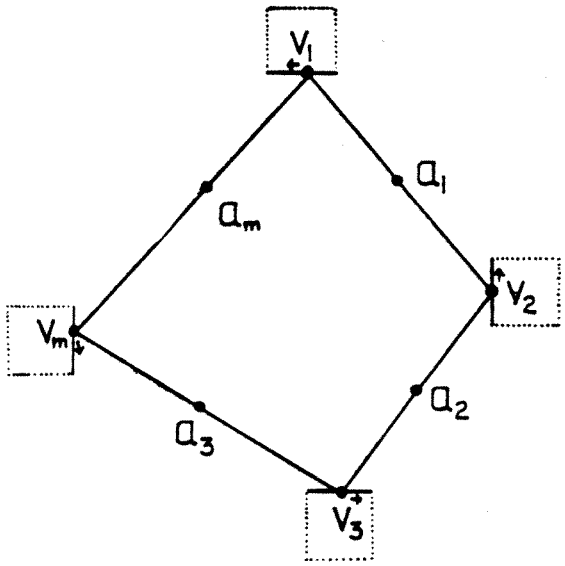


Figure 9.16: The situation if there are no static vertices.

Proof: From the previous discussion it follows that we can assume the first property. So consider such a polygon that has mobile vertices $v_1, \ldots, v_m$. We speak of the points $a_1, \ldots, a_m$ as *anchor points.*

We proceed by induction over the number m of mobile vertices. For each m, we have an induction over the number s of static vertices. We therefore consider two cases:

(CASE A) $\mathcal{P}^*$ has no static vertex:

If $\mathcal{P}^*$ has no static vertices and only one mobile vertex, the claim is trivial. Assume the claim was true for all situations with less than m mobile points.

We can change $\mathcal{P}^*$ by moving v_1 on e_1. See Figure 9.16. By Theorem 9.3.4, neither (v_m, v_1) nor (v_1, v_2) share any other point with B_1 than v_1. This means we

can move v_1 along e_1 such that the interior angle of $\mathcal{P}^*$ at v_m is increased and the interior angle of $\mathcal{P}^*$ at v_2 is decreased. (The situation where v_1 is a vertex of B_1 is part of case *1.* below.) Moving v_1 does not change the properties of $\mathcal{P}^*$ until one of the following "events" happens:

1. v_1 hits a vertex a_1 of e_1.
2. v_m ceases to be an extreme point of $\mathcal{P}^*$.
3. (v_m, v_1) or (v_1, v_2) hits a vertex a_1 of a B_i.

In Case 1., we can choose the anchor point $a_1 = v_1$. This determines the position of v_1 and we can treat v_1 like a static vertex, changing the number of mobile vertices to $m - 1$, proving the claim.

In Case 2., we can eliminate $\overline{M}_0$ from the list of nontrivial candidate sets. This reduces the number of mobile points to $m - 1$, again establishing the claim.

For Case 3., assume that (v_1, v_2) hits a_1. Add a_1 to the list of anchor points. Then continue with moving v_1 along e_1 and v_2 along e_2, such that (v_1, v_2) runs through a_1. As before, this can be done until we get another event. If v_1 or v_2 hit a vertex, we are done again; similarly, if a vertex ceases to be an extreme point of $\mathcal{P}^*$. If the edge (v_1, v_2) hits another vertex a_2, the anchor points a_1 and a_2 determine v_1 and v_2, allowing us to treat them as static points.

This leaves the situation where one of the edges (v_m, v_1) and (v_2, v_3) hits a vertex a_2; without loss of generality assume it is (v_2, v_3). Then we continue moving v_1, v_2 and v_3, such that (v_1, v_2) runs through a_1 and (v_2, v_3) runs through a_2.

Continuing in this manner, we either get an edge (v_i, v_{i+1}) that contains two anchor points, in which case the location of v_i and v_{i+1} is determined, or an anchor point a_i for each of the edges (v_i, v_{i+1}). In the latter case, we only need to discuss the situation where we cannot continue to move the v_i. Then the polygon $\mathcal{P}^*$ is determined by the information about the mobile vertices and the anchor points: Every v_i lies on an edge e_i, so it can be described by a linear variable x_i. The condition that (v_i, v_{i+1}) runs through a_i can be formulated as a linear equation involving x_i and x_{i+1}. The resulting system of m linear equations with m variables has a unique solution, since we cannot continue moving the v_i, i.e. cannot change any x_i.

(CASE 2) $\mathcal{P}^*$ has $s > 0$ static vertices:

By the induction hypothesis, we may assume that the claim is true if we have a $\mathcal{P}^*$ with less than s static vertices. Consider a chain $v_m, v_1, \ldots, v_l, v_{l+1}$ of consecutive vertices, such that v_m and v_{l+1} are static vertices and $v_1, \ldots, v_m$ are mobile vertices. (See Figure 9.17.) As before, consider moving v_1 along e_1 such that the interior angle of $\mathcal{P}^*$ at v_m increases. We can continue until one of the following events happens:

1. v_1 hits a vertex a_1 of e_1.
2. v_m ceases to be an extreme point of $\mathcal{P}^*$.
3. (v_m, v_1) hits a vertex a_1 of a B_i.
4. (v_1, v_2) hits a vertex a_1 of a B_i.

As before, Case 1. and Case 3. lead to a reduction of the number of mobile vertices,

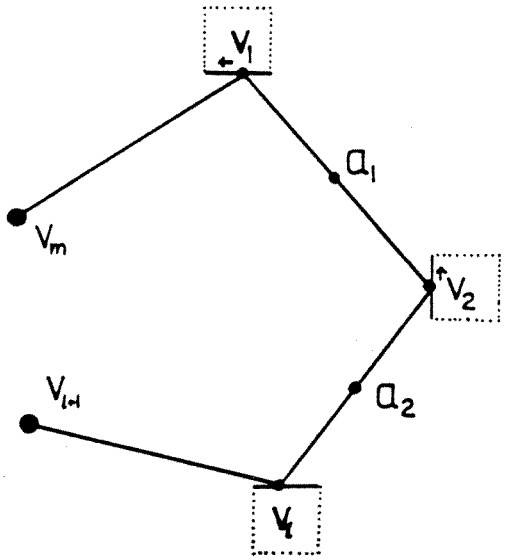


Figure 9.17: The situation if there are static vertices.

---

establishing the claim. In Case 2., we can remove v_m from the list of static vertices. Since this reduces the number of static vertices to $s-1$, the claim follows from the induction hypothesis. This leaves Case 4., which is treated similarly to Case 3. in our above discussion of $s = 0$: We add a_1 to the list of anchor points and move v_1 along e_1 and v_2 along e_2 such that (v_1, v_2) runs through a_1. Continuing this process, we eventually get one of the Cases 1.-3., or find an edge (v_i, v_{i+1}) with two anchor points or an anchor point and a static vertex.

This concludes the proof.

□

With the help of Theorem 9.3.6, we can solve the problem of convex stabbing for a family of identical convex k-gons in polynomial time:

Corollary 9.3.7 *Consider a family of n sets $B_1, \ldots, B_n$, such that $B_i = p_i + B$ for some point p_i and a convex k-gon B. Then we can decide in $O(k^{\frac{k}{3}} n^{3k+2})$ time whether there exists a convex curve that intersects all sets B_i.*

A special case occurs when B arises as the unit ball for some polygonal norm:

Corollary 9.3.8 *Let $L_{\mathcal{P}}$ be a polygonal error norm. Then we can decide in polynomial time whether each point in a set $\{p_1, \ldots, p_n\}$ can be moved by at most ε with respect to $L_{\mathcal{P}}$, such that the resulting set $\{p'_1, \ldots, p'_n\}$ forms a convex arrangement in the Euclidean sense.*

9.3.3 Euclidean Error Bounds

The previous section showed that we can achieve polynomiality when the error bounds are given by a polygonal norm. We made use of some linear properties of the resulting ε-balls around the points $p_1, \ldots, p_n$, so we cannot exploit this approach to deal with situations where the error bounds are given by the Euclidean norm. Since we achieved polynomiality for k-gons with a fixed k, we cannot use the approach of approximating the disks by a sequence of k-gons with increasing k. For Euclidean error bounds, we have to consider the following equivalent problem:

Problem 9.3.9 *Given a set of unit disks $D_1, \ldots, D_n$. Is there a convex curve that intersects all the disks?*

A detailed analysis of some of the resulting geometric aspects can be found in Salesin [87]. If we want to consider whether a polygon $\langle p_1, \ldots, p_n, p_1 \rangle$ is weakly

ε-convex, we have to consider disks of radius ε around the vertices. Salesin gives an $O(n^3)$ method for determining upper and lower bounds for the smallest ε that makes a given polygon weakly ε-convex: First approximate the smallest ε_1, for which the polygon is weakly ε_1-L_1-convex. Then approximate the smallest necessary ε_2 to transform the resulting L_1-convex polygon in one that is L_2-convex. Putting together ε_1 and ε_2 yields a lower and an upper bound for the true ε.

As already mentioned before, we are mainly interested in point sets instead of polygons. In the terminology of convex stabbing this means that we have no information about the order in which we have to encounter the disks. This seems to be a difficult aspect; but even knowledge of the order does not seem to solve the problem.

One of the reasons for this difficulty lies in the algebra that is involved: If we have to examine tangents of disks, we immediately have to consider algebraic expressions that involve square roots. It is not even clear whether the results can be expressed in polynomial space. If we need to consider solutions that contain irrational numbers, a considerable number of approaches cannot be successful.

This is similar to the problem arising from the Euclidean TSP, when we have to decide whether a given tour is smaller than some bound, i.e. decide whether a sum of square roots is smaller than a given rational number. (See Garey, Graham and Johnson [33].)

The basic question that arises for the convex stabbing of disks is whether we necessarily get solutions with irrational coordinates, assuming we start with rational input data. Superficially, it would seem so: All we have to do is construct an arrangement of disks that has a unique solution for the resulting intersection points.

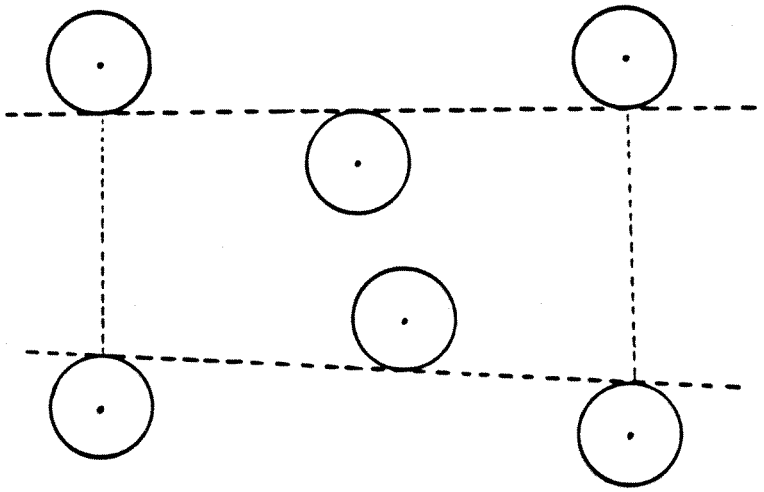


Figure 9.18: An arrangement of disks that forces a unique choice of intersection points

---

Choosing the input parameters in an appropriate manner should force irrational solutions. See Figure 9.18 for an example. (Considering arrangements without a unique solution may not be helpful, since it may be possible to perturb the intersection points and make them rational.)

The following Theorem 9.3.10, however, shows that we have to do a lot better for constructing an arrangement that has only intersection points with irrational coordinates:

Theorem 9.3.10 *Consider three r-balls for* any *norm* $\|\cdot\|$ *and* any *real* r. *Suppose there is a straight line that has exactly one point in common with each of the three* $\|\cdot\|$*-balls and separates one from the two others. If the* $\|\cdot\|$*-balls have rational centers, the touching points must also be rational.*

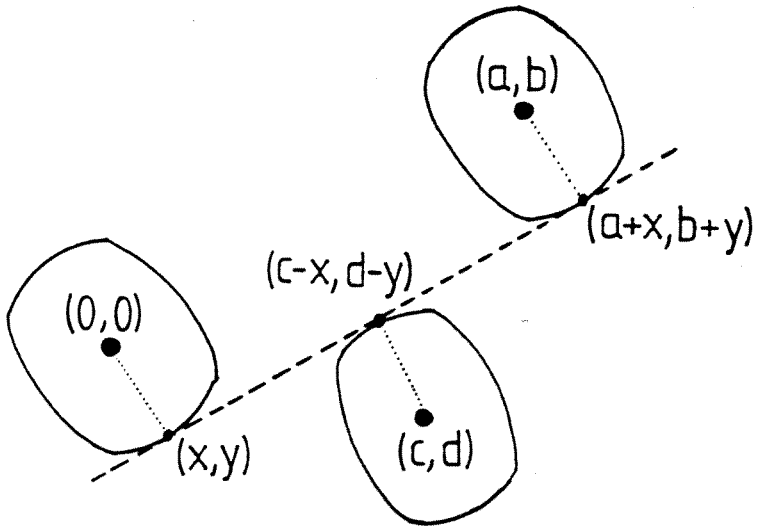


Figure 9.19: All touching points must be rational

Proof: See Figure 9.19. Suppose the centers of the balls are given as $p_1 = (0,0)$, $p_2 = (a,b)$, $p_3 = (c,d)$, with $p_1 \neq p_2$ and p_1, p_2, p_3 not collinear, i.e. $ad \neq bc$. Let the point on the first ball be denoted by $t_1 = (x,y)$. Every $\|\cdot\|$-ball is a closed convex set, so it has a continuous curve as its boundary. Therefore the other touching points must have coordinates $t_2 = (a+x, b+y)$ and $t_3 = (c-x, d-y)$ and (a,b) must be orthogonal to (x,y). This implies $bx = -ay$. Since t_1, t_2, t_3 are collinear, they form a triangle of area 0, i.e.

$$0 = \det((t_2 - t_1),(t_3 - t_1)) = \det((a,b),(c-2x, d-2y) = (ad - 2ay - bc + 2bx).$$

Since not both a and b are 0, assume $a \neq$. Then we get $4ay = ad - bc$, i.e. $y = \frac{ad-bc}{4a}$. Then b cannot be 0 and we get $x = \frac{bc-ad}{b}$, showing that all three points have rational coordinates.

□

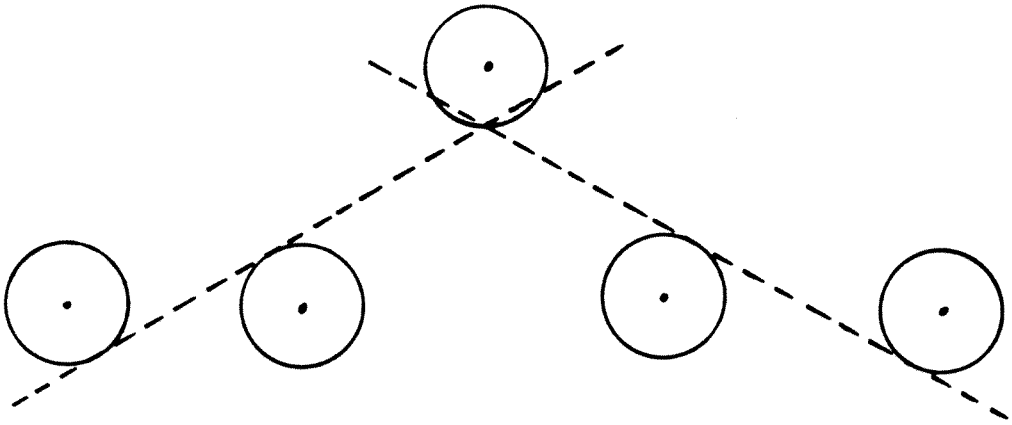


Figure 9.20: Another arrangement with a unique solution

There are more complicated arrangements with a unique solution that we can try. (See Figure 9.20 for an example.) They can always be described by a system of linear and quadratic Diophantine equations. For all arrangements that we have considered so far, extremely tedious algebraic calculations in the appropriate quadratic field extensions over the rationals show that s has to be rational if the centers of the disks are rational. This motivates the following conjecture:

Conjecture 9.3.11 *Assume that there is a convex curve that intersect all of a given sets of unit disks with rational centers. Then there is a convex curve that intersects every disk in a point with rational coordinates.*

We may consider other norms than the Euclidean L_2 norm for this rationality question. One choice may be the so-called L_p norms that are given as

$$\|(x,y)\| := (|x|^p + |y|^p)^{\frac{1}{3}} .$$

The existence of rational points (other than $(\pm 1, 0)$, $(0, \pm 1)$) on the boundary of the L_p unit balls for integer $p \geq 3$ has already attracted considerable attention: It is equivalent to Fermat's conjecture that the equation

$$x^p + y^p = z^p$$

cannot be solved with positive integers x, y and z whenever $p \geq 3$.

The margin is not wide enough to warn the reader sufficiently of all the consequences of dealing with this problem.

Appendix A

Rectilinear Planar Layouts

In Chapter 3, we need a particular way of embedding a planar graph into the Euclidean plane. In a *rectilinear planar layout*, every vertex is represented by a horizontal line segment, every edge is represented by a vertical line segment. Two vertices are connected by an edge if and only if the corresponding horizontal line segments have nonempty intersection with the vertical line segment representing the edge. (See Figure A.3 for a planar rectilinear layout for the graph shown in Figure 3.2.)

Rosenstiehl and Tarjan have described how every planar graph can be represented by such a layout. Their method takes only linear time and produces integer coordinates that are bounded by the number of vertices and faces. For our purposes, we need a modified version of their algorithm that does not produce any collinear line segments. We describe the basic steps; for more details, see Rosenstiehl and Tarjan [85].

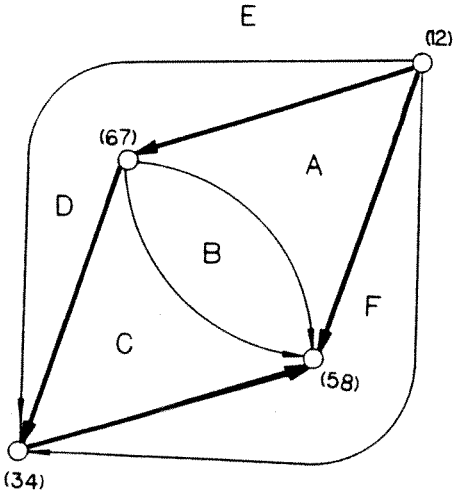


Figure A.1: A planar digraph D^*

A *bipolar order* $\mathcal{O}$ of an undirected graph G is a labelling of its vertices, such that every vertex except the one with the minimal label is adjacent to lower-labelled vertices and every vertex except the one with the maximal label is connected to higher-labelled vertices. This implies that the digraph induced by directing the edges of G from lower to higher vertices (a "bipolar orientation") has a unique source and a unique sink.

For the sake of simplicity, assume that the linear orders defined in the following are one-to-one mappings onto appropriate sets of the form $\{1, \ldots, k\}$.

1. Choose any mandatory edge $e^* = \langle v_a, v_z \rangle$. Pick any bipolar order $\mathcal{O}_V$ of the vertices of $\overline{G}$ such that v_a is the source (i.e. $\mathcal{O}_V(v_a) = 1$) and v_z is the sink (i.e. $\mathcal{O}_V(v_z) = \frac{m}{2}$.) $\mathcal{O}$ induces an acyclic digraph D^*, such that for every face f the edges having f to their left form one connected chain — see [85].

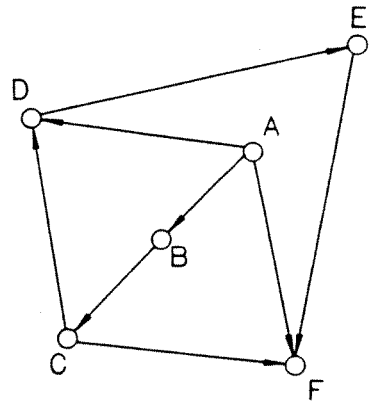


Figure A.2: The dual of D^* with edge directions induced by the linear order $\mathcal{O}_F$

(Assuming the contrary implies the existence of more than one source and sink.)

2. $\mathcal{O}_V$ induces a linear order $\mathcal{O}_F$ of the faces of D^* as follows: For two faces f_1 and f_2 that are separated by an edge $e \neq e^*$, $\mathcal{O}_F(f_1) < \mathcal{O}_F(f_2)$ if and only if f_1 lies to the left of e. If f_1 and f_2 are separated by e^*, $\mathcal{O}_F(f_1) < \mathcal{O}_F(f_2)$ if and only if f_1 lies to the right of e^*. (In the latter case, $\mathcal{O}_F(f_1) = 1$ and $\mathcal{O}_F(f_2) = \frac{m}{2} + 2$.) Since the order of the vertices is bipolar, it is not hard to see that this defines a linear order.

3. This further induces a linear order $\mathcal{O}_E$ of the edges of D^*:

   If two edges e_1 and e_2 have two different faces f_1 and f_2 to their left, let $\mathcal{O}_E(e_1) < \mathcal{O}_E(e_2)$ if and only if $\mathcal{O}_F(f_1) < \mathcal{O}_F(f_2)$. If e_1 and e_2 have the same face f_1 to their left, let $\mathcal{O}_E(e_1) < \mathcal{O}_E(e_2)$ if and only if e_1 comes before e_2 in the connected chain of directed edges that have f_1 to their left. (Note that $\mathcal{O}_E(e^*) = m$.) Again, it is not hard to see that this defines a linear order.

4. Represent any edge $e = (v_1, v_2)$ by the vertical line segment with end points

$$(\mathcal{O}_E(e), \mathcal{O}_V(v_1)) \text{ and } (\mathcal{O}_E(e), \mathcal{O}_V(v_2)).$$

5. Represent any vertex v by the horizontal line segment with end points

$$(\min(v), \mathcal{O}_V(v)) \text{ and } (\max(v), \mathcal{O}_V(v)),$$

where $\min(v) := \min\{\mathcal{O}_V(e) \mid e \text{ adjacent to } v\}$ and $\max(v) := \max\{\mathcal{O}_V(e) \mid e \text{ adjacent to } v\}$.

A proof of correctness can be found as Theorem 1 in [85]. (Lay out the faces of $\overline{G}$ one at a time in the order given by $\mathcal{O}_F$. Consider the moving *frontier* corresponding to a sequence of horizontal and vertical line segments forming the rightmost extent of the partial layout. Apply induction on the number of faces.)

Giving all edge segments the direction induced by D^* results in a directed rectilinear planar layout for D^*.

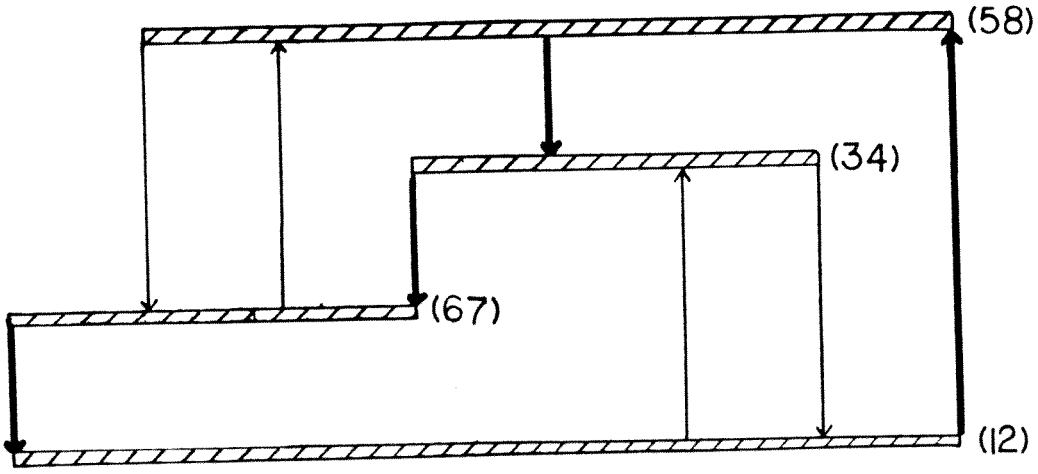


Figure A.3: A planar rectilinear layout for D^*

Bibliography

[1] A.AGGARWAL, H.IMAI, N.KATOH, S.SURI, Finding k Points with Minimum Diameter and Related Problems, *Journal of Algorithms*, 12 (1991), 38–56.

[2] M.BEN-OR, Lower bounds for algebraic computation trees, *Proceedings of the 15th Annual Symposium on Theory of Computing*, Association for Computing Machinery, 1983, 80–86.

[3] R.V.BENSON, *Euclidean Geometry and Convexity*, McGraw-Hill, New York, 1966, 188–189.

[4] L.M.BLUMENTHAL, K.MENGER, *Studies in Geometry*, W.H.Freeman and Company, 1970.

[5] C.DE BOOR, A.PINKUS Backward error analysis for totally positive linear systems, *MRC Technical Report #1620*, Mathematics Research Center, University of Wisconsin-Madison, 1976.

[6] J.E.BOYCE, D.P.DOBKIN, R.L.DRYSDALE, L.J.GUIBAS, Finding Extremal Polygons, *SIAM Journal of Comput.*, 14 (1985), 134–147.

[7] S.C.BOYD, The subtour polytope of the travelling salesman polytope, Ph.D. thesis, University of Waterloo, Waterloo, Ontario, Canada, 1986.

[8] S.C.BOYD, W.R.PULLEYBLANK, Optimization over the subtour polytope of the travelling salesman polytope, *Mathematical Programming*, 49 (1990), 163–187.

[9] H.BUSEMANN, *The Geometry of Geodesics*, Academic Press, 1955.

[10] B.CHAZELLE, Convex Partitions of Polyhedra: A Lower Bound and Worst-Case Optimal Algorithm, *SIAM Journal of Computing*, 13 (1984), 488–507.

[11] B.M.CHAZELLE, On the convex layers of a planar set, *IEEE Transactions on Information Theory* **IT-31**, 1985, 509–517.

[12] B.CHAZELLE, Triangulating a Simple Polygon in Linear Time, *Discrete and Computational Geometry*, 6 (1991), 485–524.

[13] B.CHAZELLE, L.PALIOS, Triangulating a Nonconvex Polytope, *Discrete and Computational Geometry*, 5 (1990), 505–526.

[14] N.CHRISTOFIDES, Worst-Case Analysis of a New Heuristic for the Travelling Salesman Problem, *Report 388*, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, PA, 1976.

[15] G.CORNUÉJOLS, J.FONLUPT, D.NADDEF, The travelling salesman problem on a graph and some related integer polyhedra, *Mathematical Programming*, 33 (1985), 1–27.

[16] H.S.M.COXETER, *Introduction to Geometry*, Wiley, 1969.

[17] J.CULBERSON, G.J.E.RAWLINS, Turtlegons: Generating Simple Polygons from Given Sequences of Angles, *Proceedings of the ACM Symposium on Computational Geometry*, 1985, 305–310.

[18] M.B.DILLENCOURT, Traveling salesman cycles are not always subgraphs of Delaunay triangulations or of minimum weight triangulations, *Information Processing Letters*, 24 (1987), 339–342.

[19] M.B.DILLENCOURT, A non-Hamiltonian, nondegenerate Delaunay triangulation, *Information Processing Letters*, 25 (1987), 149–151.

[20] G.B.DANTZIG, D.R.FULKERSON, S.M.JOHNSON, Solutions of a large-scale travelling salesman problem, *Operations Research*, 2 (1954), 393–410.

[21] L.DANZER, B.GRÜNBAUM, V.KLEE, Helly's theorem and its relatives in Convexity, *Proceedings of the Symposium on Pure Mathematics*, American Mathematical Society, 7 (1963), 101–180.

[22] D.P.DOBKIN, R.L.DRYSDALE, L.J.GUIBAS, Finding Extremal Polygons, in: *Advances in Computing Research*, Vol.1 (1983), 181–214.

[23] R.L.DRYSDALE, J.W.JAROMCZYK, A Note on Lower Bounds for the Maximum Area and Maximum Perimeter k-gon Problems *Information Processing Letters*, 32 (1989), 301–303.

[24] H.EDELSBRUNNER, *Algorithms in Combinatorial Geometry*, Springer Verlag, 1987.

[25] H.EDELSBRUNNER, H.A.MAURER, F.P.PREPARATA, A.L.ROSENBERG, E.WELZL, D.WOOD, Stabbing line segments, *BIT*, 22 (1982), 274–281.

[26] H.EDELSBRUNNER, F.P.PREPARATA, D.B.WEST, Tetrahedrizing point sets in three dimensions. Technical Report UIUCDCCS-R86-1310, Department of Computer Science, University of Illinois, 1986.

[27] D.EPPSTEIN, M.OVERMARS, G.ROTE, G.WOEGINGER, Finding minimum area k-gons, *Discrete and Computational Geometry*, 7 (1992), 45–58.

[28] P.ERDÖS, P.M.GRUBER, J.HAMMER, *Lattice Points* , Longman Scientific & Technical (copublished with Wiley, New York), 1989.

[29] M.ESPIE, Computing Efficiently Shortest Paths for Degenerate Metrics, *Proceedings of the 3rd Canadian Conference on Computational Geometry*, (1991), 149–152.

[30] L. FEJES TÓTH, *Lagerungen in der Ebene, auf der Kugel und im Raum*, Springer-Verlag, 1972.

[31] S.FORTUNE, Stable Maintenance of Point-Set Triangulation in Two Dimensions. In: *30th Annual Symposium on the Foundations of Computer Science*, IEEE, 1989, 494–499.

[32] J.W.GIVENS, Numerical computations of the characteristic values of a real symmetric matrix, *Oak Ridge National Laboratory Report ORNL-1574*, 1954.

[33] M.R.GAREY, R.L.GRAHAM AND D.S.JOHNSON, Some NP-complete geometric problems, *Proceedings of the 8th Annual Symposium on Theory of Computing*, Association for Computing Machinery, 1976, 10–22.

[34] M.R.GAREY AND D.S.JOHNSON, *Computers and Intractability, A guide to the theory of NP-completeness*, Freeman, San Francisco, 1979.

[35] R.E.GOMORY, T.C.HU, Multiterminal Network Flows, *Journal of SIAM*, 9 (1961), 551–570.

[36] M.T.GOODRICH, J.S.SNOYEINK, Stabbing Parallel Segments with a Convex Polygon, *in*: F.Dehne, J.-R.Sack, N.Santoro, eds., *Algorithms and Data Structures*, (WADS '89 Proceedings), Lecture Notes in Computer Science, vol. 382, Springer Verlag, 1989, 231–242.

[37] M.GRÖTSCHEL, On the symmetric travelling salesman problem: solution of a 120-city problem, *Mathematical Programming Study*, 12 (1980), 61–77.

[38] M.GRÖTSCHEL, M.PADBERG, On the symmetric travelling salesman problem I: inequalities, *Mathematical Programming*, 16 (1979), 265–180.

[39] M.GRÖTSCHEL, M.PADBERG, Polyhedral Theory, *in*: E.Lawler, J.Lenstra, A.Rinnooy Kan, D.Shmoys, eds., *The Traveling Salesman Problem*, Wiley, New York, 1985.

[40] M.GRÖTSCHEL, M.PADBERG, Polyhedral computations, *in*: E.Lawler, J.Lenstra, A.Rinnooy Kan, D.Shmoys, eds., *The Traveling Salesman Problem*, Wiley, New York, 1985.

[41] M.GRÖTSCHEL, L.LOVÁSZ, A.SCHRIJVER, The ellipsoid method and its consequences in combinatorial optimization, *Combinatorica*, 1 (1981), 169–197.

[42] B. VON HOHENBALKEN, Finding simplicial subdivisions of polytopes. *Mathematical Programming*, 21 (1981), 233–234.

[43] C.M.HOFFMAN, The problems of accuracy and robustness in geometric computation, *Computer*, 22 (1989), 31–42.

[44] C.M.HOFFMAN, J.E.HOPCROFT, M.S.KARASICK, Towards implementing robust geometric computations, *Proceedings of the Fourth Annual ACM Symposium on Computational Geometry*, 1988, 106–117.

[45] C.A.J.HURKENS, A.SCHRIJVER, É.TÁRDOS, On Fractional Multicommodity Flows and Distance Functions, *Discrete Mathematics*, 73 (1989), 99–109.

[46] A.ITAI, C.H.PAPADIMITRIOU, J.L.SZWARCFITER, Hamilton paths in grid graphs, *SIAM Journal of Computing*, 11 (1982), 676–686.

[47] D.S.JOHNSON AND C.H.PAPADIMITRIOU, Computational Complexity and the traveling salesman problem, *in*: E.Lawler, J.Lenstra, A.Rinnooy Kan, D.Shmoys, eds., *The Traveling Salesman Problem*, Wiley, New York, 1985.

[48] M.JÜNGER, W.R.PULLEYBLANK, New Primal and Dual matching Heuristics *IBM Research Report*, T.J.Watson Research Center, Yorktown Heights, NY 10598, 1991.

[49] M.JÜNGER, W.R.PULLEYBLANK, Geometric Duality and Combinatorial Optimization, *IBM Research Report RC 17863*, T.J.Watson Research Center, Yorktown Heights, NY 10598, 1992.

[50] R.KARP, C.PAPADIMITRIOU, On linear characterizations of combinatorial optimization problems, *Proceedings of the Twenty-first Annual Symposium on the Foundations of Computer Science*, IEEE Press, New York, 1980, 1–9.

[51] V.KLEE, What is a convex set?, *in*: A.K.Stehney, T.K.Milnor, J.E.D'Atri, T.F.Banchoff, eds., *Selected Papers on Geometry*, Mathematical Association of America, 1979, 218–233.

[52] R.KLEIN, *Concrete and Abstract Voronoi Diagrams*, Lecture Notes in Computer Science, Vol. 400, Springer, 1989.

[53] C.LANCZOS, *Applied Analysis*, Prentice-Hall, 1956.

[54] E.LAWLER, J.LENSTRA, A.RINNOOY KAN, D.SHMOYS, eds., *The Traveling Salesman Problem*, Wiley, New York, 1985.

[55] Z.LI, V.MILENKOVIC, Constructing Strongly Convex Hulls Using Exact or Rounded Arithmetic, *Proceedings of the Sixth Annual ACM Symposium on Computational Geometry*, 1990, 235–243.

[56] D.MCCALLUM, D.AVIS, A linear algorithm for finding the convex hull of a simple polygon, *Information Processing Letters*, 9 (1979), 201–206.

[57] K.MENGER, Untersuchungen über allgemeine Metrik, I, II, III. *Mathematische Annalen*, 100 (1928), 75–163.

[58] V.MILENKOVIC, Verifiable implementations of geometric algorithms using finite precision arithmetic, *Artificial Intelligence*, 37 (1989), 377–401.

[59] V.MILENKOVIC, *Verifiable Implementations of Geometric Algorithms Using Finite Precision Arithmetic*, Ph.D. thesis, Carnegie-Mellon University, 1988. Available as CMU report CMU-CS-88-168.

[60] V.MILENKOVIC, Calculating approximate curve arrangements using rounded arithmetic, *Proceedings of the Fifth Annual ACM Symposium on Computational Geometry*, 1989, 197–207.

[61] I.NIVEN, H.S.ZUCKERMANN, Lattice Points and polygonal Area *in:* A.K.Stehney, T.K.Milnor, J.E.D'Atri, T.F.Banchoff, eds., *Selected Papers on Geometry*, Mathematical Association of America, 1979, 218–233.

[62] J.O'ROURKE, *Polyhedral Object Models from 3D Points*, Universität Hamburg Technical Report IFI-HH-M-77/80, July 1980.

[63] J.O'ROURKE, Computational Geometry Column #3, *Computer Graphics*, 21 (1987), 314–315.

[64] J.O'ROURKE, *Art Gallery Theorems and Algorithms*, Oxford University Press, 1987.

[65] J.O'ROURKE, Uniqueness of orthogonal connect-the-dots, *in*: G.T. Toussaint ed., *Computational Morphology*, North-Holland, 1988, 97–104.

[66] J.O'ROURKE, Personal Communication, 1992.

[67] M.W.PADBERG, S.HONG, On the symmetric travelling salesman problem: a combinatorial study, *Mathematical Programming Study*, 12 (1980), 78–107.

[68] M.W.PADBERG, M.R.RAO, Odd Minimum Cut-Sets and b-Matchings, *Mathematics of Operations Research*, 7 (1982), 67–80.

[69] F.P.PREPARATA, M.I.SHAMOS, *Computational Geometry*, Springer, 1985.

[70] M.W.PADBERG, L.A.WOLSEY, Trees and Cuts, *Annals of Discrete Mathematics*, 17 (1983), 511–517.

[71] G.PICK, Geometrisches zur Zahlenlehre, *Naturwissenschaftliche Zeitschrift Lotus* (Prag) (1899), 311–319.

[72] J.PLESŃIK, The NP-Completeness of the Hamiltonian Cycle Problem in Planar Digraphs with Degree Bound Two, *Information Processing Letters*, 8 (1979), 199–201.

[73] C.H.PAPADIMITRIOU, K.STEIGLITZ, *Combinatorial Optimization: Algorithms and Complexity*, Prentice-Hall, 1982.

[74] C.H.PAPADIMITRIOU, U.V.VAZIRANI, On two geometric problems related to the travelling salesman problem, *Journal of Algorithms*, 5 (1984), 231–246.

[75] D.RAPPAPORT, Computing simple circuits from a set of line segments is NP-complete, *SIAM Journal of Comput.*, 18 (1989), 1128–1139.

[76] D.RAPPAPORT, On the complexity of computing orthogonal polygons from a set of points, *McGill University Technical report No. SOCS-86.9*, 1986.

[77] G.J.E.RAWLINS, *Explorations in Restricted Orientation Geometry*, Ph.D. thesis, Department of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 1987. (Available as Research Report CS-87-57.)

[78] G.J.E.RAWLINS, D.WOOD, On the optimal computation of finitely-oriented convex hulls, *Information and Computation*, 72 (1987) 150–166.

[79] G.J.E.RAWLINS, D.WOOD, Computational geometry with restricted orientations, in: *Proceedings of the 13th IFIP Conference on System Modelling and Optimization*, Lecture Notes in Computer Science, Springer 1988.

[80] G.J.E.RAWLINS, D.WOOD, Orthoconvexity and Its Generalizations, in: *Computational Morphology*, G.T.Toussaint ed., North-Holland, 1988, 137–152.

[81] J.E.REEVE, On the Volume of Lattice Polyhedra, *Proceedings of the London Mathematical Society*, 3 (1957), 378–395.

[82] D.REN, J.R.REAY, The Boundary Characteristic and Pick's Theorem in the Archimedean Planar Tilings, *Journal of Combinatorial Theory*, Series A, 44 (1987), 110–119.

[83] W.RINOW, *Die innere Geometrie der metrischen Räume*, Springer Verlag, 1961.

[84] W.RINOW, *Lehrbuch der Topologie*, VEB Deutscher Verlag der Wissenschaft, 1975.

[85] P.ROSENSTIEHL, R.E.TARJAN, Rectilinear Planar Layouts and Bipolar Orientations, *Discrete and Computational Geometry*, 1 (1986), 342–351.

[86] J.RUPERT, R.SEIDEL, On the Difficulty of Triangulating Three-Dimensional Nonconvex Polyhedra, *Discrete and Computational Geometry*, 7 (1992), 227–253.

[87] D.H.SALESIN, *Epsilon Geometry: Building Robust Algorithms from Imprecise Computations*, Ph.D. thesis, Stanford University, 1991.

[88] J.S.SALOWE, Selecting the kth Largest-Area Convex Polygon *in*: F.Dehne, J.-R.Sack, N.Santoro, eds., *Algorithms and Data Structures*, (WADS '89 Proceedings), Lecture Notes in Computer Science, vol. 382, Springer Verlag, 1989, 243–250.

[89] E.SCHÖNHARDT, Über die Zerlegung von Dreieckspolyedern in Tetraeder, ("On the Decomposition of Triangular Polyhedra into Tetrahedra"), *Mathematische Annalen*, 98 (1928), 309–312.

[90] S.SCHUIERER, *On Generalized Visibility*, Doctoral thesis, Albert-Ludwigs-Universität Freiburg i.Br., Germany, 1991.

[91] A.TAMIR, Problem 4-2 (New York University, Dept. of Statistics and Operations Research), *Problems Presented at the Fourth NYU Computational Geometry Day (13/3/87).*

[92] D.WOOD, Personal Communication, 1992.

[93] P.WIDMAYER, Y.F.WU, C.K.WONG, On some problems in fixed orientations, *SIAM Journal of Computing*, 16 (1987) 728–746.

[94] J.H.WILKINSON, *Rounding Errors in Algebraic Processes*, Prentice-Hall, 1963.

[95] H.WOLKOWICZ, Personal Communication, 1991.

[96] L.A.WOLSEY, Heuristic Analysis, linear programming and branch and bound, *Mathematical Programming Study*, 13 (1980) 121–134.

[97] D.M.YOUNG, R.T.GREGORY, *A Survey of Numerical Mathematics*, Addison-Wesley, 1972.