

RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN  
INSTITUT FÜR INFORMATIK I

---

Elmar Langetepe

# Bewegungsplanung für Roboter

Teil I: Offline-Bewegungsplanung

---

---

Wintersemester 2010/2011  
Dozent: Elmar Langetepe  
Skript: Tom Kamphans/Elmar Langetepe



## Einleitung

Die Vorstellung von autonomen Maschinen, die den Menschen schwierige und gefährliche Arbeiten abnehmen, hat immer wieder die Phantasie vieler Science-Fiction Autoren<sup>1</sup> beflügelt und viele Wissenschaftler unterschiedlicher Disziplinen beschäftigt. Das Feld “Robotics” involviert daher verschiedenste Bereiche, darunter:

- Elektrotechnik, Elektronik, Mechanik
- Kontroll- und Regelungstechnik
- Systemtheorie
- Softwareengineering
- Künstliche Intelligenz (Neuronale Netze, Fuzzy Logic)
- Algorithmik

Im Rahmen dieser Vorlesung werden wir uns auf den letzten Punkt konzentrieren und Probleme aus der Algorithmischen Geometrie behandeln, die im Zusammenhang mit der Steuerung von Robotern interessant sind.

Wir benötigen zur Behandlung dieser Fragestellungen einige Begriffe und Verfahren der Algorithmischen Geometrie, die hier nicht in aller Ausführlichkeit vorgestellt werden können, sondern nur kurz angerissen oder gar nur als “black-box” verwendet werden. Für ein tieferes Verständnis dieser Themen sei auf die Vorlesung Algorithmische Geometrie oder auf die einschlägige Literatur in diesem Bereich verwiesen.

Die diskutierten Themen sind Varianten einer Grundaufgabe: Gegeben ist ein Robotersystem mit  $k$  Freiheitsgraden, das sich in einer Umgebung  $U$  mit oder ohne (stationären) Hindernissen bewegt, ein Startpunkt  $s$  und ein Zielpunkt  $t$ . Stelle fest, ob es eine kollisionsfreie und kostengünstige Bewegung von  $s$  nach  $t$  gibt und falls ja, bestimme eine solche. Wichtige Kriterien für die vorgestellten Lösungen sind dabei:

- Korrektheit, d.h. es soll genau dann eine kollisionsfreie Bewegung gefunden werden, wenn eine solche existiert.
- Effizienz der Berechnung, denn natürlich soll die Bewegung schnell berechnet werden, der Roboter soll nicht vor lauter Rechenaufwand auf der Stelle stehen bleiben.
- Qualität der Lösung, d.h. die gefundene Bahn ist auch günstig bzgl. der festgelegten Kostenmaße.

---

<sup>1</sup>Der Begriff “Roboter” wurde von dem tschechischen Autor Karel Čapek (1890-1938) geprägt. [?]

Zunächst wollen wir die Betrachtung von Kollisionsproblemen zurückstellen, uns auf punktförmige Roboter beschränken und die Länge der vom Roboter zurückgelegten Bahnen als Qualitätskriterium betrachten.

## Literatur

Begleitend zu diesem Skript sei folgende Literatur empfohlen:

- Mark de Berg, Marc van Kreveld, Mark Overmars und Otfried Schwarzkopf:  
*Computational Geometry: Algorithms and Applications*.  
Springer-Verlag, Berlin 1997, [?]
- Rolf Klein:  
*Algorithmische Geometrie*.  
Addison–Welsey, Bonn 1997, [?]
- Micha Sharir und Pankaj K. Agarwal:  
*Davenport-Schinzel Sequences and Their Geometric Applications*.  
Cambridge University Press, New York 1995, [?]
- Micha Sharir:  
*Algorithmic Motion Planning*.  
Kapitel 40 in: Jacob E. Goodman und Joseph O’Rourke: *Handbook of Discrete and Computational Geometry*.  
CRC Press LLC, Boca Raton, FL, 1997, [?]
- Joseph S. B. Mitchell:  
*Geometric Shortest Paths and Network Optimization*.  
Kapitel 15 in: Jörg-Rüdiger Sack and Jorge Urrutia: *Handbook of Computational Geometry*.  
Elsevier Science Publishers B.V. North-Holland, 2000, [?]
- Jean-Claude Latombe:  
*Robot Motion Planning*.  
Kluwer Academic Publishers, Boston, 1991, [?]
- Jakob T. Schwartz und Chee-Keen Yap (Editor):  
*Advances in Robotics*.  
Lawrence Earlbaum, 1987, [?]
- Sowie viele in Journalen und Konferenzbänden erschienene Arbeiten und auch Dissertationen wie:  
A. Frank van der Stappen:  
*Motion Planning amidst Fat Obstacles*.  
Dissertation, Dept. Comput. Sci., Utrecht Univ., NL, 1994, [?]

An dieser Stelle sei auch auf das Geometrie-Labor

<http://www.geometrylab.de/>

mit Java-Applets aus dem Bereich Algorithmische Geometrie und Bewegungsplanung für Roboter hingewiesen.

## **Danksagungen**

Sehr herzlich danken wir Annette Ebbers-Baumann, Dorte Lübbert, Andrea Eubeler, Stefan Kamphans und allen Teilnehmern der Vorlesung für viele Korrekturhinweise und Verbesserungsvorschläge.

## **Korrekturhinweise**

Trotz aller Sorgfalt ist dieses Skript bestimmt nicht fehlerfrei. Bitte teilt uns Fehler aller Art oder unverständlich aufgeschriebenes per Mail an [langetep@cs.uni-bonn.de](mailto:langetep@cs.uni-bonn.de) mit. Eine Auflistung aller gefundenen Fehler findet ihr unter

[web.informatik.uni-bonn.de/I/Lehre/Vorlesungen/Bewegungsplanung/](http://web.informatik.uni-bonn.de/I/Lehre/Vorlesungen/Bewegungsplanung/)



# Kapitel 1

## Kürzeste Pfade

Die Untersuchung kürzester Wege beschäftigt die Wissenschaft bereits seit dem Altertum. In der Mathematik befaßt man sich in der Theorie der metrischen Räume und in der Differentialgeometrie damit. So zeigten bereits Joh. Bernoulli (1698) und L. Euler (1728), daß für Kurven  $C$  von  $a$  nach  $b$  auf einer glatten Fläche  $S$  gilt:  $C$  ist Kürzeste  $\implies C$  ist Geodätische, d. h.  $\forall p \in C : \text{Schmiegeebene}(C, p) \perp \text{Tangentialebene}(S, p)$ .

Dabei ist die Schmiegeebene die Ebene, an die  $C$  sich lokal anschmiegt,<sup>1</sup> und die Tangentialebene die in  $p$  tangential zu  $S$  liegende Ebene. Geodätische Kurven sind nur lokal nicht verkürzbar, im Gegensatz zu kürzesten Wegen, die auch global unverkürzbar sind. Anschaulich kann man sich geodätische Kurven als straff gespannte Gummibänder vorstellen. Die Umkehrung obigen Satzes gilt nicht: In Abbildung ?? ist die Helix von  $a'$  nach  $b'$  zwar eine Geodätische, aber keine Kürzeste.

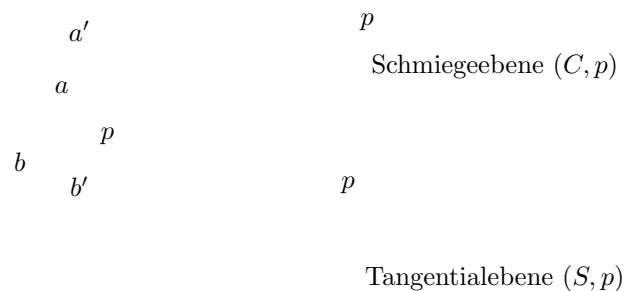


Abbildung 1.1: Kürzeste von  $a$  nach  $b$ , Geodätische von  $a'$  nach  $b'$ .

Wir interessieren uns hier für diskrete Umgebungen.

<sup>1</sup>Genauer: die Grenzlage einer durch drei benachbarte Kurvenpunkte  $p'$ ,  $p$ ,  $p''$  gegebene Ebene für die Grenzwerte  $p' \rightarrow p$  und  $p'' \rightarrow p$  [?].

## 1.1 Kürzeste Pfade in der Ebene mit polygonalen Hindernissen

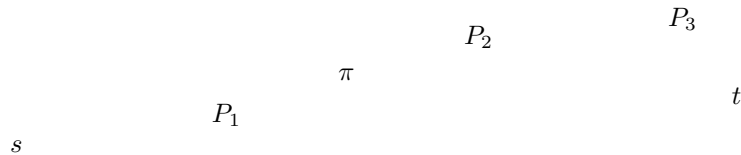


Abbildung 1.2: Umgebung mit polygonalen Hindernissen.

Die Umgebung, in der sich der Roboter bewegt, sei durch  $h$  Hindernisse in Form von Polygonen  $P_i$  mit insgesamt  $n$  Ecken gegeben. Dabei sind auch Polygone mit  $\overset{\circ}{P}_i = \emptyset$  erlaubt,<sup>2</sup> also zu einem Linienzug degenerierte Polygone. Weiterhin seien ein Startpunkt  $s$  und Zielpunkt  $t$  gegeben, die nicht im Inneren eines Polygons liegen dürfen; es muß gelten  $s, t \notin \bigcup \overset{\circ}{P}_i$ . Die Aufgabe ist nun, einen Pfad  $\pi$  minimaler Länge von  $s$  nach  $t$  zu finden, der keine Hinderniskante kreuzt.

Der kürzeste Pfad ist nicht immer eindeutig, je nach Umgebung kann es exponentiell in  $h$  viele kürzeste Wege geben, Abbildung ?? zeigt eine Umgebung mit  $h = 4m+1$  Hindernissen (Liniensegmente) und  $2^{m+1}$  kürzesten Wegen (Akman, 1987 [?]).

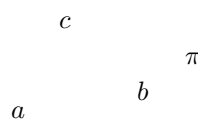


Abbildung 1.3: Es kann exponentiell viele kürzeste Wege geben.

<sup>2</sup> $\overset{\circ}{P}$  bezeichnet das Innere eines Polygons, also alle Punkte aus  $P$ , deren  $\varepsilon$ -Umgebung vollständig in  $P$  liegt.



Kürzeste Pfade haben die Gestalt einer polygonalen Kette, die nur an konvexen Polygonecken<sup>3</sup> abknicken kann, da sich die Kette andernfalls verkürzen ließe. Dies wird klar, wenn man sich den kürzesten Pfad als ein Gummiband vorstellt, das von  $s$  nach  $t$  gespannt ist: die Knicke im Gummiband müssen sich an einer Polygonecke abstützen, ohne eine Polygonecke würde sich das Gummiband strammziehen. In nebenstehender Abbildung ist  $\pi$  also kein kürzester Pfad, da sich der Knick in Punkt  $c$  nicht auf eine Polygonecke abstützt. Der kürzeste Pfad zwischen  $a$  und  $b$  wäre die gestrichelte Linie.



Genauer bestehen kürzeste Pfade aus einer Startkante von  $s$  aus, einer Zielkante zu  $t$  hin und ansonsten nur aus Kanten, die *gegenseitig sichtbare*, konvexe Hindernisecken verbinden. Diese Beobachtung führt zu einem ersten Ansatz zur Lösung des Problems, bei der wir auf eine aus der Algorithmischen Geometrie bekannte Struktur zurückgreifen können:

**Definition 1.1** Sei  $\mathcal{L}$  eine Menge sich nicht kreuzender Liniensegmente in der Ebene. Dann ist der **Sichtbarkeitsgraph** von  $\mathcal{L}$  ein Graph  $\text{VisG}(\mathcal{L}) = (V, E)$  mit

$$V = \{ \text{alle Endpunkte der Liniensegmente in } \mathcal{L} \}$$

$$E = \{ (p, q) \mid p, q \in V, \overline{pq} \text{ kreuzt kein Liniensegment aus } \mathcal{L} \}.$$

Der Sichtbarkeitsgraph einer Menge von Polygonen  $P_i$  entsteht aus  $\text{VisG}(\mathcal{L})$  mit  $\mathcal{L} = \{ \text{Kanten der } P_i \}$  durch Entfernen aller im Inneren eines Polygons liegenden Sichtbarkeitskanten.

Der Sichtbarkeitsgraph von  $n$  Liniensegmenten hat die Komplexität  $O(n^2)$ . Abbildung ??(i) zeigt ein Beispiel eines Arrangements von  $n$  Liniensegmenten, dessen Sichtbarkeitsgraph  $\Omega(n^2)$  viele Kanten enthält. Jedoch kann der Sichtbarkeitsgraph auch deutlich weniger Kanten enthalten, wie das Beispiel in Abbildung ??(ii) zeigt.

(i)

(ii)

VisG( $\mathcal{L}$ )

Abbildung 1.4: Der Sichtbarkeitsgraph kann (i)  $\Omega(n^2)$  viele Kanten enthalten, aber auch (ii) nur  $O(n)$  viele.

<sup>3</sup>Polygonecke, deren Außenwinkel  $> \pi$  ist.

---

**Algorithmus 1.1** Berechnung des kürzesten Weges mit Hilfe des Sichtbarkeitsgraphen

---

1. Berechne den Sichtbarkeitsgraph  $\text{VisG}(P)$ .
  2. Füge alle Kanten  $(s, v)$  mit  $v \in V$  und  $s$  sieht  $v$  ein:  
 $\forall v \in V$  : Teste, ob das Segment  $\overline{sv}$  von einer Randkante eines der Hindernisse geschnitten wird. Ist dies nicht der Fall, dann füge das Segment  $\overline{sv}$  zum Sichtbarkeitsgraphen hinzu.
  3. Füge alle Kanten  $(v, t)$  mit  $v \in V$  und  $t$  sieht  $v$  ein (analog).
  4. Berechne mit dem Algo. von Dijkstra in dem modifizierten Graph mit Distanzen als Kantengewichten alle kürzesten Wege von  $s$ .
- 

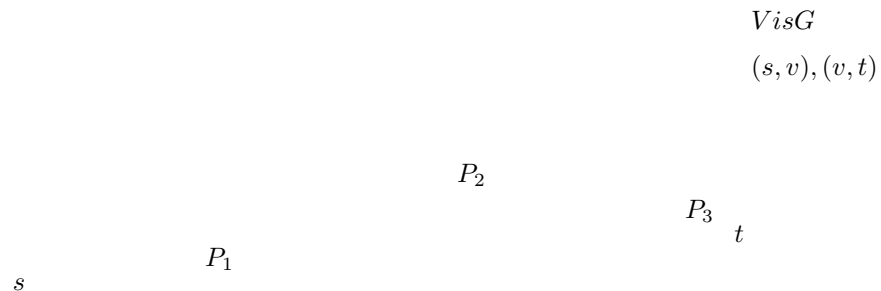


Abbildung 1.5: VisG und Kanten von  $s$  und zu  $t$ .

Damit läßt sich ein kürzester Weg wie in Algorithmus ?? beschrieben berechnen. Abbildung ?? zeigt das Ergebnis der Schritte 1 bis 3 auf der Umgebung aus Abbildung ?. Die zusätzlichen Kanten lassen sich offensichtlich in Zeit  $O(n^2)$  einfügen, die Komplexität der anderen beiden Schritte bedarf einer genaueren Analyse. Beginnen wir mit der Berechnung des Sichtbarkeitsgraphen: Dieser läßt sich naiv in Zeit  $O(n^3)$  berechnen, indem für jedes Paar  $(p, q)$  von Ecken das Segment  $\overline{pq}$  auf Kreuzung mit den  $n$  Polygonrandsegmenten getestet wird. Ein besseres Laufzeitverhalten hat der in Algorithmus ?? skizzierte radiale Sweep, der für ein festes  $p$  Zeit  $O(n \log n)$  benötigt, insgesamt also  $O(n^2 \log n)$ .

Dies kann noch verbessert werden, indem nicht jeder Punkt für sich einen radialen Sweep durchführt, sondern die Bearbeitung so synchronisiert

**Algorithmus 1.2** Sweep zur Berechnung des Sichtbarkeitsgraphen

Für jeden Endpunkt  $p$ :

- Sortiere alle Endpunkte rechts von  $p$  nach Winkeln von  $p$ .
- Drehe einen Strahl  $S$  von  $p$  gegen den Uhrzeigersinn von Süd nach Nord:
  - Führe in balanciertem Baum sortiert nach Abstand von  $p$  Buch darüber, welche Segmente von  $S$  gerade geschnitten werden.
  - Sei  $q$  der Schnittpunkt von  $S$  mit dem zu  $p$  nächsten Liniensegment auf  $S$ . Falls  $q \in V$ , berichte Sichtbarkeitskante  $\overline{pq}$ .

wird, daß Sichtinformationen weitergereicht werden können. Wenn z. B. in Abbildung ??(i) der Punkt  $f$  beim Sweep den Punkt  $a$  entdeckt, kann  $a$  die Information “ $\overline{cd}$  ist sichtbar” an  $f$  weiterreichen.

Die Wahl der Bearbeitungsreihenfolge  $R$  hängt dabei von der Steigung  $\text{slope}(p, q)$  des Segmentes zwischen  $p$  und  $q$  ab. Wir müssen dabei zwei Typen von Bedingungen erfüllen, siehe Abbildung ??(ii):

- (1)  $\text{slope}(p, q) < \text{slope}(p, r) \Rightarrow (p, q)$  vor  $(p, r)$  bearbeiten
- (2)  $\text{slope}(a, c) < \text{slope}(f, a) < \text{slope}(a, d) \Rightarrow (a, c)$  vor  $(f, a)$  vor  $(a, d)$

Man beachte, dass in Bedingung (2) weder die erste noch die zweite Ungleichung aus Bedingung (1) folgt.

Die Wahl der Bearbeitungsreihenfolge werden wir später betrachten. Nehmen wir zunächst an, wir hätten eine solche Bearbeitungsreihenfolge  $R$  bereits festgelegt. Nun wird zur Initialisierung zu jedem Punkt  $p$  ein Zeiger  $\wp$  auf das erste unterhalb von  $p$  liegende Segment berechnet; also das Liniensegment, das ein Strahl von  $p$  in negativer Y-Richtung als erstes treffen würde. Liegt unter  $p$  kein Liniensegment, wird der Zeiger auf NIL gesetzt.

Bei der Bearbeitung der Paare  $(p_i, q_i)$  aus der Liste  $R$  soll folgende Invariante erhalten bleiben:

Sei  $q$  der zuletzt bearbeitete Partner von  $p$ . Dann gibt es einen Zeiger  $\wp$  von  $p$  auf das in Sweeprichtung kurz hinter  $q$  sichtbare Segment  $\sigma$ .

Nach der Initialisierung ist die Invariante erfüllt,  $\wp$  zeigt auf das von  $p$  in Richtung Süden liegende Segment. Vor der Bearbeitung eines Paares  $(p, r)$  sei  $q$  der zuletzt bearbeitete Partner von  $p$ , und der Zeiger von  $p$  zeige auf  $\sigma$ . Es gibt drei mögliche Ausgangssituationen von  $p, q$  und  $\sigma$ , siehe Abbildung ??:

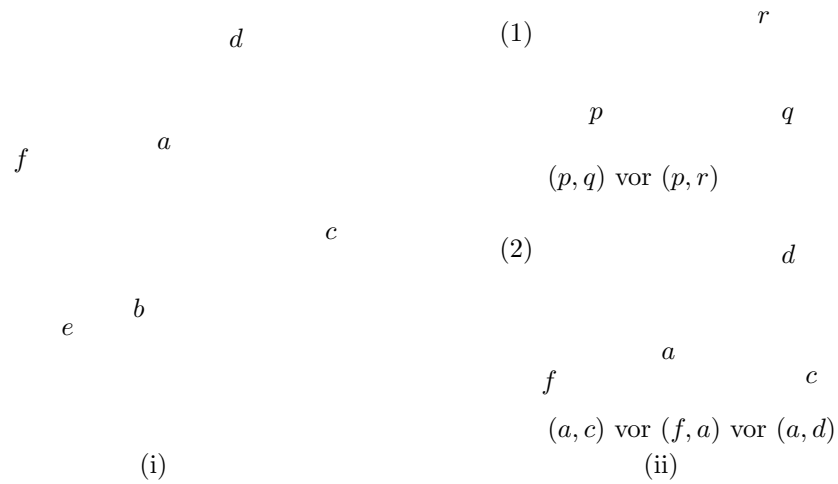


Abbildung 1.6: (i) Sichtbarkeitsinformationen können von  $a$  an  $f$  weitergereicht werden, (ii) Typen von Bedingungen an die Bearbeitungsreihenfolge.

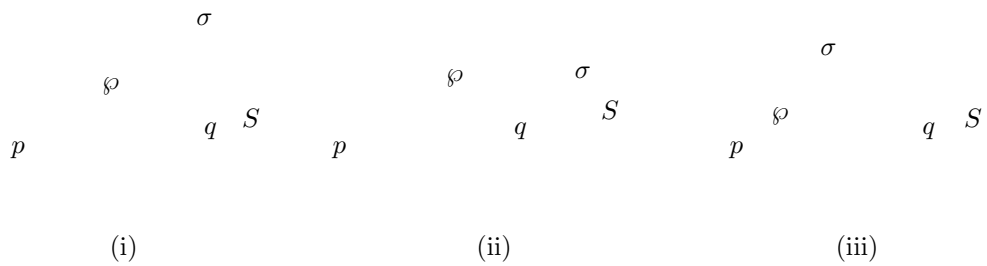


Abbildung 1.7: Es gibt drei mögliche Ausgangssituationen eines Punktes  $p$  mit zuletzt bearbeitetem Partner  $q$  und sichtbarem Segment  $\sigma$ .

- (i)  $q$  ist von  $p$  aus sichtbar und Anfangspunkt eines Segments  $\sigma$ .
- (ii)  $q$  ist von  $p$  aus sichtbar und Endpunkt eines Segments.  $\sigma$  ist das auf dem Strahl  $S$  von  $p$  durch  $q$  nächste Segment hinter  $q$ .
- (iii)  $q$  ist von  $p$  aus nicht sichtbar.  $\sigma$  ist das zu  $p$  nächste Segment auf dem Strahl  $S$  von  $p$  durch  $q$ .

Bei der Bearbeitung des Paares  $(p, r)$  können nun folgende Fälle eintreten, siehe Abbildung ??:

1.  $r$  liegt hinter  $\sigma$ . In allen drei Ausgangssituationen zeigt  $\wp$  nach wie vor auf  $\sigma$ .
2.  $r$  liegt vor  $\sigma$ . Dann ist  $r$  Anfangspunkt eines Segments  $\tau$  und  $\wp$  zeigt nun auf  $\tau$ .

3.  $r$  ist Endpunkt von  $\sigma$ . Hier muss die Sichtbarkeitsinformation für  $r$  benutzt werden. Je nach Ausgangssituation gibt es in diesem Fall folgende drei Möglichkeiten, die prinzipiell analog behandelt werden können:

- (i)  $\wp$  wird auf das erste Segment auf dem Strahl zwischen  $p$  und  $r$  hinter  $r$  gesetzt. Dies entspricht der Ausgangssituation (ii).
- (ii) In diesem Fall kann der Winkelbereich  $W$  zwischen den Strahlen von  $p$  durch  $q$  und  $r$  keinen Punkt enthalten, ansonsten wäre  $r$  nicht der nächste Partner von  $p$ . Außerdem muß  $r$  von  $p$  aus sichtbar sein. Andernfalls müßte sich ein Endpunkt in  $W$  befinden, oder  $\sigma$  wäre nicht das nach  $q$  sichtbare Segment.

Falls ein Segment  $\tau$  existiert auf das  $r$  zeigt, ist es durch die Bearbeitung des letzten Paares  $(r, v)$  bezüglich  $r$  entstanden, dass vor  $(p, r)$  dran war. Der Punkt  $v$  liegt unterhalb von  $W$  wegen Teil 2) der Bearbeitungsreihenfolge. Das Segment  $\tau = \overline{uv}$  muss vollständig den obigen Winkelbereich  $W$  durchqueren und ist somit sichtbar von  $p$  aus. Der untere Endpunkt  $u$  kann nicht oberhalb vom Strahl durch  $r$  und  $v$  liegen ( $u$  könnte allerdings mit  $v$  identisch sein), da  $\tau$  ja die Sichtbarkeitsbedingung erfüllen muss. Der obere Endpunkt  $w$  kann nicht zwischen dem Strahl durch  $r$  und  $v$  und dem Bereich  $W$  liegen, da dann  $(r, w)$  noch nach  $(r, v)$  und vor  $(p, r)$  betrachtet worden wäre. Der Endpunkt  $w$  kann wie oben bereits erwähnt nicht im Bereich  $W$  liegen. Folglich sieht  $p$  einen Punkt  $r'$  auf  $\tau$  und  $\wp$  wird also auf  $\tau$  gesetzt.

- (iii) Analog zu (ii).

Damit können wir den Sichtbarkeitsgraphen nach der Berechnung der Bearbeitungsreihenfolge  $R$  in Zeit  $O(n^2)$  bestimmen. Um  $R$  festzulegen, müssen wir die Liste der Punktpaare  $(p_i, q_i)$  so sortieren, daß die Steigungen der Liniensegmente  $\overline{p_i q_i}$  die geforderten Eigenschaften (1) und (2) haben. Dabei können wir eine aus der Algorithmischen Geometrie bekannte Eigenschaft ausnutzen, die Dualität zwischen Punkten und Geraden, die gegeben ist durch die Abbildung<sup>4</sup>

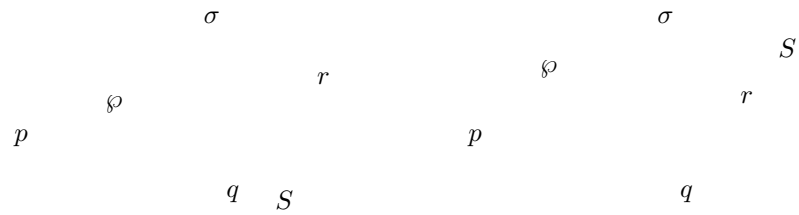
$$\text{Punkt } p = (p_x, p_y) \mapsto \text{Gerade } p^* = \{Y = p_x X - p_y\}.$$

Seien  $p = (p_x, p_y)$  und  $q = (q_x, q_y)$  o. E. mit  $p_x < q_x$  gegeben. Dann gilt für die Steigung des Segments

$$\text{slope}(p, q) = \frac{q_y - p_y}{q_x - p_x}.$$

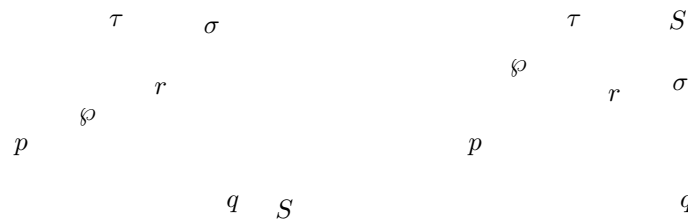
<sup>4</sup>Diese Dualisierung unterscheidet sich nur im Vorzeichen von der in [?] betrachteten Dualisierung  $p = (p_x, p_y) \mapsto p^* = \{Y = p_x X + p_y\}$ .

Fall 1:  $r$  liegt hinter  $\sigma$  (Situation (i))



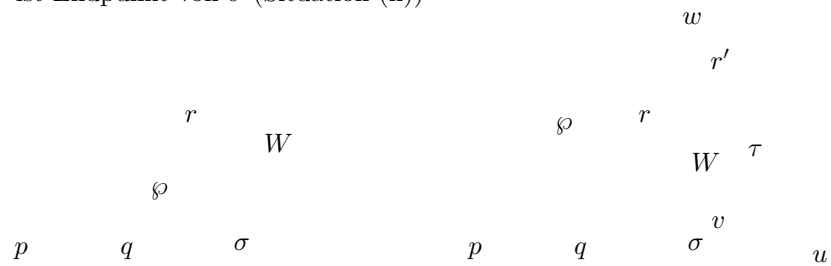
Analog Ausgangssituationen (ii) und (iii)

Fall 2:  $r$  liegt vor  $\sigma$  (Situation (i))



Analog Ausgangssituationen (ii) und (iii)

Fall 3:  $r$  ist Endpunkt von  $\sigma$  (Situation (ii))



Analog Ausgangssituationen (i) und (iii)

Abbildung 1.8: Drei Fälle bei der Bearbeitung eines Paares  $(p, r)$ .

Dem entspricht die X-Koordinate des Schnittpunktes  $(x, y)$  der zu  $p$  und  $q$  dualen Geraden  $p^*$  und  $q^*$ :

$$Y = p_x X - p_y = q_x X - q_y$$

$$\Rightarrow X = \frac{q_y - p_y}{q_x - p_x} = \text{slope}(p, q).$$

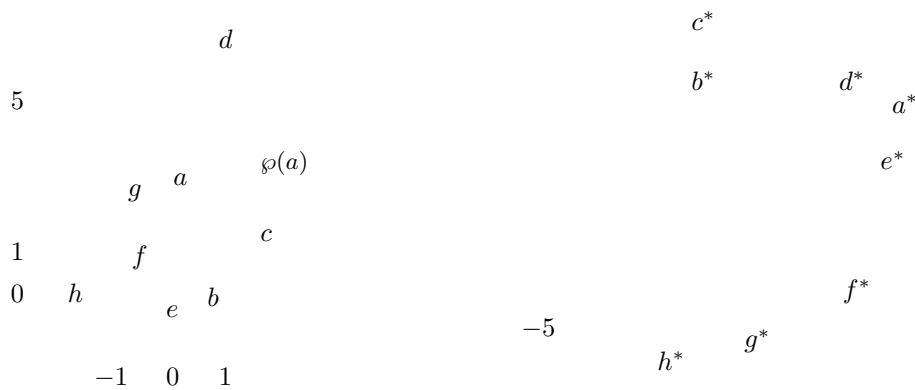


Abbildung 1.9: Dualisierung  $p = (p_x, p_y) \mapsto p^* = \{Y = p_x X - p_y\}$ .

Wir können also die Bedingungen aus Abbildung ??(ii) so übersetzen:

- (1)  $\text{slope}(p, q) < \text{slope}(p, r) \Leftrightarrow p^* \cap q^*$  links von  $p^* \cap r^*$  auf Gerade  $p^*$
- (2a)  $\text{slope}(a, c) < \text{slope}(f, a) \Leftrightarrow a^* \cap c^*$  links von  $f^* \cap a^*$  auf Gerade  $a^*$
- (2b)  $\text{slope}(f, a) < \text{slope}(a, d) \Leftrightarrow f^* \cap a^*$  links von  $a^* \cap d^*$  auf Gerade  $a^*$

Also müssen wir die Schnittpunkte im Arrangement der Geraden  $p_i^*$  so bearbeiten, daß längs jeder Geraden die richtige Reihenfolge eingehalten wird. Abbildung ?? zeigt ein Beispiel eines Arrangements mit den dazu dualen Geraden. Da  $\text{slope}(a, c) < \text{slope}(f, a)$  gilt, muß der Schnittpunkt von  $a^*$  mit  $c^*$  links des Schnittpunktes von  $f^*$  mit  $a^*$  auf  $a^*$  liegen.

Zu lösen ist nun folgendes Problem: Gegeben seien  $n$  Geraden im  $\mathbb{R}^2$  in allgemeiner Lage, die eine Zerlegung der Ebene bewirken. Gesucht ist eine Reihenfolge der Schnittpunkte, die mit der X-Ordnung längs jeder Geraden verträglich ist. Dazu sind folgende Verfahren bekannt:



Abbildung 1.10: (i) Topologischer Sweep, (ii) Elementarschritt, (iii) Die Sweepkurve kann eine Zelle nur von oben betreten und nach unten verlassen.

Inkrementelle Konstruktion des Arrangements (Edelsbrunner, O'Rourke, Seidel, 1986, [?])	Rechenzeit	Speicherplatz
Sweep mit einer Geraden	$O(n^2)$	$O(n^2)$
Sweep mit einer Kurve (topologischer Sweep, Edelsbrunner, Guibas, 1986, [?, ?])	$O(n^2 \log n)$	$O(n)$
	$O(n^2)$	$O(n)$

Im folgenden konzentrieren wir uns auf den topologischen Sweep. Die wesentliche Idee dabei ist, den Sweep in der Ebene nicht mit einer Geraden durchzuführen, sondern mit einer Kurve, die topologisch einer Geraden entspricht:

**Definition 1.2** In einem Arrangement  $\mathcal{A}$  von Geraden heißt eine Kurve  $S$  **Pseudogerade**, wenn sie jede Gerade aus  $\mathcal{A}$  in höchstens einem Punkt schneidet.

Beim Sweep wird also eine Sweepkurve  $S$  (Pseudogerade) über das Arrangement  $\mathcal{A}$  bewegt (Abbildung ??(i)). Die von  $S$  geschnittenen Geraden werden längs  $S$  sortiert in einem Verzeichnis mitgeführt. In einem Elementarschritt des Sweeps wird die Sweepkurve über genau einen Schnittpunkt hinweg geführt (Abbildung ??(ii)). Das Problem dabei ist, zwei konsekutive Geraden längs  $S$  mit Schnittpunkt rechts von  $S$  für das nächste Elementarereignis zu finden, wenn im Mittel dafür Zeit  $O(1)$  zur Verfügung steht.

Die Existenz solcher zwei Geraden ist jedenfalls sichergestellt: Seien  $G_1$  und  $G_2$  die Geraden, deren Schnittpunkt rechts von  $S$  am weitesten links liegt, dann müssen  $G_1$  und  $G_2$  entlang  $S$  benachbart sein. Ein Elementarschritt ist also stets durchführbar, wenn es auch im allgemeinen mehrere Kandidaten dafür gibt.



Ein Ansatz zur Lösung dieses Problems ist, in einer Liste  $N$  entlang  $S$  sortiert für jede Gerade  $G$  festzuhalten, von welcher anderen Geraden  $G'$  das von  $S$  durchkreuzte Segment von  $G$  rechts begrenzt wird. Für das Arrangement in Abbildung ??(i) ergibt sich also entlang der ersten eingezeichneten Sweepkurve  $S_1$  die Liste

$$N = \{ (1, 2), (2, 1), (3, 5), (4, 5), (5, 4) \}.$$

Aus dieser Liste läßt sich ablesen, daß die Geraden 1 und 2 sowie die Geraden 4 und 5 benachbart und somit Kandidaten für den nächsten Elementarschritt sind. Für die zweite Sweepkurve  $S$  ergibt sich die Liste

$$N = \{ (1, 2), (2, 1), (5, 1), (3, 4), (4, 3) \},$$

und Kandidaten für den nächsten Elementarschritt sind die Paare (1, 2) und (3, 4).

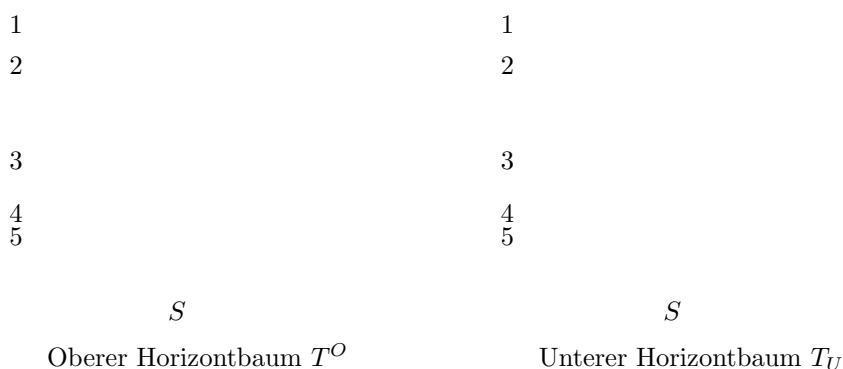


Abbildung 1.11: Oberer und unterer Horizontbaum.

Nach einem Elementarschritt ist die Liste  $N$  zu aktualisieren. Dabei helfen uns obere und untere Horizontbäume, die wie folgt definiert sind, vgl. Abbildung ??:

Oberer Horizontbaum  $T^O$ :

- Markiere jede von  $S$  geschnittene Kante.
- Setze Markierung längs der Geraden nach rechts fort.
- Wo sich zwei markierte Kanten treffen, setze diejenige mit größerer Steigung nach rechts fort.

Unterer Horizontbaum  $T_U$ :

- Markiere jede von  $S$  geschnittene Kante.
- Setze Markierung längs der Geraden nach rechts fort.
- Wo sich zwei markierte Kanten treffen, setze diejenige mit kleinerer Steigung nach rechts fort.

Eine Eigenschaft der Horizontbäume ist folgende: Sei  $e$  eine von  $S$  geschnittene Kante von  $\mathcal{A}$ , und seien  $e^O$  und  $e_U$  ihre Verlängerungen in  $T^O, T_U$ . Dann ist  $e = \min(e^O, e_U)$ , d. h. aus  $T^O$  und  $T_U$  kann schnell bestimmt werden, von welchen Geraden die Kante  $e$  begrenzt wird; die kürzere Kante liefert den Schnittpunkt mit der begrenzenden Geraden.

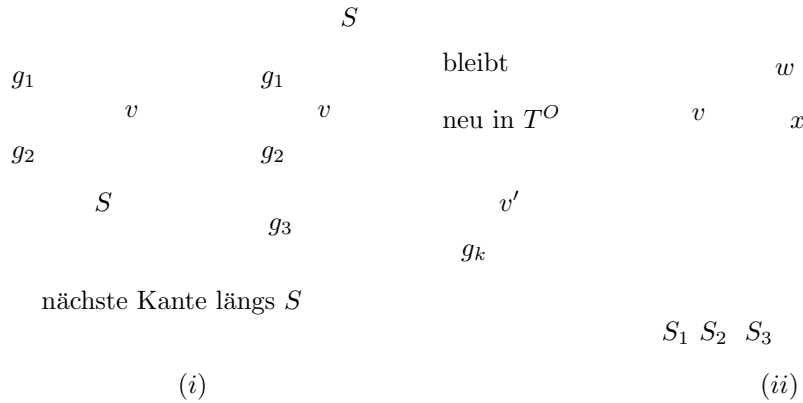


Abbildung 1.12: Aktualisierung der Horizontbäume am Beispiel von  $T^O$ .

Abbildung ??(i) zeigt die Aktualisierung der Horizontbäume am Beispiel von  $T^O$ : Nachdem die Sweepkurve  $S$  den Punkt  $v$  überschritten hat, wird die Gerade  $g_1$ , die bisher nur bis  $v$  markiert war, hinter  $v$  von  $S$  geschnitten. Somit wird das Segment  $\overline{vv'}$  der Geraden rechts von  $v$  Teil des Horizontbaumes.  $v'$  ist dabei der zu  $v$  nächstgelegene Schnittpunkt von  $g_1$  mit einer markierten Geraden  $g_k$  mit größerer Steigung. Um diese zu finden, suchen wir längs  $S$  die nächste Schnittkante  $g_3$  und folgen der konvexen Kantenfolge bis zum Schnittpunkt mit  $g_1$ . Die Suche nach einem Schnittpunkt kann dabei  $O(n)$  viele Schritte erfordern. Es kann auch vorkommen, daß dieselben Kanten mehrfach abgesucht werden müssen; Abbildung ??(ii) zeigt ein Arrangement, in dem bei den Knoten  $v$  und  $x$  dieselben Kanten durchsucht werden.

Um eine Übersicht über die Kosten zu bekommen, wollen wir die Kosten des Besuchs einer Kante  $e$  dem verursachenden Knoten  $v$  in Rechnung stellen. Ideal wäre, wenn Bilanzen nur zwischen Knoten und Kanten derselben Zelle im Arrangement aufgestellt werden müßten. Abbildung ??(ii) zeigt allerdings, daß dies nicht ohne weiteres möglich ist: die Kanten, die bei der Bearbeitung von  $v$  besucht werden, liegen nicht alle in derselben Zelle wie  $v$ . Dies wird erst bei der Bearbeitung von  $x$  festgestellt, und  $x$  liegt in derselben Zelle wie die Kanten, also transferieren wir die Kosten dieser Kanten auf  $x$ . Zur Berechnung dieser Kosten brauchen wir Folgendes:

**Definition 1.3** Sei  $\mathcal{A}$  ein Arrangement von  $n$  nicht senkrechten Geraden, und sei  $\ell$  eine beliebige Gerade. Die **Zone** von  $\ell$  in  $\mathcal{A}$  besteht aus allen

$\ell$

Abbildung 1.13: Zone einer Geraden  $\ell$  in einem Arrangement  $\mathcal{A}$ .

Zellen, deren Abschluß von  $\ell$  geschnitten wird:

$$\text{Zone}(\ell) := \{ \text{Zellen } Z \text{ in } \mathcal{A} \mid \ell \cap \bar{Z} \neq \emptyset \}$$

**Theorem 1.4** (Zonentheorem)

Sei  $\mathcal{A}$  ein Arrangement von  $n$  nicht senkrechten Geraden und sei  $\ell$  eine beliebige Gerade. Die Zone von  $\ell$  hat die Komplexität  $O(n)$ .

**Beweis.**

In jeder Zelle gibt es einen rechten, obersten Punkt und einen linken, untersten Punkt. Diese Punkte teilen den Rand einer Zelle eindeutig in einen linken und einen rechten Rand.

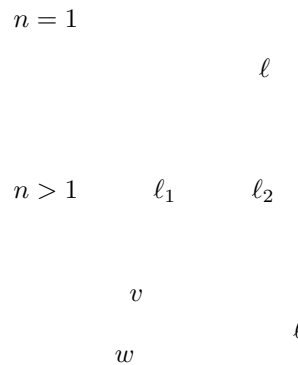
Wir zeigen durch Induktion über  $n$ , daß es höchstens  $5n$  linke Kanten auf den Rändern der Zonenzellen geben kann. Aus Symmetriegründen gibt es auch höchstens  $5n$  rechte Kanten.

Für  $n = 1$  haben wir eine linke Kante.

Für  $n > 1$  sei  $\ell_1$  die Gerade im Arrangement, deren Schnitt mit  $\ell$  am weitesten rechts liegt (wir nehmen an, diese sei eindeutig).

Sei  $v$  der niedrigste Schnitt von  $\ell_1$  mit einer Geraden aus  $\mathcal{A}$  oberhalb von  $\ell$  und  $w$  der höchste Schnitt von  $\ell_1$  mit einer Geraden aus  $\mathcal{A}$  unterhalb von  $\ell$ .

Sei  $\mathcal{A}'$  das Arrangement ohne  $\ell_1$ . Lt. Induktionsvoraussetzung hat die Zone von  $\ell$  in  $\mathcal{A}'$  höchstens  $5(n - 1)$  linke Kanten.



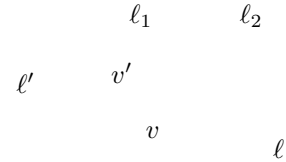
Wieviele linke Kanten werden nun beim Einfügen von  $\ell_1$  höchstens hinzugefügt?

- (i) Die Kante  $\overline{vw}$ .
- (ii) Sei  $\ell_2$  die Gerade, die  $\ell_1$  in  $v$  schneidet.  $\ell_2$  hat vor dem Einfügen von  $\ell_1$  möglicherweise mehrere linke Kanten gestellt. Höchstens eine davon kann durch Einfügen von  $\ell_1$  in  $v$  geteilt werden.

(iii) Entsprechend für  $w$ .

Es kommen also höchstens drei linke Kanten hinzu, also hat  $\text{Zone}(\ell)$  in  $\mathcal{A}$  höchstens  $5(n - 1) + 3 \leq 5n$  linke Kanten.

Zu zeigen ist noch, warum beim Einfügen von  $\ell_1$  keine weiteren linken Kanten geteilt werden können. Nehmen wir dazu an, die Gerade  $\ell'$  würde  $\ell_1$  oberhalb von  $\ell$  schneiden



$\Rightarrow \exists$  Schnittpunkt  $v'$  oberhalb von  $v$ .

$\Rightarrow$  kurz rechts von  $\ell_1$  kann kein Stück von  $\ell'$  zu  $\text{Zone}(\ell)$  gehören, weil  $\ell'$  dort in einer Region liegt, die nach unten hin von  $\ell_1$  und  $\ell_2$  begrenzt wird.

$\Rightarrow$  es kann keine Teilung stattgefunden haben.



Abbildung 1.14: Im rechtesten Schnittpunkt schneiden sich zwei oder mehrere Geraden. Die gestrichelten Kanten werden neu eingefügt.

Ist die Gerade  $\ell_1$  nicht eindeutig, d. h. in dem weitesten rechts liegenden Schnittpunkt mit  $\ell$  schneiden sich mehrere Geraden, so wählen wir eine dieser Geraden als  $\ell_1$ . Mit der gleichen Zählweise wie oben stellen wir fest, daß wir für zwei schneidende Geraden 5 neue linke Kanten bekommen können, siehe Abbildung ??(i), für drei oder mehrere Geraden 4 linke Kanten (Abbildung ??(ii)).

□

Wir können die Kosten für die Aktualisierung der Horizontbäume wie folgt abschätzen:

$$\begin{aligned}
 & \text{Kosten der Aktualisierung von } T^O, T_U \\
 & \approx \text{Summe aller Kosteneinheiten Kante–Knoten aus je einer Zelle} \\
 & \leq \sum_{\text{Zellen } Z} (\#\text{Kanten von } Z) \cdot (\#\text{Knoten von } Z) \\
 & = \sum_{\text{Zellen } Z} (\#\text{Kanten von } Z)^2
 \end{aligned}$$

$$\begin{aligned} &\leq \sum_{\text{Geraden } G} \underbrace{\sum_{\substack{\text{Zellen } Z \text{ mit} \\ \text{Kante auf } G}} (\#\text{Kanten von } Z)}_{O(n) \text{ nach Theorem ??}} \\ &\in O(n^2) \end{aligned}$$

Bemerkung: Die Kosten der Aktualisierung der Horizontbäume, nachdem die Sweep Pseudogerade über einen Schnittpunkt hinweggehoben wurde, entstehen durch entlang laufen des Randes der aktuellen Zelle  $z$ . Der Rand einer Zelle  $z$  kann also insgesamt maximale Anzahl Knoten von  $z$  mal durchlaufen werden. Die Gesamtkosten ergeben sich durch Summation über alle Zellen  $z$ .

Damit haben wir folgende Laufzeiten gezeigt:

Aktualisierung der Horizontbäume	$O(n^2)$
$\Rightarrow$ Topologischer Sweep	$O(n^2)$
$\Rightarrow$ Wahl der Bearbeitungsreihenfolge	$O(n^2)$
$\Rightarrow$ Konstruktion des Sichtbarkeitsgraphen	$O(n^2)$

Also gilt:

**Theorem 1.5** (Welzl, 1985)

Der Sichtbarkeitsgraph von  $n$  Liniensegmenten kann in Zeit  $O(n^2)$  und in Speicherplatz  $O(n)$  konstruiert werden. [?]

Dabei werden sichtbare Paare von Endpunkten nur berichtet, nicht gespeichert. Außerdem setzen wir voraus, daß sich die Punkte in allgemeiner Lage befinden, insbesondere keine drei Punkte kollinear sind.<sup>5</sup>

Den Sichtbarkeitsgraphen in Algorithmus ?? können wir also in Zeit  $O(n^2)$  konstruieren, auch das Hinzufügen der Kanten zu Start- und Zielpunkt ist in  $O(n^2)$  möglich. Damit bleibt aus Algorithmus ?? das Problem der Suche nach einem kürzesten Weg in ungerichteten Graphen mit nicht-negativen Kantengewichten; dies ist ein diskretes Problem.

Dieses Problem läßt sich mit dem Algorithmus von Dijkstra (Algorithmus ??, [?]) lösen. Dieser verwaltet eine Welle  $W$  mit einem Ausläufer  $A$  um  $a$  und aktualisiert in jedem Schritt die Knotenmarkierung

$$d(p) = \begin{cases} \text{min. Länge eines Pfads von } a \text{ nach } p & \text{für } p \in W \\ \text{min. Länge eines Pfads von } a \text{ nach} \\ p, \text{ der bis auf } p \text{ in } W \text{ verläuft} & \text{für } p \in A \\ \infty & \text{sonst} \end{cases}$$

<sup>5</sup>Ansonsten wären noch einige Spezialfälle zu beachten, die aber an Laufzeit und Speicherplatz nichts ändern würden.

**Algorithmus 1.3** Kürzeste Pfade in gewichtetem Graph (Dijkstra, 1959)

**Gegeben:** Graph  $G = (V, E)$  zusammenhängend,  $|V| = n$ ,  $|E| = m$   
 Kantengewichtung  $g : E \rightarrow \mathbb{R}^{\geq 0}$ , Knoten  $a, b \in V$ .

**Gesucht:** Kantenfolge von  $a$  nach  $b$  mit minimalem Gesamtgewicht.

Solange  $A \neq \emptyset$ :

- Entnehme  $A$  ein  $p$  mit minimalem  $d(p)$ .
- $W := W \cup \{p\}$ .
- Für alle direkten Nachbarn  $q$  von  $p$  in  $W^C$ :
  - $d(q) := \min \{ d(q), d(p) + g(p, q) \}$
  - Wenn  $q \notin A$  dann  $A := A \cup \{q\}$ .

Abbildung ?? zeigt einen Zwischenstand bei der Berechnung des kürzesten Weges mit der aktuellen Welle  $W$ , dem aktuellen Ausläufer  $A$  und den aktuellen Knotenmarkierungen (fett).

Zur Implementierung des Algorithmus von Dijkstra wird eine Prioritätswarteschlange für den Ausläufer  $A$  benötigt. Diese soll eine Menge von Paaren  $(p, d(p))$  verwalten. Die Operationen Einfügen eines Elementes, Löschen des Minimums werden insgesamt  $O(n)$  mal durchgeführt. Vermindern des Wertes  $d(p)$  kann insgesamt  $O(m)$  mal durchgeführt werden. Die Elemente befinden sich bei diesen Operationen in bekannter Position, d. h. es kann ohne vorherige Suche darauf zugegriffen werden. Zur Realisierung dieser Warteschlange gibt es folgende Möglichkeiten:

Führen wir  $i$  mal einfügen,  $j$  mal vermindern und  $k$  mal löschen des Minimums in beliebiger Reihenfolge aus.

- Lineare Liste, unsortiert: Einfügen, Vermindern in  $O(1)$ , Löschen–Min in  $O(i)$ . Gesamtlaufzeit in  $O(i^2 + j)$
- Fibonacci Heap (Fredman, Tarjan, 1987; beschrieben u.a. im Buch von Ottmann und Widmayer, [?]): Einfügen, Vermindern in  $O(1)$  und Löschen–Min  $O(\log i)$ . Gesamtlaufzeit in  $O(k \log i + i + j)$ . [?]
- Relaxed Heap (Driscoll, Gabow, Shrairaman, Tarjan, 1988): Einfügen, Vermindern in  $O(1)$ , Löschen–Min in  $O(\log i)$ . Gesamtlaufzeit in  $O(k \log i + j)$  [?]

Wird der Algorithmus von Dijkstra mit Fibonacci–Heaps implementiert, kann man in einem Graph  $G$  mit  $n$  Knoten und  $m$  Kanten die kürzesten

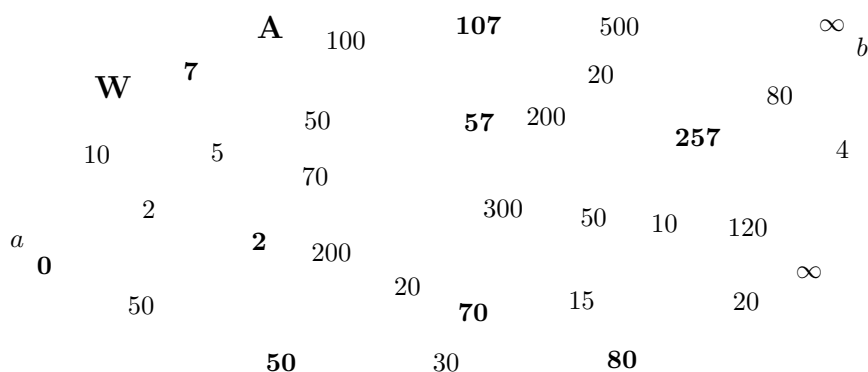


Abbildung 1.15: Welle, Ausläufer und Knotenmarkierungen während des Algorithmus von Dijkstra.

Wege von einem Knoten  $a$  zu allen übrigen Knoten in Zeit  $O(n \log n + m)$  bestimmen.

Damit ist folgendes gezeigt:

**Theorem 1.6**

Gegeben seien  $h$  Polygone mit  $n$  Ecken, ein Startpunkt  $s$  und ein Zielpunkt  $t$ . Ein kürzester Weg von  $s$  nach  $t$  kann in Zeit  $O(n^2)$  berechnet werden, die kürzesten Wege von  $s$  zu allen Ecken ebenfalls in Zeit  $O(n^2)$ .

Der Engpaß bei der bisherigen Berechnung kürzester Wege ist die Konstruktion des Sichtbarkeitsgraphen. Wie bereits erwähnt, kann dieser  $\Omega(n^2)$  viele Kanten enthalten, jedoch auch viel weniger, vgl. Abbildung ???. Beim obigen Algorithmus brauchen wir stets eine Laufzeit von  $\Omega(n^2)$  wegen der Konstruktion des Arrangements der dualen Geraden. Die Frage ist nun, ob sich der Sichtbarkeitsgraph auch output-sensitiv bzgl. seiner Größe berechnen läßt. Tatsächlich ist dies möglich: Zu einem Arrangement aus  $n$  Liniensegmenten, dessen Sichtbarkeitsgraph aus  $m$  Kanten besteht, sind folgende output-sensitiven Algorithmen bekannt:

Overmars, Welzl, 1988: Zeit  $O(m \log n)$  und Platz  $O(n)$  [?],

Ghosh, Mount, 1987: Zeit  $O(n \log n + m)$  und Platz  $O(m)$  [?, ?],

Pocchiola, Vegter, 1995: Zeit  $O(n \log n + m)$  und Platz  $O(n)$  [?].

Damit gilt:

**Theorem 1.7** Mit der Fibonacci-Heap Implementierung des Algorithmus von Dijkstra und der output-sensitiven Berechnung des Sichtbarkeitsgraphen lassen sich kürzeste Wege in Zeit  $O(n \log n + m)$  berechnen.

Bei der Laufzeit ist jedoch der Anteil von  $O(n \log n)$  unvermeidlich:

**Theorem 1.8** (untere Schranke) *In einer Umgebung von Polygonen mit insgesamt  $n$  Ecken braucht die Berechnung eines kürzesten Pfades Zeit  $\Omega(n \log n)$ .*



Abbildung 1.16: Reduktion: Konstruktion der konvexen Hülle auf Bestimmung des kürzesten Pfades.

**Beweis.**

Wir zeigen, daß sich die Konstruktion der konvexen Hülle auf die Berechnung kürzester Pfade reduzieren läßt. Da sich Sortieren auf die Konstruktion der konvexen Hülle reduzieren läßt, erhalten wir eine Laufzeit von  $\Omega(n \log n)$ .

Sei eine Menge von Punkten gegeben. Sei  $s$  der linkeste Punkt,  $t$  der rechteste Punkt und  $H$  die Höhe der vertikale Abstand zwischen niedrigsten und höchsten Punkt. Lege an jeden Punkt ein senkrecht nach unten zeigendes Liniensegment einer Länge  $> 2H$ . Dann entspricht der kürzeste Weg von  $s$  nach  $t$  der oberen konvexen Hülle, siehe Abbildung ??.

Die nächste Frage ist, ob man wirklich den Sichtbarkeitsgraphen konstruieren muß. In der Tat ist dies nicht nötig, folgender Ansatz ist als Locus Approach bekannt und sei hier nur kurz umrissen:

$$r_1 \quad \{ z \mid |zp| = |zq| + a_1 - a_2 \}$$

The diagram shows points  $q$ ,  $a_1$ ,  $p$ ,  $a_2$ ,  $s$ ,  $r_1$ , and  $r_2$ . The locus is defined as the set of points  $z$  such that  $|zp| = |zq| + a_1 - a_2$ .

Abbildung 1.17: Berechnung der Shortest Path Map.



Wir zerlegen den freien Teil der Ebene in Regionen, so daß alle Punkte einer Region kombinatorisch gleiche<sup>6</sup> Kürzeste zum Startpunkt  $s$  haben (Shortest Path Map). In Abbildung ?? sind genau für die Punkte auf der Hyperbel  $H = \{z \mid |zp| = |zq| + a_1 - a_2\}$  die Wege von  $s$  über  $p$  oder über  $q$  gleich lang. Dies entspricht einem Voronoi-Diagramm mit additiven Gewichten, siehe Abschnitt ?. Die Shortest Path Map läßt sich mit der Continuous Dijkstra Methode berechnen: Wir lassen um  $s$  einen Kreis wachsen. Wenn dieser eine Hindernisecke berührt, entsteht ein neuer Kreis, der mit der gleichen Geschwindigkeit wie die bisherigen Kreise um die Hindernisecke wächst (Wellenfront). Es wird festgehalten, von welchem Kreis ein Punkt der Ebene zuerst erreicht wird. Auf die Continuous Dijkstra Methode werden wir in Abschnitt ?? noch zurückkommen, wir halten hier nur fest:

**Theorem 1.9** (Hershberger, Suri, 1997)

*Gegeben seien polygonale Hindernisse mit insgesamt  $n$  Ecken und ein fester Startpunkt  $s$ . Dann läßt sich in Zeit  $O(n \log n)$  und Platz  $O(n \log n)$  eine Karte der kürzesten Wege berechnen. Diese braucht  $O(n)$  Platz und ermöglicht für einen beliebigen Zielpunkt  $t$*

- (i) in Zeit  $O(\log n)$  die Entfernung von  $s$  zu bestimmen und*
- (ii) in Zeit  $O(\log n + k)$  den kürzesten Weg von  $s$  nach  $t$  zu berichten, wobei  $k$  die Anzahl der Kanten auf dem Weg ist. [?]*

---

<sup>6</sup>D. h. sie berühren dieselben Ecken in derselben Reihenfolge.

## 1.2 Kürzeste Pfade im Inneren eines einfachen Polygons



Abbildung 1.18: Kürzester Pfad im Inneren eines einfachen Polygons.

Betrachten wir nun das Problem, zu einem gegebenem einfachen Polygon mit  $n$  Ecken und zwei beliebigen Punkten  $s$  und  $t$  im Inneren des Polygons den kürzesten Pfad  $\pi$  zu bestimmen, der im Inneren des Polygons verläuft. Im Gegensatz zu dem Problem in Abschnitt ?? ist der kürzeste Pfad hier eindeutig bestimmt. Man kann sich den Pfad auch in diesem Fall als straff gespanntes Gummiband zwischen  $s$  und  $t$  vorstellen, das wie in Abschnitt ?? nur an den Ecken des Polygons abknicken kann.

### 1.2.1 Der Algorithmus von Lee und Preparata

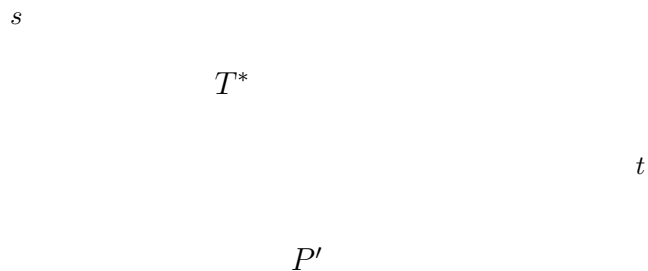


Abbildung 1.19: Triangulation  $T$  und dualer Graph  $T^*$ .

Wie in Abschnitt ?? ist unser Ansatz auch hier, das gegebene Problem in ein diskretes Problem zu verwandeln; dabei hilft uns die aus der Algorithmischen Geometrie bekannte Triangulation des Polygons:

**Definition 1.10** Zu einem einfachen Polygon  $P$  ist die **Triangulation**  $T$  eine maximale Menge von kreuzungsfreien Liniensegmenten im Innern von

$P$ , deren Endpunkte auf den Ecken von  $P$  liegen. Der zur Triangulation  $T$  duale Graph  $T^*$  enthält die Dreiecke der Triangulation als Knoten und Kanten zwischen den Knoten von Dreiecken mit einer gemeinsamen Seite.

Aus der Algorithmischen Geometrie [?] kennen wir ein randomisiertes Verfahren zur Berechnung der Triangulation in Zeit  $O(n \log n)$ , ein lineares Verfahren wurde 1991 von Chazelle vorgestellt [?]; Abbildung ?? zeigt die Triangulation und den dualen Graphen zu dem Beispiel aus Abbildung ??.

Der duale Graph  $T^*$  ist ein Baum mit Knotengrad  $\leq 3$ , also entspricht dem kürzesten Pfad  $\pi$  von  $s$  nach  $t$  ein eindeutiger Pfad  $\pi_{T^*}$  im Baum  $T^*$  von dem Dreieck, das  $s$  enthält, zu dem Dreieck, das  $t$  enthält. Nach Lokalisation der Dreiecke, die  $s$  bzw.  $t$  enthalten, läßt sich  $\pi_{T^*}$  durch Tiefensuche in  $T^*$  in linearer Zeit bestimmen.

Betrachten wir das Teilpolygon  $P'$ , das die Dreiecke längs  $\pi_{T^*}$  sowie zwei angesetzte Dreiecke mit Ecken  $s$  bzw.  $t$  enthält und sich in Zeit  $O(n)$  bestimmen läßt, so muß der kürzeste Pfad  $\pi$  in  $P'$  verlaufen. Der nicht zu  $P'$  gehörende Teil ist in Abbildung ?? schraffiert dargestellt.

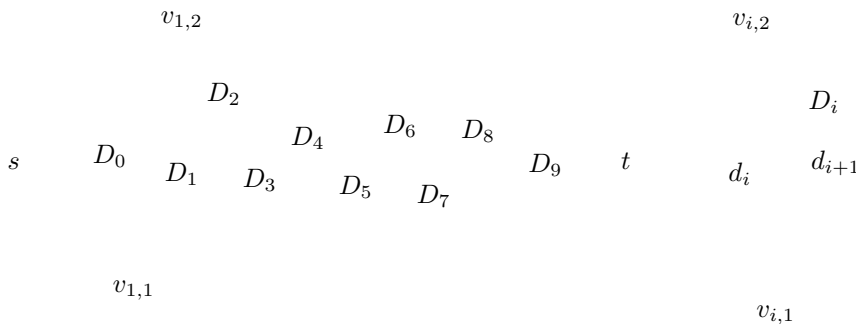


Abbildung 1.20: Kette von Dreiecken.

Damit reduziert sich das anfängliche Problem auf die Bestimmung des kürzesten Pfades von  $s$  nach  $t$  in  $P'$ . Dieses Polygon hat jedoch eine spezielle Eigenschaft: es besteht aus einer Kette von Dreiecken, bei der zwei konsekutive Diagonalen  $d_i, d_{i+1}$  immer eine gemeinsame Ecke haben, vgl. Abbildung ???. Das erlaubt uns, induktiv die Kürzesten  $\pi(s, v_{i,1})$  und  $\pi(s, v_{i,2})$ ,  $1 \leq i \leq n$  von  $s$  zu den Endpunkten  $v_{i,1}$  und  $v_{i,2}$  der Diagonalen  $d_i$  zu bestimmen: der Fall  $i = 1$  ist trivial, der kürzeste Weg besteht aus den beiden Kanten des Polygons mit Ecke  $s$ . Nehmen wir also per Induktion an, der kürzeste Pfad von  $s$  zu  $v_{i,1}$  und  $v_{i,2}$  sei bekannt. Die Gestalt der Pfade  $\pi(s, v_{i,1})$  und  $\pi(s, v_{i,2})$  ist in Abbildung ?? dargestellt: nach einem eventuellen gemeinsamen Teil vom Startpunkt  $s$  bilden sie einen Trichter mit der Diagonalen  $d_i$  als Oberseite.

Sei o. B. d. A.  $v_{i+1,1} = v_{i,1}$ . Zu berechnen ist also der kürzeste Weg von  $s$  nach  $v_{i+1,2}$ . Dazu starte in  $v_{i,2}$  und suche die Wand des alten Trichters

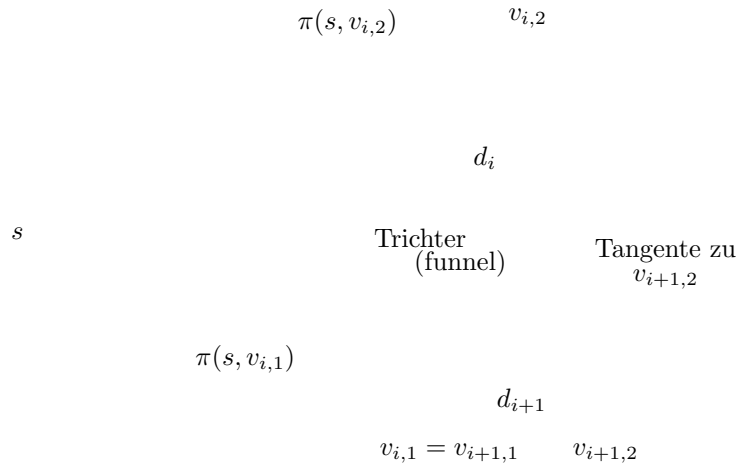


Abbildung 1.21: Trichter.

gegen den Uhrzeigersinn nach einem Tangentenpunkt von  $v_{i+1,2}$  ab. Diese Tangente gehört zum kürzesten Weg von  $s$  nach  $v_{i+1,2}$ . Dabei durchlaufene Kanten fallen entweder weg oder gehören zu einem gemeinsamen Stück des neuen Trichters und werden nie wieder besucht, so daß sich eine Laufzeit von  $O(n)$  für die Bestimmung des kürzesten Weges von  $s$  nach  $t$  in  $P'$  ergibt. Damit haben wir folgendes Theorem gezeigt, vgl. Algorithmus ??:

**Theorem 1.11** (Lee, Preparata, 1984)

Der kürzeste Pfad zwischen zwei Punkten in einem einfachen Polygon kann in Zeit  $O(n)$  und Platz  $O(n)$  bestimmt werden. [?]

---

**Algorithmus 1.4** Kürzester Pfad in einem Polygon (Lee, Preparata)

---

- Berechne Triangulation  $T$  und dualen Graphen  $T^*$  des Polygons  $P$ .  $O(n)$
  - Bestimme die Dreiecke  $D_s$  und  $D_t$  die  $s$  bzw.  $t$  enthalten.  $O(n)$
  - Bestimme das Teilpolygon  $P'$ , das nur Dreiecke längs des Pfades von  $D_s$  nach  $D_t$  in  $T^*$  enthält (Tiefensuche).  $O(n)$
  - Bestimme induktiv die Kürzesten  $\pi(s, v_{i,1})$  und  $\pi(s, v_{i,2})$ ,  $1 \leq i \leq n$  von  $s$  zu den Endpunkten  $v_{i,1}$  und  $v_{i,2}$  der Diagonale  $d_i$ .  $O(n)$
- 

### 1.2.2 Der Algorithmus von Guibas und Hershberger

Das Ergebnis von Lee und Preparata läßt sich verbessern; wir benötigen dazu folgendes Theorem:

**Theorem 1.12** (*Cutting–Theorem, Chazelle, 1982*)

Sei  $T$  eine Triangulation eines einfachen Polygons  $P$ , dann existiert eine Diagonale  $d \in T$ , so daß auf jeder Seite von  $d$  mindestens  $\frac{n}{3} - 1$  Dreiecke liegen. [?]

**Beweis.** Übungsaufgabe. Hinweis: Man nutze die Tatsache aus, daß der duale Graph von  $T$  ein Baum ist.  $\square$

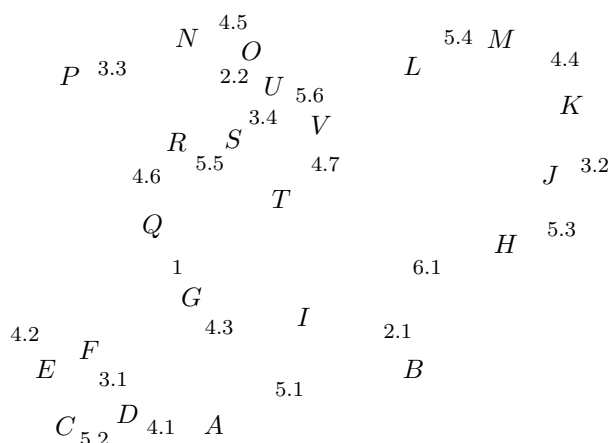


Abbildung 1.22: Balancierte hierarchische Triangulation.

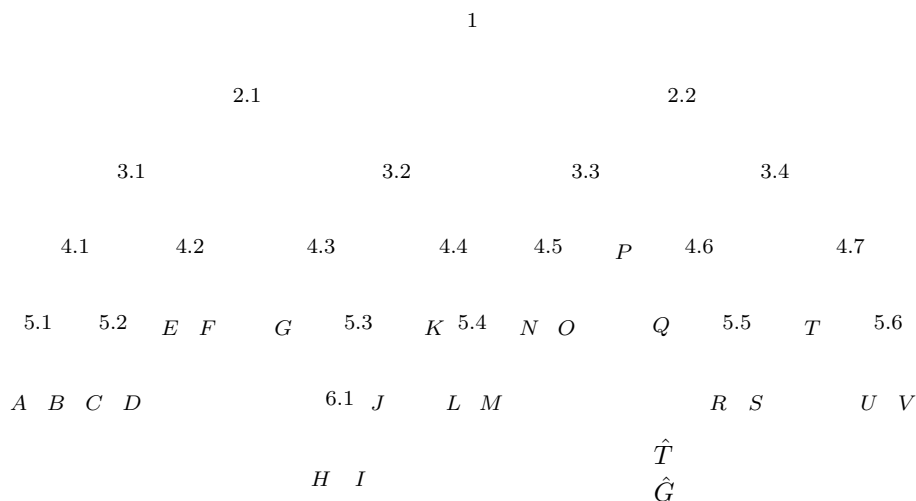


Abbildung 1.23: Baum der balancierten hierarchischen Triangulation.

Durch iterierte Anwendung von Theorem ?? auf die Teilpolygone zu beiden Seiten von  $d$  entsteht eine balancierte, hierarchische Zerlegung von  $P$



Abbildung ???. So entsteht aus dem Hierarchiebaum<sup>7</sup>  $\hat{T}$  der Triangulation  $T$  der Schichtengraph  $\hat{G}$ , mit folgenden Eigenschaften:

**Lemma 1.13** *Eigenschaften von  $\hat{G}$ :*

(i) *Der Weg zwischen zwei beliebigen Dreiecken in  $T$  ergibt sich durch Konkatenation von  $O(\log n)$  vielen Kanten im Graphen  $\hat{G}$ .*

(ii) *Diese Kanten lassen sich in Zeit  $O(\log n)$  in  $\hat{G}$  bestimmen.*

(iii) *Die Größe von  $\hat{G}$  ist  $O(n)$ .*

**Beweis.**

(i) Sei  $\pi^*$  der Weg im dualen Graphen  $T^*$ . Betrachte die von  $\pi^*$  überquerten Diagonalen der Triangulation und ihre Tiefe (Schicht) in  $\hat{T}$ :

Wir konstruieren einen Pfad  $\pi' = d_1, d_2, \dots, d_l, d_{l+1}, \dots, d_{\text{first}}, \dots, d_k$  aus  $\pi^*$ .  $\pi'$  entsteht aus  $\pi^*$  durch Streichen von Diagonalen. Es soll gelten, dass zwischen zwei Diagonalen  $d_l$  und  $d_{l+1}$  in  $\pi^*$  keine Diagonale existiert mit Tiefe kleiner als  $d_l$ . Jedoch soll  $d_{l+1}$  eine kleinere Tiefe haben als  $d_l$ . Bis  $d_{\text{first}}$  wird so die Tiefe immer geringer und ab  $d_{\text{first}}$  entsprechend wieder größer.  $\pi'$  entsteht aus einem Durchlauf von  $d_1$  nach  $d_{\text{first}}$ , bei dem die minimale Tiefe festgehalten wird und alle Diagonalen verworfen werden, die tiefer liegen als das bisherige Minimum, entsprechend für den Weg von  $d_k$  nach  $d_{\text{first}}$ .

$$\pi^* = d_1 \dots d_2 \dots d_\ell \dots d_{\ell+1} \dots d_{\text{first}} \dots d_{l+2} \dots d_k$$

$$\xrightarrow{\text{tiefer als}} \xrightarrow{\text{tiefer als}} \xrightarrow{\text{tiefer als}} \xrightarrow{\text{tiefer als}} \xrightarrow{\text{kleinste Tiefe}} \xleftarrow{\text{tiefer als}} \xleftarrow{\text{tiefer als}}$$

Im Beispiel ergibt sich also für den Weg von Dreieck  $P$  nach Dreieck  $M$  der Pfad  $\pi' = 3.3 \quad 2.2 \quad 1 \quad 3.2 \quad 4.4$ .

$\pi'$  besteht bis  $d_{\text{first}}$  aus Kanten mit strikt abnehmender Tiefe, ab  $d_{\text{first}}$  aus Kanten mit strikt zunehmender Tiefe, also können nur  $O(\log n)$  viele Kanten in  $\pi'$  liegen.

**Beh.:** Im Baum  $\hat{T}$  ist  $d_{i+1}$  Vorfahre von  $d_i$ , und im Graphen  $\hat{G}$  sind  $d_i, d_{i+1}$  mit einer Kante verbunden.

**Bew.:**

Unmittelbar nach dem Einfügen von  $d_i$  in die hierarchische Triangulation sind  $d_i$  und  $d_{i+1}$  längs  $\pi'$  direkt benachbart, denn die Diagonalen auf  $\pi'$  zwischen  $d_i$  und  $d_{i+1}$  sind noch nicht eingefügt

$$\Rightarrow d_i \subseteq P(d_{i+1})$$

$\Rightarrow$  in  $\hat{T}$  ist  $d_i$  im Teilbaum mit Wurzel  $d_{i+1}$  enthalten, d. h.  $d_{i+1}$  ist Vorgänger von  $d_i$ .

Außerdem ist  $d_{i+1} \in \partial P(d_i)$ , also sind  $d_i, d_{i+1}$  in  $\hat{G}$  verbunden. ◇

<sup>7</sup>Nicht zu verwechseln mit dem dualen Graph  $T^*$ !

(ii) Seien  $d_1$  und  $d_m$  die erste bzw. letzte Diagonale auf dem Weg vom Start zum Ziel in  $T$ . Sei  $d_{\text{first}}$  der tiefste gemeinsame Vorfahre in  $\hat{T}$  von  $d_1$  und  $d_m$ . Dieser läßt sich in Zeit  $O(\log n)$  bestimmen, da  $\hat{T}$  balanciert ist.

Laufe in  $\hat{T}$  von  $d_1$  aufwärts nach  $d_{\text{first}}$  und teste für jeden Knoten mit Diagonale  $d$  auf dem Weg in  $\hat{T}$ , ob  $d$  in der Triangulation  $T$  zwischen  $d_1$  und  $d_{\text{first}}$  liegt. Falls ja, nimm die letzte Kante zu  $d$  in  $\hat{G}$  in den Weg auf (d. h. die Kante vom aktuellen Knoten zu dem Knoten, in dem zuletzt eine Kante aufgenommen wurde). Verfahre ebenso für  $d_m$  bis  $d_{\text{first}}$  rückwärts.

Um diesen Test durchführen zu können, betrachte die Kanten des dualen Graphen  $T^*$  und zeichne einen inneren Knoten als Wurzel aus.<sup>8</sup> Offenbar liegt die Diagonale  $d$  genau dann auf dem Pfad von  $d_1$  nach  $d_{\text{first}}$ , wenn die Kante  $d$  in  $T^*$  mit Wurzel  $w$  Vorgänger von *genau einer* der beiden Kanten  $d_1$  oder  $d_{\text{first}}$  ist. Vergleiche hierzu  $d \in \{j, k, l\}$  in Abbildung ???. Die Diagonale  $j$  ist weder Vorgänger von  $d_1$  noch von  $d_{\text{first}}$ . Die Diagonale  $k$  ist Vorgänger von  $d_1$  und  $l$  ist Vorgänger von beiden Diagonalen  $d_1, d_{\text{first}}$ .

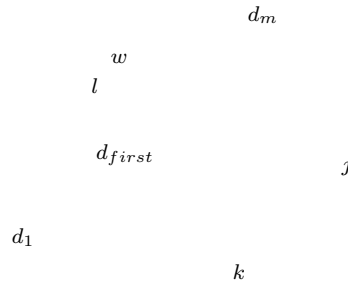


Abbildung 1.25: Diagonalen  $j, k, l, d_1, d_{\text{first}}$  im dualen Baum  $T^*$  mit Wurzel  $w$ .

Identifiziere nun jede Kante  $d$  mit ihrem “unteren”, d. h. von der Wurzel  $w$  weiter entfernten Knoten  $\delta(d)$ .

$$\begin{aligned} \delta : \quad & \text{Kanten} \xrightarrow{\text{bijektiv}} \text{Knoten} \setminus \{w\} \\ \delta : \quad & \text{Kante } d \text{ von } T^* \longmapsto \text{Knoten } \delta(d) \text{ der weiter von } w \text{ entfernt ist.} \end{aligned}$$

Es gilt:

$$\forall \text{ Kanten } d, d' : d \text{ ist Vorgänger von } d' \iff \delta(d) \text{ ist Vorgänger von } \delta(d').$$

<sup>8</sup>Zur Erinnerung: In  $T^*$  entsprechen die Kanten den Diagonalen der Triangulation; im Gegensatz zu  $\hat{T}$ , wo die Knoten den Diagonalen entsprechen!



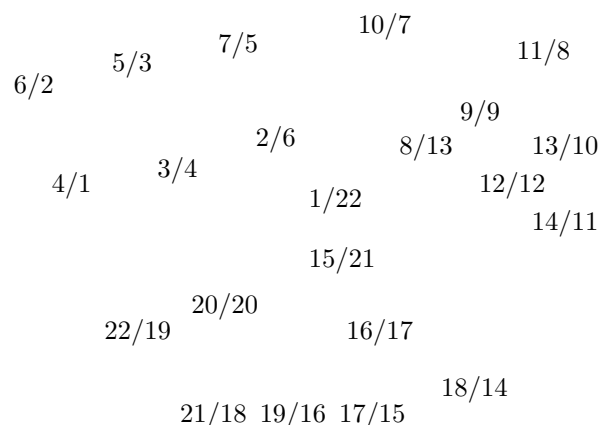


Abbildung 1.26: Prä- und Postorderdurchlauf.

---

**Algorithmus 1.5** Schneller Test auf Vorgängerrelation im Baum
 

---

**Preprocessing:**

Durchlaufe  $T^*$  einmal in Präorder und einmal in Postorder und beschrifte jeden Knoten  $a$  mit seinen Werten  $\text{pre}(a)$  und  $\text{post}(a)$ .

$O(n)$

**Nun gilt:**

$$a \text{ ist Vorgänger von } b \iff \text{pre}(a) < \text{pre}(b) \wedge \text{post}(a) > \text{post}(b)$$


---

Damit reduziert sich das Problem auf folgende Frage: Wie entscheidet man schnell für zwei Knoten  $a, b$  eines Baumes  $T^*$  mit Wurzel  $w$ , ob  $a$  Vorgänger von  $b$  ist. Dies kann wie in Algorithmus ?? beschrieben nach einem Preprocessing in Zeit  $O(n)$  für jedes Knotenpaar in Zeit  $O(1)$  entschieden werden.

(iii) Beim Einfügen berandet jede Diagonale zu jedem Zeitpunkt genau zwei Teilpolygone. Daraus folgt, daß von jedem Knoten in  $\hat{G}$  maximal zwei Kanten zu jeder tieferen Schicht führen. Weiter ist die Anzahl der Schichten unterhalb von  $d$  gleich der Höhe  $h$  von  $d$  in  $\hat{T}$ . Der Teilbaum von  $d$  in  $\hat{T}$  enthält wegen der Balancierung mindestens  $\left(\frac{3}{2}\right)^h$  Blätter (Übungsaufgabe).

Da die Teilbäume disjunkt sind, kann es höchstens  $\frac{n}{\left(\frac{3}{2}\right)^h} = \left(\frac{2}{3}\right)^h n$  solcher Knoten der Höhe  $h$  geben. Also gilt:

$$|\hat{G}| \leq 2 \sum_d \text{Höhe von } d \leq 2 \sum_{h=1}^{\log_{\frac{3}{2}} n} h \left(\frac{2}{3}\right)^h n \in O(n),$$

da die Reihe  $\sum_h hX^h$  für  $X < 1$  den Konvergenzradius 1 hat.  $\square$

Wir wissen nun, daß jeder Pfad in  $P$  zwischen zwei Punkten nur eine logarithmische Anzahl von Subpolygonen passiert. Anstatt Wege der Länge  $O(n)$  im dualen Graphen der Triangulation zu suchen, können wir also Wege der Länge  $O(\log n)$  in  $\hat{G}$  betrachten und daraus mit Hilfe vorab berechneter Informationen den kürzesten Weg ableiten.

Dazu benutzen wir “Sanduhren”, eine Verallgemeinerung der Trichter (Abbildung ??) ohne festen Startpunkt, als Struktur zur Repräsentation aller kürzesten Wege zwischen zwei Diagonalen. Mit jeder Kante in  $\hat{G}$  ist eine Sanduhr verbunden. Durch Konkatenation der Sanduhren aller Diagonalen entlang eines Weges  $\pi'$  in  $\hat{G}$  erhält man eine Repräsentation aller kürzesten Wege zwischen den gewünschten Diagonalen.

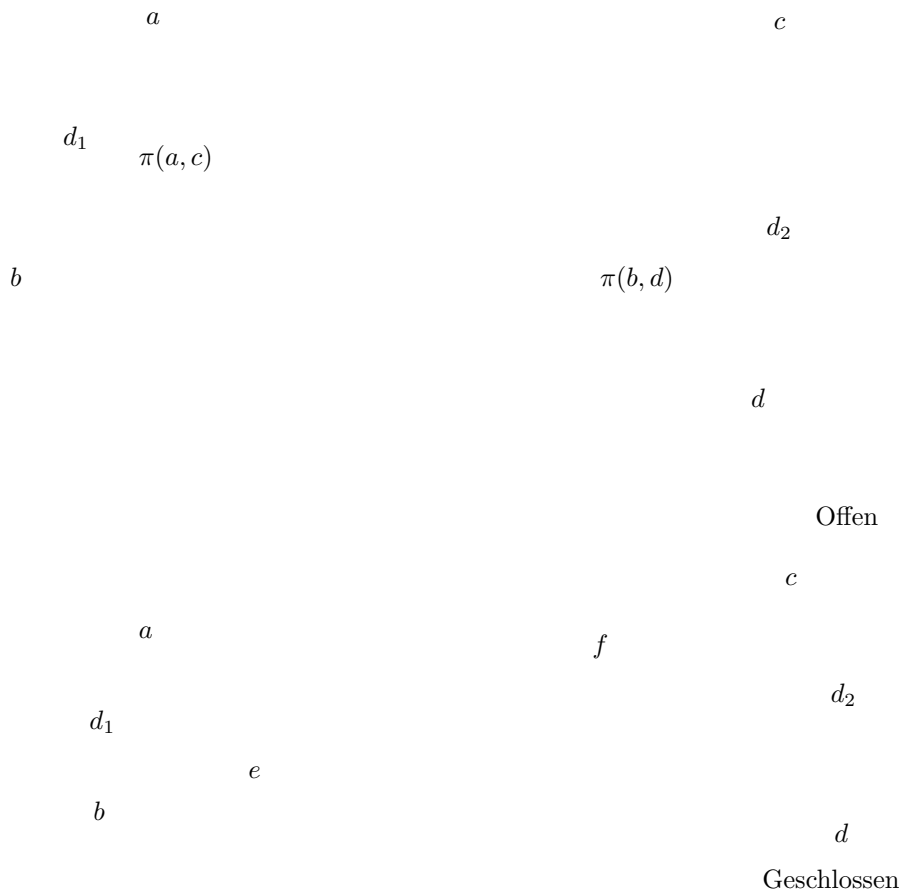


Abbildung 1.27: Offene und geschlossene Sanduhr zwischen  $d_1 = \overline{ab}$  und  $d_2 = \overline{cd}$ .

Eine Sanduhr zwischen zwei Diagonalen  $d_1 = \overline{ab}$  und  $d_2 = \overline{cd}$  —  $a, c, d, b$  seien Eckpunkte von  $P$  im Uhrzeigersinn — wird von den beiden Diagonalen und den kürzesten Pfaden  $\pi(a, c)$  und  $\pi(b, d)$  berandet, vgl. Abbildung ??.

Wir nennen eine Sanduhr *geschlossen*, falls  $\pi(a, c)$  und  $\pi(b, d)$  gemeinsame Segmente haben. In diesem Fall besteht die Sanduhr aus zwei Trichtern mit einer polygonalen Kette zwischen den Trichteröffnungen  $e$  und  $f$ . Ansonsten nennen wir die Sanduhr *offen*. Offene Sanduhren werden von zwei konvexen, polygonalen Ketten berandet, und die Diagonalen  $d_1$  und  $d_2$  sind gegenseitig sichtbar, d. h. es existiert ein Liniensegment  $\overline{gh}$  mit  $g \in d_1$  und  $h \in d_2$ .

Über die Komplexität der Sanduhren gibt folgendes Lemma Auskunft:

**Lemma 1.14** *Die Gesamtkomplexität aller Sanduhren liegt in  $O(n)$ .*

**Beweis.** Die Komplexität einer Sanduhr  $S(d_i, d_j)$  hängt von der Höhe  $h$  der in  $\hat{T}$  tieferen Diagonalen ab. Sei  $d_j$  die tiefere Diagonale, dann läßt sich die Komplexität der Sanduhr  $S(d_i, d_j)$  durch die Anzahl der Knoten unterhalb  $d_j$  abschätzen, also durch  $(\text{Höhe von } d_j)^2$ . Da wir beim Einfügen der Diagonalen  $d_j$  den Knoten  $d_j$  in  $\hat{G}$  mit den benachbarten Diagonalen verbinden, gibt es 2 Sanduhren, in denen  $d_j$  der tiefere Knoten ist. Analog zu Lemma ??(iii) gibt es  $\left(\frac{2}{3}\right)^h$   $n$  Knoten der Höhe  $h$ . Insgesamt erhalten wir:

$$\begin{aligned}
 2 \sum_{\text{Diagonale } d} (\text{Höhe von } d)^2 &\leq 2 \sum_{h=1}^{\log_{\frac{2}{3}} n} (\#\text{Knoten der Höhe } h) \cdot h^2 \\
 &\leq 2n \sum_{h=1}^{\log_{\frac{2}{3}} n} \left(\frac{2}{3}\right)^h \cdot h^2 = 2nC \in O(n)
 \end{aligned}$$

□

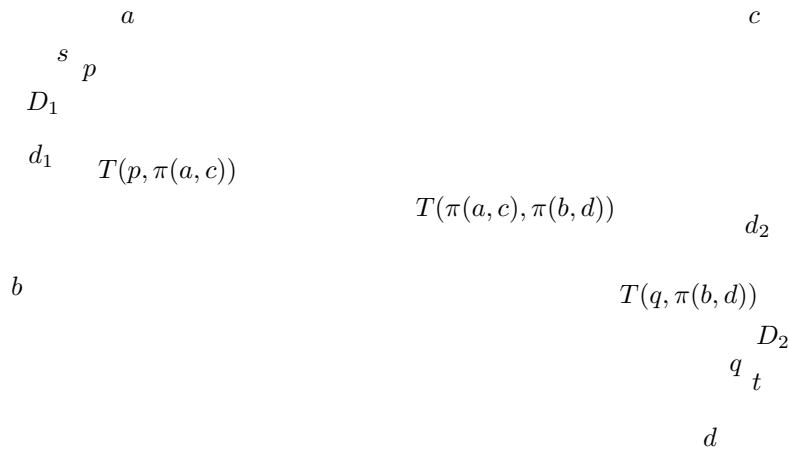


Abbildung 1.28: Kürzester Pfad in offener Sanduhr.

Für den kürzesten Pfad zwischen zwei gegebenen Anfragepunkten  $p \in d_1$  und  $q \in d_2$  sind folgende Fälle zu betrachten:

- Im geschlossenen Fall besteht  $\pi(p, q)$  aus der Konkatenation der Kürzesten  $\pi(p, e)$ , der polygonalen Kette  $\pi(e, f)$  und  $\pi(f, q)$ .
- Ist die Sanduhr offen, sind folgende Fälle zu betrachten:
  - Sind  $p$  und  $q$  gegenseitig sichtbar, so ist  $\pi(p, q) = \overline{pq}$ .
  - Andernfalls betrachte  $\mathcal{L} = \{ \overline{rs} \mid r \in d_1, s \in d_2 \}$ . Wenn ein  $\ell \in \mathcal{L}$  existiert, so daß  $p$  und  $q$  auf derselben Seite von  $\ell$  liegen (o. E. auf der Seite von  $b$  und  $d$ ), so besteht  $\pi(p, q)$  aus der Tangente von  $p$  an  $\pi(b, d)$ , der Tangente von  $q$  an  $\pi(b, d)$  und dem Stück von  $\pi(b, d)$ , das beide Tangentenendpunkte verbindet. Die Menge  $\mathcal{L}$  dient dabei nur der Anschauung, natürlich können nicht alle Liniensegmente explizit berechnet werden. Die Lage von  $p$  und  $q$  läßt sich auch durch sukzessive Tangentenbildung bestimmen.
  - Wenn jedes  $\ell \in \mathcal{L}$  zwischen  $p$  und  $q$  liegt — o. E. liege  $p$  auf der Seite von  $a$ ,  $q$  auf der Seite von  $d$  — setzt sich  $\pi(p, q)$  zusammen aus der Tangente von  $p$  an  $\pi(a, c)$ , der Tangente von  $q$  an  $\pi(b, d)$ , der gemeinsamen Tangente zwischen  $\pi(a, c)$  und  $\pi(b, d)$  und den Teilen von  $\pi(a, c)$  und  $\pi(b, d)$ , die die Tangentenendpunkte verbinden, siehe Abbildung ??.

Analog lassen sich kürzeste Wege zwischen zwei Punkten  $s \in D_1$  und  $t \in D_2$  finden, die in den  $d_1$  und  $d_2$  aufgesetzten Dreiecken  $D_1$  und  $D_2$  liegen.

Damit reduziert sich die Bestimmung kürzester Wege zwischen Punkten in  $d_1$  und  $d_2$  oder in den angrenzenden Dreiecken auf die Berechnung eines Tangentenpunktes auf konvexen Ketten.

Dies läßt sich mit binärer Suche in Zeit  $O(\log n)$  lösen; ein Verfahren, das auch bei der Konstruktion der konvexen Hülle benutzt wird, siehe Abbildung ??.

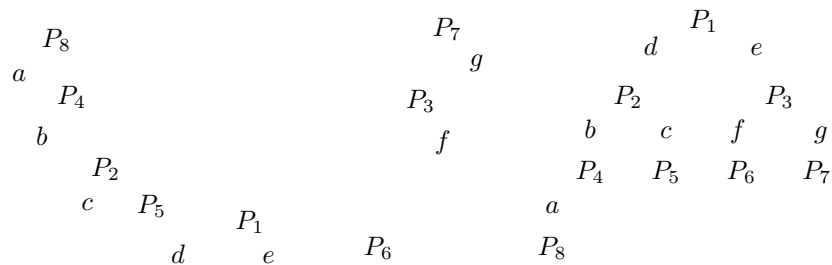


Abbildung 1.29: Speicherung einer Sanduhr in einem binären Baum.

Eine entscheidende Eigenschaft der Sanduhren ist, daß sie sich in logarithmischer Zeit konkatenieren lassen. Dabei sind einige Fallunterscheidungen zu treffen und Details zu beachten, z. B. kann die Konkatenation zweier offener Sanduhren geschlossen sein. Für Details sei hier auf die Arbeit von Guibas und Hershberger [?] verwiesen. Es gilt:

**Lemma 1.15** *Eine Sanduhr zwischen zwei Diagonalen  $d_i$  und  $d_j$  werde von  $m_{ij}$  Dreiecken der Triangulation überdeckt. Es existiert eine Datenstruktur zur Speicherung der Sanduhren  $S(d_i, d_j)$ , die folgende Operationen unterstützt:*

- (i) *Entfernung zwischen zwei Punkten in  $D_i$  und  $D_j$  in Zeit  $O(\log m_{ij})$ .*
- (ii) *Kürzeste zwischen zwei Punkten in  $D_i$  und  $D_j$  in Zeit  $O(\log m_{ij} + k)$ .*
- (iii) *Zusammensetzen zweier Sanduhren  $S(d_i, d_j)$  und  $S(d_j, d_\ell)$  zu einer Sanduhr  $S(d_i, d_\ell)$  in Zeit und zusätzlichem Platz  $O(\log m_{ij} + \log m_{j\ell})$ .*

Zusammenfassend erhalten wir Algorithmus ?? und folgendes Theorem:

**Theorem 1.16** *(Guibas, Hershberger, 1987)*

*Zu einem Polygon  $P$  mit  $n$  Ecken kann nach Triangulation und Vorbereitungszeit  $O(n)$  zwischen zwei beliebigen Punkten  $a$  und  $b$  im Inneren von  $P$  die Entfernung in Zeit  $O(\log n)$  und ein kürzester Pfad in Zeit  $O(\log n + k)$  bestimmt werden, wobei  $k$  die Anzahl der Segmente auf dem kürzesten Pfad ist. Der Speicherbedarf liegt in  $O(n)$ .* [?, ?]

Eine einfachere Datenstruktur für Sanduhren findet sich in [?]. Offen ist die Frage, ob sich kürzeste Wege im Polygon in Zeit  $O(n)$  ohne Triangulation berechnen lassen.

---

**Algorithmus 1.6** Kürzeste Pfade in einem Polygon (Guibas, Hershberger)
 

---

**Vorbereitungsphase:**

- Berechne Triangulation  $T$ , dualen Graphen  $T^*$ , hierarchische Triangulation  $\hat{T}$  und Schichtengraph  $\hat{G}$ .
- Bestimme Prä- und Postorderbeschriftungen der Knoten in  $T^*$ .
- Berechne für jede Kante  $(d_i, d_j)$  in  $\hat{G}$  die Sanduhr  $S(d_i, d_j)$ .
- Baue eine Suchstruktur auf, um Punkte schnell in den Dreiecken lokalisieren zu können (Point-Location).
- Bereite den Graphen  $\hat{G}$  für Pfadsuche vor.

Zeit und Platz  $O(n)$ **Anfrage für Punkte  $p, q$ :**

- Bestimme die Dreiecke  $D_p$  und  $D_q$ .  $O(\log n)$
  - Bestimme in  $\hat{G}$  den Pfad zwischen den Diagonalen  $d_p$  und  $d_q$ :
    - Finde tiefsten gemeinsamen Vorfahren  $d_{\text{first}}$
    - Betrachte Pfad  $\pi_{\hat{T}}$  von  $D_p$  nach  $D_q$  in  $\hat{T}$ :  
 Teste für jede Diagonale  $d$  in  $\pi_{\hat{T}}$ , ob  $d$  zwischen  $D_p$  und  $D_q$  in  $T$  liegt. Dies ist dann der Fall, wenn  $d$  Vorgänger von *entweder*  $D_p$  *oder*  $D_q$  in  $T^*$  ist.
  - Liefert Pfad  $\pi_{\hat{G}}$  mit Länge  $O(\log n)$  in Zeit  $O(\log n)$
  - Konstruiere  $S(d_p, d_q)$  aus den Sanduhren längs des Pfades  $\pi_{\hat{G}}$  von  $D_p$  nach  $D_q$  im Graphen  $\hat{G}$ .  $O(\log^2 n)$   
 Durch Berechnung zusätzlicher Sanduhren im voraus in Zeit und Platz  $O(n)$  auch dies auch möglich in Zeit  $O(\log n)$
  - Bestimme die Kürzeste von  $p$  nach  $q$  mit Hilfe von  $S(d_p, d_q)$ .  $O(\log n + k)$
-

### 1.2.3 Anwendung: Berechnung des geodätischen Durchmessers



Abbildung 1.30: Ein Polygon und sein geodätischer Durchmesser.

Eine Anwendung von kürzesten Pfaden im Inneren eines Polygons ist die Bestimmung des geodätischen Durchmessers, der im Bereich der Mustererkennung eine Rolle spielt; Abbildung ?? zeigt ein Polygon mit seinem geodätischen Durchmesser.

**Definition 1.17** Der geodätische Durchmesser  $\pi_{\max}$  eines Polygons  $P$  ist der längste kürzeste Pfad zwischen zwei Ecken im Inneren von  $P$ :

$$\pi_{\max} := \max_{p_i, p_j \text{ Ecken von } P} d(p_i, p_j).$$

Dabei bezeichnet  $d(p_i, p_j)$  den geodätischen Abstand von  $p_i$  und  $p_j$ : die Länge des kürzesten Pfades von  $p_i$  nach  $p_j$ .

Ein naiver Ansatz wäre, für jedes Paar  $(p_i, p_j)$  den geodätischen Abstand  $d(p_i, p_j)$  nach Theorem ?? in Zeit  $O(\log n)$  und Vorbereitungszeit  $O(n)$  zu berechnen und dann das Maximum dieser Abstände zu bestimmen; damit könnten wir den geodätischen Durchmesser in Zeit  $O(n^2 \log n)$  berechnen.

Auch hier läßt sich mehr erreichen, die folgenden beiden Ideen gehen auf Aggarwal, Klawe, Moran, Shor und Wilber (1986) zurück. [?]

#### 1. Idee:

Nutze aus, daß die paarweisen Abstände  $d(p_i, p_j)$  nicht voneinander unabhängig sind. Konkret gilt für alle Ecken  $p_i, p_j, p_\ell, p_m$ , die in dieser Reihenfolge auf dem Rand des Polygons liegen:

$$d(p_i, p_m) + d(p_j, p_\ell) < d(p_i, p_\ell) + d(p_j, p_m).$$

$p_\ell$

Dies folgt aus den Dreiecksungleichungen

$d(p_i, p_m) \leq d(p_i, q) + d(q, p_m)$  und

$d(p_j, p_\ell) \leq d(p_j, q) + d(q, p_\ell)$  mit

$d(p_j, p_m) = d(p_j, q) + d(q, p_m)$  und

$d(p_i, p_\ell) = d(p_i, q) + d(q, p_\ell)$ .

$p_m$

$q$

$p_j$

$p_i$

Seien  $p_1, p_2, \dots, p_n$  die Ecken des Polygons gegen den Uhrzeigersinn. Definiere eine  $n \times n$  Matrix  $A = (a_{i\ell})$  durch:

$$a_{i\ell} = \begin{cases} \ell - n & \text{für } 1 \leq \ell \leq i \\ d(i, \ell) & \text{für } i + 1 \leq \ell \leq n \end{cases}$$

Dabei sei  $d(i, j) := d(p_i, p_j)$ . Die Matrix  $A$  besteht im oberen Dreieck aus Pfadlängen, im unteren Dreieck aus negativem Füllwerk.

$$\begin{matrix} & 1 & 2 & 3 & & n-1 & n \\ \begin{matrix} 1 \\ 2 \\ 3 \\ \vdots \\ n-1 \\ n \end{matrix} & \left( \begin{array}{cccccc} 1-n & d(1,2) & d(1,3) & \dots & d(1,n-1) & d(1,n) \\ 1-n & 2-n & d(2,3) & \dots & d(2,n-1) & d(2,n) \\ 1-n & 2-n & 3-n & \dots & d(3,n-1) & d(3,n) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1-n & 2-n & 3-n & \dots & \dots & d(n-1,n) \\ 1-n & 2-n & 3-n & \dots & \dots & 0 \end{array} \right) \end{matrix}$$

**Definition 1.18** Eine  $n \times m$ -Matrix  $A = (a_{i\ell})$  heißt **monoton**, falls

$$\forall 1 \leq i < j \leq n, 1 \leq k < \ell \leq m : (a_{ik} < a_{i\ell} \Rightarrow a_{jk} < a_{j\ell}).$$

Anschaulich:

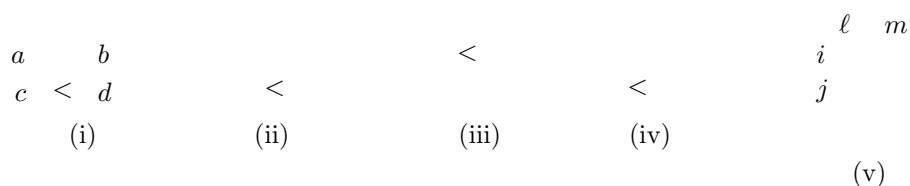
$$\begin{matrix} & k & \ell \\ i & \left( \begin{array}{c} a_{ik} < a_{i\ell} \\ \Downarrow \\ a_{jk} < a_{j\ell} \end{array} \right) \\ j & \end{matrix}$$

**Bemerkung.** Sei  $A$  eine monotone Matrix. Dann gilt: je tiefer die Zeile, desto weiter rechts liegt das linkeste Zeilenmaximum.

So kann in folgender monotonen Matrix das Maximum der dritten Zeile nicht links von der vierten Spalte stehen — dem Maximum der zweiten Zeile.

$$\begin{pmatrix} 5 & 4 & 3 & 2 & 1 & 0 \\ 1 & 7 & 3 & \boxed{9} & 4 & 8 \\ 4 & 6 & 5 & \boxed{8} & 8 & 9 \end{pmatrix}$$



Abbildung 1.31: Fallunterscheidung nach Lage von  $A$  in  $T$ .

**Lemma 1.19** Die oben definierte Matrix  $A$  ist monoton.

**Beweis.**

Sei  $T = \begin{pmatrix} a_{i\ell} & a_{im} \\ a_{j\ell} & a_{jm} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  mit  $i < j, \ell < m$  eine Untermatrix von  $A$ . Je nach Lage von  $T$  in  $A$  unterscheiden wir folgende Fälle, siehe Abbildung ?? : In Fall (i) liegt  $T$  unterhalb der Treppe, in diesem Fall—und ebenso in den Fällen (ii) bis (iv) gilt  $c < d$ .

Als einzig interessanter Fall bleibt (v). In diesem Fall folgt aus der Definition der Matrix  $A$ :  $i < j < \ell < m$ . Die Ungleichung  $j < l$  gilt, da sich  $a_{i\ell}$  oberhalb der Diagonalen der Matrix befindet. Insgesamt folgt aus den Annahmen  $i < j, \ell < m$ , dass  $1 \leq i < j < l < m \leq n$  gilt und sich somit die Punkte  $i, j, l, m$  in dieser Reihenfolge auf dem Rand des Polygons befinden. Für die Knoten auf dem Rand des Polygons gilt deshalb  $d(i, m) + d(j, \ell) \leq d(i, \ell) + d(j, m)$ , aus  $d(i, \ell) < d(i, m)$  folgt also  $d(j, \ell) < d(j, m)$ .

□

## 2. Idee:

**Theorem 1.20** Sei  $A$  eine monotone  $n \times m$ -Matrix,  $n \leq m$ . Dann läßt sich das maximale Element in Zeit  $O(m)$ .

**Beweis.** Berechne mit Algorithmus ?? für jede Zeile  $i$  das am weitesten links stehende Maximum  $\max(i)$ . Die Anzahl der Vergleiche in Algorithmus ?? ist dabei beschränkt durch das Maximum der Anzahl Zeilen (Anzahl der Elemente auf der Diagonalen) und dem der Anzahl der Spalten (maximale Anzahl der Spaltenstreichungen), also  $m$ .

Damit ist nur noch die Laufzeit von Algorithmus ?? zu bestimmen: Sei  $T(n)$  die Anzahl der Zugriffe auf Matrixelemente der Prozedur Zeilenmaxima. Sei  $n = 2^k$ :

$$\begin{aligned} T(n) &\leq T\left(\frac{n}{2}\right) + C \cdot n + O(m) \\ &\leq T\left(\frac{n}{4}\right) + C \cdot \frac{n}{2} + C \cdot n + O(m) \\ &\leq \dots \end{aligned}$$

$$\begin{aligned}
&\leq T(1) + C \left( \frac{1}{2^{k-1}} + \frac{1}{2^{k-2}} + \dots + \frac{1}{2} + 1 \right) \cdot n + O(m) \\
&\in O(n + m) \\
&\in O(m)
\end{aligned}$$

□

Bemerkung: Im ersten Rekursionsschritt entstehen  $O(m)$  Kosten für die Spaltenreduktion zu einer quadratischen Matrix. In den weiteren Rekursionsschritten sind die Kosten der Spaltenreduktion in der Konstante  $C$  enthalten. Analog entstehen am Ende einmalig zusätzlich  $O(m)$  Kosten für die Berechnung der Maxima der ungeraden Zeilen.

Da wir nach Theorem ?? die Zeilenmaxima in Zeit  $O(m)$  bestimmen können, folgt direkt, dass wir nur  $O(m)$  Matrixelemente miteinander vergleichen müssen.

Mit Theorem ?? können wir zu einer *gegebenen* Matrix schnell das maximale Element ausrechnen. Allerdings würde es uns schon Zeit  $O(n^2)$  kosten, die Matrix  $A$  aufzustellen! Hier hilft uns die Tatsache, daß wir gar nicht alle  $O(n^2)$  Einträge der Matrix brauchen, sondern mit  $O(n)$  Vergleichen auskommen. Wir berechnen die Matrix  $A$  also nicht explizit, sondern die benötigten Matrixelemente erst dann, wenn sie gebraucht werden. Damit gilt:

**Theorem 1.21** *Der geodätische Durchmesser eines einfachen Polygons mit  $n$  Ecken läßt sich in Zeit  $O(n \log n)$  bestimmen. Es läßt sich sogar für jede Ecke der am weitesten entfernte Partner in dieser Zeit feststellen.* [?]

Durch bessere Kombination von Matrixsuche und Kürzestenbestimmung ist dies sogar in Zeit  $O(n)$  möglich (Hershberger, Suri, 1993, [?, ?]).

---

**Algorithmus 1.7** Spaltenreduktion

---

**Input:** monotone  $n \times m$ -Matrix  $A$ ,  $n \leq m$ .**Output:** monotone  $n \times n$ -Matrix  $A'$  mit denselben Zeilenmaxima wie  $A$ ;  
aus der Position von  $\max(i)$  in  $A'$  ist  $\max(i)$  in  $A$  rekonstruierbar.

Beginnend mit  $a_{1,1}$  teste, ob  $a_{i,i} < a_{i,i+1}$  ist. Ist dies der Fall, streiche die  $i$ -te Spalte, da die Elemente der  $i$ -ten Spalte wegen der Monotonieeigenschaft keine Zeilenmaxima mehr sein können. Vergleiche dann  $a_{i-1,i-1}$  mit  $a'_{i-1,i} = a_{i-1,i+1}$  usw. Ist  $a_{i,i} \geq a_{i,i+1}$ , inkrementiere  $i$ ; vergleiche also als nächstes  $a_{i+1,i+1}$  mit  $a_{i+1,i+2}$ .

Lautet die Antwort für alle Elemente auf der Diagonalen "nein", so sind alle Elemente der  $n$ -ten Spalte größer gleich den Elementen der Spalte  $(n+1)$ , also können wir die  $(n+1)$ -ten Spalte streichen und mit dem Test  $a_{n,n} < a_{n,n+2}$  fortfahren.

Als Output erhalten wir eine quadratische Matrix mit der Eigenschaft, dass die Elemente rechts von der Diagonalen in jeder Zeile eine monoton fallende Folge bilden, d.h.  $a_{i,j} \geq a_{i,j+1} \forall j \geq i$ .

---



---

**Algorithmus 1.8** Zeilenmaxima

---

**Input:** monotone  $n \times m$ -Matrix  $B$ ,  $n \leq m$ **Output:** alle linkesten Zeilenmaxima  $\max(i), 1 \leq i \leq n$ 

- $C :=$  alle Zeilen von  $B$  mit geradem Zeilenindex;  $O(n)$
  - $C' :=$  Spaltenreduktion( $C$ );  $O(m)$
  - Zeilenmaxima( $C'$ );  $T(\frac{n}{2})$   
/\*  $C'$  ist  $\frac{n}{2} \times \frac{n}{2}$ -Matrix \*/
  - Rekonstruiere Zeilenmaxima von  $C$ ;  $O(n)$   
/\* = gerade Zeilen von  $B$  \*/
  - Berechne die Maxima der ungeraden Zeilen von  $B$ :  
Wegen Bemerkung nach Definition ?? liegt  $\max(2i+1)$  zwischen den Spalten, in denen  $\max(2i)$  und  $\max(2i+2)$  liegen. Da sich die Intervalle, in denen nach den Maxima der ungeraden Zeilen gesucht werden muß, nicht überlappen, ist dies möglich in Zeit  $O(m)$
-



### 1.2.4 Shortest Watchman Routes

Eine andere Anwendung von kürzesten Pfaden innerhalb eines Polygons ist die Berechnung der kürzesten Wächterroute oder Shortest Watchman Route (SWR). Gesucht ist der kürzeste Weg in einem Polygon, den ein Wachmann oder Roboter mit unbeschränkter Sicht gehen kann und auf dem Weg jeden Punkt in dem Polygon mindestens einmal sieht.<sup>9</sup> Eine Route ist stets ein geschlossener Weg, d. h. der Roboter kehrt zum Startpunkt zurück. Bei der einfacheren Variante des Shortest Watchman Problems ist der Startpunkt des Roboters und damit der Route gegeben, bei der schwierigeren Variante sucht man die kürzeste Route, die in einem beliebigen Punkt des Polygons startet. Ebenfalls schwieriger zu berechnen als eine SWR ist eine Shortest Watchman Tour, also ein Pfad, der jeden Punkt im Polygon einmal sieht, jedoch nicht zum Startpunkt zurückkehrt. Das Shortest Watchman Problem wurde 1986 von Chin und Ntafos vorgestellt [?]. Seitdem wurden zahlreiche Arbeiten zur Berechnung der SWR veröffentlicht, von denen sich jedoch viele als fehlerhaft herausgestellt haben oder bzgl. der Laufzeit verbessert wurden, siehe z. B. [?, ?, ?, ?, ?, ?, ?].

s

Abbildung 1.33: Die Berechnung einer SWR liefert eine Besuchsreihenfolge der Punkte.

Im allgemeinen ist die Berechnung einer SWR schwierig:<sup>10</sup>

**Theorem 1.22** (*Chin, Ntafos, 1986*)

*Die Berechnung einer Shortest Watchman Route in einer Szene mit polygonalen Hindernissen (Polygon mit Löchern) ist NP-hard.* [?, ?]

**Beweis.** In Polygonen mit Löchern läßt sich das Traveling Salesman Problem auf das Shortest Watchman Problem reduzieren: die Bestimmung der kürzesten Besuchsreihenfolge der polygonalen Löcher entspricht dem TSP. Wir umgeben alle Punkte aus der Eingabe für das TSP mit einem Hindernis, das die Sicht auf alle anderen Punkte abschneidet, siehe Abbildung ???. Jedes dieser

<sup>9</sup>Ein Weg  $\pi$  sieht das gesamte Polygon  $P$   $:\Leftrightarrow$  zu jedem Punkt  $p \in P$  existiert ein Punkt  $p' \in \pi$ , so daß das Liniensegment  $\overline{pp'}$  vollständig in  $P$  liegt.

<sup>10</sup>Zu den Begriffen NP-hard und NP-vollständig siehe Anhang.

Hindernisse hat  $O(1)$  Kanten und die Konstruktion der gesamten Szene ist in Zeit  $O(n)$  möglich. Eine Lösung für das Shortest Watchman Problem in dieser Szene liefert eine kürzeste Besuchsreihenfolge der Punkte und damit eine Lösung für das TSP.  $\square$

Wir beschränken uns daher auf einfache Polygone, d. h. Polygone ohne Selbstschnitte und Löcher, und auch in dieser Klasse von Polygonen zunächst nur auf zwei Spezialfälle, siehe Abbildung ??.

(i) (ii) (iii)

Abbildung 1.34: (i) X-monotones Polygon, (ii) nicht-monotones Polygon, (iii) rechtwinkliges Polygon.

**Definition 1.23** Ein Polygon  $P$  heißt **monoton**, wenn eine Gerade  $\ell$  existiert, so daß für jede zu  $\ell$  orthogonale Gerade  $\ell'$  der Schnitt von  $P$  mit  $\ell'$  zusammenhängend ist, d. h. der Schnitt ist *ein* Liniensegment, *ein* Punkt oder leer. Ist die Gerade  $\ell$  eine Parallele zur Y-Achse, so heißt  $P$  **y-monoton**.

$P$  heißt **rechtwinklig**, wenn jeder innere Winkel  $90^\circ$  oder  $270^\circ$  beträgt.

Den einfachsten Fall stellen rechtwinklige, monotone Polygone dar:

**Theorem 1.24** (*Chin, Ntafos, 1986*)

*In einem rechtwinkligen, monotonen Polygon kann die Shortest Watchman Route in Zeit  $O(n)$  berechnet werden.* [?, ?]

**Beweis.** Übungsaufgabe.  $\square$

Auch in rechtwinkligen, nicht-monotonen Polygonen kann die SWR recht einfach berechnet werden. Wir brauchen dazu einige Begriffe, die wir gleich für allgemeine — d. h. nicht-rechtwinklige — Polygone einführen.

**Definition 1.25** Die Verlängerungen der Kanten von reflexen Ecken<sup>11</sup> in das Polygon hinein sind die **Cuts** des Polygons. Von den beiden Cuts einer reflexen Ecke ist nur einer interessant, nämlich derjenige, der bzgl. des Startpunktes  $s$  die Sicht blockiert.<sup>12</sup> Wir nennen diese Cuts **notwendige**

<sup>11</sup>Polygonecke, deren Innenwinkel  $> \pi$  ist.

<sup>12</sup> Anschaulich der Cut, den wir überschreiten müssen, um hinter die Ecke schauen zu können.

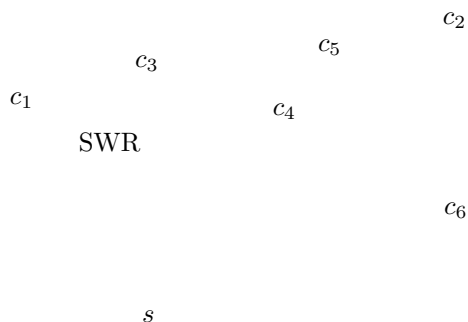


Abbildung 1.35: Polygon mit notwendigen Cuts (gepunktet), wesentlichen Cuts (gestrichelt) und Shortest Watchman Route.

**Cuts.** Ein Cut  $c_1$  **dominiert** einen Cut  $c_2$ , wenn jeder Weg von  $s$  zu  $c_1$  den Cut  $c_2$  schneidet. Notwendige Cuts, die von keinem anderen notwendigen Cut dominiert werden, heißen **wesentliche Cuts**.

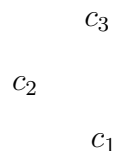
Ein Cut  $c_i$  teilt das Polygon in zwei Teile. Die zu  $c_i$  gehörende **Tasche**  $P_{c_i}$  ist der Teil, der den Startpunkt  $s$  nicht enthält. Damit können wir definieren — und man kann leicht sehen, daß beide Definitionen äquivalent sind —, daß ein Cut  $c_1$  einen anderen Cut  $c_2$  dominiert, wenn  $P_{c_1} \subset P_{c_2}$  gilt.

Cuts, die von anderen dominiert werden, werden sozusagen nebenbei exploriert und müssen nicht weiter berücksichtigt werden. Abbildung ?? zeigt ein Beispiel für ein Polygon mit seinen notwendigen und wesentlichen Cuts.  $c_3$  und  $c_4$  sind nicht wesentlich, da jeder Weg zu  $c_5$  über diese Cuts führt.  $c_5$  dominiert  $c_3$  und  $c_4$ . Die SWR (und jede andere Explorationstour) berührt alle wesentlichen Cuts und die Menge der wesentlichen Cuts ist die kleinste Menge von Cuts, die besucht werden müssen, um das ganze Polygon zu sehen.

In einem Polygon mit gegebenem Startpunkt  $s$  können wir die Cuts entsprechend dem im Uhrzeigersinn ersten Schnittpunkt zwischen Cut und Polygonrand ordnen, siehe Abbildung ?? . Es gilt:

**Lemma 1.26** *In einem rechtwinkligen Polygon existiert stets eine SWR, die die wesentlichen Cuts gemäß der Ordnung auf dem Rand besucht.*

**Beweis.** Wir zeigen zunächst, dass eine SWR nie in eine Tasche hineinläuft bzw. es wird kein wesentlicher Cut überschritten. Kritisch sind nur Situationen, wo sich wesentliche Cuts schneiden. Dies können



bei rechtwinkligen Polygonen maximal vier in Folge sein. Wir betrachten die Situation mit drei Cuts in Folge. Für jeden Cut  $c_i$ ,  $i = 1, 2, 3$  gilt: Wenn man den Weg in der Tasche von  $c_i$  durch die orthogonale Projektion auf  $c_i$  ersetzt, werden die gleichen Cuts besucht wie vorher. Der Weg wird dabei nur kürzer. Daher reicht es, die Cuts  $c'_1 = ab$ ,  $c'_2 = bc$ ,  $c'_3 = cd$  zu besuchen, ohne in die Taschen zu laufen.

Eine kürzeste Tour darf keine Selbstschnitte haben, sonst könnte man die Tour lokal abkürzen. Es folgt somit, dass nur durch Besuchen der Cuts entlang der Ordnung auf dem Rand eine schnittfreie und somit kürzeste Tour entsteht.  $\square$

---

**Algorithmus 1.9** Shortest Watchman Route in einfachen Polygonen
 

---

- Bestimme die wesentlichen Cuts  $c_1, \dots, c_k$ .  $O(n)$
  - Schneide die Teile des Polygons ab, die bzgl. des Startpunktes hinter wesentlichen Cuts liegen.  $O(n)$
  - Trianguliere das resultierende Polygon  $P'$ .  $O(n)$
  - Konstruiere  $P''$  durch “Roll-Out” von  $P'$ :  $O(n)$ 
    - $P^{(1)}$  seien die relevanten Dreiecke von  $P'$  auf dem Weg im dualen Graphen  $T^*$  von  $s$  nach  $c_1$ .
    - Für jeden Cut  $c_i$ ,  $i = 2, \dots, k$ : Erweitere  $P^{(i-1)}$  zu  $P^{(i)}$  durch alle relevanten Dreiecke entlang des Randes von  $P'$  auf dem Weg von  $c_{i-1}$  zu  $c_i$ , gespiegelt an den Cuts  $c_1, \dots, c_{i-1}$  und angesetzt an  $c_{i-1}$ .
    - Erweitere  $P^{(k)}$  zu  $P''$  wie oben durch alle relevanten Dreiecke auf dem Weg von  $c_k$  nach  $s$ . Dabei entsteht eine gespiegelte Kopie  $s'$  des Zielpunktes.
  - Berechne in  $P''$  den kürzesten Weg  $\pi$  von  $s$  nach  $s'$ .  $O(n)$
  - Die SWR entsteht durch “zurückfalten” der Liniensegmente in  $\pi$ .
- 

Lemma ?? liefert uns den Algorithmus für die Berechnung einer SWR:

**Theorem 1.27** (Chin, Ntafos, 1986)

In einem rechtwinkligen Polygon kann die Shortest Watchman Route in Zeit  $O(n)$  berechnet werden.  $[?, ?]$

**Beweis.** Algorithmus ?? berechnet die SWR in einem rechtwinkligen Polygon. Abbildung ?? zeigt ein Beispiel.

Die wesentlichen Cuts lassen sich in Zeit  $O(n)$  berechnen (ohne Beweis). Damit bleibt zu zeigen, daß  $P''$  nicht mehr als  $O(n)$  Kanten enthält. Dazu



betrachten wir den zur Triangulation dualen Graphen  $T^*$ . Dem sukzessiven Besuch der Cuts  $c_1, \dots, c_n$  entspricht ein Depth-First Durchlauf von  $T^*$ . Jede Kante von  $T^*$  wird dabei zweimal durchlaufen. Damit wird auch jedes Dreieck der Triangulation maximal zweimal in  $P''$  übernommen und die Anzahl der Kanten von  $P''$  liegt in  $O(n)$ .  $\square$

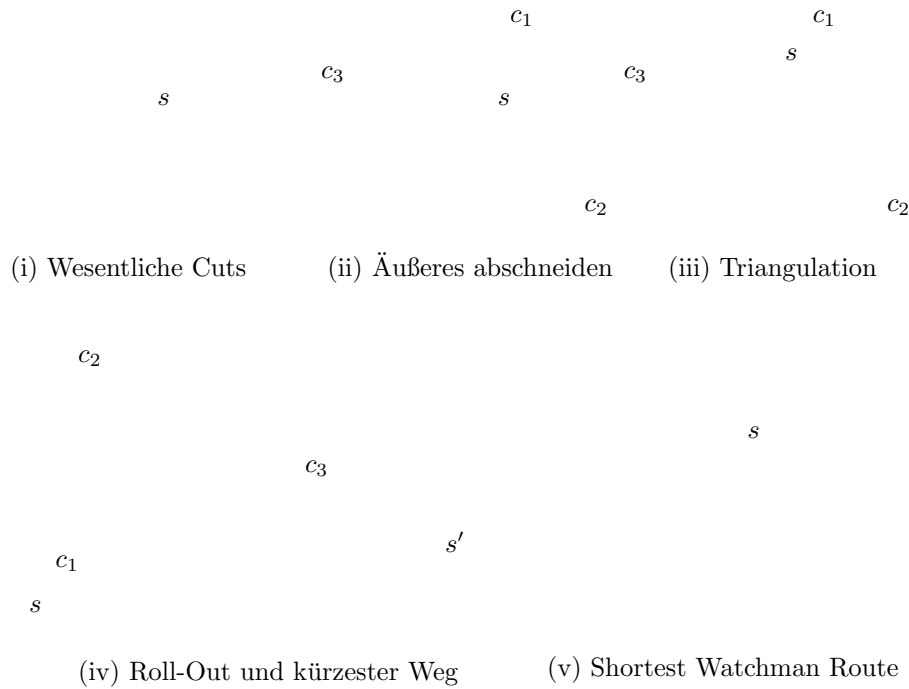


Abbildung 1.36: Berechnung der SWR in einem rechtwinkligen Polygon.

Algorithmus ?? wäre mit leichten Modifikationen bei gleicher Laufzeit in Polygonen anwendbar, in denen sich je maximal zwei wesentliche Cuts gegenseitig schneiden, da auch Polygone dieser Art Lemma ?? erfüllen. In allgemeinen Polygonen können sich allerdings mehr als zwei Cuts gegenseitig schneiden. Wir nennen solche Mengen von Cuts eine “Corner”. In solchen Situationen ist nicht von vornherein klar, in welcher Reihenfolge die Cuts besucht werden.

**Definition 1.28** Eine “Corner” ist eine geordnete Menge von Cuts

$c_i, c_{i+1}, \dots, c_j$ , für die gilt

- (i)  $c_k$  schneidet  $c_{k+1}$  für  $k \in \{i, \dots, j - 1\}$ ,
- (ii)  $c_i$  schneidet  $c_{i-1}$  nicht und
- (iii)  $c_j$  schneidet  $c_{j+1}$  nicht.

Betrachte Abbildung ??: die Cuts  $c_1, c_2, c_3$  und  $c_4$  bilden eine Corner und die Reihenfolge der Cuts wird von der SWR nicht eingehalten:  $c_3$  wird

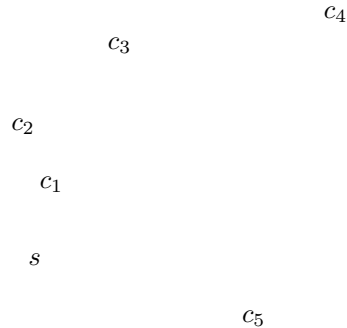


Abbildung 1.37: Eine "Corner": die Cuts  $c_1, c_2, c_3$  und  $c_4$  schneiden sich gegenseitig.

überschritten, bevor  $c_2$  erreicht wird. Die Besuchsreihenfolge der Corners ist jedoch fest:

**Lemma 1.29** *Die Shortest Watchman Route besucht die Corners eines Polygons gemäß der Ordnung entlang des Randes.*

Die von der SWR berührten Abschnitte der Cuts zwischen zwei Schnittpunkten werden aktive Segmente genannt. Sind die aktiven Segmente gegeben, kann man die SWR wie oben berechnen, das große Problem ist jedoch, die Menge der aktiven Segmente zu berechnen. Erste Ansätze waren, mit einer Menge von aktiven Segmenten eine Watchman Route  $R'$  zu berechnen und diese durch lokale Anpassungen, d. h. durch Austausch der aktiven Segmente, zu verkürzen. Einen anderen Ansatz verfolgen Dror et. al. die die folgende Verallgemeinerung des Shortest Watchman Problems und ähnlicher Probleme lösen.

**Definition 1.30**

- (i) Beim einfachen **Touring Polygon Problem** (TPP) ist eine Sequenz von einfachen, konvexen, disjunkten Polygonen  $P_1, P_2, \dots, P_k$  mit insgesamt  $n$  Kanten, ein Startpunkt  $s$  und ein Zielpunkt  $t$  gegeben. Gesucht ist der kürzeste Pfad  $\pi$ , der in  $s$  startet, in  $t$  endet und alle Polygone  $P_i$  in der gegebenen Reihenfolge besucht.
- (ii) Beim allgemeinen TPP verlangen wir zusätzlich, daß die Pfade zwischen zwei Polygonen  $P_i$  und  $P_{i+1}$  ( $i = 0, \dots, k; P_0 := s; P_{k+1} := t$ ) innerhalb des **Zaunes**  $F_i$  verlaufen. Ein Zaun  $F_i$  ist gegeben durch ein einfaches Polygon, das  $P_i$  und  $P_{i+1}$  enthält. Die Polygone dürfen sich überlappen und nur die **Fassade** eines Polygons  $P_i$  muß konvex sein. Die Fassade eines Polygons  $P_i$  ist der Teil des Randes von  $P_i$ , der nicht auf dem Rand des Zaunes  $F_{i-1}$  liegt:  $\text{Fassade}(P_i) := \partial P_i \setminus \partial F_{i-1}$ .

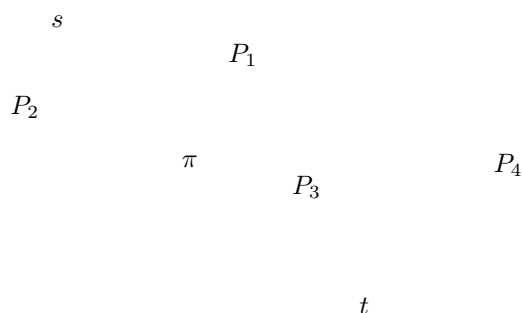


Abbildung 1.38: Ein Beispiel für das einfache Touring Polygon Problem.

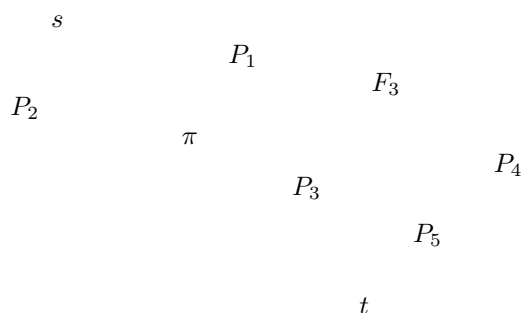


Abbildung 1.39: Ein Beispiel für das allgemeine Touring Polygon Problem.

Beim TPP wird nicht ausgeschlossen, daß ein Polygon  $P_i$  vor einem Polygon  $P_j$  mit  $j < i$  besucht wird, allerdings muß das Polygon  $P_i$  dann ein weiteres mal besucht werden. Der Besuch eines Polygons  $P_i$  ist erst dann gültig, wenn vorher alle anderen Polygone  $P_1, \dots, P_{i-1}$  besucht wurden, andernfalls gilt das Polygon nach wie vor als unbesucht. Abbildung ?? zeigt ein Beispiel für das einfache TPP, Abbildung ?? für das allgemeine TPP. Der gestrichelte Teil des Randes von  $P_4$  ist die Fassade von  $P_4$ . Beachte, daß  $P_5$  vor  $P_4$  besucht wird, aber “gewertet” wird erst der Besuch von  $P_5$  nach dem Besuch von  $P_4$ .

Beschränken wir uns zunächst auf das einfache TPP. Klar ist, daß der kürzeste Weg eine polygonale Kette sein muß, andernfalls ließe sich der Weg verkürzen. Die erste wichtige Idee zur Berechnung des Pfades  $\pi$  ist, lokale Optimalität auszunutzen; eine Idee, die wir schon bei der Berechnung der kürzesten Pfade in Abschnitt ?? benutzt haben:

### Lemma 1.31

- (i) Ein Pfad  $\pi$  ist lokal optimal, wenn die Ecken von  $\pi$  auf dem Rand eines der Polygone  $P_i$  liegen. Liegt die Ecke  $v$  von  $\pi$  auf einer Polygonkante,

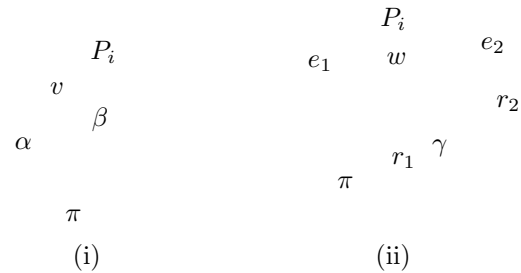


Abbildung 1.40: Lokale Optimalität: (i) bei Reflektion an einer Kante gilt  $\alpha = \beta$ , (ii) bei Reflektion an einer Ecke liegt die Ausgangskante im Winkelbereich  $\gamma$ .

so sind die Winkel zwischen der Polygonkante und den beiden zu  $v$  inzidenten Kanten gleich (Einfallswinkel = Ausfallswinkel, siehe Abbildung ?? (i)).  
 Liegt die Ecke  $v$  auf einer Polygonecke  $w$ , so liegt die ausgehende Kante von  $\pi$  in dem Winkelbereich  $\gamma$ , der durch die Reflexionen  $r_1$  und  $r_2$  der eingehenden Kante an den beiden zu  $w$  inzidenten Polygonkanten  $e_1$  und  $e_2$  begrenzt wird, siehe Abbildung ?? (ii).

- (ii) Für jedes  $p \in \mathbb{R}^2$  und jedes  $i \in \{1, \dots, k\}$  gibt es einen eindeutigen, lokal optimalen Pfad  $\pi_i(p)$ , der in  $s$  startet, die Polygone  $P_1$  bis  $P_i$  in der gegebenen Reihenfolge besucht und in  $p$  endet. Einen solchen Pfad nennen wir im folgenden den **kürzesten  $i$ -Pfad** zu  $p$ .
- (iii) Ein lokal optimaler  $i$ -Pfad ist global optimal.

### Beweis.

- (i) Ist die Bedingung verletzt, wäre der Pfad lokal verkürzbar.  
 (ii) Dies folgt später aus der Konstruktion der lokal optimalen Pfade.  
 (iii) Folgt aus (ii).

□

Die zweite Idee ist, die Ebene in Regionen mit kombinatorisch gleichen Pfaden einzuteilen, d. h. wir zerlegen die Ebene  $\mathbb{R}^2$  disjunkt in Regionen  $\mathcal{R}_j$ , so daß für alle Punkte  $p \in \mathcal{R}_j$  der kürzeste  $k$ -Pfad von  $p$  zu  $s$  über die gleichen Polygonecken/Polygonkanten führt bzw. an den gleichen Polygonecken/Polygonkanten reflektiert wird. Auch diese Idee haben wir bereits in Form der Shortest Path Map benutzt. Abbildung ?? zeigt ein Beispiel für eine vollständige kombinatorische Shortest Path Map für unser Problem. Die Beschriftungen der Regionen geben an, an welchen Kanten bzw. Ecken der Szene der kürzeste Weg zu  $s$  reflektiert wird. Das Problem ist, daß eine vollständige kombinatorische Shortest Path Map in einer Szene mit  $k$  Polygonen und insgesamt  $n$  Ecken eine Komplexität von  $\Omega((n-k)2^k)$  haben kann (Übungsaufgabe). Andererseits brauchen wir gar nicht die gesamte Karte zu speichern. Es reicht aus, wenn wir für jedes Hindernis  $P_i$  eine Karte speichern in der festgehalten wird, über welche Kante bzw. Ecke von  $P_i$  das letzte Segment eines kürzesten  $i$ -Pfades von  $s$  zu  $p$  verläuft. Die Fortsetzung des Weges von  $P_i$  zu  $P_{i-1}$  erhalten wir

dann durch Nachschauen in der Karte zu  $P_{i-1}$  usw.

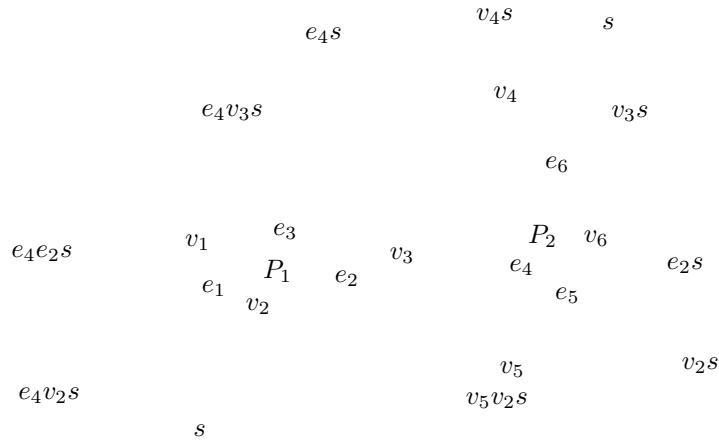


Abbildung 1.41: Vollständige kombinatorische Shortest Path Map.

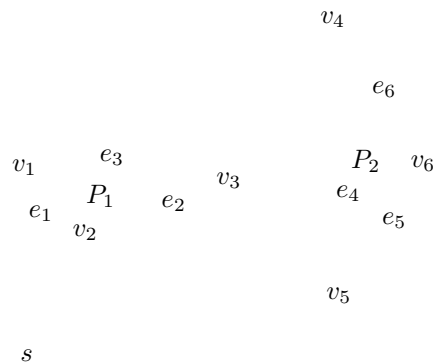
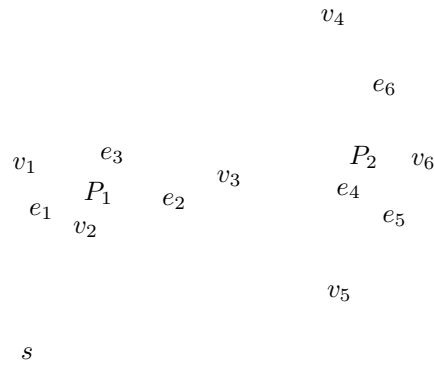


Abbildung 1.42: Last Step Shortest Path Map zu  $P_1$ .

In einer solchen Karte gibt es zwei Klassen von Zellen: Zellen, bei denen der kürzeste Weg über  $P_i$  zu  $s$  an einer Kante oder Ecke von  $P_i$  gemäß den Bedingungen in Lemma ??(i) reflektiert wird nennen wir **Reflektionszellen**. Zellen, bei denen der kürzeste Weg über die Kante/Ecke hinweg führt nennen wir **Durchgangszellen**. Wir können alle Durchgangszellen zur **Durchgangsregion** zusammenfassen.

Die Reflektionszellen lassen sich wiederum in zwei Klassen einteilen, abhängig davon, ob der kürzeste Pfad an einer Kante oder an einer Ecke reflektiert wird. Unsere Karte, die **Last Step Shortest Path Map  $\mathcal{S}_i$**  zu  $P_i$ , besteht dann aus Zellen, die von den an Polygonecken reflektierten Eingangskanten begrenzt werden. Die Eingangskante zu einem Knoten  $v$  von  $P_i$  ist das letzte Segment des kürzesten  $i - 1$ -Pfades zu  $v$ . Damit lassen sich die  $\mathcal{S}_i$  sukzessive aufbauen. Abbildung ?? und Abbildung ?? zeigen Beispiele

Abbildung 1.43: Last Step Shortest Path Map zu  $P_2$ .

für die Last Step Shortest Path Maps  $\mathcal{S}_1$  und  $\mathcal{S}_2$ . Abbildung ?? zeigt ein weiteres Beispiel. Die hell gefärbten Bereiche stellen die Durchgangsregionen dar.

*Durchgangsregion*

$P$

$s$

Abbildung 1.44: Last Step Shortest Path Map zu  $P$ .

Wir bezeichnen den Teil von  $P_i$ , in dem eingehende Pfadsegmente gemäß Lemma ??(i) reflektiert werden, den **Reflektionsbereich**  $T_i$  von  $P_i$ . Gleichzeitig ist  $T_i$  der Bereich, in dem ein kürzester  $i$ -Pfad das Polygon  $P_i$  nach dem Besuch von  $P_1, \dots, P_{i-1}$  zum ersten Mal besuchen kann. Die von  $T_i$  ausgehenden, reflektierten Strahlen bezeichnen wir als den **Stern**  $R_i$  von  $T_i$ .

Im Gegensatz zur vollständigen kombinatorischen Shortest Path Map kann eine Last Step Shortest Path Map keine hohe Komplexität haben, wie folgendes Lemma zeigt:

**Lemma 1.32**

- (i)  $T_i$  ist zusammenhängend.
- (ii) Die Strahlen  $R_i$  sind disjunkt und jeder Punkt  $p \in \mathbb{R}^2$  außerhalb der Durchgangsregion wird von genau einem Strahl von  $R_i$  getroffen.
- (iii) Die Komplexität einer  $\mathcal{S}_i$  liegt in  $O(|P_i|)$ .

**Beweis.**

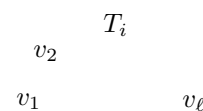
- (i) Übungsaufgabe
- (ii) Wir zeigen induktiv, daß die Strahlen  $R_i$  disjunkt sind. Für  $\mathcal{S}_1$  gilt dies offensichtlich. Nehmen wir an, die Aussage gilt für  $\mathcal{S}_1, \dots, \mathcal{S}_{i-1}$ . Zu jedem Knoten  $v$  von  $P_i$  läßt sich eindeutig die Region von  $\mathcal{S}_{i-1}$  bestimmen, in der  $v$  liegt. Damit ist auch der kürzeste  $(i-1)$ -Pfad zu  $v$  eindeutig, insbesondere die letzte Kante dieses Pfades. Damit sind alle letzten Kanten zu Knoten aus  $P_i$  disjunkt bis auf einen gemeinsamen Startpunkt. Es ist bekannt, daß die Reflektionen disjunkter Liniensegmente an einem konvexen Polygon ebenfalls disjunkt sind.
- (iii) Die Kanten in  $\mathcal{S}_i$  lassen sich in zwei Klassen unterteilen: die Kanten von  $T_i$  und die Strahlen, die von den Ecken von  $T_i$  ausgehen.  $T_i$  ist zusammenhängend und — da die  $P_j$  disjunkt sind — Teil des Randes von  $P_i$ . Daher besteht  $T_i$  aus maximal  $|P_i|$  Kanten. Die von den Ecken von  $T_i$  ausgehenden Strahlen sind Teile von  $R_i$  und damit disjunkt. Ihr Beitrag zur Komplexität von  $\mathcal{S}_i$  liegt damit ebenfalls in  $O(|P_i|)$ . □

Damit haben wir eine Vorstellung, wie eine Karte aussehen kann, mit Hilfe derer wir unsere Pfade berechnen können. Es stellen sich nun die Fragen, wie wir die  $\mathcal{S}_i$  effizient berechnen und wie wir aus den  $\mathcal{S}_i$  einen kürzesten Pfad konstruieren können. Die letzte Frage beantwortet folgendes Lemma:

**Lemma 1.33** Sind für das einfache TPP die Last Step Shortest Path Maps  $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_i$  gegeben, so kann zu allen  $q \in \mathbb{R}^2$  der kürzeste  $i$ -Pfad  $\pi_i(q)$  in Zeit  $O(k \cdot \log \frac{n}{k})$  berechnet werden.

**Beweis.** Algorithmus ?? liefert zu einem Anfragepunkt  $q \in \mathbb{R}^2$  den Pfad  $\pi_i(q)$ , Abbildung ?? zeigt einige Beispiele für kürzeste  $i$ -Pfade.

Wir wissen aus Lemma ??, daß die Komplexität einer LSSPM  $\mathcal{S}_i$  in  $O(|P_i|)$  liegt. Die Karte kann daher in einer Datenstruktur gespeichert werden, die in Zeit  $O(\log |P_i|)$  die Zelle liefert, in der ein Anfragepunkt liegt. Wir könnten dazu eine Point-Location



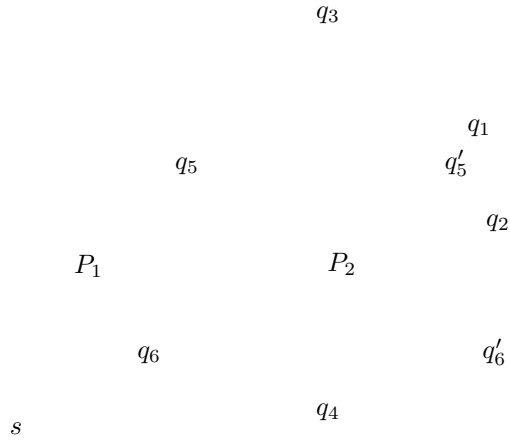


Abbildung 1.45: Queries beim einfachen TPP:  $q_1$  und  $q_2$  liegen in der Durchgangsregion,  $q_3$  und  $q_4$  in einer Zelle zu einer Ecke,  $q_5$  und  $q_6$  in der Zelle einer Kante.

Datenstruktur benutzen; es reicht jedoch, die Knoten der konvexen Kette  $T_i$  in einem Suchbaum oder Array zu speichern. Zu jedem Knoten  $v$  halten wir die beiden von  $v$  ausgehenden Strahlen ebenfalls fest.

Durch die rekursiven Anfragen liegt die Zeit, um einen kürzesten  $i$ -Pfad zu bestimmen, in  $O(\sum_{j=1}^k \log |P_j|)$ . Diese Summe nimmt bei  $|P_j| = \frac{n}{k}$  ihr Maximum an.  $\square$

Mit Hilfe der Queries lassen sich die Last Step Shortest Path Maps recht einfach sukzessiv konstruieren, siehe Algorithmus ???. Die Zeit zur Konstruktion einer einzelnen  $\mathcal{S}_i$  liegt in  $O(|P_i|k \log \frac{n}{k})$ , insgesamt also

$$O\left(\sum_{i=1}^k |P_i|k \log \frac{n}{k}\right) = O(nk \log \frac{n}{k}).$$

Mit Lemma ??? und Lemma ??? folgt:

**Theorem 1.34** (Dror, Efrat, Lubiw, Mitchell, 2003)

Für das einfache TPP mit  $k$  Polygonen mit insgesamt  $n$  Kanten läßt sich in Zeit  $O(kn \log \frac{n}{k})$  eine Suchstruktur der Größe  $O(n)$  aufbauen, mit Hilfe derer der kürzeste  $i$ -Pfad zu einem Punkt  $q \in \mathbb{R}^2$  in Zeit  $O(k \log \frac{n}{k})$  berechnet werden kann. [?]

Im allgemeinen TPP müssen wir die Zäune  $F_i$  berücksichtigen, d. h. der kürzeste  $i$ -Pfad zu einem Punkt  $p$  muß zwischen  $P_j$  und  $P_{j+1}$  in  $F_j$  liegen. Kürzeste Pfade können nun zusätzlich an reflexen Ecken eines Zaunes abknicken, wie z. B. der Weg von  $s$  nach  $p$  in Abbildung ???(i). Dadurch gibt es eine weitere Klasse von Zellen in der Zerlegung der Ebene: Zellen, die einer reflexen Ecke eines Zaunes zugeordnet sind. Z. B. führen kürzeste



---

**Algorithmus 1.10** Einfaches TPP: Query

---

Bestimmt zu gegebenen Last Step Shortest Path Maps  $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_i$  und Anfragepunkt  $q \in \mathbb{R}^2$  den kürzesten  $i$ -Pfad von  $s$  zu  $q$ .

- Bestimme die Region von  $q$  in  $\mathcal{S}_i$ :  
Durch prüfen der Lage (rechts oder links) von  $q$  bzgl. der konvexen Kette aus  $T_i$  und den beiden Strahlen am Rand der Durchgangsregion wird bestimmt, ob  $q$  in der Durchgangsregion oder in einer Reflektionszelle liegt. Liegt  $q$  in einer Reflektionszelle, so kann diese durch binäre Suche mit Rechts/Linkstests in den Strahlen, die von den Knoten von  $T_i$  ausgehen, bestimmt werden.
  - Falls  $q$  in der Durchgangsregion liegt: es gilt  $\pi_i(q) = \pi_{i-1}(q)$ , bestimme rekursiv  $\pi_{i-1}(q)$ .
  - Falls  $q$  in der Zelle zu einer Ecke  $v$  liegt: das letzte Segment des kürzesten  $i$ -Pfades von  $s$  zu  $q$  ist  $\overline{vq}$ . Bestimme rekursiv  $\pi_{i-1}(v)$ .
  - Falls  $q$  in der Zelle zu einer Kante  $e$  liegt: sei  $q'$  die Spiegelung von  $q$  an  $e$ . Berechne rekursiv  $\pi_{i-1}(q')$  und ersetze in diesem Pfad das Segment von  $e$  zu  $q'$  durch das Segment von  $e$  zu  $q$ .
- 

---

**Algorithmus 1.11** Einfaches TPP: Konstruktion der  $\mathcal{S}_i$ 

---

- Berechne  $\mathcal{S}_1$  durch Reflektionen an den Knoten von  $P_1$ .
  - Berechne sukzessiv  $\mathcal{S}_i$  aus  $\{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_{i-1}\}$  wie folgt: Berechne zu jedem Knoten  $v \in P_i$  den kürzesten  $(i-1)$ -Pfad  $\pi_{i-1}(v)$ . Schneidet das letzte Stück dieses Pfades das Innere von  $P_i$ , so liegt  $v$  in der Durchgangsregion von  $\mathcal{S}_i$ . Andernfalls liegt  $v$  in  $T_i$  und die beiden Reflektionen des letzten Segmentes von  $\pi_{i-1}(v)$  an den zu  $v$  inzidenten Kanten bestimmen die Strahlen, die die Zelle von  $v$  begrenzen.
-

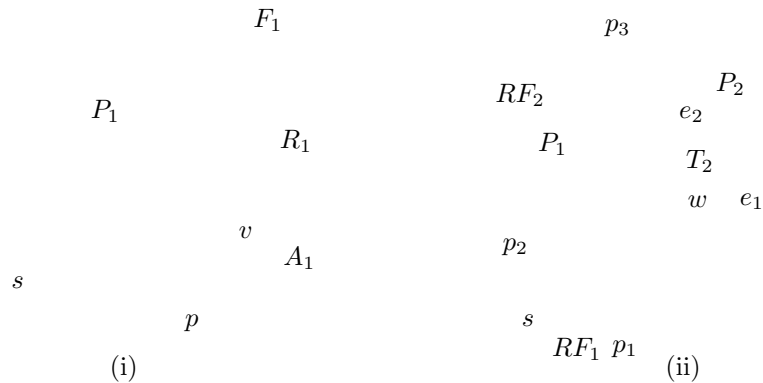


Abbildung 1.46: (i) Kürzeste Wege können an reflexen Ecken der Zäune abknicken, (ii) mögliche Kombinationen bei überlappenden Polygonen.

$i$ -Pfade aus Punkten der Zelle unterhalb der strichpunktierten Linien in Abbildung ??(i) über den Knoten  $v$ .

Die zweite Verallgemeinerung ist, daß sich die Polygone überlappen dürfen, wodurch weitere Arten von kombinatorisch kürzesten Wegen ins Spiel kommen. Betrachte Abbildung ??(ii): der Strahl  $RF_1$  entsteht durch Punktspiegelung von  $\overline{sw}$  an  $w$  und anschließender Spiegelung an den Kanten  $e_1$  und  $e_2$ , also den Kanten von  $P_1$  und  $P_2$ , auf denen der Schnittpunkt  $w$  liegt.  $RF_2$  entsteht durch Reflektion von  $\overline{sw}$  an  $e_2$ . Der Punkt  $p_1$  liegt unterhalb des Strahles  $RF_1$  und der kürzeste Weg von  $s$  nach  $p_1$  wird erst an  $P_1$ , danach an  $P_2$  reflektiert.  $p_2$  liegt zwischen  $RF_1$  und  $RF_2$  und sein kürzester Weg verläuft direkt über  $w$ . Schließlich liegt  $p_3$  oberhalb von  $RF_2$  und in der Durchgangsregion von  $P_1$  und der kürzeste Weg wird lediglich durch eine Reflektion an  $P_2$  gebildet.

Wie bereits im einfachen Fall sei  $T_i$  der Reflektionsbereich, also der Bereich von  $P_i$ , in dem die letzten Kanten eingehender  $i$ -Pfade reflektiert werden. Da sich die  $P_j$  überlappen können, ist  $T_i$  hier nicht mehr unbedingt ein Teil des Randes von  $P_i$  sondern kann auch im Inneren von  $P_i$  liegen, siehe z. B.  $T_2$  (gepunktet) in Abbildung ??(ii).  $R_i$  bezeichne wie oben die Menge der Strahlen, die von  $T_i$  ausgehen. Berücksichtigen wir auch den Zaun  $Z_j$ , so müssen wir folgendes beachten: Die Strahlen  $R_j$ , die an einer Ecke des Zaunes abknicken, um eine sonst unerreichbare Region zu durchqueren nennen wir  $A_j$ , siehe Abbildung ?. Wir können folgendes für das allgemeine TPP festhalten (ohne Beweis):

### Lemma 1.35

- (i) Für ein Polygon  $P_i$  ist der Reflektionsbereich  $T_i$  ein Baum.
- (ii) Der Stern  $R_i$  besteht aus disjunkten Strahlen außerhalb der Durchgangsregion (ohne Berücksichtigung der  $F_i$ ) und jeder Punkt  $p \in \mathbb{R}^2$  wird von einem eindeutigen Strahl getroffen.

(iii) Ein lokal optimaler Pfad ist eindeutig und lokale Optimalität garantiert globale Optimalität.

Wie beim einfachen TPP können wir nun sukzessive die  $T_i$ ,  $R_i$  und  $A_i$  sowie die zugehörigen Last Step Shortest Path Maps berechnen. Wir wollen hier nur das Ergebnis festhalten:

**Theorem 1.36** (Dror, Efrat, Lubiw, Mitchell, 2003)

Für das allgemeine TPP mit  $k$  Polygonen und  $k+1$  Zäunen mit insgesamt  $n$  Kanten können wir in Zeit  $O(k^2 n \log n)$  eine Suchstruktur der Größe  $O(kn)$  aufbauen, mit Hilfe derer der kürzeste  $i$ -Pfad zu einem Punkt  $q \in \mathbb{R}^2$  in Zeit  $O(i \log n + m)$  berechnet werden kann, wobei  $m$  die Anzahl der Segmente des Pfades angibt. [?]

Kommen wir nun zurück zu unserem eigentlichen Problem, die Shortest Watchman Route zu berechnen.

**Theorem 1.37** (Dror, Efrat, Lubiw, Mitchell, 2003)

In einem einfachen Polygon mit  $n$  Ecken und Startpunkt  $s$  läßt sich die SWR in Zeit  $O(n^3 \log n)$  berechnen. [?]

**Beweis.** Aus der Eingabe  $(P, s)$  zum SWR können wir wie folgt eine Eingabe  $(P_1, \dots, P_k, F_1, \dots, F_k, s', t)$  für das allgemeine TPP berechnen: Start- und Zielpunkt des TPP-Pfades ist der Startpunkt  $s$  der SWR. Als Polygone  $P_i$  wählen wir die Taschen  $P_{c_i}$  von  $P$ , siehe Definition ??, als Zäune legen wir  $P$  fest ( $F_i := P$ ). Für die Laufzeit des TPP ist nur die Komplexität der Fassade der  $P_i$  entscheidend, und diese besteht in unserem Fall nur aus 2 Ecken. Somit liegt die Komplexität aller Fassaden in  $O(n)$ . Die Komplexität aller Zäune liegt in  $O(n^2)$  und wir erhalten eine Laufzeit von  $O(n^3 \log n)$ .  $\square$

### 1.3 Kürzeste Pfade im Raum

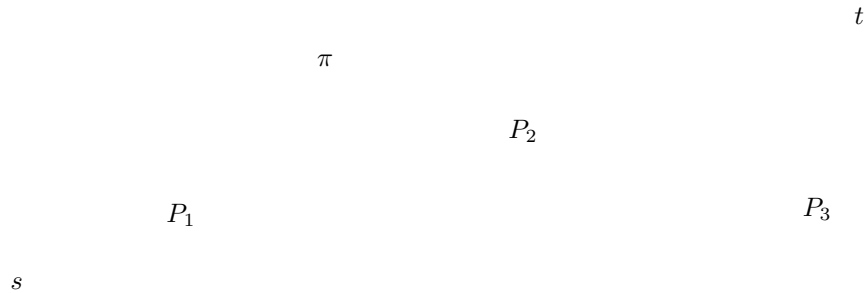


Abbildung 1.47: Kürzester Pfad im 3D.

Verlassen wir nun die Ebene und betrachten kürzeste Pfade im dreidimensionalen Raum: Gegeben seien  $h$  Polyeder im Raum mit insgesamt  $n$  Ecken, ein Startpunkt  $s$  und ein Endpunkt  $t$ . Gesucht ist der kürzeste Pfad  $\pi$  von  $s$  nach  $t$ , der außerhalb der Polyeder verläuft.

Eine erste (ernüchternde) Erkenntnis ist, daß dieses Problem nicht mehr diskret ist, da kürzeste Pfade über innere Punkte von Kanten verlaufen können. Wir haben zwei Teilprobleme zu lösen:

1. Bestimme die Folge der Kanten bzw. Ecken, die von einer Kürzesten berührt werden. (kombinatorisches Problem)
2. Bestimme die Berührungspunkte auf den Kanten. (algebraisch-numerisches Problem)

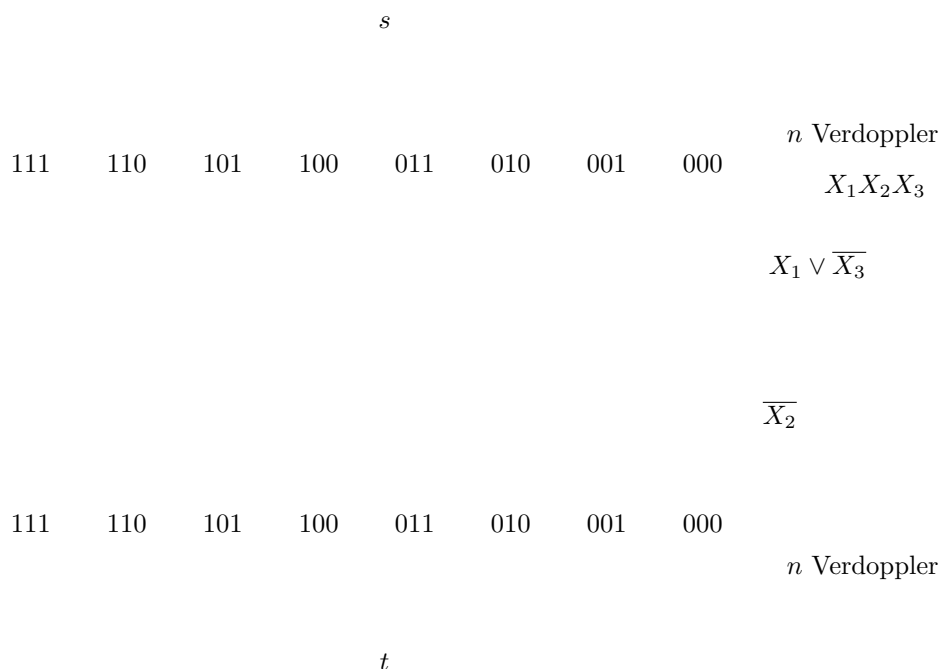
Leider ist bereits das erste Teilproblem NP-hard:

**Theorem 1.38** (Canny, Reif, 1987)

Die Bestimmung der Kantenfolge einer Kürzesten im  $\mathbb{R}^3$  ist NP-hard. Bereits die Frage, ob es zu gegebenem  $k$  eine Kantenfolge gibt, die einen Weg  $\pi$  mit  $|\pi| \leq k$  liefert, ist NP-hard. [?]

**Beweis.**

Zeige: 3-SAT ist auf die Kantenbestimmung reduzierbar. Dazu bilden wir in polynomieller Zeit den booleschen Ausdruck  $\alpha = \bigwedge_{i=1}^m (\Pi_{i1} \vee \Pi_{i2} \vee \Pi_{i3})$  mit  $\Pi_{ij} \in \{X_k, \overline{X_k}\}$ ,  $1 \leq k \leq n$ , auf einen Parcours  $P_\alpha$  im Raum mit Startpunkt

Abbildung 1.48: Beispiel: Filter für Klausel  $(X_1 \vee X_3) \wedge \overline{X_2}$ .

$s$  und Endpunkt  $t$  ab. Die Geodätischen von  $s$  nach  $t$  in  $P_\alpha$  zerfallen dabei in Klassen  $G_w$ , die den  $2^n$  möglichen Wahrheitswertbelegungen  $w$  der Variablen  $X_i$  entsprechen. Die Klasse einer Geodätischen läßt sich aus der Folge der von ihr besuchten Kanten in  $P_\alpha$  ablesen.  $P_\alpha$  wird so konstruiert, daß für eine Kürzeste  $\pi$  von  $s$  nach  $t$ , die in einer Klasse  $G_w$  liegt, gilt: entweder die Belegung  $w$  erfüllt den Ausdruck  $\alpha$ , oder  $\alpha$  ist unerfüllbar.

Abbildung ?? zeigt die wesentliche Idee: die vom Startpunkt  $s$  ausgehende Geodätische wird  $n$  mal verdoppelt, so daß für jede der  $2^n$  möglichen Variablenbelegungen ein Pfad existiert. Diejenigen Pfade, die die Klausel nicht erfüllen, werden verlängert — in der Abbildung durch  $\sim\sim\sim$  angedeutet — und schließlich wieder zu einer Geodätischen zusammengefaßt, wobei darauf zu achten ist, daß die Konstruktion des Parcours in polynomieller Zeit möglich sein soll!

Der Parcours besteht im wesentlichen aus Platten der Dicke  $\varepsilon$  mit Schlitzern, durch die die Geodätischen verlaufen können. Legt man mehrere Platten in geeigneter Weise mit einem Abstand  $\varepsilon$  übereinander, läßt sich der Verlauf der Geodätischen beeinflussen. Zur Konstruktion eines solchen Parcours benötigen wir die im folgenden erklärten Bausteine. In jedem Baustein — bis auf die Barrieren im Literalfilter — werden alle Klassen von Geodätischen um den gleichen Betrag verlängert.

### Verdoppler

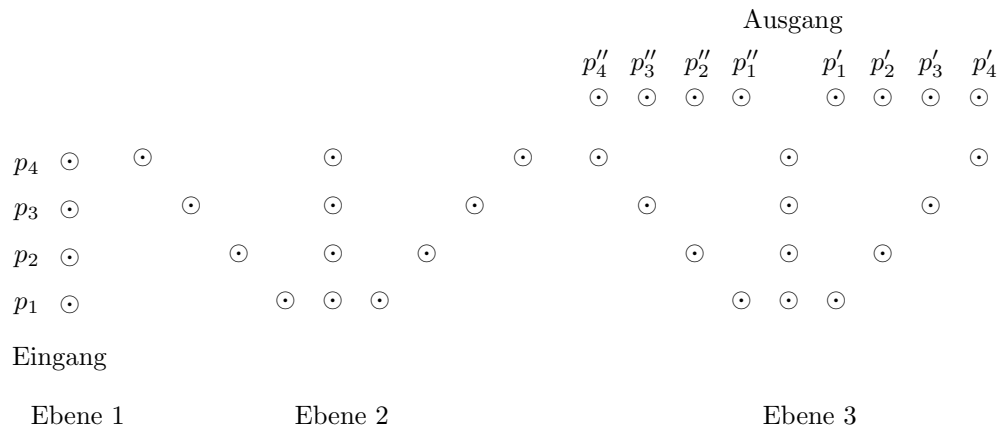


Abbildung 1.49: Die drei Ebenen eines Verdopplers.

Ein Verdoppler dient dazu, seine aus  $n$  Klassen von Geodätischen bestehende Eingabe in  $2n$  Klassen aufzuspalten. Er besteht aus drei Ebenen: die unterste Ebene enthält einen vertikalen Schlitz, durch den die Eingabe in den Verdoppler geführt wird. Die Eingabe verzweigt nach rechts und links und wird über zwei diagonale Schlitzte in der zweiten Ebene zu einem horizontalen Schlitz in der dritten Ebene geführt, siehe Abbildung ???. Das Zeichen  $\odot$  symbolisiert in den Abbildungen eine Geodätische, die von einer unteren Ebene in eine obere Ebene läuft.

### Mischer

Ein Mischer besteht aus vier Ebenen. Die Eingabe wird durch einen horizontalen Schlitz in der untersten Ebene geführt, in zwei Hälften geteilt und durch zwei horizontale Schlitzte in der zweiten Ebene geführt, vgl. Abbildung ???. In der dritten Ebene liegen zwei Reihen mit diagonalen Schlitzten, die die gesplittete Eingabe in die Vertikale umlenken, wo sie durch eine Reihe vertikaler Schlitzte in der vierten Ebene ausgegeben wird. Die Schlitzte in den Diagonalen sind dabei so versetzt, daß sich jeweils eine Klasse der ersten Hälfte und eine Klasse der zweiten Hälfte abwechseln. Die Eingabe  $X_i X_j X_k \dots$  wird derart permutiert, daß die Klassen mit  $X_j = 1$  in einer Hälfte liegen, die Klassen mit  $X_j = 0$  in der anderen Hälfte. Anders ausgedrückt ergibt sich eine Linksrotation der Variablen,  $X_j$  wird zum signifikantesten Bit.

### Literalfilter

Eine Klasse von Geodätischen kodiert eine Binärzahl  $w = (w_1, \dots, w_n) \in \{0, 1\}^n$ , die einer Belegung der Variablen entspricht. Der Literalfilter dient nun dazu, die Klassen von Geodätischen zu verlängern, deren Belegung das Literal nicht erfüllt. Um also alle Pfade herauszufiltern,

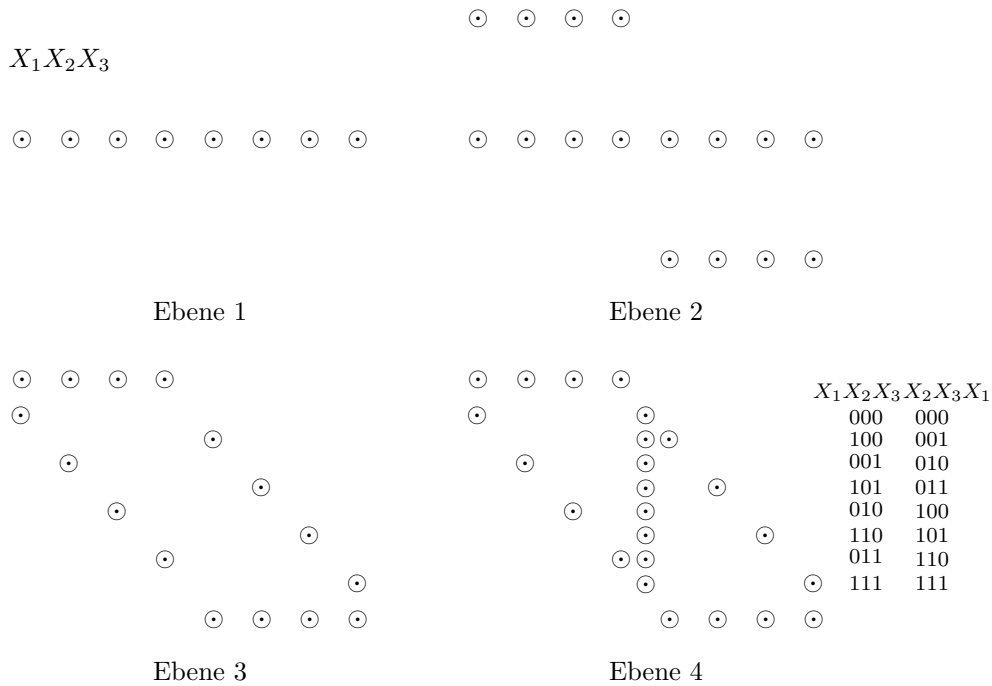


Abbildung 1.50: Die vier Ebenen eines Mischers.

bei denen  $w_i = 0$  den Ausdruck erfüllt, müssen alle Pfade mit  $w_i = 1$  mit einer Barriere verlängert werden. Dabei dürfen wir die Barrieren allerdings nicht in die ursprüngliche Kodierung einfügen, wie dies in Abbildung ?? gezeigt ist, da dies evtl. zu exponentiell vielen Barrieren führen würde; um z. B.  $w_n$  zu filtern, wäre eine Barriere in jedem zweiten Pfad nötig! Also setzen wir Mischer ein, um die zu filternde Variable zur signifikantesten Variable zu machen, und können dann mit *einer* Barriere auf der rechten bzw. linken Seite alle Pfade mit  $w_i = 0$  bzw.  $w_i = 1$  verlängern, siehe Abbildung ??. Da ein Mischer eine Linksrotation der Variablen bewirkt, ist  $w_i$  nach  $i-1$  Mischerdurchläufen das signifikanteste Bit. Anschließend benötigen wir noch  $n - i + 1$  Mischer, um wieder die ursprüngliche Kodierung zu erhalten.

**Klauselfilter**

Der Klauselfilter dient schließlich zur Umsetzung einer Klausel der Form  $C_i = \Pi_{i1} \vee \Pi_{i2} \vee \Pi_{i3}$  mit  $\Pi_{ij} \in \{X_k, \overline{X_k}\}, 1 \leq k \leq n$ . Dazu benötigen wir drei “parallel geschaltete” Literalfilter: die Eingabe von  $2^n$  Klassen von Geodätischen wird allen Literalfiltern zugeführt. In den Literalfiltern werden die Geodätischen verlängert, die die Klausel nicht erfüllen. Durch die Parallelschaltung ergibt sich eine Disjunktion der Literale, siehe Abbildung ??.

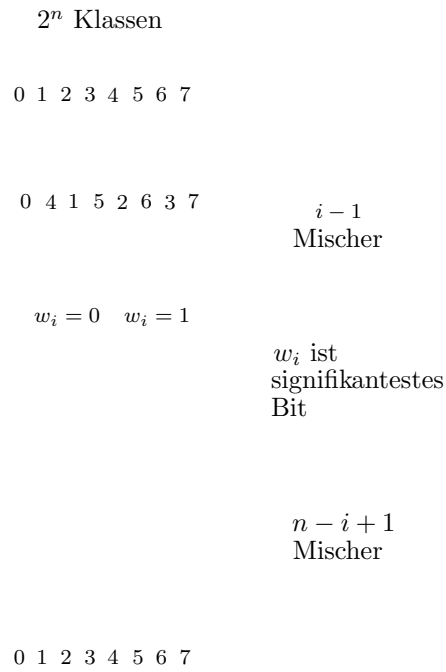


Abbildung 1.51: Literalfilter.

 $S_{\text{in}}$   
 $2^n$  Geodätische
 $\Pi_{i1}$  $\Pi_{i2}$  $\Pi_{i3}$ 
 $S_{\text{out}}$   
 $2^n$  Geodätische

Abbildung 1.52: Klauselfilter.



Zur Konstruktion eines Parcours aus einer 3-SAT Klauselmenge

$$\alpha = \bigwedge_{i=1}^m C_i$$

über  $n$  Variablen benötigen wir zunächst einen Block aus  $n$  Verdopplern, um aus der vom Startpunkt ausgehenden Geodätischen  $2^n$  Klassen für jede mögliche Variablenbelegung zu erzeugen. Diese Klassen werden durch  $m$  hintereinander geschaltete Klauselfilter und schließlich durch  $n$  invers benutzte Verdoppler geführt, die die gefilterten Klassen von Geodätischen wieder zu einer zusammenfassen, die zum Zielpunkt  $t$  führt.  $\square$

Damit bleibt das zweite Teilproblem: Wie bestimmt man zu gegebener Kantenfolge die richtigen Berührungspunkte, also den kürzesten Pfad, der die Kanten in dieser Reihenfolge besucht?

Schon das einfachere Problem, zu Geraden  $g_i$  im Raum und fest vorgegebener Besuchsreihenfolge die Berührungspunkte des Pfades so zu verschieben, daß seine Länge minimal wird, ist schwierig. Ein lokales Optimalitätskriterium erhält man durch Hochklappen der durch  $g_i$  und  $g_{i+1}$  gegebenen Ebene in die durch  $g_i$  und  $g_{i-1}$  gegebene Ebene: liegen  $b_{i-1}$ ,  $g_i$  und  $b_{i+1}$  in einer Ebene, muß der kürzeste Pfad zwischen  $b_{i-1}$  und  $b_{i+1}$  ein Liniensegment durch  $b_i$  sein. Dies gilt genau dann, wenn die beiden Winkel  $\alpha_i$  und  $\beta_i$  gleich sind. Wären beide Winkel nicht gleich, so ließe sich der Pfad zwischen  $b_{i-1}$  und  $b_{i+1}$  verkürzen.

Sei  $\vec{u}_i$  ein Vektor in Richtung  $g_i$  mit  $|\vec{u}_i| = 1$ , dann gilt:<sup>13</sup>

$$\alpha_i = \beta_i \iff \frac{\overrightarrow{b_{i-1}b_i} \cdot \vec{u}_i}{|\overrightarrow{b_{i-1}b_i}|} = \frac{\overrightarrow{b_i b_{i+1}} \cdot \vec{u}_i}{|\overrightarrow{b_i b_{i+1}}|} \quad (*)$$

Die Gleichung (\*) muß für jede Gerade  $g_i$  mit  $1 \leq i \leq n$  gelten. Ein naiver Ansatz wäre nun, den ersten Berührungspunkt  $b_1$  hinter  $s$  zu raten und mit der Winkelbedingung die Fortsetzung des Pfades zu bestimmen. Das Problem dabei ist, daß  $g_2$  den  $\alpha_1$ -Kegel um  $g_1$  zweimal schneiden kann (siehe Abbildung), die Fortsetzung des Pfades ist also nicht eindeutig, es bleiben  $2^{n-1}$  Möglichkeiten.

Der Versuch, die Gleichungen (\*) mit algebraischen Methoden exakt zu lösen, führt zu Weglängen, die durch

<sup>13</sup> $\vec{a} \cdot \vec{b}$  bezeichnet dabei das Skalarprodukt zweier Vektoren mit eingeschlossenem Winkel  $\gamma$ :  $\vec{a} \cdot \vec{b} = |\vec{a}||\vec{b}| \cos \gamma$ .

algebraische Zahlen<sup>14</sup> mit exponentiell in  $n$  wachsendem Grad dargestellt werden. [?]

Das Problem zu entscheiden, ob ein Pfad einer Länge  $\leq k$  in einer Umgebung mit  $n$  Ecken existiert, kann in Zeit  $2^{n^{O(1)}}$  und Platz  $n^{O(\log n)}$  Platz gelöst werden. [?]

Es gibt polynomiale  $(1 + \varepsilon)$ -Approximationsschemata<sup>15</sup> in Zeit  $O(n^2 \log^c n \cdot \frac{1}{\varepsilon^d})$ . [?]

---

<sup>14</sup>Eine algebraische Zahl ist eine komplexe Zahl, die Nullstelle eines Polynoms  $a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$  mit rationalen oder ganzzahligen Koeffizienten ist. Nicht algebraisch sind z. B.  $\pi$  oder  $e$ .

<sup>15</sup>D. h. der kürzeste Weg läßt sich beliebig genau approximieren, je kleiner das  $\varepsilon$ , desto genauer die Approximation, aber desto höher die Rechenzeit.

## 1.4 Kürzeste Pfade auf der Oberfläche eines Polyeders

Da sich die Bestimmung kürzester Pfade im Raum im allgemeinen als sehr schwierig erwiesen hat, betrachten wir den Spezialfall, bei dem sowohl  $s$  als auch  $t$  auf der Oberfläche desselben Polyeders mit  $n$  Kanten im Raum liegen. Auch der kürzeste Weg soll auf der Oberfläche des Polyeders verlaufen und nicht frei durch den Raum, auch wenn die ‐Luftlinie‐ kürzer w re. Wir interessieren uns also z. B. f r den Weg, den ein Wanderer vom Gipfel eines Berges zum n chsten Gipfel nehmen w rde: ohne weitere Hilfsmittel mu  er durch das Tal zwischen beiden Gipfeln wandern.

Ein Polyeder habe dabei folgende Eigenschaften:

- Es ist eine kompakte, zusammenh ngende Teilmenge des  $\mathbb{R}^3$ .
- Der Rand besteht aus Polygonen.
- In jeder Umgebung eines Randpunktes liegen sowohl innere als auch  u ere Punkte. Obige Abbildung zeigt also *keinen* Polyeder in unserem Sinn.

Weiterhin nehmen wir an, die Oberfl che des Polyeders sei bereits trianguliert;<sup>16</sup> aus der Euler-Formel ergibt sich, da  die Anzahl der Dreiecke in  $O(n)$  liegt.

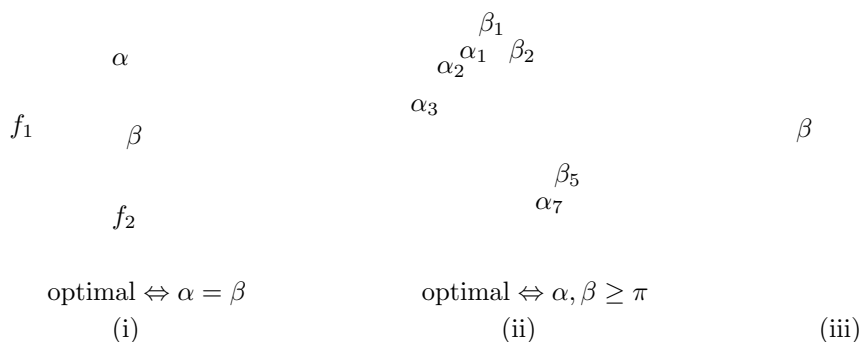


Abbildung 1.53: Eigenschaften von K rztzesten auf Polyedern.

Halten wir zun chst einige Eigenschaften von k rztzesten Pfaden auf Polyedern fest:

Eine lokale Eigenschaft ist, da  in Abbildung ??(i) der Pfad genau dann optimal ist, wenn  $\alpha = \beta$  gilt. Dreht man n mlich eine der Fl chen um ihre gemeinsame Kante in die von der anderen Fl che aufgespannte Ebene, so da 

<sup>16</sup>Ansonsten lie e sich das Polyeder in Zeit  $O(n)$  triangulieren.

die Flächen rechts und links von der Kante liegen und nicht übereinander, so muß der kürzeste Pfad eine Gerade in der Ebene sein. Dies ist mit der Bedingung  $\alpha = \beta$  äquivalent.

In Abbildung ??(ii) verläuft die Kürzeste (fett gezeichnet) durch einen Eckpunkt. Dies ist dann und nur dann möglich, wenn sowohl die Summe der Winkel  $\alpha = \sum \alpha_i$  als auch die Summe der Winkel  $\beta = \sum \beta_i$  größer gleich<sup>17</sup>  $\pi$  ist. Wäre z. B.  $\beta < \pi$ , wie in der Abwicklung in Abbildung ??(iii) gezeigt, kann die fett gezeichnete Linie nicht die Kürzeste sein, da die gestrichelte Linie kürzer ist. Da zudem gilt:  $p$  konvex<sup>18</sup>  $\Rightarrow \alpha + \beta < 2\pi$  haben wir folgendes gezeigt:

**Lemma 1.39** *Eine Geodätische  $\pi$  auf einem Polyeder hat folgende lokale Eigenschaften:*

- (i) *An jeder von  $\pi$  überschrittenen Polyederkante müssen die Scheitelwinkel gleich sein.*
- (ii)  *$\pi$  kann nicht über eine konvexe Ecke von  $P$  führen.*

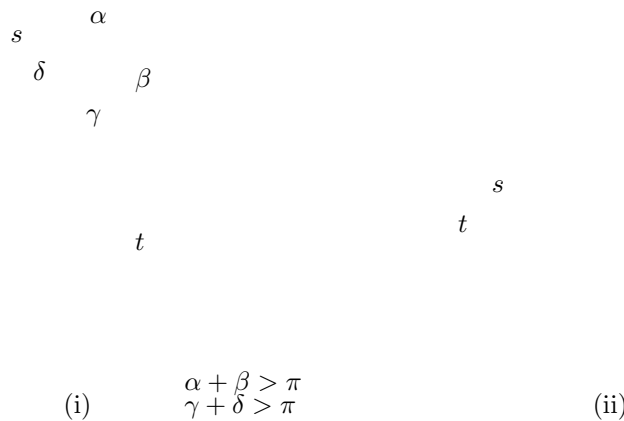


Abbildung 1.54: (i) Kürzeste über nichtkonvexe Ecke, (ii) Geodätische von  $s$  nach  $t$ , aber keine Kürzeste.

Wie das Beispiel in Abbildung ??(i) zeigt, kann eine Kürzeste sehr wohl über nichtkonvexe Ecken laufen.

Neben den lokalen Eigenschaften können wir auch einige globale Eigenschaften festhalten, die jedoch *nicht* für Geodätische gelten. Abbildung ??(ii) zeigt eine Geodätische, für die alle drei Eigenschaften nicht gelten!<sup>19</sup>

<sup>17</sup>An dieser Stelle ist natürlich die Zahl  $\pi = 3,1415\dots$  gemeint, nicht der Pfad  $\pi$ .

<sup>18</sup>Eine Ecke  $v$  im  $\mathbb{R}^3$  ist konvex, wenn eine Ebene existiert an der das Polyeder 'durchgeschnitten' werden kann, so daß der  $v$  enthaltende Teil des Polyeders ein konvexes Polyeder ist.

<sup>19</sup>Zur Erinnerung: eine Geodätische ist lokal unverkürzbar, eine Kürzeste auch global.



viele Wege zu berechnen! Abhilfe schafft die Idee, Regionen von Punkten mit kombinatorisch gleichen Kürzesten zum Punkt  $x$  zu bilden. Es genügt, die Kantenfolge  $\mathcal{E} = e_{k,1}, \dots, e_{k,m_k}$  die der Weg ab der "letzten Ecke"  $v = v_{k-1}$  durchlaufen hat zu betrachten, da wir nur *eine* Kürzeste berechnen wollen. Die verschiedenen Kürzesten von  $a$  nach  $v$  brauchen nicht unterschieden zu werden, wir müssen uns nur eine solche merken.

Wegen der Winkelbedingung in Lemma ??(i) lassen sich  $v$  und alle Dreiecke ab  $v$ , die der Weg  $\pi_v^x$  durchquert, in die  $f$ -Ebene aufklappen.<sup>20</sup> Dabei wird  $\pi_v^x$  zu einem Liniensegment zwischen  $a$  und  $x$ . Im folgenden bezeichne  $\bar{v}$  das in die  $f$ -Ebene aufgeklappte Bild des Punktes  $v$ .

Fassen wir also alle  $x \in e$  zusammen, für die ein kürzester Weg von  $a$  über  $v$  und  $\mathcal{E}$  den Punkt  $x$  erreicht:

**Definition 1.42** Sei  $e$  eine fest gewählte Kante des Dreiecks  $f$ ,  $v$  eine beliebige Ecke von  $P$  und  $\mathcal{E} = (e_1, \dots, e_m)$  eine beliebige Folge von Kanten von  $P$ . Das Optimalitätsintervall von  $v$  und  $\mathcal{E}$  bzgl.  $e$  und  $f$  ist definiert als

$$I(v, \mathcal{E}) := \left\{ x \in e \mid \exists \text{ Kürzeste } \delta \text{ von } a \text{ nach } x \text{ mit} \right. \\ \left. \begin{array}{l} \bullet \delta \cap \overset{\circ}{f} = \emptyset \\ \bullet \delta \text{ endet mit } v, e_1, \dots, e_m, x \end{array} \right\} .$$

**Lemma 1.43** *Eigenschaften von  $I(v, \mathcal{E})$ :*

- (i) Jede solche Menge  $I$  ist ein Intervall auf  $e$  (evtl. leer).
- (ii) Zwei verschiedene Intervalle können sich nicht überlappen.
- (iii)  $e$  wird von Intervallen  $I(v, \mathcal{E})$  ganz überdeckt.

**Beweis.**

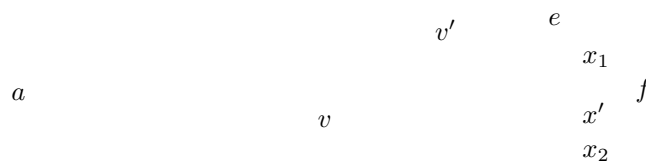


Abbildung 1.56:  $I(v, \mathcal{E})$  muß zusammenhängend sein.

- (i) Zu zeigen:  $I(v, \mathcal{E})$  ist zusammenhängend. Seien  $x_1, x_2 \in e$  in  $I(v, \mathcal{E})$  enthalten, und ein Punkt  $x'$  liege zwischen  $x_1$  und  $x_2$  auf  $e$ , vgl. Abbildung ??. Zeige:  $x' \in I(v, \mathcal{E})$ .  
Angenommen,  $x' \notin I(v, \mathcal{E})$ , dann klappe die Dreiecke entlang der Kantenfolge  $\mathcal{E}$  in die Fläche von  $f$  auf

<sup>20</sup>D. h. eines von je zwei Dreiecken wird solange um die gemeinsame Kante gedreht, bis beide in einer Ebene, eins rechts und eins links der Kante liegen.

$\Rightarrow v' \neq v$ , sogar liegt  $v'$  außerhalb des Dreiecks  $\triangle(v, x_1, x_2)$   
 $\Rightarrow$  die beiden Kürzesten von  $a$  nach  $x_1/x_2$  und  $x'$  kreuzen sich im Inneren einer Fläche  $\zeta$ .

$f$   
 $e$

$\bar{w}_{\ell-1}$

Hyperbel

$\bar{v}_{k-1}$

Abbildung 1.57: Der Bisektor von  $\bar{v}_{k-1}$  und  $\bar{w}_{\ell-1}$  ist eine Hyperbel.

- (ii) Angenommen, zwei Intervalle überlappen sich. Sei  $x$  im Durchschnitt der Intervalle, dann ergibt sich beim Weg von  $a$  nach  $x$  über  $\bar{v}_{k-1}$  eine Weglänge von  $|\pi_a^{v_{k-1}}| + |\bar{v}_{k-1} - x|$ , beim Weg über  $\bar{w}_{\ell-1}$  eine von  $|\pi_a^{w_{\ell-1}}| + |\bar{w}_{\ell-1} - x|$ .  
Für jeden Punkt  $x$  im Durchschnitt der beiden Intervalle müssen beide Weglängen gleich sein, also muß gelten:

$$|\pi_a^{v_{k-1}}| + |\bar{v}_{k-1} - x| = |\pi_a^{w_{\ell-1}}| + |\bar{w}_{\ell-1} - x|.$$

- Durch Umformen dieser Gleichung erhält man die Gleichung einer Hyperbel, die den Bisektor zwischen  $\bar{v}_{k-1}$  und  $\bar{w}_{\ell-1}$  beschreibt, also die Menge der Punkte, für die beide Alternativen gleich sind. Diese Hyperbel hat mit  $e$  jedoch nur einen Schnittpunkt, nämlich den Punkt, an dem sich beide Intervalle berühren. Zwei Intervalle können sich also nicht überlappen.  
(iii) Dies ist klar, da es für jedes  $x \in e$  eine Kürzeste von  $a$  gibt.

□

**Achtung:** Es kann mehrere Intervalle  $I(v_i, \mathcal{E}_j)$  mit derselben "letzten" Ecke  $v_i$  geben, d. h.  $v_i$  kann mehrere unterschiedlich abgewinkelte Kopien  $\bar{v}_{i,k}$  haben.

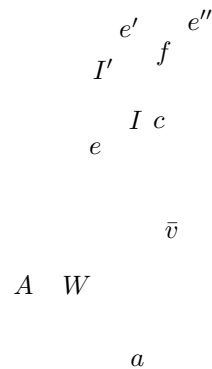
**Algorithmus 1.12** Continuous Dijkstra zur Berechnung der  $I(v, \mathcal{E})$ 

**Datenstruktur:** Prioritätswarteschlange  $W$  von Intervallen mit bisher bekannten Kürzesten von  $a$ .

**Iterationsschritt:**

- Wähle das Intervall  $I = I(v, \mathcal{E})$  aus  $W$ , für das  $d(I) = d(a, v) + |\bar{v} - c|$  minimal ist, wobei  $c$  der zu  $\bar{v}$  nächstgelegene Punkt auf  $I$  ist.
- Propagiere dieses Intervall  $I$  durch das Dreieck auf die beiden gegenüberliegenden Kanten  $e'$  und  $e''$ : Berechne das Intervall  $I'$  auf Kante  $e'$  mit  $I' = \{x \in e' \mid \text{der Weg von } a \text{ über } v \text{ und } I \text{ durch das Dreieck } f \text{ ist kürzer als der kürzeste bisher bekannte Weg zu dem Intervall } I'(v', \mathcal{E}') \text{ auf Kante } e', \text{ das } x \text{ enthält}\}$  und füge dies in die Intervallliste von  $e'$  ein; analog mit  $e''$ .

Im ersten Schritt werden also zu jeder Fläche  $f$  die Intervalle  $I(v, \mathcal{E})$  aller drei Kanten von  $f$  berechnet. Dazu benutzen wir eine Methode, die als Continuous Dijkstra bekannt ist. Allgemein wird bei dieser Methode wie in Algorithmus ?? (Dijkstra) auf Seite ?? eine Welle mit einem Ausläufer verwaltet, die sich vom Startpunkt ausbreitet. Erreicht die Welle einen Punkt  $x$  zum ersten Mal, wird  $x$  beschriftet, und die Welle breitet sich von



$x$  weiter aus und propagiert dabei den Wert von  $x$ . In unserem Fall genügen eine endliche Zahl solcher Wellenausbreitungen. Die Welle enthält dabei alle Intervalle mit bekannten Kürzesten, der Ausläufer die Intervalle mit vorläufigen Kürzesten. Propagiert wird das "nächstgelegene" Intervall auf die anderen Dreiecksseiten, dort konkurriert es mit den vorhandenen vorläufigen Intervallen, siehe Algorithmus ??,  $I'$  ist dabei wirklich nur *ein* Intervall!

Für jede Kante  $e$  werden die Intervalle  $I$  in einem balancierten Blattsuchbaum gespeichert. Dies sind maximal  $O(n)$  Intervalle je Kante. Um ein neues Intervall  $I'$  der Kante  $e'$  hinzuzufügen, benötigen wir zunächst  $O(\log n)$  Zeit um ein Intervall  $J$  zu finden, für das es Punkte in  $J$  gibt, so dass der kürzeste Weg über  $I(v, \mathcal{E})$  kürzer ist als der bisherige Weg zu diesen Punkten. Von diesem Intervall aus suchen wir alle Nachbarintervalle, die sich ändern bzw. verschwinden. Das können  $O(k)$  viele sein. Weil wir einen Blattsuchbaum verwenden benötigen wir dafür  $O(k)$  Schritte und für das Entfernen jeweils  $O(\log n)$  Zeit. Am Ende wird ein neues Intervall eingefügt.



Der Dijkstra-Algorithmus wählt stets das Intervall mit dem momentan kürzesten Weg zu einem Intervall aus, und dieses Intervall wird nie mehr verschwinden. Pro Schritt werden nun maximal 2 neue Intervalle eingefügt. Insgesamt kann es nur  $O(n^2)$  Intervalle geben, also kann der Dijkstra insgesamt nicht mehr als  $O(n^2)$  Intervalle erzeugen. Das Einfügen und Entfernen von Intervallen wie oben beschrieben benötigt also insgesamt  $O(n^2 \log n)$  Zeit. Das Ablaufen der Intervalllisten der Blattsuchbäume benötigt nicht mehr als  $O(n^2)$  Zeit. Alle Blattsuchbäume zusammen haben einen Speicherplatzbedarf von  $O(n^2)$ .

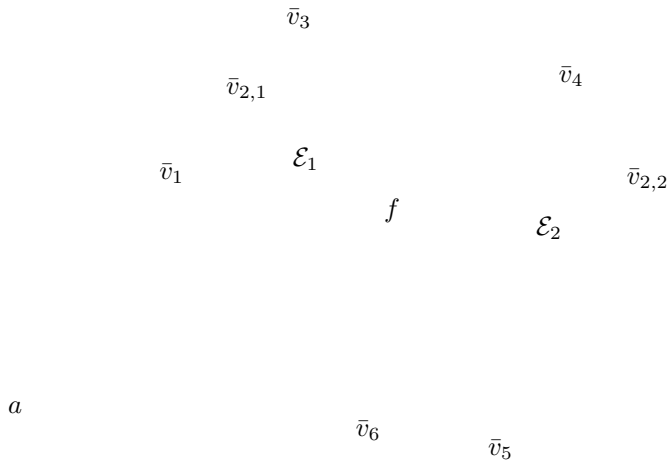


Abbildung 1.58: Dreiecksfläche  $f$  mit Intervallen  $I(v, \mathcal{E})$ .

Im zweiten Schritt werden die durch diese Intervalle vorhandenen Weginformationen in das Innere der Dreiecke fortgesetzt. Die Intervalle auf den drei Kanten von  $f$  — genauer: ihre “Quellen”  $\bar{v}_i$  — teilen sich das Innere von  $f$  nach einer gewichteten Distanzfunktion auf: Setze  $a_i := d(a, v_i)$ . Ein Punkt  $x$  auf der Dreiecksfläche wird dem Punkt  $\bar{v}_i$  zugeordnet, über den der Weg von  $a$  nach  $x$  minimal wird, für den also gilt:

$$a_i + |\bar{v}_{i,k} - x| < \min_{\substack{j \neq i \\ \forall \ell \neq k}} \left\{ a_j + |\bar{v}_{j,\ell} - x| \right\}.$$

Die  $\bar{v}_{i,k}$  teilen also die Fläche  $f$  untereinander auf. Die Grenzen der von den  $\bar{v}_{i,k}$  definierten Regionen von  $f$ , die Bisektoren, sind Hyperbelbögen, da die Punkte auf dem Bisektor zwischen  $\bar{v}_{i,k}$  und  $\bar{v}_{j,\ell}$  gleiche Weglängen haben müssen; durch die Gleichung  $a_i + |\bar{v}_{i,k} - x| = a_j + |\bar{v}_{j,\ell} - x|$  wird eine Hyperbel beschrieben.

Diese Unterteilung der Fläche  $f$  entspricht einem Voronoi-Diagramm der Ecken  $\bar{v}_{i,k}$  mit additiven Gewichten  $a_i$ , siehe Anhang. Dies kann für  $\ell$  Intervalle auf den Kanten von  $f$  in Zeit  $O(\ell \log \ell)$  und Speicherplatz  $O(\ell)$

konstruiert werden. Doch wieviele nicht-leere Intervalle kann es auf einer Kante geben?

**Lemma 1.44** *Die Anzahl  $\ell$  der nicht-leeren Intervalle auf einer Kante  $e$  ist in  $O(n)$ .*

**Beweis.**

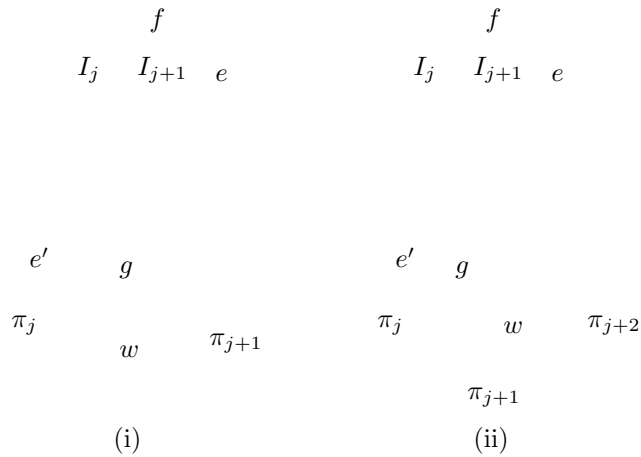


Abbildung 1.59: Ein Dreieck  $g$  mit Ecke  $w$  fungiert höchstens zweimal als Trenner für zwei Intervalle.

Seien  $I_j = I_j(v_j, \mathcal{E}_j)$  und  $I_{j+1} = I_{j+1}(v_{j+1}, \mathcal{E}_{j+1})$  zwei benachbarte Intervalle auf  $e$ . Wähle je einen inneren Punkt pro Intervall und verfolge die definierenden Kürzesten  $\pi_j$  und  $\pi_{j+1}$ , die zu  $I_j$  und  $I_{j+1}$  gehören, rückwärts.  $\pi_j$  und  $\pi_{j+1}$  können zunächst dieselben Kanten im Inneren kreuzen, d. h.  $\mathcal{E}_j$  und  $\mathcal{E}_{j+1}$  können gemeinsame Endstücke haben, müssen sich jedoch wegen  $(v_j, \mathcal{E}_j) \neq (v_{j+1}, \mathcal{E}_{j+1})$  einmal trennen. Betrachte das Dreieck  $g$  und die Ecke  $w$ , an der sich die Trennung vollzieht: das Paar  $(g, w)$  fungiert höchstens zweimal als Trenner, siehe Abbildung ??(ii), und jedem Paar läßt sich eindeutig eine Kante  $e'$  zuordnen, die  $g$  berandet und im Uhrzeigersinn  $w$  verläßt. Wir haben also eine Abbildung  $(I_j, I_{j+1}) \mapsto (g, w)$ , die “fast injektiv” ist — jedes  $(g, w)$  hat höchstens 2 Urbilder — und eine injektive Abbildung  $(g, w) \mapsto e'$ . Also gilt  $\ell \leq 2n \in O(n)$ .  $\square$

Der Beweis zu Lemma ?? benutzt die häufig angewandte Technik, die Kosten für eine Berechnung den geometrischen Objekten, hier den Kanten, zuzuordnen.

---

**Algorithmus 1.13** Berechnung kürzester Wege auf Polyedern
 

---

**Input:** Polyeder  $P$  mit  $n$  Ecken, Oberfläche sei trianguliert, Startecke  $a$ .

**Output:** Für beliebigen Anfragepunkt  $b \in \partial P$ : Entfernung und kürzesten Weg von  $a$  nach  $b$ .

**Vorbereitungsphase:**

- Berechne für jede Dreieckskante die Intervalle  $I(v, \mathcal{E})$  mit Continuous Dijkstra.  $O(n^2 \log n)$
- Berechne für jede Dreiecksfläche das durch die Intervalle  $I(v, \mathcal{E})$  der drei Kanten bestimmte Voronoi-Diagramm mit additiven Gewichten (siehe Anhang ??).  $O(n^2 \log n)$
- Bereite die Voronoi-Diagramme auf den Flächen für Lokalisierung von Punkten vor (z. B.: Trapezzerlegung, [?]).  $O(n^2)$ , Platz  $O(n^2)$

**Query** für Zielpunkt  $b$  in Dreieck  $f$ :

- Bestimme die Voronoi-Region in  $f$ , die den Punkt  $b$  enthält  
 $\implies$  Abstand zu  $a$  ist bekannt.  $O(\log n)$
- Setze die Kürzeste von  $a$  nach  $b$  zusammen aus:
  - einer Kürzesten von  $a$  nach  $v$  (Nebenprodukt des Continuous Dijkstra: in  $v$  verankerte, verkettete Liste)
  - dem "geraden" Weg durch die Kantenfolge  $e$  zum Punkt  $x \in I$
  - dem Liniensegment  $\overline{xb}$ .

Bei  $k$  Segmenten auf dem kürzesten Weg:

$O(k)$

---

Zusammenfassend erhalten wir Algorithmus ?? und folgendes Theorem:

**Theorem 1.45** (Mitchell, Mount, Papadimitriou, 1986)

Sei  $s \in P$  fest, o. E. eine Ecke von  $P$ , gegeben  $b \in P$ . Die Entfernung bzw. die Kürzeste von  $a$  nach  $b$  in  $P$  läßt sich nach Vorbereitungszeit  $O(n^2 \log n)$  in Zeit  $O(\log n + k)$  und Platz  $O(n^2)$  berechnen; dabei ist  $k$  die Anzahl der Segmente auf dem kürzesten Pfad. [?]

Ein praxistauglicher Ansatz zur  $(1 + \varepsilon)$ -Approximation kürzester Wege auf Polyedern stammt von J. Sack.<sup>21</sup> Hierbei werden in die Kanten der Dreiecke geeignet viele und adäquat verteilte Stützpunkte eingefügt und daraus ein vollständiger Graph gebildet, auf dem der Algorithmus von Dijkstra angewendet wird. Dies liefert gute Approximationsergebnisse und erlaubt die Einhaltung von Nebenbedingungen wie maximale Steigung, maximaler Böschungswinkel, z. B. bei der Berechnung von Wegen für Fahrzeuge. [?]

---

<sup>21</sup><http://www.scs.carleton.ca/~sack/>

## 1.5 Kürzeste Pfade für Liniensegmente

Bisher hatten wir nur kürzeste Pfade für Punkte betrachtet. Zum Abschluß des ersten Kapitels wollen wir kurz auf die Probleme eingehen, die bei der Berechnung kürzester Pfade für ausgedehnte Objekte entstehen. Schon für Liniensegmente im  $\mathbb{R}^2$  ohne Hindernisse ist das Problem nicht trivial. Die Bewegung von Liniensegmenten im  $\mathbb{R}^2$  kann man als Modell für einen Balken oder eine Leiter ansehen, die von zwei Trägern an je einem Ende getragen wird.



Abbildung 1.60: Bewegung eines Liniensegmentes.

Bei ausgedehnten Objekten ist zunächst einmal festzulegen, was eine Kürzeste ist bzw. wie sich die Kosten für eine Bewegung berechnen. In der Anschauung mit dem Balken sind sicherlich die Wege interessant, die die beiden Träger zurücklegen.<sup>22</sup> Eine Möglichkeit ist also, die Längen der von den Endpunkten zurückgelegten Bahnen zu summieren und diese Summe zu minimieren. Für Liniensegmente gilt damit für eine Bewegung  $B$ :

$$\text{Kosten}(B) := \text{Länge}(\pi_1) + \text{Länge}(\pi_2).$$

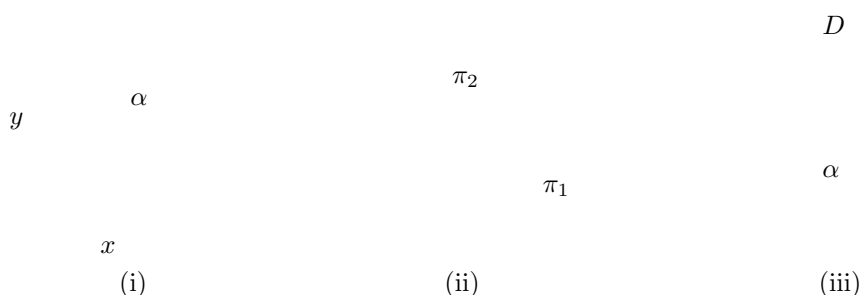


Abbildung 1.61: (i) Freiheitsgrade eines Liniensegmentes, (ii) geradlinige Bewegung (iii) Rotation.

Ein Liniensegment hat drei Freiheitsgrade, siehe Abbildung ??(i): die Koordinaten seines Startpunktes und sein Drehwinkel; seine Position läßt

<sup>22</sup>Je schwerer der Balken ist, umso größer dürfte das Interesse der beiden Träger an kürzesten Wegen sein.

sich also durch das Tripel  $A = (x, y, \alpha) \in \mathcal{C}$  mit  $\mathcal{C} = \mathbb{R} \times \mathbb{R} \times [0, 2\pi)$  beschreiben.

Unser Problem läßt sich damit folgendermaßen beschreiben: gegeben zwei Positionen  $A, A' \in \mathcal{C}$ , bestimme eine Bewegung von  $A$  nach  $A'$  mit minimalen Kosten.

Doch welche Bewegungen sind minimal? Die geradlinige Bewegung in Abbildung ??(ii) ist sicherlich optimal, da  $\pi_1$  und  $\pi_2$  Geraden sind und nicht kürzer sein können. Auch die Rotation in Abbildung ??(iii) ist optimal, zum Beweis dieser vermeintlich einfachen Tatsache brauchen wir das folgende Theorem:

**Theorem 1.46** (*Surface–Area Formel von Cauchy*)

Sei  $C$  eine geschlossene konvexe Kurve in der Ebene. Mit

$$h_C(\alpha) := \sup \{ x \cos \alpha + y \sin \alpha \mid (x, y) \in C \},$$

$$\text{dia}_C(\alpha) := h_C(\alpha) + h_C(\alpha + \pi), \quad \alpha \in [0, \pi)$$

gilt für die Länge von  $C$ :

$$\text{Länge}(C) = \int_0^\pi \text{dia}_C(\alpha) d\alpha.$$

Dabei ist  $\text{dia}_C(\alpha)$  anschaulich die Messung von  $C$  mit einer um den Winkel  $\alpha$  zur X–Achse gedrehten Schieblehre, siehe Abbildung ??(i).

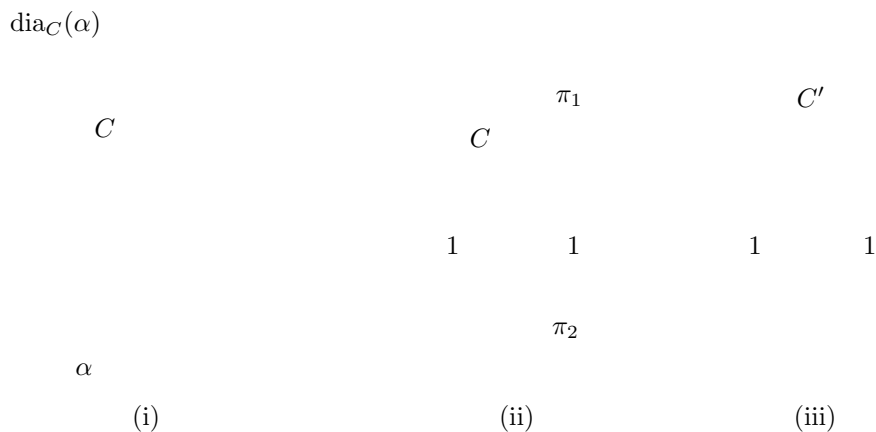


Abbildung 1.62: (i)  $\text{dia}_C(\alpha)$ , (ii) Bewegung  $B$ , (iii) konvexe Hülle von  $C$ .

Nehmen wir nun an, eine beliebige andere Bewegung  $B$  mit Pfaden  $\pi_1$  und  $\pi_2$  sei optimal, vgl. Abbildung ??(ii). Sei  $C$  die Verknüpfung der beiden Kurven  $\pi_1$  und  $\pi_2$  und der beiden Liniensegmente (o. E. sei die Länge der

Liniensegmente 1).  $C'$  sei die konvexe Hülle von  $C$ , siehe Abbildung ??(iii).  $D$  sei die Kurve der Rotation (Abbildung ??(iii)). Dann gilt:

$$\begin{aligned}
 2 + \text{Kosten}(B) &= \text{Länge}(C) \\
 &\geq \text{Länge}(C') \\
 &= \int_0^\pi \text{dia}_{C'}(\alpha) d\alpha \quad \text{nach Theorem ??} \\
 &\geq \int_0^\pi \text{dia}_D(\alpha) d\alpha \quad \text{wegen } \text{dia}_{C'}(\alpha) \geq \text{dia}_D(\alpha) (*) \\
 &= \text{Länge}(D) \\
 &= 2 + \text{Kosten}(R)
 \end{aligned}$$

Begründung (\*): Im *kritischen* Winkelbereich für  $\alpha$  gilt  $\text{dia}_D(\alpha) = 1$ . Da auch bei beliebiger Bewegung  $C$  der Winkel  $\alpha$  (oder  $-\alpha$ ) angenommen werden muss (irgendwo muss das Segment ja liegen), gilt für alle Winkel in diesem Bereich  $\text{dia}_{C'}(\alpha) \geq 1$ , siehe Abbildung ??.

Abbildung 1.63:  $\text{dia}_D$  im *kritischen* Winkelbereich.

Also ist eine Rotation um einen Winkel  $\leq \pi$  optimal.

Für sich betrachtet sind also Rotation um einen Punkt auf dem Liniensegment und Translation optimal, doch wie verhält es sich mit Kombinationen davon? Es gilt:

**Theorem 1.47** (*Icking, Rote, Welzl, Yap, 1989*)

Zwischen je zwei Positionen von Liniensegmenten gibt es eine optimale Bewegung von einem der folgenden Typen:

- (i) maximal drei Rotationen,
- (ii) maximal zwei Rotationen und eine geradlinige Bewegung,
- (iii) eine Rotation zwischen zwei geradlinigen Bewegungen.

Die Teilbewegungen lassen sich effektiv berechnen.

[?]

Das Problem der kürzesten Pfade von Liniensegmenten wurde von Icking et. al. wiederentdeckt. Erwähnt wurde es bereits von Dubovitskij [?] und Gurevich [?].

Fassen wir abschließend die Ergebnisse aus Kapitel 1 zusammen:

- Kürzeste Pfade für Punkte im  $\mathbb{R}^2$  lassen sich effizient berechnen.
- Kürzeste Pfade für Punkte im  $\mathbb{R}^d$ ,  $d > 2$  zu finden ist NP-hard.
- Kürzeste Pfade für ausgedehnte Objekte zu finden ist schwierig.

Deshalb werden wir uns im folgenden Kapitel auf Kollisionsfreiheit konzentrieren.



## Kapitel 2

# Kollisionsfreie Bahnen für polygonale Roboter

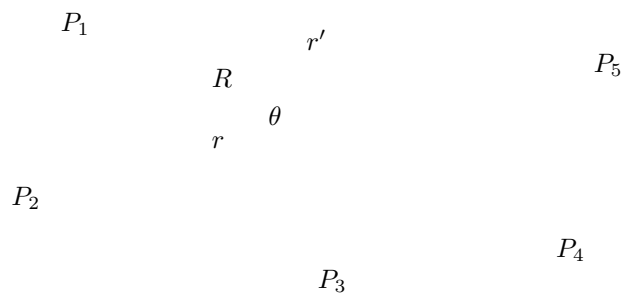


Abbildung 2.1: Roboter  $R$  in einem Arbeitsraum.

Wir betrachten nun nicht länger punktförmige Roboter, sondern ausgedehnte Objekte. Wir modellieren den Roboter  $R$  durch ein abgeschlossenes Polygon — konvex oder allgemein — mit  $m$  Kanten. Der Arbeitsraum des Roboters sei eine beschränkte Umgebung mit  $h$  polygonalen Hindernissen  $P_i$  mit insgesamt  $n$  Ecken, siehe Abbildung ??.

Zur Diskussion kollisionsfreier Bahnen wollen wir einige Begriffe definieren:

### Definition 2.1

- (i) Eine **Plazierung** oder **Konfiguration** ist eine Position des Roboters im Arbeitsraum. Zur eindeutigen Beschreibung der Position legen wir

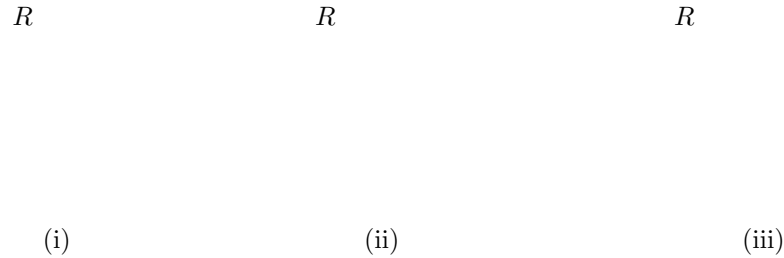


Abbildung 2.2: (i) Verbotene Konfiguration, (ii) freie Platzierung (erlaubt), (iii) halbfreie Platzierung (erlaubt).

einen Referenzpunkt  $r = (x, y)$  fest, sowie einen zweiten Punkt  $r'$  als Winkelreferenz. Die Platzierung läßt sich dann durch die Angabe des Referenzpunktes  $r$  und des Winkels  $\theta$  zwischen der X-Achse und dem Vektor  $\overrightarrow{rr'}$  als  $R(x, y, \theta)$  mit  $(x, y, \theta) \in \mathbb{R} \times \mathbb{R} \times [0, 2\pi)$  angeben. Der Roboter hat damit drei Freiheitsgrade (Degree of Freedom, DOF).

- (ii) Der **Konfigurationsraum** ist der Raum der möglichen Platzierungen des Roboters, hier  $\mathcal{C} = \mathbb{R} \times \mathbb{R} \times [0, 2\pi)$ .
- (iii) Die **verbotenen Konfigurationen** sind solche, bei denen der Roboter in ein Hindernis eindringt. Das Berühren einer Hinderniswand dagegen ist erlaubt (**halbfreie Platzierung**), vgl. Abbildung ??.<sup>1</sup>

$$\mathcal{C}_{\text{verb}} := \left\{ c \in \mathcal{C} \mid R(c) \cap \bigcup \overset{\circ}{P}_i \neq \emptyset \right\}$$

- (iv) Der **Raum der freien Platzierungen** enthält alle erlaubten Konfigurationen

$$\mathcal{C}_{\text{frei}} := \mathcal{C} \setminus \mathcal{C}_{\text{verb}}$$

Damit haben wir folgendes Bahnplanungsproblem: Gegeben  $s, t \in \mathcal{C}_{\text{frei}}$ . Gibt es einen Weg von  $s$  nach  $t$ , der ganz in  $\mathcal{C}_{\text{frei}}$  verläuft? Wenn ja, bestimme einen.

Die Frage nach der Existenz eines Weges ist äquivalent zu der Frage, ob  $s$  und  $t$  in derselben Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}$  liegen. Zur Erinnerung:

**Definition 2.2** Eine Menge  $Z$  (Teilmenge eines topologischen Raumes  $T$ ) heißt genau dann **weg-zusammenhängend**, wenn

$$\forall a, b \in Z : \exists \text{ Pfad } \pi \text{ von } a \text{ nach } b \text{ in } Z.$$

Der Pfad  $\pi$  ist dabei gegeben durch eine stetige Abbildung  $\pi : [0, 1] \rightarrow Z$  mit  $\pi(0) = a$  und  $\pi(1) = b$ .

<sup>1</sup>Zur Erinnerung:  $\overset{\circ}{P}$  bezeichnet das Innere des Polygons.

**Lemma 2.3** Sei  $A \subset T$  beliebig, dann wird für alle  $a, b \in A$  durch

$$a \sim b :\Leftrightarrow \exists \text{ Pfad } \pi \subset A \text{ von } a \text{ nach } b \text{ in } A \text{ mit } \pi(0) = a \text{ und } \pi(1) = b$$

eine Äquivalenzrelation definiert. Die Äquivalenzklassen  $Z_i$  von  $A$  bzgl.  $\sim$  heißen **Zusammenhangskomponenten** von  $A$ . Sie sind die größten zusammenhängenden Teilmengen und überdecken  $A$ :  $A = \dot{\bigcup} Z_i$ .<sup>2</sup>

Natürlich kann  $\mathcal{C}_{\text{frei}}$  aus mehreren Zusammenhangskomponenten bestehen. In nebenstehender Abbildung kann  $R$  sich in einem der beiden Räume bewegen,  $\mathcal{C}_{\text{frei}}$  hat zwei Zusammenhangskomponenten.  $R$  Das Bahnplanungsproblem ist unlösbar, wenn  $s \in Z_i$  und  $t \in Z_j$  mit  $i \neq j$ .

Somit stellen sich uns folgende algorithmische Probleme:

- Bestimme die Zusammenhangskomponenten von  $\mathcal{C}_{\text{frei}}$  bzw. die Komponente  $Z_s$ , die  $s$  enthält.
- Bestimme, ob  $t$  in  $Z_s$  liegt. Falls ja, berechne einen Weg  $\pi \subseteq Z_s$  mit  $\pi(0) = a$  und  $\pi(1) = b$ .

---

<sup>2</sup> $A \dot{\cup} B$  bezeichne die **disjunkte Vereinigung**  $A \dot{\cup} B = A \cup B$  mit  $A \cap B = \emptyset$ .

## 2.1 Reine Translationsbewegungen für konvexe Roboter

Beschränken wir uns zunächst auf den Spezialfall eines *konvexen* Roboters mit  $m$  Kanten, der sich in einer Umgebung aus  $h$  Polygonen  $P_i$  mit insgesamt  $n$  Kanten und  $\forall i \neq j : \overset{\circ}{P}_i \cap \overset{\circ}{P}_j = \emptyset$  bewegt. Außerdem beschränken wir uns auf reine Translationsbewegungen, der Roboter darf also nur verschoben, aber nicht gedreht werden. Die Ausrichtung des Roboters ist stets gleich, damit hat der Roboter hier nur zwei Freiheitsgrade, und seine Plazierung wird durch  $R(x, y) \in \mathcal{C} = \mathbb{R} \times \mathbb{R}$  eindeutig beschrieben.

### 2.1.1 Voronoi–Diagramme zur Planung von Translationsbewegungen mit Sicherheitsabstand

Sei zunächst der Roboter  $R$  durch einen Kreis gegeben. Kreisförmige Roboter haben die Vorteile, daß man keine Rotationsbewegungen zu betrachten hat, wenn als Referenzpunkt der Kreismittelpunkt gewählt wird, und sie außerdem als Approximation für beliebige Roboter dienen können, siehe nebenstehende Abbildung. Dabei werden allerdings nicht alle Wege gefunden; man kann sich leicht Umgebungen vorstellen, in denen der Roboter  $R$  eine Lücke zwischen zwei Hindernissen passieren kann, während der kreisförmige Roboter zu groß dafür ist.

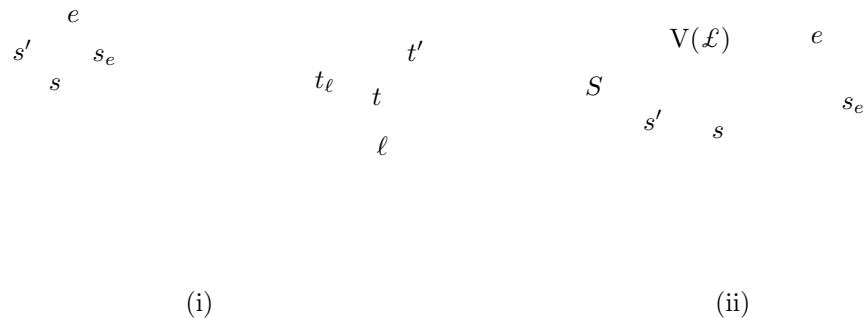


Abbildung 2.3: (i) Bahnplanung im Voronoi–Diagramm für Liniensegmente (ii) Konstruktion von  $s'$ .

Das Problem läßt sich lösen, indem wir, wie in Abschnitt ?? beschrieben, das Voronoi–Diagramm von Liniensegmenten bzgl. der  $L_2$ –Metrik mit den Kanten der Hindernisse als Liniensegmente konstruieren, siehe Abbildung ?. Wie in Algorithmus ?? beschrieben reduziert sich damit die Suche nach einem kollisionsfreien Weg von  $s$  nach  $t$  auf die Suche in dem planaren Graphen  $V(\mathcal{L})$  von  $s'$  nach  $t'$ . Wegen  $s, t \in \mathcal{C}_{\text{frei}}$  existiert eine Bewegung von  $s$  nach  $s'$  und von  $t'$  nach  $t$ .

---

**Algorithmus 2.1** Bahnplanung mit Sicherheitsabstand für kreisförmige Roboter

---

**Vorbereitung:**

- Konstruiere das Voronoi–Diagramm  $V(\mathcal{L})$  mit den Kanten der Hindernisse als Liniensegmente.  $O(n \log n)$
- Bereite das Voronoi–Diagramm für Point-Location vor.  
mit Trapezzerlegung  $O(n \log n)$

**Query** für zwei Punkte  $s, t$ :

- Lokalisier  $s, t$  im Voronoi–Diagramm.  $O(\log n)$
  - Sei  $s \in \text{VR}(e, \mathcal{L})$ ,  $t \in \text{VR}(\ell, \mathcal{L})$ :
  - Bestimme den zu  $s$  nächstgelegenen Punkt  $s_e$  auf  $e$ .
  - Bestimme Strahl  $S$  von  $s_e$  durch  $s$ .
  - $s'$  sei Schnittpunkt von  $S$  mit  $V(\mathcal{L})$ ; analog  $t'$ .
  - Berechne mit Breitensuche in  $V(\mathcal{L})$  einen Pfad von  $s'$  nach  $t'$  mit zulässigem Mindestabstand oder berichte, daß es keinen solchen gibt.  $O(n)$
- 

Mit dieser Methode erhält man Pfade mit maximalem Sicherheitsabstand. Außerdem läßt sie sich auf die Translation konvexer Roboter verallgemeinern, indem man mit dem Spiegelbild von  $R$  eine konvexe Distanzfunktion definiert und das Voronoi–Diagramm von Liniensegmenten bzgl. dieser konvexen Distanzfunktion berechnet. Nach einer Vorbereitungszeit von  $O(mn \log n)$  ermöglicht dies, eine Anfrage in Zeit  $O(mn)$  zu beantworten, siehe [?]. Auf nicht–konvexe Roboter, Roboter mit mehreren Freiheitsgraden oder höhere Dimensionen läßt sich diese Methode nicht effektiv verallgemeinern. Das Prinzip der Reduktion auf einen eindimensionalen Graphen läßt sich jedoch verallgemeinern, wir werden später darauf zurückkommen. Der Ansatz, in einer erzeugten Karte einen Pfad zu suchen, ist als Roadmap–Approach bekannt.

### 2.1.2 Berechnung von $\mathcal{C}_{\text{frei}}$

Ein anderer Ansatz zur Berechnung reiner Translationsbewegungen konvexer Roboter ist, den Raum der freien Konfigurationen  $\mathcal{C}_{\text{frei}}$  explizit zu berechnen. Nehmen wir zunächst an, die Hindernisse seien konvex. Betrachten wir ein Hindernis  $P_i$  und untersuchen, welche Positionen  $R(x, y)$  aufgrund der  $P_i$  verboten sind. Von Lozano–Pérez stammt die Idee, die Hindernisse so zu expandieren, daß in der entstehenden Szene das Bahnplanungsproblem für einen Punkt zu lösen ist [?].

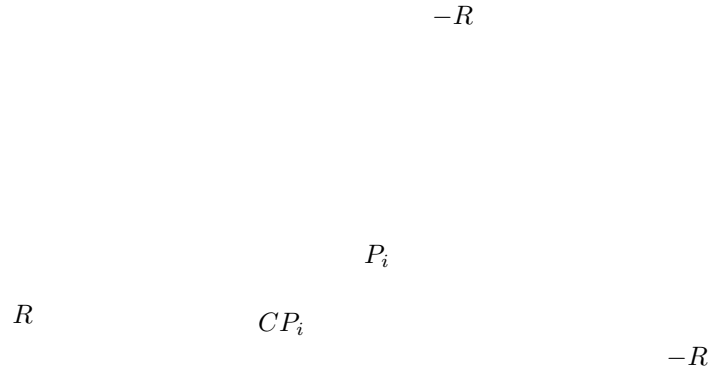


Abbildung 2.4: Hindernis, Roboter und Konfigurationsraum-Hindernis.

Das zu  $P_i$  gehörende expandierte Hindernis  $CP_i$  (Konfigurationsraum-Hindernis), ist definiert als

$$CP_i := \{c \in \mathcal{C} \mid R(c) \cap P_i \neq \emptyset\}.$$

Anschaulich ist der Rand von  $CP_i$  die Bahn des Referenzpunktes die entsteht, wenn der Roboter  $R$  um  $P_i$  wandert, vgl. Abbildung ??.<sup>3</sup> Abbildung ?? zeigt die Konfigurationsraum-Hindernisse zu Abbildung ???. Man beachte, daß  $\mathcal{C}_{\text{frei}}$  hier aus zwei Komponenten besteht! Anstatt den Rand von  $R$  um den Rand von  $P_i$  herumschieben, kann man auch den Referenzpunkt des am Ursprung gespiegelten Roboters  $-R = -R(0,0)$  um den Rand von  $P_i$  schieben. Das Konfigurationsraum-Hindernis zu  $P_i$  ergibt sich dann aus der Vereinigung von  $P_i$  und allen Positionen von  $-R$ , bei denen der Referenzpunkt auf dem Rand von  $P_i$  liegt. Dies entspricht genau der Minkowski-Summe von  $P_i$  und  $-R$ .

**Definition 2.4** Die **Minkowski-Summe** zweier Mengen  $A, B \subset \mathbb{R}^2$  ist

$$A \oplus B := \{a + b \mid a \in A, b \in B\}.$$

**Bemerkung 2.5** *Minkowski-Summen sind*

- (i) *kommutativ*:  $A \oplus B = B \oplus A$ ,
- (ii) *assoziativ*:  $A \oplus (B \oplus C) = (A \oplus B) \oplus C$ ,

<sup>3</sup>In Anhang B findet sich eine Seite mit Robotern zum Ausschneiden und Nachvollziehen der in diesem Kapitel vorgestellten Beispiele.

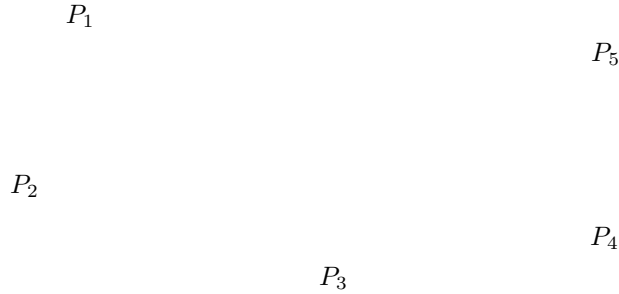


Abbildung 2.5: Konfigurationsraum-Hindernisse zu Abbildung ??.

(iii) distributiv bzgl. der Vereinigung:  $A \oplus (B \cup C) = (A \oplus B) \cup (A \oplus C)$ .

**Lemma 2.6** Das zu  $P_i$  gehörende Konfigurationsraum-Hindernis bzgl.  $R$  ist

$$CP_i = P_i \oplus (-R(0, 0)).$$

**Beweis.**

Zeige:  $\forall x, y : R(x, y) \cap P_i \neq \emptyset \iff (x, y) \in P_i \oplus (-R(0, 0))$ .

“ $\implies$ ” Sei  $q = (q_x, q_y) \in R(x, y) \cap P_i$

$$\begin{aligned} \Rightarrow (q_x, q_y) &\in R(x, y) \\ \Rightarrow (q_x - x, q_y - y) &\in R(0, 0) \\ \Rightarrow (-q_x + x, -q_y + y) &\in -R(0, 0). \end{aligned}$$

Außerdem gilt  $(q_x, q_y) \in P_i$ , insgesamt also

$$(x, y) = (-q_x + x, -q_y + y) + (q_x, q_y) \in -R(0, 0) \oplus P_i.$$

“ $\impliedby$ ” Sei  $(x, y) \in P_i \oplus -R(0, 0)$

$$\begin{aligned} \Rightarrow \exists (q_x, q_y) \in P_i, (r_x, r_y) \in R(0, 0) : \\ (x, y) &= (q_x, q_y) - (r_x, r_y) = (q_x - r_x, q_y - r_y) \\ \Rightarrow (q_x, q_y) &= (x, y) + \underbrace{(r_x, r_y)}_{\in R(0, 0)} \in R(x, y) \cap P_i. \quad \square \end{aligned}$$

Minkowski-Summen treten also bei der Bewegungsplanung in natürlicher Weise auf. Betrachten wir deshalb einige Eigenschaften von Minkowski-Summen.

**Definition 2.7** Gegeben sei ein Richtungsvektor  $d$  mit  $|d| = 1$ . Ein Punkt  $p \in P$  ist **Extremalpunkt** von  $P$  in Richtung  $d$ , wenn folgendes gilt: Die zu  $d$  orthogonale Gerade  $g$  durch  $p$  teilt die Ebene in zwei Halbebenen, von denen die eine zusammen mit  $g$  ganz  $P$  enthält und die Andere den Punkt  $p + d$ .

**Bemerkung 2.8** *Extremalpunkte der Minkowski-Summe  $P \oplus R$  sind Summen von Extremalpunkten von  $P$  und  $R$ .*

$$\begin{array}{ccc} P \oplus R & & p + r \\ & r & \\ & R & \end{array}$$

$$\begin{array}{ccc} P & & p \end{array}$$

**Lemma 2.9** *Seien  $P$  und  $R$  konvexe Polygone mit  $n$  bzw.  $m$  Kanten. Dann ist  $P \oplus R$  konvex, hat maximal  $n + m$  Kanten und kann in Zeit  $O(n + m)$  berechnet werden.*

**Beweis.**

Die Konvexität folgt aus der Definition der Minkowski-Summe.

Rotiere simultan zwei parallele Tangenten  $T_P, T_R$  um  $P$  und

$R$ . Wenn  $T_P$  eine Ecke  $p$  von  $P$  berührt und gleichzeitig  $T_R$  eine Ecke  $r$  von  $R$ , dann ist

$p + r$  eine Ecke von  $P \oplus R$ .

Alle Ecken von  $P \oplus R$  entstehen

auf diese Weise. Da alle Ecken von  $P$  und  $R$  genau einmal betrachtet werden, braucht die Rotation Zeit  $O(n + m)$ .

□

**Definition 2.10** Zwei durch Jordankurven<sup>4</sup> berandete Mengen  $A, B \subseteq \mathbb{R}^2$  heißen ein Paar **Pseudokreise** (pseudo disks), wenn ihre Ränder entweder höchstens zwei echte Kreuzungen oder höchstens einen Berührungspunkt aufweisen.

Abbildung ??(i) und (ii) veranschaulicht die Definition. Beispiele für Pseudokreise sind Kreise, Quadrate und achsenparallele Rechtecke mit festem Seitenverhältnis (Abbildung ??(i)–(iii)). Beliebige Rechtecke sind im allgemeinen keine Pseudokreise (Abbildung ??(iv)). Folgendes Lemma gibt eine äquivalente Charakterisierung von Pseudokreisen; die Konvexität ist dabei eine notwendige Forderung, wie Abbildung ??(iii) zeigt.

<sup>4</sup>Eine geschlossene Kurve, die die Ebene in zwei Gebiete unterteilt.



(i) (ii) (iii)

Abbildung 2.6: (i) Pseudokreise, (ii) keine Pseudokreise, (iii) die Konvexität ist eine notwendige Forderung.

**Lemma 2.11** *Seien  $A, B$  konvex.  $A, B$  sind genau dann Pseudokreise, wenn  $A \setminus B$  und  $B \setminus A$  zusammenhängend sind.*

**Beweis.** Übungsaufgabe. □

(i) (ii) (iii) (iv)

Abbildung 2.7: (i)–(iii) Beispiele für Pseudokreise, (iv) Rechtecke sind im allgemeinen keine Pseudokreise.

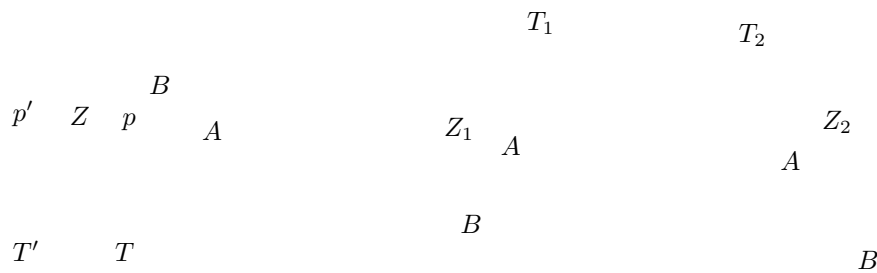


Abbildung 2.8: (i)  $p$  muß Element von  $Z$  sein, (ii)  $A \cup B$  kann nicht aus zwei Zusammenhangskomponenten bestehen.

**Lemma 2.12** *Seien  $P_1, P_2, R$  konvexe Polygone mit  $\overset{\circ}{P}_1 \cap \overset{\circ}{P}_2 = \emptyset$ . Dann sind  $A = P_1 \oplus R$  und  $B = P_2 \oplus R$  ein Paar konvexer Pseudokreise. [?]*

**Beweis.**  
Die Konvexität folgt aus der Definition von  $\oplus$ . Wegen der Symmetrie genügt

es zu zeigen:  $A \setminus B$  ist zusammenhängend. Jede Zusammenhangskomponente  $Z$  von  $A \setminus B$  trägt zum Rand der konvexen Hülle  $\text{ch}(A \cup B)$  bei: Falls  $Z = A$  ist, also  $A \cap B = \emptyset$ , müssen auch Punkte von  $A$  auf  $\partial \text{ch}(A \cup B)$  liegen; sonst wäre  $A \subset \text{ch}(A \cup B) = \text{ch}(B) = B$ . Ist  $Z$  eine *echte* Teilmenge von  $A$ , so muß  $\partial Z$  einen Punkt  $p \in \partial B$  enthalten. Lege die Tangente  $T$  durch  $p$  an  $B$ , dann liegt  $B$  vollständig auf einer Seite von  $T$ ; siehe Abbildung ??(i). Verschiebe jetzt  $T$  parallel in die  $B$  entgegengesetzte Richtung, bis der letzte Randpunkt  $p'$  von  $A$  erreicht wird. Die Tangente  $T'$  durch  $p'$  an  $A$  ist auch Tangente an  $A \cup B$ , also gilt  $p' \in \partial A \cap \partial \text{ch}(A \cup B)$ .

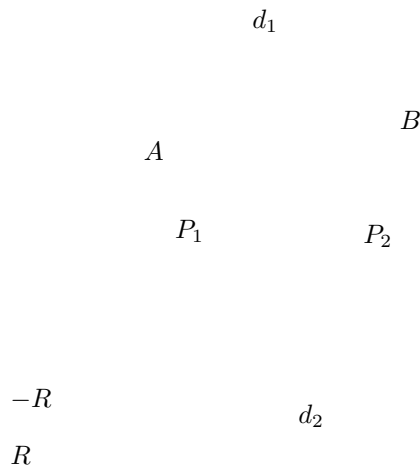


Abbildung 2.9: Um  $A \cup B$  rotierende Tangenten wechseln zweimal zwischen  $A$  und  $B$ , weil  $P_1$  und  $P_2$  disjunkt sind.

Nehmen wir nun an, es gäbe zwei Zusammenhangskomponenten  $Z_1, Z_2$  in  $A \setminus B$ .

- $\Rightarrow$  Es gibt zwei Richtungen  $\vec{d}_1, \vec{d}_2$ , in denen  $A$  extremer als  $B$  ist, d. h. die Tangente an  $\text{ch}(A \cup B)$  berührt  $A$ , siehe Abbildung ??(ii).
- $\Rightarrow$  Es gibt zwei Richtungen  $\vec{d}_1, \vec{d}_2$ , in denen  $P_1$  extremer als  $P_2$  ist.
- Bem. ??  $\Rightarrow$  Da  $P_1, P_2$  konvex und nicht leer sind, und  $\overset{\circ}{P}_1 \cap \overset{\circ}{P}_2 = \emptyset$  gilt, ist  $P_1$  entweder in allen Richtungen zwischen  $\vec{d}_1$  und  $\vec{d}_2$  oder in allen Richtungen zwischen  $\vec{d}_2$  und  $\vec{d}_1$  extremer als  $P_2$ , siehe Abbildung ??.
- $\Rightarrow$   $A$  ist entweder in allen Richtungen zwischen  $\vec{d}_1$  und  $\vec{d}_2$  oder in allen Richtungen zwischen  $\vec{d}_2$  und  $\vec{d}_1$  extremer als  $B$ .
- Bem. ??  $\Rightarrow$  Eine um  $A \cup B$  rotierende Tangente berührt erst  $A$ , dann  $A \cup B, B, A \cup B$  und schließlich wieder  $A$ , siehe Abbildung ??.

⇒ Es gibt nur eine Zusammenhangskomponente  $\zeta$ .

□

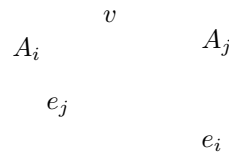
**Theorem 2.13** (Kedem, Livne, Pach, Sharir, 1986)

Sei  $A_1, \dots, A_k$  eine Familie von polygonalen Pseudokreisen mit insgesamt  $n$  Kanten. Dann hat  $\partial \cup A_i$  die Komplexität  $O(n)$ . [?]

**Beweis.**

Auf  $\partial \cup A_i$  gibt es zwei Typen von Ecken: Ecken von Pseudokreisen — davon gibt es maximal  $n$  Stück — und Schnittpunkte von Kanten von Pseudokreisen. Sei  $v$  ein Schnittpunkt zweier Pseudokreis-Kanten  $e_i$  und  $e_j$ . Folge der Kante  $e_i$  ins Innere von  $A_j$ , dann gibt es wiederum zwei Fälle:  
1. Fall:  $e_i$  endet in  $A_j$ , dann stelle  $v$  dem Endpunkt von  $e_i$  in  $A_j$  in Rechnung.

2. Fall:  $e_i$  geht durch  $A_j$  hindurch. Auf  $e_i$  liegen dann zwei Schnittpunkte von  $A_i$  und  $A_j$ . Da  $A_i$  und  $A_j$  Pseudokreise sind, kann es keinen weiteren Schnittpunkt geben. Also muß  $e_j$  im Inneren von  $A_i$  enden, und wir können  $v$  dem Endpunkt von  $e_j$  in  $A_i$  in Rechnung stellen.



Auf diese Weise wird jede Ecke  $w$  höchstens zweimal belastet, je einmal von jeder von  $w$  ausgehenden Kante, falls diese sich mit den Rändern anderer  $A_\ell$  kreuzt. □

Theorem ?? gilt auch allgemeiner, ist jedoch in dieser Form viel schwieriger zu beweisen:

**Theorem 2.14** Sei  $A_1, \dots, A_k$  eine Familie von Pseudokreisen. Dann hat  $\partial \cup A_i$  die Komplexität  $O(n)$ . [?]

Welche Komplexität hat nun  $P \oplus R$ , wenn nicht beide Polygone konvex sind?

**Lemma 2.15** Seien  $P_1, P_2$  Polygone mit  $n$  bzw.  $m$  Ecken.

- (i) Sind  $P_1, P_2$  konvex, so ist  $P_1 \oplus P_2$  konvex und hat die Komplexität  $O(m+n)$ .
  - (ii) Ist nur  $P_2$  konvex, so hat  $P_1 \oplus P_2$  die Komplexität  $O(mn)$ .
  - (iii) Ist kein  $P_i$  konvex, so hat  $P_1 \oplus P_2$  die Komplexität  $O(m^2n^2)$ .
- Alle Schranken sind scharf.

(ii)

(iii)

Abbildung 2.10: (ii)  $P_1 \oplus P_2$  hat  $\Theta(mn)$  Ecken, wenn nur eines der Polygone konvex ist, (iii)  $P_1 \oplus P_2$  hat  $\Theta(m^2n^2)$  Ecken, wenn keines der Polygone konvex ist.

**Beweis.**

- (i) Bereits gezeigt.  
(ii) Trianguliert man  $P_1$ , so entstehen  $n - 2$  Dreiecke  $D_i$  mit paarweise disjunktem Inneren und  $P_1 = \bigcup_{i=1}^{n-2} D_i$ .

$$P_1 \oplus P_2 = \left( \bigcup_{i=1}^{n-2} D_i \right) \oplus P_2 = \bigcup_{i=1}^{n-2} \underbrace{(D_i \oplus P_2)}_{O(m) \text{ Kanten}} .$$

Dies ist nach Lemma ?? eine Familie von Pseudokreisen mit  $O(mn)$  Kanten, also hat  $P_1 \oplus P_2$  die Komplexität  $O(mn)$  nach Theorem ??.

- (iii) Trianguliere  $P_1$  und  $P_2$ ,  $P_1 = \bigcup_{i=1}^{n-2} D_i$  und  $P_2 = \bigcup_{j=1}^{m-2} D'_j$ .

$$P_1 \oplus P_2 = \bigcup_{i=1}^{n-2} \bigcup_{j=1}^{m-2} \underbrace{D_i \oplus D'_j}_{O(1) \text{ Kanten}} .$$

Insgesamt erhalten wir  $O(mn)$  Kanten von denen sich jedes Paar kreuzen kann, also  $O(m^2 n^2)$  viele Schnittpunkte. Theorem ?? ist hier nicht anwendbar, da  $\bigcup D_i \oplus D'_j$ ,  $1 \leq i \leq n - 2, 1 \leq j \leq m - 2$  keine Familie von Pseudokreisen ist (Übungsaufgabe).

Abbildung ?? zeigt Beispiele, in denen die Schranken scharf sind.  $\square$

Damit können wir folgende Aussage beweisen:

**Theorem 2.16** *Sei  $R$  ein polygonaler, konvexer Roboter mit  $m$  Ecken, und seien  $P_i$  polygonale Hindernisse mit insgesamt  $n$  Ecken; dann hat  $\mathcal{C}_{\text{frei}}$  die Komplexität  $O(mn)$  und kann in Zeit  $O(mn \log^2(mn))$  berechnet werden.*

**Beweis.**

Es genügt,  $\mathcal{C}_{\text{verb}}$  zu berechnen.  $P_i$  habe  $n_i$  Ecken. Triangulieren wir die  $P_i$ , so entsteht eine Familie von Pseudokreisen, für die gilt:

$$\mathcal{C}_{\text{verb}} = \underbrace{\bigcup_i \underbrace{P_i \oplus -R(0,0)}_{\text{Komplexität } O(m \cdot n_i)}}_{\text{Komplexität } O(m \cdot \sum n_i) = O(mn)} .$$

$\mathcal{C}_{\text{verb}}$  kann mit dem Divide-and-Conquer Algorithmus ?? berechnet werden. Die Lösung der dabei entstehenden Rekursion  $T(k) \leq 2T(\frac{k}{2}) + Ck \log k$  mit  $k = mn$  liegt in  $O(k \log^2 k)$ :

$$\begin{aligned} T(k) &\leq 2T\left(\frac{k}{2}\right) + Ck \log k \\ &\leq 4T\left(\frac{k}{4}\right) + 2C\frac{k}{2} \log \frac{k}{2} + Ck \log k \end{aligned}$$

$$\begin{aligned}
&\leq \dots \\
&\leq 2^{\log k} T(1) + \sum_{i=0}^{\log k-1} Ck \log \frac{k}{2^i} \\
&\leq Ck + \sum_{i=0}^{\log k-1} Ck \log k \\
&= Ck + Ck \log^2 k \in O(k \log^2 k)
\end{aligned}$$

□

---

**Algorithmus 2.2** Berechnung von  $\mathcal{C}_{\text{verb}}$ 


---

- Unterteile die Hindernisse beliebig in gleichgroße Teilmengen  $\mathcal{P}_1$  und  $\mathcal{P}_2$ .
  - Berechne rekursiv  $\mathcal{C}_{\text{verb}}(\mathcal{P}_1)$  und  $\mathcal{C}_{\text{verb}}(\mathcal{P}_2)$  mit je  $O(mn)$  Kanten.
  - Berechne mit Sweep die Vereinigung  $\mathcal{C}_{\text{verb}}(\mathcal{P}_1) \cup \mathcal{C}_{\text{verb}}(\mathcal{P}_2)$  in Zeit  $O(mn \log(mn))$ .
- 

Damit haben wir das Bahnplanungsproblem für konvexe Roboter und Translationen zwischen polygonalen Hindernissen auf das Finden von Pfaden für Punkte im zweidimensionalen Raum  $\mathcal{C}_{\text{frei}}$  — d. h. zwischen polygonalen Hindernissen mit Gesamtgröße  $O(mn)$  — reduziert. Dies läßt sich mit Algorithmen zur Berechnung kürzester Pfade oder mit dem in Algorithmus ?? beschriebenen Roadmap-Verfahren lösen.

Abbildung 2.11: Konfigurationsraum-Hindernisse und Roadmap.

Insgesamt erhalten wir folgendes Theorem:

---

**Algorithmus 2.3** Roadmap-Verfahren

---

**Vorbereitung:**

- Berechne die Trapezzerlegung von  $C_{\text{frei}}$  und die zugehörige Suchstruktur für Point-Location. randomisiert  $O(mn \log(mn))$
- Entferne die Trapeze, die im Inneren eines Hindernisses liegen.
- Platziere je einen Knoten im Inneren eines jeden Trapez und auf jeder vertikalen Verlängerung.  $O(mn)$
- Verbinde adjazente Knoten (siehe Abbildung ??)  
     $\leadsto$  Zusammenhangsgraph von  $C_{\text{frei}}$  (Roadmap).

**Query:**

- Lokalisier  $s$  und  $t$  in den Trapezen.  $O(\log(mn))$
  - Falls  $s$  und  $t$  im selben Trapez liegen: verbinde  $s$  und  $t$ .
  - Ansonsten verbinde  $s$  und  $t$  mit den in den Trapezen liegenden Knoten und führe Breitensuche in der Roadmap durch.  $O(mn)$
- 

**Theorem 2.17** *Translationsbewegungen eines konvexen, polygonalen Roboters mit  $m$  Ecken in einer Umgebung mit polygonalen Hindernissen mit insgesamt  $n$  Ecken können nach  $O(mn \log^2(mn))$  (randomisiert) Vorbereitungszeit in Zeit  $O(mn)$  geplant werden.*

Algorithmus ?? ist vergleichsweise einfach, jedoch nicht optimal.

## 2.2 Reine Translationsbewegungen für beliebige Roboter

Betrachten wir nun einen allgemeinen, polygonalen Roboter mit  $m$  Ecken, der sich wie gehabt in einer Umgebung mit polygonalen Hindernissen  $P_i$  mit insgesamt  $n$  Ecken bewegt.

Eine mögliche Vorgehensweise wäre, die konvexe Hülle des Roboters zu betrachten und wie in Abschnitt ?? vorzugehen. Das Problem dabei ist uns schon bei der Approximation durch kreisförmige Roboter begegnet: es werden evtl. keine Bahnen gefunden, obwohl welche existieren.

Da der Roboter  $R$  nicht mehr konvex ist, hat  $\mathcal{C}_{\text{frei}}$  hier nach Lemma ??(iii) die Komplexität  $O(m^2n^2)$ . Die Berechnung des gesamten Raumes  $\mathcal{C}_{\text{frei}}$  wäre also sehr aufwendig. Die Idee ist nun, nur noch diejenige Zusammenhangskomponente  $Z_s$  von  $\mathcal{C}_{\text{frei}}$  zu berechnen, die die Startposition  $s$  enthält. Dies genügt, denn wenn die Zielposition  $t$  nicht in  $Z_s$  liegt, ist das Bahnplanungsproblem nicht lösbar, andernfalls ist ein Pfad von  $s$  nach  $t$  in  $Z_s$  zu berechnen.

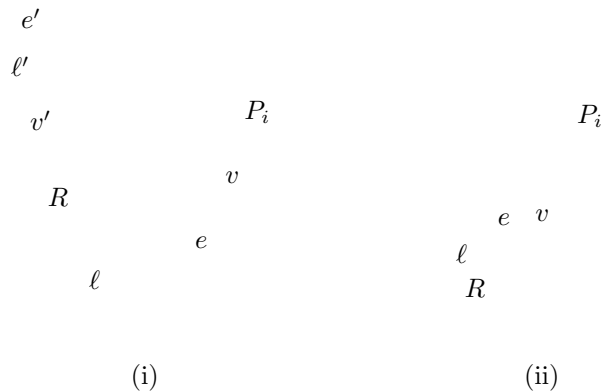


Abbildung 2.12: Paare von Ecken und Kanten von  $P_i$  und  $R$  erzeugen im Konfigurationsraum ein Liniensegment.

Untersuchen wir einmal genauer, wodurch  $\mathcal{C}_{\text{frei}}$  eingeschränkt wird: Für jede Ecke  $v$  des Roboters und jede Kante  $e$  eines Hindernisses erzeugt das Ecke/Kante-Paar  $(v, e)$  im Konfigurationsraum ein Liniensegment  $\ell$ , über das der Referenzpunkt des Roboters nicht hinweg bewegt werden darf. Dieses Liniensegment entsteht durch Bewegung der Roboterecke  $v$  entlang der Hinderniskante  $e$ . Andere Ecke/Kante-Paare werden dabei außer acht gelassen. In Abbildung ??(i) erzeugen die Ecke/Kante-Paare  $(v, e)$  und  $(v', e')$  auf diese Weise die Liniensegmente  $\ell$  und  $\ell'$ . Weiterhin erzeugt jedes Kante/Ecke-Paar  $(e, v)$  einer Kante  $e$  des Roboters und einer Hindernisecke  $v$  ein Liniensegment  $\ell$ , siehe Abbildung ??(ii).

Wir brauchen daher nicht zu jedem Hindernis das Konfigurationsraumhindernis zu berechnen, sondern können jedes Ecke/Kante- und Kante/Ecke-Paar einzeln betrachten. Für jedes dieser Paare erhalten wir ein Liniensegment.  $Z_s$  ist dann eine Zelle in einem Arrangement von  $2mn$  Liniensegmenten.



### 2.2.1 Komplexität und Berechnung einer einzelnen Zelle in einem Arrangement von Kurvenstücken

Wie kompliziert kann eine solche Zelle sein? Das folgende Theorem beantwortet die Frage etwas allgemeiner, nämlich für die Komplexität einer Zelle in einem Arrangement von Kurvenstücken ohne Selbstschnitte:

$Z$

Abbildung 2.13: Zelle in Arrangement von Kurvenstücken.

**Theorem 2.18** (Guibas, Sharir, Sifrony, 1989)

Sei  $A$  ein Arrangement von  $n$  Kurvenstücken, von denen sich je zwei höchstens  $s$  mal schneiden. Dann hat jede Zelle in  $A$  die Komplexität  $O(\lambda_{s+2}(n)) \subseteq O(n \log^* n)$ . [?]

$\lambda_s(n)$  bezeichnet die maximale Länge einer Davenport–Schinzel–Sequenz der Ordnung  $s$  über einem Alphabet mit  $n$  Buchstaben, siehe Abschnitt ??.

Für den Beweis von Theorem ?? betrachten wir eine Zelle. Der Rand  $\partial Z$  von  $Z$  zerfällt in Zykel. Wir betrachten nur den äußeren Zyklus,  $C$ , und nehmen an, daß  $C$  entgegen den Uhrzeigersinn orientiert ist. Wie man sich leicht überlegen kann, reicht für die Komplexitätsbetrachtung der Zelle die Betrachtung des äußeren Zyklus aus (Übungsaufgabe).

Nun orientieren wir jeden Bogen, z. B. von links nach rechts, wodurch ein positiver und negativer Durchlaufsinne festgelegt wird. Den Zyklus  $C$  aus Abbildung ?? können wir so rein kombinatorisch durch die zyklische Folge

$$S = \langle \gamma_4^+ \gamma_2^- \gamma_1^- \gamma_1^+ \gamma_2^- \gamma_2^+ \gamma_1^+ \gamma_7^- \gamma_3^- \gamma_6^+ \gamma_6^- \gamma_3^- \gamma_5^+ \gamma_5^- \gamma_3^- \rangle$$

beschreiben.

**Lemma 2.19** (Konsistenzlemma)

Die Segmente eines orientierten Bogens  $\gamma_i^+$  ( $\gamma_i^-$ ) erscheinen in  $C$  in derselben Reihenfolge wie längs  $\gamma_i^+$  ( $\gamma_i^-$ ).

**Beweis.** Zur Repräsentation von  $\gamma_i^+$  und  $\gamma_i^-$  erweitern wir  $\gamma_i$  zu einem Schlauch ohne die Topologie von  $C$  zu verändern. Diesen Schlauch nennen wir nun ebenfalls  $\gamma_i$ .

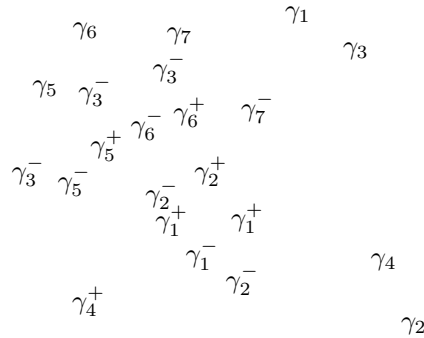


Abbildung 2.14: Die orientierten Segmente von  $C$ .

Seien  $\zeta$  und  $\eta$  zwei Segmente auf einer Seite von  $\gamma_i$ , die in  $C$  konsekutiv sind, das heißt, kein Stück von  $\gamma_i$  taucht zwischen  $\zeta$  und  $\eta$  entlang  $C$  auf. Unterscheiden wir je nach Verlauf der Kurve  $C$ , so können wir einige Fälle ausschließen. In Abbildung ?? liegt die Zelle  $Z$  auf der linken Seite von  $C$ . Da sie in  $\eta$  auf der Außenseite des Schlauches liegen muss, ist dieser Fall nicht möglich. Die Zelle  $Z$  läge hier auf beiden Seiten von  $C$ .



Abbildung 2.15: Die Zelle  $Z$  läge hier auf beiden Seiten von  $C$ .

Dadurch ergeben sich zwei Fälle, siehe Abbildung ??.

In beiden Fällen können wir innerhalb des Schlauches eine Kurve  $\alpha$  von  $x \in \zeta$  nach  $y \in \eta$  legen, die  $\zeta$  und  $\eta$  verbindet. Das Kurvenstück  $C_x^y$  nennen wir im folgenden  $\beta$ . Die geschlossene Kurve  $\alpha \cup \beta$  liegt im Komplement des Inneren von  $Z$ .

In beiden Fällen liegt  $Z$  links von der gegen den Uhrzeigersinn orientierten, geschlossenen Kurve  $C$ . Das bedeutet, daß  $\alpha \cup \beta$  die Kurve  $C$  vom Segment  $\kappa$  zwischen  $\zeta$  und  $\eta$  separiert. Kein Punkt  $z \in \kappa$  liegt somit auf  $C$ .  $\square$

Um die Komplexitäten der im Anhang beschriebenen Davenport–Schinzel–Sequenzen anwenden zu können, müssen wir uns von der zyklischen Betrachtung lösen und eine lineare Ordnung der betrachteten Segmente erschaffen.

Dabei kann der Fall auftreten, daß in der Sequenz  $s$  bezüglich eines Startpunktes  $p$  einige orientierte Bogensegmente nicht mehr in der Reihenfolge entlang der Segmente besucht werden, siehe zum Beispiel  $\gamma_3^+$  in Abbildung ??, daher verdoppeln wir diese Segmente einfach. Aus  $\gamma_3^-$  wird  $\gamma_{3.1}^-$  und  $\gamma_{3.2}^-$ . Es entsteht dabei eine lineare Sequenz  $s^*$  mit  $4n$  Segmenten.

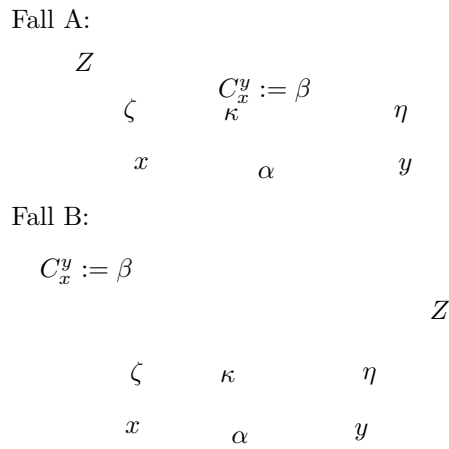


Abbildung 2.16: Kein Stück von  $\gamma_i$  taucht zwischen  $\zeta$  und  $\eta$  entlang  $C$  auf.

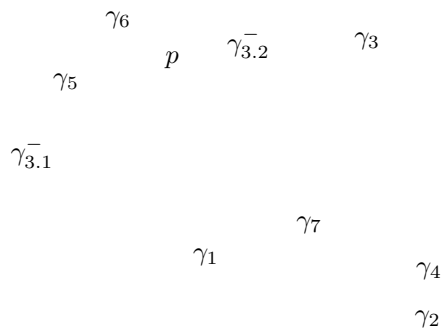


Abbildung 2.17: Das Segment  $\gamma_3^-$  wird in der nicht-zyklischen Sequenz nicht in der Reihenfolge gemäß der Orientierung besucht und wird deshalb am Startpunkt  $p$  verdoppelt.

Bemerkung: Es kann tatsächlich vorkommen, dass wir jedes Segment teilen müssen, siehe Abbildung ??.

Mit diesen Vorbereitungen zeigen wir die folgende Behauptung, mit der Theorem ?? bewiesen wird.

**Lemma 2.20** Die Sequenz  $s^*$  ist eine  $(4n, s + 2)$  Davenport-Schinzel-Sequenz.

**Beweis.** Es ist klar, daß nie zwei Segmente  $\gamma_{i,j}^+$  ( $\gamma_{i,j}^-$ ) hintereinander stehen, da ein Schnittpunkt dazwischen sein muß. Es bleibt zu zeigen, daß sich keine zwei verschiedenen Bogensegmente  $\zeta$  und  $\eta$  in der Sequenz  $s^*$  mehr als  $s + 2$  mal abwechseln. Dazu nehmen wir an, es gäbe eine solche (Teil-)Sequenz  $s^*$

$p$

Abbildung 2.18: Wird der Startpunkt  $p$  gewählt, so müssen alle Segmente geteilt werden.

mit  $s+3$  Wechseln und führen dies zu einem Widerspruch, indem wir zeigen, daß sich  $\zeta$  und  $\eta$  dann mehr als  $s$  mal schneiden.

Für eine gegebene Sequenz

$$s^* = (\cdots \zeta_1 \cdots \eta_1 \cdots \zeta_2 \cdots \eta_2 \cdots \cdots \zeta_j \cdots \eta_j \cdots \zeta_k \cdots (\eta_k \cdots))$$

mit den signifikanten  $\zeta_i$  und  $\eta_i$  fassen wir je 4 sukzessive Segmente zu einem Quadrupel wie folgt zusammen.

$$\begin{array}{c} (\zeta_1, \eta_1, \zeta_2, \eta_2) \\ (\eta_1, \zeta_2, \eta_2, \zeta_3) \\ (\zeta_2, \eta_2, \zeta_3, \eta_3) \\ \vdots \end{array}$$

Der Index  $k$  und die Existenz von  $\eta_k$  hängt von  $s$  ab: wenn  $s+3$  ungerade ist, dann existiert  $\eta_k$  und es gilt  $k = \frac{s+2}{2} + 1$ . Wenn  $s+3$  gerade ist, entfällt das letzte  $\eta_k$  und der Wert von  $k$  ist eine leichte Übungsaufgabe.

Sei  $s+3$  ungerade. In diesem Fall erhalten wir insgesamt  $s+1$  Quadrupel:

$$(\zeta_1, \eta_1, \zeta_2, \eta_2), (\eta_1, \zeta_2, \eta_2, \zeta_3), (\zeta_2, \eta_2, \zeta_3, \eta_3), \dots, (\zeta_{\frac{s+2}{2}}, \eta_{\frac{s+2}{2}}, \zeta_{\frac{s+2}{2}+1}, \eta_{\frac{s+2}{2}+1}).$$

Die Anzahl der Quadrupel ergibt sich aus  $2 \cdot (\frac{s+2}{2}) - 1$  Quadrupeln angeführt von je  $\frac{s+2}{2}$  Elementen aus  $\zeta$  und  $\frac{s+2}{2} - 1$  Elementen aus  $\eta$ .

Wir zeigen, daß für jedes Quadrupel der Form  $(\zeta_i, \eta_i, \zeta_{i+1}, \eta_{i+1})$  (bzw.  $(\eta_i, \zeta_{i+1}, \eta_{i+1}, \zeta_{i+2})$ ) genau ein Schnittpunkt zwischen  $\zeta_i$  und  $\eta_i$  (bzw. zwischen  $\eta_i$  und  $\zeta_{i+1}$ ) liegen muß. Damit hätten wir insgesamt  $s+1$  verschiedene Schnittpunkte im Widerspruch zu der Annahme, daß sich  $\zeta$  und  $\eta$  nicht mehr als  $s$  mal schneiden. Ergo existieren höchstens  $s+2$  Wechsel.

Es bleibt zu zeigen, daß für ein beliebiges Quadrupel, nennen wir es o. B. d. A.  $(\zeta_i, \eta_i, \zeta_{i+1}, \eta_{i+1})$  der besagte Schnittpunkt existiert. Dazu wählen wir  $x \in \zeta_i$ ,  $y \in \zeta_{i+1}$ ,  $z \in \eta_i$  und  $w \in \eta_{i+1}$  so, daß  $C$  diese Punkte in der Reihenfolge  $x, z, y, w$  besucht. Die Teilbögen  $\beta_x^z, \beta_z^y, \beta_y^w$  und  $\beta_w^x$  aus  $C$

schneiden sich nur in den Endpunkten, weil  $C$  eine geschlossene, einfache Kurve ist. Wir betrachten weiterhin die Bögen  $\beta_x^y \in \zeta$  und  $\beta_z^w \in \eta$ , siehe Abbildung ??.

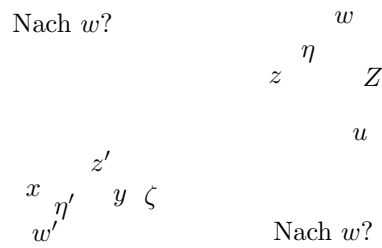


Abbildung 2.19: Die paarweise disjunkten Bögen  $\beta_x^z, \beta_z^y, \beta_y^w, \beta_w^x, \beta_x^y$  und  $\beta_z^w$ .

Wir wollen zeigen, daß sich  $\beta_x^y$  und  $\beta_z^w$ , wie Abbildung ?? durch den Bogen  $\beta_{z'}^{w'}$  angedeutet wird.

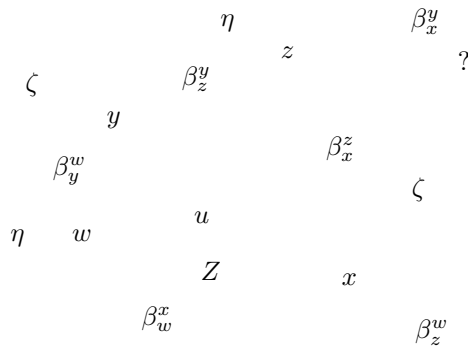


Abbildung 2.20: Eine kreuzungsfreie Einbettung, wie sie von der angenommenen Lage der Bögen induziert wird, existiert nicht.

Wir wollen einen formalen Widerspruchsbeweis führen. Dazu nehmen wir an, daß sich  $\beta_x^y$  und  $\beta_z^w$  nicht schneiden. Wählen wir einen Punkt  $u$  im Inneren von  $Z$ , dann können wir  $x, y, z$  und  $w$  kreuzungsfrei mit  $u$  verbinden. Da die Bögen  $\beta_x^z, \beta_z^y, \beta_y^w, \beta_w^x, \beta_x^y$  und  $\beta_z^w$  allesamt kreuzungsfrei sind, hätten wir den vollständigen Graphen  $K_5$  kreuzungsfrei in die Ebene eingebettet. Das ist aber nicht möglich: wie in Abbildung ?? zeigt, kann man nicht, jeden der Punkte  $u, w, x, y, z$  mit jedem anderen verbinden, ohne daß sich zwei Bögen kreuzen. Diese Tatsache läßt sich auch mit Hilfe der Eulerformel und einer weiteren Abschätzung formal beweisen, siehe z. B. [?].  $\square$

Wie können wir nun eine einzelne Zelle  $Z$  im Arrangement von  $n$  Kurvenstücken berechnen, die durch einen in  $Z$  enthaltenen Punkt — den Referenzpunkt des Roboters — bestimmt ist?

Zur Berechnung gehen wir von folgenden Annahmen aus:<sup>5</sup>

(i) Je zwei Bögen schneiden sich höchstens  $s$  mal.

(ii) Jeder Bogen hat nur konstant viele vertikale Tangenten. Damit können wir ohne Einschränkung annehmen, daß alle Bögen  $X$ -monoton sind, nicht  $X$ -monotone Bögen können an den Tangentenpunkten der vertikalen Tangenten in  $X$ -monotone Teilbögen zerlegt werden, vgl. Abbildung.

Kurve hier  
zerlegen

---

**Algorithmus 2.4** Berechnung einer Zelle im Arrangement von  $n$  Bögen

---

**Gegeben:** Punkt  $x$ ,  $n$  Bögen.

**Gesucht:** Die Zelle  $Z$  im Arrangement der Bögen, die  $x$  enthält.

- Zerlege die Menge der Bögen in gleichgroße Teilmengen  $R$  und  $B$ .
  - Berechne rekursiv im Arrangement  $\mathcal{A}(R)$  die Zelle  $Z_R$ , die  $x$  enthält.
  - Berechne rekursiv im Arrangement  $\mathcal{A}(B)$  die Zelle  $Z_B$ , die  $x$  enthält.
  - Berechne die Zusammenhangskomponente  $Z$  von  $Z_R \cap Z_B$ , die  $x$  enthält und berichte diese (Red-Blue-Merge).
- 

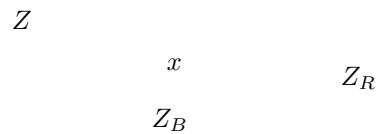


Abbildung 2.21:  $Z_R \cap Z_B$  kann sehr komplex sein, obwohl  $Z$  recht einfach ist.

Ein erster Ansatz zur Lösung dieses Problems ist der Divide-and-Conquer Algorithmus ???. Das Problem dabei ist, daß  $Z_R \cap Z_B$  sehr komplex werden kann und die explizite Berechnung zu hohe Kosten verursacht. Abbildung ??? zeigt ein Beispiel, in dem  $Z_R \cap Z_B$  recht komplex ist,  $Z$  selbst jedoch recht einfach. Wir brauchen also einen effizienteren Ansatz für den Merge-Schritt;

<sup>5</sup>Diese Forderungen sind erfüllt, wenn die Bögen durch algebraische Gleichungen vom Grad  $O(1)$  beschrieben werden.

die Frage ist nun, wie man die Zelle  $Z$  berechnen kann, ohne alle Zusammenhangskomponenten von  $Z_R \cap Z_B$  zu berechnen.

Betrachten wir einen etwas allgemeineren Fall; gegeben sind:

- Ein “rotes” Arrangement  $\mathcal{R}$  mit ausgezeichneten Zellen  $R_1, \dots, R_{m_R}$  mit insgesamt  $r$  Ecken (Bogensegmenten).
- Ein “blaues” Arrangement  $\mathcal{B}$  mit ausgezeichneten Zellen  $B_1, \dots, B_{m_B}$  mit insgesamt  $b$  Ecken (Bogensegmenten).<sup>6</sup>
- Eine Menge von Punkten  $p_i \in R_{\mu_i} \cap B_{\nu_i}$ ,  $1 \leq i \leq k$  — d. h. jeder Punkt ist sowohl in einer roten als auch in einer blauen Zelle enthalten —, wobei nicht jede Menge  $R_{\mu} \cap B_{\nu}$  ein  $p_i$  enthalten muß und verschiedene  $p_i$  in derselben Menge  $R_{\mu} \cap B_{\nu}$  liegen können.

Abbildung 2.22: Arrangement von roten und blauen Zellen und Überlagerung. Komplexität der linken blauen (linken roten) Zelle: 6 (5) Segmente.

Betrachte nun das “violette” Arrangement  $\mathcal{V}$ , das durch Überlagerung des roten Arrangements  $\mathcal{R}$  und des blauen Arrangements  $\mathcal{B}$  entsteht. In  $\mathcal{V}$  liegt jeder Punkt  $p_i$  in einer Zelle  $Z_i$ . Wir wollen nun diese  $Z_i$  berechnen, jedoch jede nur einmal, auch wenn sie mehrere Punkte enthält.

Über die Komplexität dieser Zellen, d. h. die Anzahl Segmente auf dem Rand, gibt das folgende Lemma Auskunft:

**Lemma 2.21** (*Kombinationslemma; Guibas, Sharir, Sifrony, 1989*)

Für die Komplexität der Zellen  $Z_1, \dots, Z_\ell$ ,  $\ell \leq k$  gilt:<sup>7</sup> [?]

$$|Z_1| + |Z_2| + \dots + |Z_\ell| \in O(r + b + k)$$

Eine stärkere Version dieses Lemmas stammt von Edelsbrunner, Guibas und Sharir [?], zu finden auch in [?]. Die Aussage von Lemma ?? ist keineswegs

<sup>6</sup>Wir bezeichnen im Folgenden die Segmente der Zellen  $R_i$  und  $B_j$  als Bögen, um sie von den neuen Segmenten zu unterscheiden, die beim Schnitt der Zellen entstehen.

<sup>7</sup> $\ell \leq k$ , da mehrere Punkte in derselben Zelle liegen können.



Abbildung 2.23: Der Bogen  $\ell$  trägt zu vielen violetten Zellen bei.

trivial: ein Bogen kann zu vielen violetten Zellen, ja sogar mehrere Stücke zum Rand einer Zelle beitragen, siehe Abbildung ???. Wir geben hier keinen Beweis für Lemma ???.

Wie lassen sich nun die Zellen  $Z_i$  berechnen?

**Theorem 2.22** *Seien  $Z_1, \dots, Z_\ell$  diejenigen Zellen, welche die Punkte  $p_1, \dots, p_k$  enthalten. Dann lassen sich die Zellen  $Z_1, \dots, Z_\ell$  in Zeit  $O((r + b + k) \log(r + b + k))$  berechnen.*

**Beweis.** Algorithmus ?? berechnet etwas mehr als verlangt, nämlich alle violetten Zellen, die einen Punkt  $q \in Q$  im Abschluß enthalten. Wie bereits erwähnt, liegt die Schwierigkeit darin, *nur* diese Zellen zu berechnen, nicht alle Zusammenhangskomponenten von  $R_\mu \cap B_\nu$ . Die beiden Sweeps sind notwendig, da z. B. in Abbildung ??(i) während des ersten Sweeps nicht erkannt werden kann, daß in  $a$  eine violette Region beginnt, da noch kein  $q \in Q$  entdeckt wurde.  $a$  liegt zwar auf dem Rand einer Zelle, ist jedoch kein Endpunkt eines Bogens. Erst wenn die Sweepline  $q_1$  erreicht, wird eine violette Teilregion begonnen. Analog kann während des zweiten Sweeps der Beginn der violetten Region noch nicht im Punkt  $b$  sondern erst bei  $q_3$  erkannt werden.

Die Idee bei dem Sweep (Algorithmus ??) ist die Verwendung von je einem oberen und unteren “Scout” je violetter Teilregion. Diese Scouts bewegen sich mit der Sweepline weiter und helfen bei der Erkennung und Behandlung von Ereignissen, die Einfluß auf den Verlauf des Randes der violetten Region haben. Für die Abschätzung des Arbeitsaufwands bei der Ereignisbehandlung ist wichtig, daß jeder Bogen von höchstens einem oberen Scout und höchstens einem unteren Scout bewacht wird, der sich auf einem Bogen der anderen Farbe bewegt. Es treten zwei Arten von Aktualisierungen der Ereignisstruktur auf:

**A:** berechne den in X-Richtung nächsten Schnittpunkt zwischen Scout und

<sup>8</sup>Die Schnittpunkte gleichfarbiger Bögen gehören zwar zur Ereignisstruktur, jedoch nicht zu  $Q$ . Ansonsten würde in Abbildung ??(i) bereits in Punkt  $a$  eine Region beginnen! Diese Schnittpunkte sind aus dem letzten Rekursionsschritt bekannt, müssen also nicht explizit berechnet werden.



**Algorithmus 2.5** Red–Blue–Merge

- Sei  $Q := \{p_1, \dots, p_k\} \cup \{ \text{Endpunkte von auf dem Rand der roten oder blauen Zellen liegenden Bögen} \}$ .  $n := r + b + k$ ,  $|Q| \in O(n)$ .
- Für jedes  $q \in Q$ : berechne mit gewöhnlichem Sweep in den Arrangements  $\mathcal{R}$  und  $\mathcal{B}$  die vertikalen Segmente von  $q$  nach oben und unten zum nächsten roten und blauen Bogen. Diese werden wie zusätzliche Bögen behandelt (Abbildung ??(i)).  $O(n \log n)$
- Sortiere die  $q \in Q$  und alle Ecken<sup>8</sup> der Zellen  $R_\mu$  und  $B_\nu$  nach auf- und absteigenden X-Koordinaten (Ereignisstrukturen).  $O(n \log n)$
- Sweep von links nach rechts:  
Berechne die Teile der violetten Teilzellen, die rechts vom linkensten darin enthaltenen  $q \in Q$  liegen.
- Sweep von rechts nach links:  
Berechne die Teile der violetten Teilzellen, die links vom rechtensten darin enthaltenen  $q \in Q$  liegen.
- Setze die Teilregionen zusammen.  $O(n)$

beobachtetem Bogen.

**B**: berechne den in X-Richtung nächsten Schnittpkt. zwischen beiden Scouts.  
Folgende Ereignisse können auftreten:

1. Eine neue Region startet. Zeit  $O(1)$  für das Anlegen einer neuen Region und zwei Scouts, Aktualisierungen **A** und **B**.
2. Schnittpunkt von Bögen gleicher Farbe oder Endpunkt eines Bogens. Evtl. muß ein beobachtender Scout einen neuen Bogen beobachten, Aktualisierung **A**.
3. Der Scoutbogen und der beobachtete Bogen treffen aufeinander. Der Scout wechselt auf den beobachteten Bogen und beobachtet den ehemaligen Scoutbogen, Aktualisierung **A** und **B**.
4. Die Scouts treffen aufeinander. Eine violette Teilzelle endet, die Scouts verschwinden, Zeit  $O(1)$ .
5. Ein Punkt  $q \in Q$  wird in einer aktiven Region angetroffen. Die aktive Region endet und es entstehen ein oder zwei neue Teilzellen. Wir brauchen  $O(1)$  Zeit um festzustellen, daß  $q$  zwischen den Scouts liegt und  $O(1)$  Zeit für den Abschluß der alten Teilzelle und den Beginn der neuen Teilzelle(n) wie unter 1.

Pro Ereignis tragen wir konstant viele neue Schnittpunkte in die Ereignisstruktur ein, das Einfügen und die Aktualisierung der Sweep-Status-Struktur kostet uns jeweils  $O(\log n)$ . Wir haben nach Lemma ??  $O(r + b + k) = O(n)$  viele Ereignisse. Damit können alle violetten Regionen in Zeit  $O(n \log n)$  berechnet werden.  $\square$

Abbildung ?? zeigt zwei Beispiele für die Aufgabenteilung der Scouts:

- (i) In Punkt  $a$  wechselt der Scout  $O_1$  von einem roten auf einen blauen Bogen und bewacht jetzt den roten Bogen, auf dem er vorher gelaufen ist.  $O_2$  muß aktiv werden und den blauen Rand bewachen.
- (ii) In Punkt  $b$  ist das Ende einer violetten Teilregion erreicht. Der obere und untere Scout der Teilregion treffen zusammen und verschwinden, der Scout  $O_2$  wird aktiv und bewacht den blauen Rand.

In beiden Fällen wird der aktiv werdende Scout  $O_2$  durch binäre Suche in der Sweep-Status-Struktur gefunden.

Insgesamt können wir festhalten:

**Theorem 2.23** *Gegeben seien  $n$  Bögen mit  $O(1)$  senkrechten Tangenten; je zwei Bögen schneiden sich höchstens  $s$  mal. Die Zelle  $Z$  im Arrangement, die einen gegebenen Punkt  $x$  enthält, kann in Zeit  $O(\lambda_{s+2}(n) \log^2 n)$  berechnet werden.<sup>9</sup>*

**Beweis.** Die Laufzeit ergibt sich aus der Rekursionsgleichung  $T(n) = 2T(\frac{n}{2}) + \text{Merge}(n)$  für den Divide-and-Conquer Ansatz. Die Kosten für einen Merge betragen nach Theorem ??  $O((r + b + k) \log(r + b + k))$ . In unserem Fall gilt  $k = 1$  und mit Theorem ??  $r + b \in O(\lambda_{s+2}(n))$ . Insgesamt also  $\text{Merge}(n) \in O(\lambda_{s+2}(n) \log(\lambda_{s+2}(n))) \subseteq O(\lambda_{s+2}(n) \log n)$ . Damit ergibt sich

$$T(n) = 2^\ell T(1) + C \cdot \sum_{i=1}^{\ell} \lambda_{s+2}(n) \log n \in O(\lambda_{s+2}(n) \log^2 n).$$

$\square$

---

<sup>9</sup> $\log^2(n) = \log(n) \cdot \log(n)$ .

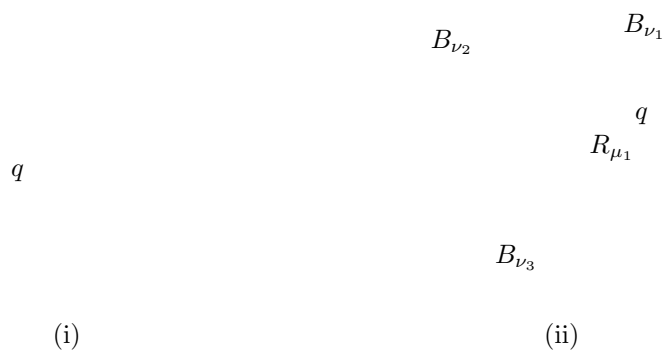


Abbildung 2.24: (i) Zusätzliche vertikale Segmente, (ii) bei  $q$  beginnt keine violette Region.



Abbildung 2.25: (i) Sweep von links nach rechts, (ii) Sweep von rechts nach links.

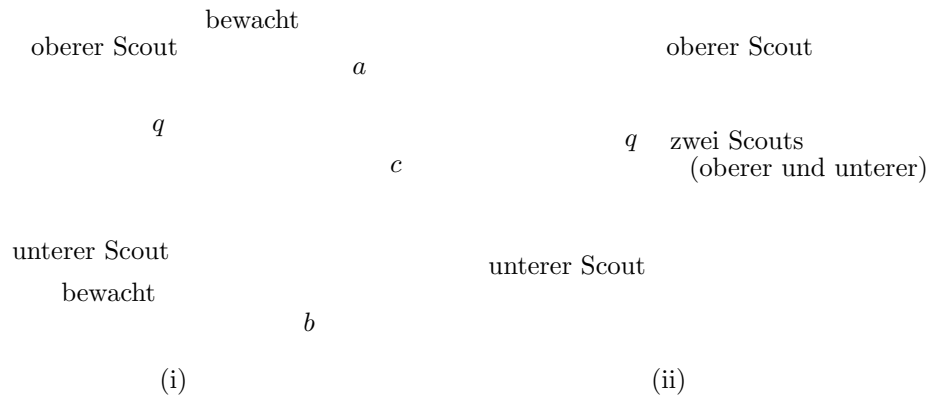
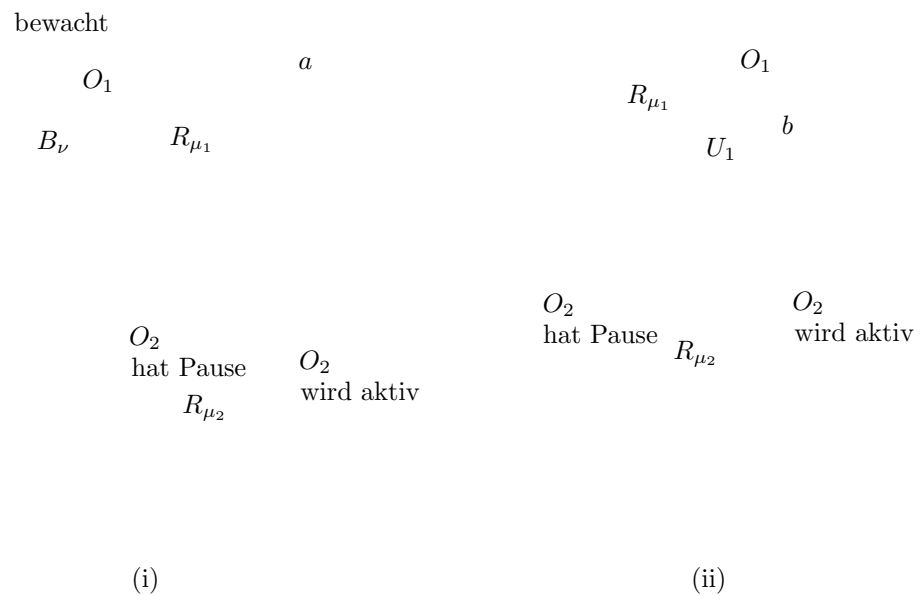
Abbildung 2.26: In  $q$  starten (i) eine oder (ii) zwei violette Teilregionen mit Scouts.

Abbildung 2.27: Beispiele für die Aufgaben der Scouts.

---

**Algorithmus 2.6** Sweep zur Berechnung der Teilzellen

---

(hier: Sweep von links nach rechts, rechts nach links analog)

**Ereignisstruktur:** Priority-Queue  $W$  sortiert nach X-Koordinaten, initialisiert mit den Punkten aus  $Q$  und allen Knoten auf dem Rand der blauen und roten Regionen.

**Sweep-Status-Struktur:** enthält die Bögen entlang der Sweepline, nach Y-Koordinaten sortiert. Für jeden Bogen wird ein Zeiger auf den beobachtenden Scout gespeichert und pro Scout ein Zeiger auf den Bogen, auf dem sich der Scout bewegt und den Bogen, den der Scout beobachtet.

**Ereignis:** Wenn ein Punkt  $q \in Q$  angetroffen wird, benutze die vertikalen Segmente von  $q$  um festzustellen, ob bei  $q$  eine violette Teilregion beginnt. Dies ist genau dann der Fall, wenn  $q$  sowohl in einem  $R_\mu$  als auch in einem  $B_\nu$  enthalten ist. In Abbildung ??(ii) beginnt in  $q$  keine violette Region, da  $q \notin \cup B_{\nu_j}$ . Falls ja, beginne bei  $q$  eine oder zwei violette Teilregionen, und starte am oberen und unteren Rand jeder Teilregion je einen Scout, siehe Abbildung ??.

**Aufgaben der Scouts:**

Hier die Aufgaben des oberen Scouts  $O$ , der untere Scout verhält sich symmetrisch.

- Der Scout startet auf dem zu  $q$  nächsthöheren Bogen. Hier starte  $O$  o. B. d. A. auf einem roten Bogen  $\varrho$ .
  - Der Scout bewegt sich mit der Sweepline.
  - Der Scout bewacht den nächsthöheren blauen Bogen  $\beta$ :
    - Berechne den in X-Richtung nächsten Schnittpunkt  $x$  zwischen  $\varrho$  und  $\beta$  und trage ihn in  $W$  ein (Abbildung ??(i), Punkte  $a$  und  $b$ ).
    - Wenn die Sweepline den Punkt  $x$  erreicht, wechselt  $O$  auf  $\beta$  und bewacht  $\varrho$ .
    - Wenn  $\beta$  endet oder einen anderen blauen Bogen schneidet, beobachte analog den nächsthöheren Bogen  $\beta'$ .
  - “Arbeitsteilung”: falls zwei Scouts denselben Bogen bewachen, ist nur der oberste aktiv. Längs der Sweepline gilt also: der Rand einer blauen Region wird nur von dem obersten darunter befindlichen Scout auf einem roten Bogen bewacht. Diese Invariante muß bei jedem Ereignis aktualisiert werden.
  - $O$  beobachtet den unteren Scout, d.h. ihr Schnittpunkt wird berechnet und falls existent in  $W$  eingetragen. Wenn beide sich treffen, ist das Ende der violetten Region erreicht und beide Scouts verschwinden (Abbildung ??(i), Punkt  $c$ ). Die violette Region endet evtl. schon eher, falls nämlich ein weiterer Punkt aus  $Q$  zwischen den Scouts erreicht wird.
-

### 2.2.2 Anwendung: Translation eines beliebigen Polygons

Mit den Ergebnissen des vorherigen Abschnittes können wir nun unser Bahnplanungsproblem lösen:

**Theorem 2.24** *Sei  $R$  ein polygonaler Roboter mit  $m$  Ecken, der sich in einer Umgebung mit polygonalen Hindernissen  $P_i$  mit insgesamt  $n$  Ecken bewegt. Gegeben seien Start- und Zielposition  $s, t$ . Dann läßt sich in Zeit  $O(mn \alpha(mn) \log^2(mn))$  eine kollisionsfreie Translation von  $s$  nach  $t$  bestimmen oder feststellen, daß keine solche existiert.*

**Beweis.** Die Laufzeit von Algorithmus ?? folgt aus Theorem ?? mit  $\lambda_3(mn) \in O(mn \alpha(mn))$  (Theorem ?? auf Seite ??).  $\square$

Wenn man bedenkt, daß  $\alpha(mn)$  fast konstant ist und auch  $\log^2(mn)$  nur sehr langsam wächst, so bleibt im wesentlichen eine Laufzeit von  $O(mn)$ . Angesichts dieses doch recht komplexen Problems ein bemerkenswertes Ergebnis!

---

#### Algorithmus 2.7 Translation eines beliebigen Roboters

---

##### Vorbereitung:

- Betrachte das Arrangement  $\mathcal{A}$  der  $2mn$  Liniensegmente, die sich aus den Bedingungen der Ecke/Kante-Paare ergeben.
- Konstruiere in  $\mathcal{A}$  die Zelle  $Z$ , die  $s$  enthält.  
Komplexität  $O(mn \alpha(mn))$ , Laufzeit  $O(mn \alpha(mn) \log^2(mn))$
- Führe Trapezzerlegung in  $Z$  aus, und konstruiere den Zusammenhangsgraph (Sweep).  $O(mn \alpha(mn) \log(mn))$

**Query:** für gegebenes  $t$ :

- Bestimme das Trapez, das  $t$  enthält.  $O(\log(mn))$
  - Finde einen Pfad von  $s$  nach  $t$  im Zusammenhangsgraph.  
 $O(mn \alpha(mn))$
- 

Abbildung 2.28: Trapezzerlegung und Zusammenhangsgraph einer Zelle.

### 2.2.3 Anwendung: Allgemeine Systeme mit zwei Freiheitsgraden

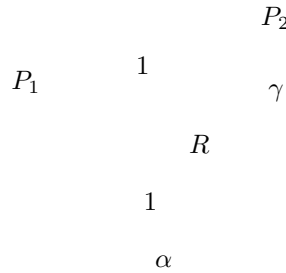
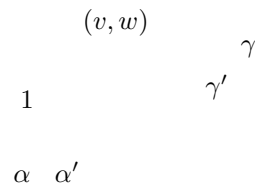


Abbildung 2.29: Roboter mit zwei Armen und Freiheitsgraden  $\alpha$  und  $\gamma$ .

Bisher haben wir nur die Translation von Polygonen betrachtet, die eindeutig durch die Position ihres Referenzpunktes bestimmt sind, also zwei Freiheitsgrade (DOF) haben. Der Ansatz des vorherigen Abschnittes ist jedoch allgemein genug, um andere Systeme mit zwei Freiheitsgraden zu behandeln! Abbildung ?? zeigt einen Roboter mit zwei Armen, dessen Position durch die beiden Winkel  $\alpha$  und  $\gamma$  eindeutig beschrieben wird. Sein Konfigurationsraum ist also  $\mathcal{C} = [0, 2\pi) \times [0, 2\pi)$ .

Wie sieht hier der Rand des Raumes der freien Konfigurationen  $\mathcal{C}_{\text{frei}}$  aus?

Ein Segment auf dem Rand von  $\mathcal{C}_{\text{frei}}$  entsteht z. B. durch den Kontakt einer Hindernisecke mit dem vorderen Arm (siehe Abbildung). Da die Hindernisecke nicht den Endpunkt von  $R$  berührt, muß ein  $\lambda \in [0, 1]$  existieren, so daß für den Kontaktpunkt  $(v, w)$  gilt:



$$\begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} + \lambda \begin{pmatrix} \cos \gamma \\ \sin \gamma \end{pmatrix}$$

$$\Rightarrow \lambda = \frac{v - \cos \alpha}{\cos \gamma} = \frac{w - \sin \alpha}{\sin \gamma}$$

mit  $x := \cos \alpha, y := \sin \alpha$ :

$$(v - x)\sqrt{1 - y^2} = (w - \sqrt{1 - x^2})y$$

$$\Leftrightarrow (v - x)^2(1 - y^2) = (w^2 - 2w\sqrt{1 - x^2} + 1 - x^2)y^2.$$

Aus der letzten Gleichung hebt sich  $x^2y^2$  fort. Löst man nach der Wurzel auf und quadriert die Gleichung, so ergibt sich eine Gleichung vom Grad<sup>10</sup>

<sup>10</sup>Der Grad eines Polynoms mit mehreren Variablen ist die Summe der Grade (d. h. der

6 in  $x, y$ . Wegen  $0 \leq \lambda \leq 1$  beschreibt diese ein Bogenstück vom Grad 6. Es handelt sich dabei um Bögen im Konfigurationsraum des Roboters, die Bögen sind daher nicht direkt in der Anschauung ablesbar.

Ein anderer Typ von Segmenten auf dem Rand von  $C_{\text{frei}}$  entsteht durch den Kontakt des Roboterkopfes mit einer Wand. Die Wand läßt sich durch eine Gerade  $Y = aX + b$  beschreiben, und für den Kontaktpunkt  $(v, w)$  gilt:

$$\begin{array}{r} Y = aX + b \\ (v, w) \\ \gamma \\ 1 \quad \alpha \end{array}$$

$$\begin{aligned} (v, w) &= (\cos \alpha + \cos \gamma, \sin \alpha + \sin \gamma) \in \{Y = aX + b\} \\ &\Leftrightarrow a(\cos \alpha + \cos \gamma) + b = \sin \alpha + \sin \gamma \\ &\Leftrightarrow a(x + y) + b = \sqrt{1 - x^2} + \sqrt{1 - y^2} \\ &\Leftrightarrow a(x + y)^2 + b - 2 + x^2 + y^2 = \sqrt{1 - x^2} \sqrt{1 - y^2} \end{aligned}$$

Durch Quadrieren der letzten Gleichung ergibt sich ein Bogenstück vom Grad 4.

Alle Konfliktbögen haben einen Grad kleiner gleich 6, mit Bezouts Theorem<sup>11</sup> folgt, daß je zwei Bögen höchstens  $6^2$  Schnitte haben. Mit Theorem ?? auf Seite ?? folgt, daß für dieses Problem kollisionsfreie Bewegungen in Zeit  $O(\lambda_{38}(n) \log^2 n) \subseteq O(n \log^{2+\varepsilon} n)$  geplant werden können.

---

höchsten Potenzen) der einzelnen Variablen.

<sup>11</sup>Zwei algebraische Kurven von Grad  $m$  und  $n$  schneiden sich in höchstens  $m \cdot n$  Punkten.



## 2.3 Rotations- und Translationsbewegungen für konvexe Roboter

*s*

*t*

Abbildung 2.30: Bahn eines Roboters mit Rotationen und Translationen.

Bisher hatten wir nur Translationsbewegungen des Roboters betrachtet. Für kreisförmige Roboter ist dies keine Einschränkung, um jedoch einen langgezogenen Roboter zu bewegen, kann es sehr wohl nötig sein, diesen zu drehen, um enge Passagen zu überwinden oder Hindernisecken zu umschiffen.

In diesem Abschnitt wollen wir Bewegungen betrachten, die sowohl aus Translationen als auch aus Rotationen bestehen können, jedoch beschränken wir uns auf den Fall eines konvexen Roboters. Ein Beispiel für Roboter, die beliebige Translations- und Rotationsbewegungen ausführen können, sind Fahrzeuge mit Mecanum-Rädern.<sup>12</sup>

$R$  sei — wie gehabt — durch ein konvexes Polygon mit  $m$  Ecken gegeben und bewege sich in einer Umgebung mit polygonalen Hindernissen  $P_i$  mit insgesamt  $n$  Ecken. Ein klassischer Spezialfall dieses Problems ist die Bewegung einer Leiter (Piano-Movers-Problem, siehe Abbildung ??); im Gegensatz zu Abschnitt ?? interessieren uns hier jedoch keine kürzesten Pfade.

**Theorem 2.25** (*Ke, O'Rourke, 1987*)

*Die Bewegung eines Liniensegmentes mit Translation und Rotation in einer Umgebung mit polygonalen Hindernissen mit insgesamt  $n$  Ecken kann  $\Omega(n^2)$  viele Schritte erfordern.* [?]

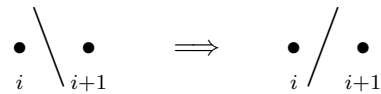
<sup>12</sup>Räder, auf deren Umfang mehrere Rollen in einem Winkel von  $45^\circ$  zur Orientierung angeordnet sind. Durch geeignete Ansteuerung aller Räder sind beliebige Translations- und Rotationsbewegungen möglich. [<http://prt.fernuni-hagen.de/forschung/mecanum-de.html>]

$t$  $s$ 

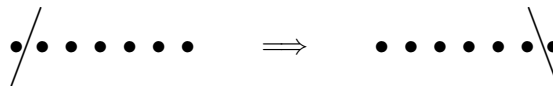
Abbildung 2.31: Kann die Leiter von  $s$  nach  $t$  bewegt werden?

**Beweis.** Abbildung ?? zeigt eine Umgebung, in der  $\Omega(n^2)$  viele Schritte gebraucht werden, um das Liniensegment  $L$  von  $s$  nach  $t$  zu bewegen:<sup>13</sup>

- Sei  $(i, i + 1)$  eine feste Lücke zwischen den B-Pflocken. Um den B-Pflock  $i + 1$  zu überwinden, muß  $L$  in die Kerbe  $i + 1$  hineingeschoben werden.
- Die Kerben sind so konstruiert, dass das Überwinden eines B-Pflocks mit ungeradem (geradem) Index nur in der ersten (letzten) Lücke zwischen den A-Pflocken möglich ist.
- Schiebt man  $L$  ganz nach unten, so kann man eine konstante Anzahl von A-Pflocken überwinden, bis  $L$  den nächsten C-Pflock berührt.
- Um einen C-Pflock zu überwinden, schiebt man  $L$  hoch bis zur oberen Wand.
- Um in einer festen Lücke zwischen den B-Pflocken die Orientierung von  $L$  folgendermaßen zu ändern, sind  $\Omega(n)$  Züge nötig:



- Zur Überwindung aller B-Pflocke von  $s$  nach  $t$



müssen abwechselnd rechte und linke Kerben aufgesucht werden, was je einen Durchlauf durch die Reihe der A-Pflocke erfordert, insgesamt also  $\Omega(n^2)$  Züge.  $\square$

<sup>13</sup>Besser als jede Erklärung ist es sich das Applet *Bewegung einer Leiter in 2 Dimensionen* anzuschauen, dass unter <http://www.geometrylab.de/LowerBound2D/index.html> (Stand 2009) zu finden ist.

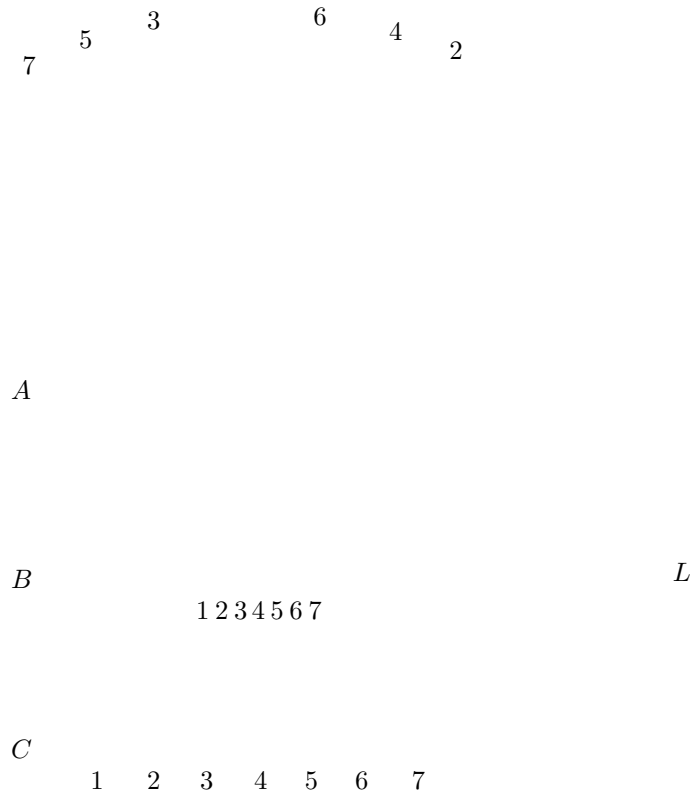


Abbildung 2.32: Die Bewegung eines Liniensegments kann  $\Omega(n^2)$  viele Schritte erfordern.

Bemerkenswert ist, daß in Theorem ?? nicht mit der Komplexität von  $\mathcal{C}_{\text{frei}}$  argumentiert wird, die Aussage bezieht sich vielmehr auf die Anzahl der notwendigen Schritte. Die Planung der Bahn kann nicht einfacher sein.

**Bemerkung 2.26** *Im  $\mathbb{R}^3$  kann die Bewegung eines Liniensegments  $\Omega(n^4)$  Züge erfordern und in Zeit  $O(n^6 \log n)$  geplant werden.* [?]

Wir wollen nun im  $\mathbb{R}^2$  Rotations- und Translationsbewegungen für beliebige konvexe Roboter planen. Der Roboter hat drei Freiheitsgrade:  $(x, y, \theta) \in \mathcal{C} = \mathbb{R}^2 \times [0, 2\pi)$ , vereinfacht auch  $\mathcal{C} = \mathbb{R}^3$ .

Die Konfigurationsraum-Hindernisse  $CP_i$  sind hier keine Polyeder mehr, sie haben gekrümmte Flächen und Kanten. Es sind durch Regelflächen<sup>14</sup> berandete, dreidimensionale Objekte. Betrachtet man den Schnitt eines  $CP_i$

<sup>14</sup>Flächen, die bei der Bewegung eines Liniensegments im Raum entstehen.

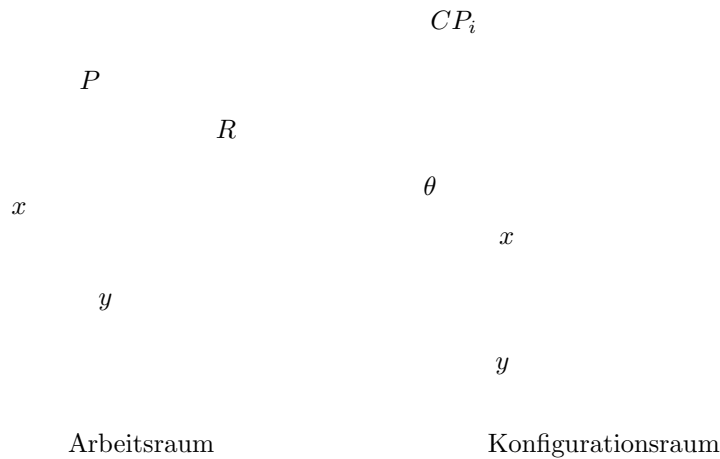


Abbildung 2.33: Konfigurationsraum-Hindernis bei Rotation und Translation.

mit einer Ebene mit konstantem  $\theta_i$ , so ist der Schnitt genau die Minkowski-Summe  $P_i \oplus -R(0, 0, \theta_i)$ . Man kann sich ein Konfigurationsraum-Hindernis also als den Stapel der Minkowski-Summen über alle  $\theta_i$  vorstellen, siehe Abbildung ??.

Im Prinzip würde der Ansatz, den Raum  $\mathcal{C}_{\text{frei}}$  in Zellen zu zerlegen und einen Zusammenhangsgraph zu konstruieren, hier auch funktionieren, wäre aber sehr schwierig zu implementieren.

Ein naiver Ansatz ist, mit fester Orientierung  $\theta_i$  für  $\theta_i = i \cdot \frac{360^\circ}{k}$ ,  $0 \leq i \leq k - 1$  mit bisher bekannten Verfahren Translationsbahnen zu planen und die 2-DOF-Lösungen geeignet zusammensetzen. Dies liefert jedoch nicht immer korrekte Lösungen: die Rotation zwischen zwei Schichten  $\theta_i$  und  $\theta_{i+1}$  kann eine Kollision mit einem Hindernis verursachen, wenn nämlich die Konfigurationen  $(x, y, \theta_i)$  und  $(x, y, \theta_{i+1})$  in  $\mathcal{C}_{\text{frei}}$  liegen, die Konfiguration  $(x, y, \theta')$  mit einem Winkel  $\theta'$  zwischen  $\theta_i$  und  $\theta_{i+1}$  jedoch verboten ist.

Eine mögliche Abhilfe wäre, den Roboter  $R$  um  $\frac{360^\circ}{k}$  zu drehen und die dabei überstrichene Fläche als Roboter  $R'$  für die Translationsplanung zu betrachten. Die mit diesem Ansatz berichteten Bahnen sind dann stets kollisionsfrei, jedoch wird für  $R'$  evtl. keine Bahn gefunden, obwohl für  $R$

eine Bahn existiert.

Der naive Ansatz widerspricht also unserem Ziel korrekter Bahnplanung. Trotzdem werden wir die Idee, nur endlich viele Orientierungen zu verwenden, weiter verfolgen; wir müssen nur die richtige Auswahl an Orientierungen finden.

Zur Vorbereitung betrachten wir im folgenden Abschnitt kritische Plazierungen des Roboters.

### 2.3.1 Kritische Plazierungen

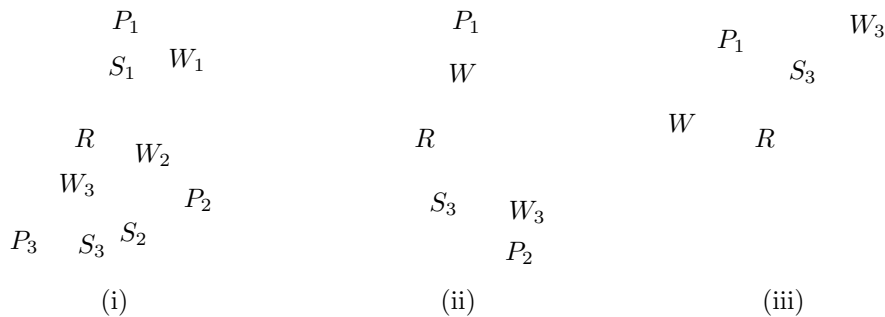


Abbildung 2.34: Kritische Plazierungen.

**Definition 2.27** Sei  $R$  ein konvexer Roboter mit  $m$  Ecken, und seien  $P_i$  polygonale Hindernisse. Ein **Kontaktpaar**  $O = (W, S)$  liegt vor, wenn

- (a)  $W$  ist eine Hinderniskante und  $S$  eine Roboterecke (Typ I) oder
- (b)  $W$  ist eine Hindernisecke und  $S$  eine Roboterkante (Typ II) oder
- (c)  $W$  ist eine Hindernisecke und  $S$  eine Roboterecke (Typ III)

und  $W$  berührt  $S$ .

Eine halbfreie<sup>15</sup> Platzierung  $(x, y, \theta)$  von  $R$  heißt **kritische Platzierung** genau dann, wenn an der Platzierung

- (A) drei paarweise verschiedene Kontaktpaare vom Typ I oder II oder
- (B) ein Kontaktpaar vom Typ III und ein Kontaktpaar vom Typ I oder II beteiligt sind.

Abbildung ?? zeigt Beispiele für kritische Plazierungen. (i) zeigt den “klassischen” Fall, in dem der Roboter an drei Stellen ein Hindernis berührt, also drei Kontaktpaare vom Typ I oder II beteiligt sind. (ii) und (iii) zeigen Beispiele für kritische Plazierungen, an denen ein Kontaktpaar vom Typ III beteiligt ist. Kritische Plazierungen spielen eine wichtige Rolle, wie folgendes Lemma zeigt:

<sup>15</sup>Zur Erinnerung: Eine Platzierung heißt **halbfreie Platzierung** wenn der Roboter eine Hinderniswand berührt.

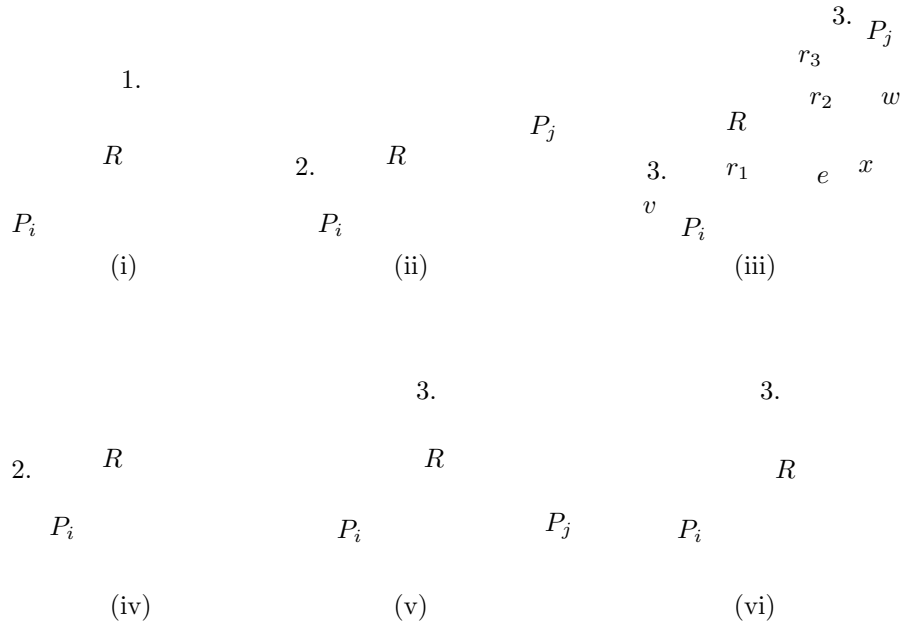


Abbildung 2.35: Beweis: In jeder Zusammenhangskomponente gibt es mindestens eine kritische Platzierung.

**Lemma 2.28** *In jeder Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}$  gibt es mindestens eine kritische Platzierung.*

**Beweis.** Starte mit freier Platzierung von  $R$  und bewege  $R$  nach unten, bis  $R$  ein Hindernis berührt. Nehmen wir o. E. an, eine Ecke von  $R$  berühre die Kante eines  $P_i$ , siehe Abbildung ??(i). Verschiebe dann  $R$  längs dieser Kante, bis einer der folgenden Fälle eintritt:

1. Es entsteht ein weiterer Kontakt zu einem Hindernis  $P_j$  (Abb. (ii)). Falls eine Ecke  $R$  eine Ecke von  $P_j$  berührt, hat  $R$  eine kritische Platzierung erreicht, ansonsten bewege  $R$  unter Aufrechterhaltung der Kontakte zu  $P_i$  und  $P_j$ . Man kann sich leicht klarmachen, daß dabei einer der folgenden Fälle eintreten muß (Abbildung ??(iii)):
  - (a) Eine der beiden Kontaktecken von  $R$  erreicht eine Hindernisecke ( $r_1$  erreicht  $v$  oder  $r_3$  erreicht  $w$ ).
  - (b) Eine Kante von  $R$  erreicht eine Hinderniskante ( $r_2$  erreicht  $e$ ).
  - (c) Es entsteht ein weiterer Hinderniskontakt ( $R$  berührt die Hindernisecke  $x$ . Diese kann zu einem dritten Hindernis  $P_k$  gehören oder zu  $P_i$  oder  $P_j$ ).

In allen Fällen ist ein dritter Kontakt entstanden und  $R$  hat eine kritische Platzierung erreicht.

2. Es wird ein Endpunkt  $v$  erreicht (Abbildung ??(iv)). Rotiere dann  $R$  um  $v$ , bis
  - (a)  $R$  hat Kontakt zu weiterem Hindernis  $P_j$  (Abb. (v)).
  - (b)  $R$  hat weiteren Kontakt zu Hindernis  $P_i$  (Abb. (vi)).

In beiden Fällen hat  $R$  eine kritische Platzierung erreicht. □

Im folgenden wollen wir die Anzahl der kritischen Platzierungen vom Typ A abschätzen. Die Anzahl möglicher Platzierungen vom Typ B ist dagegen mit  $O(m^2n^2)$  sehr einfach anzugeben (Übungsaufgabe).

**Bemerkung 2.29** Wenn man  $R$  unter Aufrechterhaltung von zwei Kontaktpaaren bewegt, beschreibt der Referenzpunkt eine Kurve vom Grad  $\leq 4$ .

**Beweis.** Siehe [?]. □

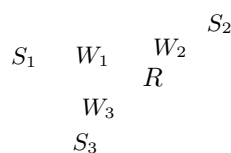
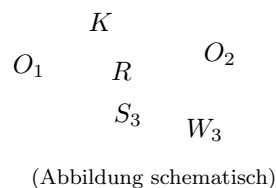
**Korollar 2.30** Für je drei Kante/Ecke-Paare<sup>16</sup>  $O_i = (W_i, S_i)$  gibt es höchstens vier halbfreie Platzierungen von  $R$ , bei denen die Kontakte  $O_i$  entstehen. [?]

**Beweis.** Angenommen, bei  $O_3 = (W_3, S_3)$  ist  $S_3$  eine Ecke von  $R$ . Lege den Referenzpunkt auf  $S_3$ . Halte  $O_1$  und  $O_2$  fest und bewege  $R$ .

$\Rightarrow S_3$  beschreibt Kurve  $K$  vom Grad  $\leq 4$ .

$\Rightarrow K$  hat  $\leq 4$  Schnitte mit  $W_3$ .

Angenommen, bei allen  $O_i$  ist  $S_i$  eine Kante von  $R$ . Dann halte  $R$  fest und bewege die Hindernisse, d. h. bewege das Dreieck  $W_1W_2W_3$  der Hindernisecken unter Aufrechterhaltung der Kontakte mit den Roboterkanten  $S_1, S_2$  und  $S_3$ . Damit kann wie oben gefolgert werden. □



Es gibt  $2mn$  mögliche Kontaktpaare  $O_i$  vom Typ I oder II, von denen wir drei auswählen. Insgesamt kann es also höchstens  $4 \cdot \binom{2mn}{3} \in \Theta(m^3n^3)$  viele kritische Platzierungen geben. Von der Konvexität des Roboters haben wir dabei noch keinen Gebrauch gemacht!

**Theorem 2.31** Für nichtkonvexe Roboter kann es  $\Theta(m^3n^3)$  viele kritische Platzierungen geben.

**Beweis.** Siehe Abbildung ???. □

Für konvexe Roboter gilt jedoch:

<sup>16</sup>Zur Erinnerung:  $W_i$  bezeichnet den am Kontakt beteiligten Teil des Hindernisses,  $S_i$  den des Roboters.

Abbildung 2.36: Ein nichtkonvexer Roboter mit  $\Theta(m^3n^3)$  kritischen Plazierungen.

**Theorem 2.32** (Leven, Sharir, 1987)

Die Anzahl der kritischen Plazierungen eines konvexen  $m$ -Ecks  $R$  zwischen Polygonen mit insgesamt  $n$  Ecken liegt in  $O(mn \lambda_6(mn))$ . [?]

Dabei gilt  $\lambda_6(mn) \in O(mn \log^*(mn))$ , die Anzahl der kritischen Plazierungen ist also kaum mehr als quadratisch.

Zum Beweis dieses Theorems sind einige Vorüberlegungen nötig. Betrachten wir zunächst alle möglichen Arten von zwei Kontaktpaaren  $O_1, O_2$  zwischen  $R$  und den Hindernissen, siehe Abbildung ???. Nehmen wir an, die beiden Tangenten  $T_1$  und  $T_2$ , die durch Verlängerung der beiden Kontaktpaare entstehen, seien nicht parallel. Verschiebt man nun  $R$  längs der Tangente  $T_1$  (o. E.) auf den Schnittpunkt  $S$  mit der anderen Tangente zu, so entsteht an der Kontaktstelle  $O_2$  ein verbotener Schnitt. Falls dieser verbotene Durchschnitt bis zum Ende der Translation entlang der Kante andauert, so sagen wir,  $O_1$  wird von  $O_2$  bei Winkel  $\theta$  beschränkt. Genauer:

**Definition 2.33** Seien  $O_1, O_2$  zwei Kontaktpaare,  $T_1$  und  $T_2$  die Verlängerungen der beteiligten Kanten.  $O_2$  **beschränkt**  $O_1$  bei Winkel  $\theta$ , wenn eine Plazierung  $R(x, y, \theta)$  mit den Kontaktpaaren  $O_1$  und  $O_2$  existiert, so daß  $R' = \text{ch}(S_1 \cup S_2)$  stets  $W_2$  schneidet, wenn  $R$  bei fester Orientierung  $\theta$  entlang der Tangente  $T_1$  unter Beibehaltung des Kontaktes  $O_1$  in Richtung des Schnittpunktes  $S$  von  $T_1$  und  $T_2$  verschoben wird.

Die konvexe Hülle  $\text{ch}(S_1 \cup S_2)$  ist in Abbildung ??? hellgrau dargestellt, in (i) gilt  $\text{ch}(S_1 \cup S_2) = \ell$ . Im folgenden bezeichnen  $u_{1/2}$  die Endpunkte der am



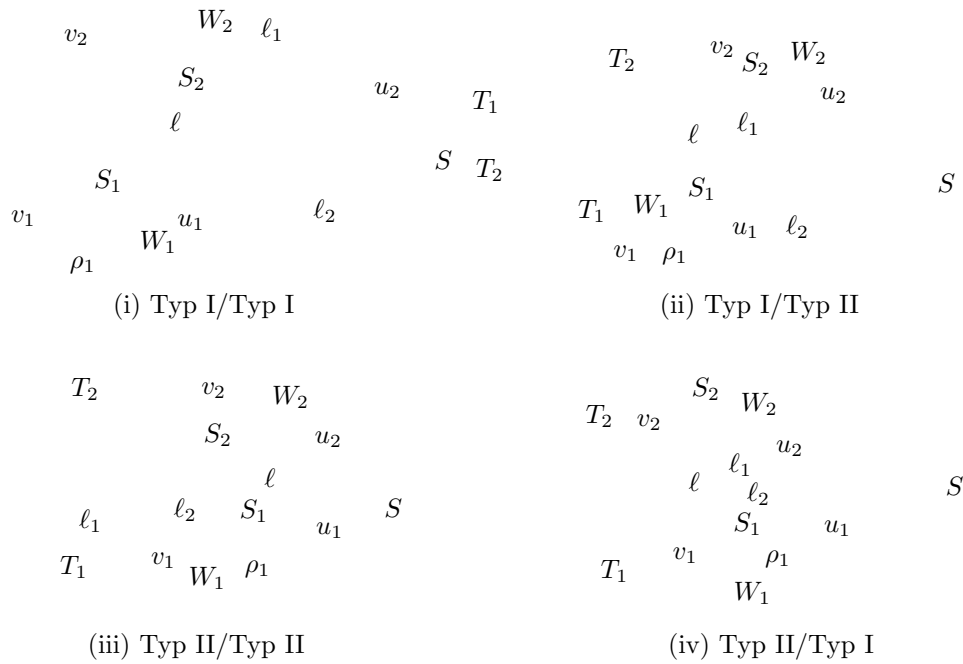


Abbildung 2.37:  $O_2$  beschränkt  $O_1$ .

Kontakt beteiligten Kanten in Richtung  $S$ ,  $v_{1/2}$  die Endpunkte der Kanten in die andere Richtung und  $x_i$  sei der Berührungspunkt des Kontaktpaares.

Die Definition der Beschränkung setzt voraus, daß sich die Tangenten  $T_1$  und  $T_2$  schneiden. Es gibt jedoch Fälle, in denen  $T_1$  und  $T_2$  parallel sind, z. B. wenn eine Roboterkante ( $S_1 = S_2$ ) zwei Hindernisecken berührt oder beide involvierten Kanten parallel sind. In allen anderen Fällen gilt:

**Lemma 2.34** Für zwei Kontaktpaare  $O_1$  und  $O_2$  gilt — von degenerierten Fällen abgesehen — mindestens eine der beiden Aussagen:

- (i)  $O_1$  beschränkt  $O_2$  bei  $\theta$ .
- (ii)  $O_2$  beschränkt  $O_1$  bei  $\theta$ .

**Beweis.**(Skizze)

Betrachte die Fallunterscheidung in Abbildung ???. In den Fällen (i) und (iii) sei  $\ell$  die Verbindung zwischen den Kontaktpunkten  $x_1$  und  $x_2$ , in den Fällen (ii) und (iv) die Verbindung zwischen der beteiligten Roboterecke  $S_1$  bzw.  $S_2$  und der Ecke  $v_2$  bzw.  $v_1$  der beteiligten Hinderniskante. Weiter sei

- (i)  $\ell_1$  die Parallele zu  $\ell$  durch  $u_1$ ,  $\ell_2$  die durch  $u_2$ .
- (ii)  $\ell_1$  die Parallele zu  $\ell$  durch  $u_1$ ,  $\ell_2$  die durch  $x_2$ .
- (iii)  $\ell_1$  die Parallele zu  $\ell$  durch  $v_2$ ,  $\ell_2$  die durch  $v_1$ .
- (iv)  $\ell_1$  die Parallele zu  $\ell$  durch  $x_1$ ,  $\ell_2$  die durch  $u_2$ .

Dann gilt

- (i)  $\ell_1$  schneidet  $W_2$  oder  $\ell_2$  schneidet  $W_1$ .
- (ii)  $\ell_1$  schneidet  $\overline{x_2v_2}$  oder  $\ell_2$  schneidet  $\overline{x_1u_1}$ .
- (iii)  $\ell_1$  schneidet  $S_1$  oder  $\ell_2$  schneidet  $S_2$ .
- (iv)  $\ell_1$  schneidet  $\overline{x_2u_2}$  oder  $\ell_2$  schneidet  $\overline{x_1u_1}$ .

In allen Fällen schneidet  $R'$  während der Bewegung  $W_1$  oder  $W_2$  und es gilt:  $O_2$  beschränkt  $O_1$  oder  $O_1$  beschränkt  $O_2$ .  $\square$

Dabei ist die Konvexität von  $R$  entscheidend.

In nebenstehender Abbildung mit nichtkonvexem

$R$  ist nicht klar, ob  $O_1$  von  $O_2$  beschränkt wird oder  $O_2$  von  $O_1$ .

Es gibt auch Fälle, in denen beides gilt, allerdings ist diese Platzierung nicht notwendigerweise frei:

**Bemerkung 2.35** Seien  $O_1, O_2$  fest. Dann gibt es genau eine Orientierung  $\theta$  von  $R$ , für die  $\ell_1$  und  $\ell_2$  zusammenfallen; für die also gilt:

$O_1$  beschränkt  $O_2$  bei  $\theta$  **und**  $O_2$  beschränkt  $O_1$  bei  $\theta$ .

Damit können wir unser Theorem beweisen:

**Beweis.** (Theorem ??)

Wir wollen einen Zusammenhang zwischen der Anzahl kritischer Platzierungen und der unteren Kontur von Funktionen<sup>17</sup> herstellen. Sei

$$\Pi_{O_1O_2} := \{ \theta \mid O_2 \text{ beschränkt } O_1 \text{ bei Winkel } \theta \}.$$

Mit den oben eingeführten Bezeichnungen  $u_{1/2}, v_{1/2}$  und  $x_i$  sei

$$\rho_i := \begin{cases} |\overline{x_i u_i}|, & \text{falls } S_i \text{ Roboterseite} \\ |\overline{v_i x_i}|, & \text{falls } S_i \text{ Roboterecke} \end{cases}$$

$\rho_i$  gibt die Distanz an, um die in Definition ?? bei konstantem  $\theta$  und unter Aufrechterhaltung des Kontaktes  $O_i$  geschoben werden kann, ohne daß ein verbotener Schnitt entsteht, vgl. auch Abbildung ??.

Definiere damit die Funktion

$$\begin{aligned} f_{O_1O_2} : \Pi_{O_1O_2} &\longrightarrow \mathbf{R}^{>0} \\ \theta &\longmapsto \rho_1 \end{aligned}$$

Eine entscheidende Eigenschaft ist: Sei  $(x, y, \theta)$  eine halbfreie Platzierung von  $R$  mit Kontaktpaaren  $O_1$  und  $O_2$  mit  $O_2$  beschränkt  $O_1$  mit Abstand  $\rho_1 = f_{O_1O_2}(\theta)$ . Für  $\rho > \rho_1$  ist keine (halb)freie Platzierung mit Kontaktpaaren  $O_1, O_2$  und Abstand  $\rho$  möglich, da  $O_2$  einen verbotenen Schnitt liefert.

<sup>17</sup>Siehe Abschnitt ?? auf Seite ??.

Sei nun  $(x, y, \theta)$  eine kritische Plazierung mit Kontaktpaaren  $O_1, O_2, O_3$  und  $\mathcal{A}(O_i)$  das Arrangement der Funktionsgraphen  $f_{O_i O_k}$  mit  $O_k$  beschränkt  $O_i$ . Folgende Fälle können eintreten:

**1. Fall:** Zwei Kontaktpaare beschränken das dritte, o.E.  $O_2$  und  $O_3$  beschränken  $O_1$ .

Betrachte das Arrangement  $\mathcal{A}(O_1)$  der Funktionsgraphen von  $f_{O_1 O}$  mit  $O$  beschränkt  $O_1$ . Dann ist der Punkt  $(\theta, \rho_1)$  mit  $\rho_1 = f_{O_1 O_2}(\theta) = f_{O_1 O_3}(\theta)$  ein Eckpunkt der unteren Kontur im Arrangement  $\mathcal{A}(O_1)$ . Denn wäre  $f_{O_1 O_2}(\theta) \neq \min_O f_{O_1 O}(\theta)$ , z. B.  $f_{O_1 O_2}(\theta) > f_{O_1 O'}(\theta)$ , so könnte  $(x, y, \theta)$  wegen zu großem  $\rho$ -Wert im Vergleich zu  $f_{O_1 O'}(\theta)$  nicht erlaubt sein. Ebenso für  $f_{O_1 O_3}(\theta)$ .

Wir haben hierbei angenommen, daß bei  $O_1$  stets derselbe Endpunkt der involvierten Kanten zur Definition von  $\rho_1 = f_{O_1 O}(\theta)$  dient. Es kann aber auch der andere Endpunkt sein, in diesem Fall muß man zwei Klassen  $\mathcal{A}_L(O_1)$  und  $\mathcal{A}_R(O_1)$  bilden und eine Fallunterscheidung durchführen.

Das Arrangement  $\mathcal{A}(O_1)$  besteht aus  $2mn$  Funktionen, von denen je zwei höchstens vier Schnittpunkte haben. Nach Theorem ?? auf Seite ?? hat die untere Kontur die Komplexität  $O(\lambda_6(mn))$ . Für ein festes  $O_1$  gibt es also höchstens  $O(\lambda_6(mn))$  solcher Plazierungen, damit insgesamt höchstens  $O(mn \lambda_6(mn))$ .

**2. Fall:** In  $O_1, O_2, O_3$  gibt es keine zwei Paare, die das dritte beschränken. Dann liegt nach Lemma ?? ein Zyklus vor, nehmen wir o.E. an  $O_1$  beschränkt  $O_3$ ,  $O_3$  beschränkt  $O_2$  und  $O_2$  beschränkt  $O_1$ .

Betrachte alle Knoten der unteren Konturen in sämtlichen Arrangements  $\mathcal{A}(O)$ . Diese liefern eine Zerlegung des Intervalls  $[0, 2\pi)$  in  $O(mn \lambda_6(mn))$  viele Teilintervalle  $I_k$ , innerhalb derer die untere Kontur eines jeden Arrangements eindeutig von einer der Funktionen bestimmt wird.

$$0 \qquad I_k \quad I_{k'} \qquad 2\pi$$

Sei  $O_1$  ein beliebiges Kontaktpaar. Die Frage ist nun: gibt es über  $I_k$  ein zyklisches Tripel von Kontaktpaaren, an dem  $O_1$  beteiligt ist? Und falls ja, bilden sie auch eine kritische Plazierung?

Zur Beantwortung dieser Frage, gehe wie folgt vor:

1. Sehe in  $\mathcal{A}(O_1)$  nach, welche Funktion  $f_{O_1O_2}$  die untere Kontur über  $I_k$  bildet.
2. Sehe in  $\mathcal{A}(O_2)$  nach, welche Funktion  $f_{O_2O_3}$  die untere Kontur über  $I_k$  bildet.
3. Sehe in  $\mathcal{A}(O_3)$  nach, ob die Funktion  $f_{O_3O_1}$  die untere Kontur über  $I_k$  bildet. Wenn ja, ist ein Zyklus gefunden.

Dies zeigt: in einem Intervall  $I_k$  kann es für jedes  $O_1$  höchstens ein zyklisches Tripel geben, an dem  $O_1$  beteiligt ist, insgesamt also  $O(mn)$  viele.

Sei nun  $I_{k'}$  ein Nachbarintervall von  $I_k$  und  $\theta^*$  der gemeinsame Endpunkt. In  $\theta^*$  sind drei Kontaktpaare  $O_1, O_2$  und  $O_3$  beteiligt. Nur in den Arrangements dieser Kontaktpaare können sich die unteren Konturen beim Übergang von  $I_k$  zu  $I_{k'}$  ändern, also kann es über  $I_{k'}$  höchstens  $O(1)$  neue zyklische Tripel geben. Wir haben  $O(mn \lambda_6(mn))$  viele Teilintervalle, in jedem kommen  $O(1)$  viele Tripel hinzu. Für jedes zyklische Tripel kann es höchstens vier kritische Plazierungen geben, insgesamt also  $O(mn \lambda_6(mn))$ .  $\square$

### Bemerkung 2.36

- (i) Wir haben für  $m = 2$  (Leiter) in Abbildung ?? auf Seite ?? ein Beispiel mit  $\Theta(n^2)$  möglichen kritischen Plazierungen gesehen.
- (ii) Man kann auch für konvexe Roboter mit  $m > 2$  Ecken Hindernisse mit  $n$  Ecken so gestalten, daß  $\Theta(m^2n^2)$  viele kritische Plazierungen entstehen.
- (iii) Die Frage, ob es sogar  $\Theta(mn \lambda_6(mn))$  viele kritische Plazierungen geben kann, ist offen.

$$\rho \quad \mathcal{A}(O_1) \\ f_{O_1 O_2}(\theta)$$

**Definition 2.37**

Sei  $T^+$  die Menge aller Plazierungen  $(x, y, \theta)$ ,

- (i) die eindeutig den Ecken der unteren Konturen der  $O(mn)$  vielen Arrangements  $\mathcal{A}(O_i)$  entsprechen oder
- (ii) die einem zyklischen Tripel von Kontaktpaaren  $(O_1, O_2, O_3)$  entsprechen, mit  $O_1$  beschränkt  $O_3$ ,  $O_3$  beschränkt  $O_2$  und  $O_2$  beschränkt  $O_1$ .

Jedes Tripel liefert  $\leq 4$  Plazierungen, und jede der 4 möglichen Plazierungen ist durch  $(\rho_1, \theta)$  (oder  $(\rho_2, \theta)$ ,  $(\rho_3, \theta)$ ) eindeutig bestimmt. Dies sind Nullstellen einer Kurve 4. Grades.

$$\rho_1$$

$$\rho \quad \mathcal{A}(O_2) \quad f_{O_2 O_3}(\theta)$$

$$\rho_2$$

$$\rho \quad \mathcal{A}(O_3) \\ f_{O_3 O_1}(\theta)$$

$$\rho_3$$

$$\theta$$

$T^+$  enthält alle kritischen Plazierungen, aber eventuell auch verbotene Plazierungen: es könnte z. B. ein Kontaktpaar  $O_4$  einen verbotenen Schnitt bei Kontakt in  $O_1$  liefern, der in  $\mathcal{A}(O_1)$  nicht sichtbar ist, weil  $O_4$  nicht  $O_1$  beschränkt. Wir werden diese Plazierungen später herausfiltern; vorerst sind wir daran interessiert, die Menge  $T^+$  zu berechnen.

**Theorem 2.38** (Leven, Sharir, 1987)

Die Menge  $T^+$  läßt sich in Zeit  $O(mn \lambda_6(mn) \log(mn))$  berechnen. [?]

**Beweis.** Algorithmus ?? betrachtet insgesamt  $O(mn \lambda_6(mn))$  Intervalle mit je  $O(\log(mn))$  Einfügeaufwand und insgesamt  $O(mn \lambda_6(mn))$  Löschooperationen aus  $W \Rightarrow$  Zeit  $O(mn \lambda_6(mn) \log(mn))$ .  $\square$

Im nächsten Abschnitt werden wir den Konfigurationsraum des Roboters für feste Orientierungen  $\theta$  betrachten und beobachten, wie dieser sich für fortlaufende  $\theta$  ändert. Die Orientierungen an den kritischen Plazierungen sind dabei genau die Orientierungen, an denen sich die Struktur des Konfigurationsraums ändert.

**Algorithmus 2.8** Berechnung von  $T^+$ 

- Berechne für jedes Kontaktpaar  $O_i$ , über welchen Intervallen  $O_i$  von  $O_k$  beschränkt wird. Bilde daraus die Arrangements  $\mathcal{A}_{L/R}(O_i)$ .  $O(n^2m^2)$
- Für jedes der  $O(mn)$  Arrangements  $\mathcal{A}(O_i)$  berechne die untere Kontur (Divide-and-Conquer mit Sweep im Merge-Schritt). Da sich zwei Funktionsgraphen  $f_{O_1O_2}, f_{O_1O_3}$  höchstens viermal schneiden:  
 $O(\lambda_6(mn) \log(mn))$  pro Arrangement
- Kritische Plazierungen, Fall 1:  
 Berichte die Schnittpunkte von Funktionen auf den unteren Konturen in den Arrangements.
- Kritische Plazierungen, Fall 2:
  - Sortiere die  $O(mn)$  Listen der Eckpunkte der unteren Konturen zu einer Liste  $L = I_0, I_1, \dots$ .  
 Wegen  $\log(\lambda_6(mn)) \leq \log(m^2n^2) = 2 \log(mn)$ :  
 $mn \lambda_6(mn) \log(mn \lambda_6(mn)) \in O(mn \lambda_6(mn) \log(mn))$
  - Beginne mit dem linken Intervall  $I_0 \in L$ .  
 Für jedes  $O_1$  teste, ob über  $I_0$  ein Zyklus  $(O_1, O_2, O_3)$  existiert.  
 Wenn ja, berechne die kritischen Plazierungen wie folgt:
    - Berechne die Winkel  $\theta_{11}, \theta_{12}, \dots$  der maximal vier zugehörigen kritischen Plazierungen, an denen der Zyklus beteiligt ist.
    - Berichte sofort für die in  $I_0$  enthaltenen Orientierungen  $\theta_i$  die Plazierungen  $(\theta_i, (O_1, O_2, O_3))$ .
    - Füge die übrigen Plazierungen  $(\theta_j, (O_1, O_2, O_3))$  in eine nach Winkeln sortierte Warteschlange  $W$  ein.
  - Bearbeite von links nach rechts die übrigen Intervalle  $I_k \in L$ .
    - Beim Übergang von  $I_{k-1}$  nach  $I_k$  kann nur in einem Arrangement  $\mathcal{A}(O_1)$  ein Wechsel der unteren Kontur stattfinden.  $f_{O_1O_2}$  werde von  $f_{O_1O_4}$  abgelöst.
    - Falls  $O_1$  über  $I_{k-1}$  einen Zyklus mit  $O_2$  hatte, entferne aus  $W$  die zu  $(O_1, O_2, O_3)$  gehörenden kritischen Plazierungen; diese sind nicht mehr aktuell.
    - Teste, ob  $O_1, O_4$  einen Zyklus  $(O_1, O_4, O_5)$  über  $I_k$  bilden. Wenn ja, berechne wie oben die kritischen Plazierungen.
    - Berichte alle in  $W$  gespeicherten Plazierungen  $(\theta, (O_i, O_k, O_\ell))$  mit  $\theta \in I_k$ , und entferne sie aus  $W$ .

$$O(mn \lambda_6(mn) \log(mn))$$

### 2.3.2 Knotengraph und Kantengraph

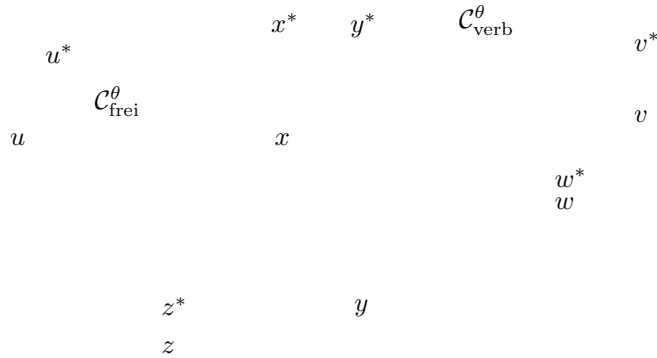


Abbildung 2.38:  $C_{\text{frei}}^\theta$ ,  $C_{\text{verb}}^\theta$  und Hilfsknoten von  $V^\theta$ .

Sei zunächst die Orientierung  $\theta$  fest. Wir betrachten den Schnitt  $C_{\text{frei}}^\theta$  durch den Konfigurationsraum  $C_{\text{frei}}$ . In  $C_{\text{frei}}^\theta$  brauchen wir nur Translationen zu betrachten, und wir wissen aus den vorherigen Abschnitten, daß für einen konvexen Roboter  $C_{\text{frei}}^\theta$  die Komplexität  $O(mn)$  hat und sich in Zeit  $O(mn \log^2(mn))$  berechnen läßt. Dabei hatten wir die Hindernisse in konvexe Teile, z. B. Dreiecke zerlegt. Zur diskreten Darstellung des Konfigurationsraumes hatten wir die Trapezzerlegung benutzt; diese Methode läßt sich jedoch nicht gut auf drei Dimensionen erweitern. Deshalb benutzen wir hier zur Darstellung von  $C_{\text{frei}}^\theta$  den Knotengraph  $V^\theta$ .

Wir gehen im folgenden davon aus, daß alle Zusammenhangskomponenten von  $C_{\text{frei}}$  beschränkt sind, sich also Wände um die Szene befinden; siehe Abbildung ???. Der Rand von  $C_{\text{frei}}^\theta$  besteht dann aus endlich vielen Polygonen.

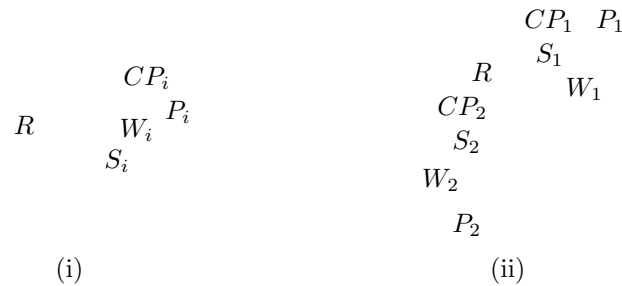


Abbildung 2.39: Konvexe und konkave Ecken in  $\partial C_{\text{verb}}^\theta$ .

In  $\mathcal{C}_{\text{verb}}^\theta$  gibt es zwei Typen von Ecken: Konvexe Ecken entstehen durch Ecke/Ecke-Kontakte  $O_i = (W_i, S_i)$  von  $R$  mit einer konvexen Hindernisecke (Kontaktpaar vom Typ III), konkave Ecken durch zwei Ecke/Kante-Kontaktpaare (Typ I oder II), siehe Abbildung ??.

**Definition 2.39** Zu einer festen Orientierung  $\theta$  enthält der **Knotengraph**  $V^\theta$ :

- Alle Knoten der Polygone in  $\mathcal{C}_{\text{verb}}^\theta$ .
- Für jeden lokal  $y$ -maximalen Knoten  $u$  auf  $\partial\mathcal{C}_{\text{verb}}^\theta$  den ersten Punkt  $u^*$  (“Hilfsknoten”) oberhalb  $u$ .
- Kanten zwischen adjazenten Knoten (einschließlich Hilfsknoten) und zwischen  $y$ -maximalen Knoten und entsprechenden Hilfsknoten, vgl. Abbildung ??.
- Folgende Beschriftungen:
  - Konvexer Knoten von  $\mathcal{C}_{\text{verb}}^\theta$ : Ecke/Ecke-Paar  $O_i = (W_i, S_i)$ .
  - Konkaver Knoten von  $\mathcal{C}_{\text{verb}}^\theta$ : zwei Ecke/Kante-Paare  $O_i, O_j$ .
  - Hilfsknoten  $u^*$ : die Bezeichnung “Hilfsknoten” und die Beschriftung des erzeugenden Knoten  $u$ .
  - Kante von  $\mathcal{C}_{\text{verb}}^\theta$ : Ecke/Kante-Paar  $O_i = (W_i, S_i)$ . Vergleiche  $(S_1, W_1)$  bzw.  $(S_2, W_2)$  in Abbildung ?? (ii).

Bei festem  $\theta$  sind die Knoten durch ihre Beschriftungen eindeutig beschrieben. Wir verwalten zusätzlich eine Suchstruktur, die es erlaubt, anhand einer Beschriftung den zugehörigen Knoten aufzufinden.

**Bemerkung 2.40**

- (i)  $V^\theta$  ist ein planarer Graph mit  $O(mn)$  Knoten und Kanten.
- (ii) Jeder Knoten hat einen Grad  $\leq 3$ , falls sich die Ecken der Hindernisse in allgemeiner Lage befinden, d.h. keine Ecken mit gleichen  $X$ -Koordinaten.
- (iii)  $V^\theta$  läßt sich in Zeit  $O(mn \log(mn))$  konstruieren.
- (iv) Zu gegebener Beschriftung läßt sich der zugehörige Knoten in Zeit  $O(\log(mn))$  finden.

Den Zusammenhang zwischen dem Knotengraph und der Bewegungsplanung stellt folgendes Lemma her:

**Lemma 2.41** Seien  $u, v$  zwei Knoten von  $V^\theta$ . Dann gilt:

- $u, v$  liegen in derselben Zusammenhangskomponente von  $V^\theta$
- $\iff u, v$  liegen auf dem Rand derselben Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}^\theta$ .



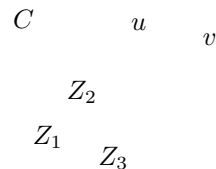
**Beweis.**

“ $\implies$ ” Falls  $(u, v)$  eine Kante von  $V^\theta$  bildet, liegen nach Definition ??  $u$  und  $v$  auf dem Rand derselben Zusammenhangskomponente von  $C_{\text{frei}}^\theta$ ; auch falls  $v = u^*$ .

Falls  $u, v$  in  $V^\theta$  über mehrere Kanten verbunden sind, folgt die Behauptung durch Induktion.

“ $\impliedby$ ” Angenommen,  $u$  und  $v$  liegen auf dem Rand derselben Zusammenhangskomponente  $C$  von  $C_{\text{frei}}^\theta$ .

Dieser Rand kann selbst in mehrere Komponenten  $Z_1, \dots, Z_s$  zerfallen, siehe Abbildung. Liegen  $u$  und  $v$  in derselben Komponente, so sind sie nach Definition ?? in  $V^\theta$  durch einen Kantenzug verbunden. Ansonsten nehmen wir an,  $u \in Z_2$ ,  $v \in Z_3$  und  $Z_1$  sei die äußere Komponente von  $\partial C$ . Seien  $V_u^\theta, V_v^\theta$  die Zusammenhangskomponenten von  $u$  und  $v$  in  $V^\theta$ .



**Behauptung:**  $V_u^\theta, V_v^\theta$  enthalten Knoten von  $Z_1$ .

**Beweis:**

Angenommen,  $V_u^\theta$  enthält keinen Knoten von  $Z_1$ . Sei  $z$  der höchstgelegene Knoten in  $V_u^\theta$ .

- $\implies$   $z$  ist lokal  $y$ -maximal.
- $\implies$  Der Strahl von  $z$  aufwärts muß durch  $C_{\text{frei}}^\theta$  verlaufen, da  $z$  nicht auf der äußeren Komponente von  $\partial C$  liegt.
- $\implies$  Es gibt einen Hilfsknoten  $z^*$ .
- $\implies$   $z^* \in V_u^\theta \quad \nabla$  Maximalität von  $z$ .  $\diamond$

Also enthalten beide Komponenten Knoten von  $Z_1 \implies V_u^\theta = V_v^\theta$ .  $\square$

Die Frage ist nun, was mit  $V^\theta$  passiert, wenn sich  $\theta$  ändert. Lassen wir  $\theta$  stetig wachsen, so verändern sich auch die Knoten und Kanten von  $V^\theta$  stetig, bis ein bestimmtes  $\theta_k$  erreicht wird, an dem sich auch die Struktur von  $V^\theta$  ändert. Eine Orientierung, an der in  $V^\theta$  eine Strukturänderung auftritt, nennen wir eine **kritische Orientierung**.

Die Abbildungen ?? bis ?? zeigen kritische Orientierungen. Die dunkel schraffierten Teile sind Hindernisse, die hell schraffierten Teile Konfigurationsraum-Hindernisse für die Translation von  $R$ . Die gestrichelten Segmente sind die Kanten der Konfigurationsraum-Hindernisse. Dargestellt ist, wie sich  $C_{\text{frei}}$

ändert, wenn  $R$  gedreht wird. Dabei können alle Ereignisse auch “rückwärts” ablaufen, z. B. kann in (i) auch eine neue Kante von  $\partial \mathcal{C}_{\text{frei}}^\theta$  entstehen. Die kritischen Orientierungen (i)–(iv) entstehen aus den Orientierungen der kritischen Plazierungen, (v) und (vi) müssen wegen der Aktualisierung von Hilfsknoten beachtet werden.

Kante von  $\mathcal{C}_{\text{frei}}$  (i)

Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}$  (ii)

(iii)

Abbildung 2.40: (i) eine Kante verschwindet, (ii) eine Komponente von  $\mathcal{C}_{\text{frei}}$  verschwindet, (iii) ein lokaler Zusammenhang verschwindet.

- (i) Drei Ecke/Kante–Kontaktpaare: eine Kante von  $\partial \mathcal{C}_{\text{frei}}^\theta$  verschwindet, ein “wanderndes” Kontaktsegment verändert  $\mathcal{C}_{\text{frei}}^\theta$ .
- (ii) Drei Ecke/Kante–Kontaktpaare: eine Komponente von  $\mathcal{C}_{\text{frei}}^\theta$  verschwindet; ein “wanderndes” Kontaktsegment hat  $\mathcal{C}_{\text{frei}}^\theta$  verkleinert.
- (iii) Ein Ecke/Ecke–Kontaktpaar, ein Ecke/Kante–Kontaktpaar: ein lokaler

Zusammenhang von  $\mathcal{C}_{\text{frei}}^\theta$  verschwindet, eine “wandernde” konvexe Ecke hat den Durchgang abgeschnitten.

Die Anzahl der kritischen Orientierungen von Typ (i)–(iii) ist kleiner gleich der Anzahl der kritischen Plazierungen, also in  $O(mn \lambda_6(mn))$ . Die Aktualisierung von  $V^\theta$  ist pro Ereignis in  $O(\log(mn))$  Suchzeit +  $O(1)$  Änderungsaufwand möglich. Die folgenden Fälle verursachen mehr Aufwand.

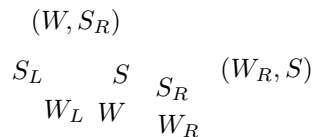


Abbildung 2.41: Zwei benachbarte Kanten werden kollinear.

- (iv) Ein Ecke/Ecke-Kontaktpaar und ein Ecke/Kante-Kontaktpaar (inzident): eine Roboterkante wird parallel zu einer Hinderniskante und zwei benachbarte Kanten werden kollinear, siehe Abbildung ??.

In diesem Fall ändern sich die Beschriftungen der Kanten des erweiterten Hindernisses wie folgt:

$$\begin{aligned} (W, S_R) &\longmapsto (W_L, S) \\ (W_R, S) &\longmapsto (W, S_L) \end{aligned}$$

- ⇒ Jede Hinderniskante kann  $O(mn)$  Kanten zu  $V^\theta$  beitragen
- ⇒ Such- und Aktualisierungsaufwand in  $V^\theta$ :  $O(mn \log(mn))$ .
- ⇒ Dieses Ereignis kann maximal  $O(mn)$  mal auftreten, der Gesamtaufwand liegt also in  $O((mn)^2 \log(mn))$ .

Abbildung 2.42: Eine Kante wird waagrecht, das lokale Y–Maximum ändert sich.

- (v) Ein Ecke/Kante–Kontaktpaar: eine Kante von  $R$  und damit eine Kante eines Konfigurationsraumhindernisses wird waagrecht. Dabei kann sich die Menge der konvexen,  $y$ –maximalen Knoten ändern.
- Es können sich maximal  $n$  Knoten ändern.
  - Das Ereignis kann  $m$  mal auftreten (für jede Kante von  $R$ ).
  - Für jeden Knoten  $u$ , der nicht länger  $y$ –maximal ist, ist der Hilfsknoten  $u^*$  aus  $V^\theta$  zu entfernen sowie die neuen  $y$ –maximalen Knoten  $u'$  und ihre Hilfsknoten  $u'^*$  zu bestimmen und in  $V^\theta$  einzufügen (Sweep über  $\mathcal{C}_{\text{frei}}^\theta$  in Zeit  $O(mn \log(mn))$  ).

Der Zeitbedarf ist also insgesamt  $O(m^2n \log(mn))$ .



Abbildung 2.43: Ein Hilfsknoten wandert auf eine Ecke.

- (vi) Ein Kontaktpaar Ecke/Ecke: eine konvexe,  $y$ -maximale Ecke liegt direkt unterhalb einer Ecke. Der Hilfsknoten wandert auf die Nachbarkante. Die Aktualisierung des Hilfsknotens ist nach  $O(\log(mn))$  Suchzeit in Zeit  $O(1)$  möglich. Dieses Ereignis tritt maximal  $O(mn \lambda_6(mn))$  mal auf; dies läßt sich mit einem ähnlichen Beweis wie für die Anzahl der kritischen Plazierungen zeigen.

Damit wissen wir:

**Bemerkung 2.42** *Der Gesamtaufwand für die Aktualisierung von  $V^\theta$  beträgt  $O(mn \lambda_6(mn) \log(mn))$ .*

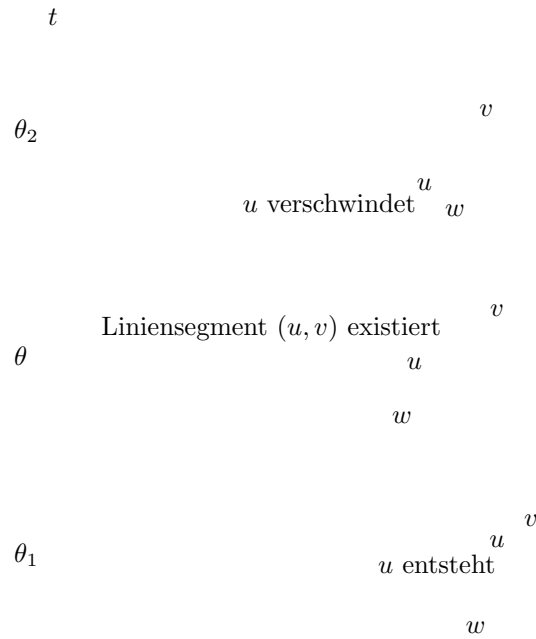
Wir haben jetzt einen Graph  $V^\theta$  definiert, der für die Berechnung von Translationsbewegungen benutzt werden kann, kritische Orientierungen  $\theta$  untersucht, bei denen sich die Struktur von  $V^\theta$  ändert, und betrachtet, wie  $V^\theta$  aktualisiert werden muß. Jetzt wollen wir einen weiteren Graphen definieren, den Kantengraph  $E$ , der die gesamten Informationen über alle  $V^\theta, \theta \in [0, 2\pi)$ , enthält und zur Bewegungsplanung verwendet werden kann.

Die Idee dabei ist diese: für wachsende Orientierung  $\theta$  beschreibt ein Knoten von  $C_{\text{frei}}^\theta$  eine — möglicherweise gekrümmte — Kurve im Raum. Stelle nun in  $E$  diese Kurve als erweiterten Knoten mit seiner “Lebensdauer” als  $\theta$ -Intervall dar. Genauer:

**Definition 2.43** Sei  $\theta$  eine nicht-kritische Orientierung; sei  $u$  ein Knoten in  $V^\theta$  (bzw. seine Beschriftung, also die Namen der verursachenden Kontaktpaare).

- (i) Die **Lebensdauer** von  $u$  sei das maximale Intervall  $L(u) = (\theta_1, \theta_2)$ , in dem der Knoten  $u$  in jedem Knotengraph  $V^\theta, \theta \in L(u)$ , vorkommt und nicht mit einem anderen Knoten  $u'$  zusammenfällt.

- (ii) Der **Kantengraph**  $E$  bestehe aus den erweiterten Knoten  $(u, L(u))$ . Zwei Knoten  $(u, L(u))$  und  $(v, L(v))$  sind in  $E$  genau dann durch eine Kante verbunden, wenn es ein  $\theta \in L(u) \cap L(v)$  gibt, so daß in  $V^\theta$  die Kante  $(u, v)$  existiert. Auch für eine Kante  $e$  kann eine Lebensdauer  $L(e)$  definiert werden.

Abbildung 2.44: Kantengraph  $E$ .

Anschaulich entspricht jeder Kante von  $\mathcal{C}_{\text{frei}}$  ein erweiterter Knoten  $(u, L(u))$  in  $E$  und jedem Stück einer Fläche von  $\mathcal{C}_{\text{frei}}$  (zwischen den  $\theta$ -Werten von zwei Knoten) eine Kante  $[(u, L(u)), (v, L(v))]$  von  $E$  in folgendem Sinn: Seien  $e_u, e_v$  die Kanten von  $\mathcal{C}_{\text{frei}}$ , die den Knoten  $(u, L(u))$  und  $(v, L(v))$  entsprechen, dann verbindet für jedes  $\theta \in L((u, v))$  die Kante  $(u, v)$  in  $\mathcal{C}_{\text{frei}}^\theta$  zwei Punkte von  $e_u$  und  $e_v$ . Wir haben also eine komplexe, räumliche Struktur auf eine einfachere, kombinatorische Struktur abgebildet.

### 2.3.3 Berechnung der kritischen Orientierungen

Bevor wir uns überlegen, wie wir das Bahnplanungsproblem mit Hilfe des Kantengraphen lösen können, müssen wir uns überlegen, wie wir die kritischen Orientierungen berechnen können.

Wir berechnen dazu die Menge  $T^*$  aller kritischen Orientierungen.  $T^*$  enthält die Orientierungen der kritischen Plazierungen  $T^+$  (siehe Algorithmus ??), sowie die Orientierungen, an denen sich die Hilfsknoten von  $V^\theta$  ändern, also:

- Allen Eckpunkten von unteren Konturen der Arrangements  $\mathcal{A}(O)$  von Kontaktpaaren  $O$ .
- Alle Orientierungen, die zu zyklischen Tripeln  $O_1, O_2, O_3$  gehören.
- Allen Orientierungen, bei denen die Kante eines erweiterten Hindernisses waagrecht wird.
- Allen Orientierungen, bei denen eine konvexe Ecke von  $\mathcal{C}_{\text{verb}}^\theta$  direkt unterhalb einer anderen Ecke liegt.

Das Problem dabei ist, daß  $T^*$  eine Obermenge der gültigen kritischen Orientierungen ist. Es kann in  $T^*$  Orientierungen geben, zu denen keine halbfreien Plazierungen von  $R$  existieren. Wir müssen also einige "falsche" Orientierungen  $\theta_i \in T^*$  ignorieren.

Ein erster Ansatz dazu ist folgender: Angenommen, bei  $\theta_i$  liegt z. B. ein dreifacher Ecke/Kante-Kontakt  $(O_1, O_2, O_3)$  vor. Falls dadurch eine halbfreie Plazierung gegeben ist, muß bei  $\theta_i - \varepsilon$  jedes der drei Kontaktpaare eine Kante zum Rand von  $\mathcal{C}_{\text{frei}}^\theta$ , also auch zu  $V^\theta$ , beigetragen haben (Stetigkeitsargument). Dies läßt sich in Zeit  $O(\log(mn))$  testen. Dazu verwalten wir zu jedem Kontakt  $O$  eine Liste  $\mathcal{L}_O(\theta)$  aller Kontakte  $O'$  die zusammen mit  $O$  einen Knoten in  $\mathcal{C}_{\text{frei}}^\theta$  erzeugen. Ist  $\theta$  eine kritische Orientierung mit der Stetigkeitseigenschaft, in die die Kontaktpaare  $O, O', O''$  involviert sind, so muß mindestens eine der Listen  $\mathcal{L}_O(\theta), \mathcal{L}_{O'}(\theta), \mathcal{L}_{O''}(\theta)$  einen der beiden anderen Kontakte enthalten.

Damit wird jedoch folgender Fall nicht gelöst: Bei  $\theta_i$  entsteht eine neue Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}^\theta$  — zunächst als Punkt (Abbildung ??(ii)) —, aber diese Plazierung  $P_i$  ist nicht frei, weil  $R$  ein weiteres Hindernis  $P_i$  enthält.

Das Stetigkeitsargument ist in diesem Fall nicht anwendbar. Außerdem ist der Konflikt zwischen  $R$  und  $P_i$  bei  $\theta_i$  nicht bekannt, und wir können es uns nicht leisten, für jede dieser  $O(mn \lambda_6(mn))$  möglichen Plazierungen alle  $n$  Hindernisdreiecke zu testen. Hinzu kommt das technische Problem, bei der Aktualisierung von  $V^\theta$  zu entscheiden, in welcher Komponente von  $\mathcal{C}_{\text{verb}}^\theta$  die neue Komponente von  $\mathcal{C}_{\text{frei}}^\theta$  liegt. Dies erfordert eine Hilfsstruktur zur Lokalisierung des ersten Punktes der neuen Komponente.

**Algorithmus 2.9** Berechnung von  $T'$ 

**Output:** Die Menge  $T' \subseteq T^*$  aller kritischen Orientierungen, an denen eine konvexe Ecke von  $\mathcal{C}_{\text{verb}}^\theta$  entsteht.

Für jedes Ecke/Ecke Kontaktpaar mit Kontaktecke  $v$ :

1. Rotiere  $R$  um  $v$  und berechne, durch Schnitttest von jeder Roboterecke bzw. Roboterseite mit jeder Hindernisecke bzw. Hindernisseite, in welchen Winkelbereichen der Roboter ein Hindernis schneidet.

Wir erhalten  $O(mn)$  verbotene Winkelintervalle in Zeit  $O(mn)$ .

2. Berechne die Vereinigung dieser Intervalle durch Sortieren.  
 $O(mn \log(mn))$

3.  $T'$  enthält die Endpunkte der vereinigten Intervalle.

Um dieses Problem zu lösen, berechnen wir nur noch die Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}$  korrekt, die die Startposition  $s$  enthält — für die Bahnplanung genügt dies bekanntlich. Die Lösung basiert auf zwei Beobachtungen:

- (i) Eine neue Zusammenhangskomponente ist nur dann interessant, wenn sie sich irgendwann mit der Zusammenhangskomponente  $Z_s$  vereinigt, die  $s$  enthält — nicht erreichbare Plazierungen sind uninteressant.
- (ii) Wenn sich zwei Komponenten vereinigen, ist eine konvexe Ecke von  $\mathcal{C}_{\text{verb}}^\theta$  beteiligt, siehe Abbildung ?? oder Abbildung ??(iii) auf Seite ??.

Solche konvexen Ecken entstehen durch Ecke/Ecke-Kontakte (Kontaktpaare vom Typ III, siehe Definition ?? auf Seite ??), von denen es nur  $O(mn)$  viele gibt. Unter Aufrechterhaltung eines Ecke/Ecke-Kontaktes hat  $R$  nur noch einen Freiheitsgrad: die Rotation um die Kontaktecke. Wir können für jeden Ecke/Ecke-Kontakt in Zeit  $O(mn \log(mn))$  alle zugehörigen, freien kritischen Plazierungen von  $R$  finden. Insgesamt benötigen wir also Zeit  $O(m^2 n^2 \log(mn))$ , um mit Algorithmus ?? alle kritischen Orientierungen zu berechnen, an denen eine konvexe Ecke von  $\mathcal{C}_{\text{verb}}^\theta$  entsteht.

Mit Hilfe der Menge  $T'$  können wir die Orientierungen berechnen, an denen eine für die Bahnplanung interessante Komponente entsteht. Komponenten, die in jedem  $\mathcal{C}_{\text{frei}}^\theta$  konvex sind und im initialen  $\mathcal{C}_{\text{frei}}^{\theta_0}$  nicht existieren, werden nicht mehr berechnet.

Nehmen wir an, bei  $\theta_i$  vereinige sich eine neue Komponente  $N$  mit der Komponente  $Z_s$ , siehe Abbildung ?. Verfolgt man das Schicksal von  $v_1$  und  $v_2$  für abnehmende Werte von  $\theta$ , so muß eine der beiden Komponenten, auf deren Rändern  $v_1$  und  $v_2$  liegen, verschwinden. Glücklicherweise ist für diesen Test nicht der gesamte Konfigurationsraum  $\mathcal{C}_{\text{frei}}$  zu betrachten, wir



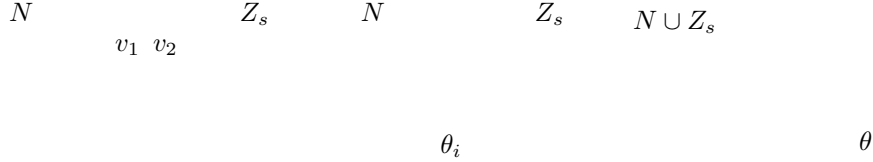


Abbildung 2.45: Vereinigung zweier Komponenten  $N$  und  $Z_s$ .

können diesen Test auch mit Hilfe der Listen  $\mathcal{L}_O(\theta)$  durchführen, wie in Algorithmus ?? beschrieben. Wir verwalten dabei *nur* gültige Kontaktpaare in den Listen  $\mathcal{L}_O(\theta)$  und verfolgen den Pfad von konvexen Ecken von  $\mathcal{C}_{\text{verb}}^\theta$  von ihrer Entstehung zurück bis zur Entstehung der Komponente, in der sie enthalten sind. Wir beobachten somit nur Komponenten, die sich von  $Z_s$  abgespalten haben und wissen zum Zeitpunkt ihres Verschwindens, daß in Richtung zunehmender  $\theta$ -Werte gesehen an dieser Stelle eine für die Bahnplanung wichtige Komponente entsteht, dieser Zeitpunkt wird also in die Outputmenge  $T''$  eingetragen.  $T''$  enthält dann alle gültigen kritischen Orientierungen, bei denen eine neue interessante Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}$  entsteht.

Algorithmus ?? berechnet schließlich mit Hilfe von  $T^*, T'$  und  $T''$  die Menge  $T$  aller gültigen kritischen Orientierungen. Wie schon in Algorithmus ?? verwalten wir nur gültige Kontaktpaare in den Listen  $\mathcal{L}_O(\theta)$  und ignorieren die Orientierungen, an denen eine Komponente entsteht, die für die Bahnplanung uninteressant oder ungültig ist.

Fassen wir die zur Konstruktion von  $T$  benötigten Schritte und deren Laufzeiten zusammen:

Kritische Orientierungen $T^*$	$O(mn \lambda_6(mn) \log(mn))$
Orientierungen $T'$ (konvexe Ecken von $\mathcal{C}_{\text{verb}}^\theta$ )	$O(m^2 n^2 \log(mn))$
$T''$ (interessante Komponenten)	$O(mn \lambda_6(mn) \log(mn))$
$T$ aus $T^*, T'$ und $T''$	$O(mn \lambda_6(mn) \log(mn))$

Wir wissen aus Theorem ?? auf Seite ?? und Bemerkung ?? auf Seite ??, daß  $|T^+|$  in  $O(mn \lambda_6(mn))$  liegt und sich in Zeit  $O(mn \lambda_6(mn) \log(mn))$  berechnen läßt. In dieser Zeit lassen sich auch die gültigen kritischen Orientierungen berechnen.

---

**Algorithmus 2.10** Berechnung von  $T''$ 


---

**Output:** Die Menge  $T'' \subseteq T^*$  aller kritischen Orientierungen, bei denen lokal ein neuer Zusammenhang entsteht.

**Datenstruktur:**

Für jedes Kontaktpaar  $O$  unterhalte eine nach  $\theta$  sortierte Liste  $\mathcal{L}_O(\theta)$  aller Kontakte  $O'$  die zusammen mit  $O$  einen Knoten in  $\mathcal{C}_{\text{frei}}^\theta$  erzeugen.

**Initialisierung:**

Berechne  $\mathcal{C}_{\text{frei}}^{\theta_0}$  für die Startorientierung  $\theta_0$  und initialisiere die Listen  $\mathcal{L}_O(\theta_0)$  mit den Kontaktpaaren  $O, O'$ , die in  $\mathcal{C}_{\text{frei}}^{\theta_0}$  einen Knoten erzeugen.

Bearbeite nach fallendem Winkel alle kritischen Orientierungen  $\theta_i \in T^*$ :

- Falls  $\theta_i \in T'$ :
    - Seien  $O_1, O'_1$  die Kontaktpaare von  $v_1$  und  $O_2, O'_2$  die von  $v_2$ . Füge ein:
      - $O'_1$  in  $\mathcal{L}_{O_1}(\theta_i)$
      - $O_1$  in  $\mathcal{L}_{O'_1}(\theta_i)$
      - $O'_2$  in  $\mathcal{L}_{O_2}(\theta_i)$
      - $O_2$  in  $\mathcal{L}_{O'_2}(\theta_i)$
  - Falls  $\theta_i \notin T'$ :
    - Seien  $O, O', O''$  die zugehörigen Kontaktpaare. Suche
      - nach  $O$  und  $O'$  in  $\mathcal{L}_{O''}(\theta_i)$
      - nach  $O$  und  $O''$  in  $\mathcal{L}_{O'}(\theta_i)$
      - nach  $O'$  und  $O''$  in  $\mathcal{L}_O(\theta_i)$
    - Falls Suche erfolglos: ignoriere  $\theta_i$ .
    - Falls z. B.  $O'$  in  $\mathcal{L}_O(\theta_i)$  und  $O$  in  $\mathcal{L}_{O'}(\theta_i)$  gefunden — in diesem Fall muß  $\theta_i$  wegen des Stetigkeitsarguments gültig (d. h. frei) sein —, berechne die lokalen Änderungen in  $\mathcal{C}_{\text{frei}}^\theta$  und aktualisiere die Listen  $\mathcal{L}_O(\theta_i), \mathcal{L}_{O'}(\theta_i), \mathcal{L}_{O''}(\theta_i)$  entsprechend.
    - Falls bei  $\theta_i$  eine Komponente verschwindet: füge  $\theta_i$  in die sortierte Outputmenge  $T''$  ein.
-

---

**Algorithmus 2.11** Berechnung von  $T$ 

---

**Output:** Die Menge  $T$  aller gültigen kritischen Orientierungen.

**Datenstruktur:**

Für jedes Kontaktpaar  $O$  unterhalte eine nach  $\theta$  sortierte Liste  $\mathcal{L}_O(\theta)$  aller Kontakte  $O'$  die zusammen mit  $O$  einen Knoten in  $\mathcal{C}_{\text{frei}}^\theta$  erzeugen.

**Initialisierung:**

Berechne  $\mathcal{C}_{\text{frei}}^{\theta_0}$  für die Startorientierung  $\theta_0$  und initialisiere die Listen  $\mathcal{L}_O(\theta_0)$  mit den Kontaktpaaren  $O, O'$ , die in  $\mathcal{C}_{\text{frei}}^{\theta_0}$  einen Knoten erzeugen.

Bearbeite nach steigendem Winkel alle kritischen Orientierungen  $\theta_i \in T^*$ :

- Teste, ob bei  $\theta_i$  evtl. eine neue Komponente entsteht.
  - Falls ja:
    - Falls  $\theta_i \notin T'$ : ignoriere  $\theta_i$ .
    - Ansonsten aktualisiere die Listen  $\mathcal{L}_O(\theta_i)$ .
  - Falls nein:
    - Seien  $O, O', O''$  die zugehörigen Kontaktpaare. Suche
      - nach  $O$  und  $O'$  in  $\mathcal{L}_{O''}(\theta_i)$
      - nach  $O$  und  $O''$  in  $\mathcal{L}_{O'}(\theta_i)$
      - nach  $O'$  und  $O''$  in  $\mathcal{L}_O(\theta_i)$
    - Falls z. B.  $O'$  in  $\mathcal{L}_O(\theta_i)$  und  $O$  in  $\mathcal{L}_{O'}(\theta_i)$  gefunden — in diesem Fall muß  $\theta_i$  wegen des Stetigkeitsarguments gültig (d. h. frei) sein —, berechne die lokalen Änderungen in  $\mathcal{C}_{\text{frei}}^\theta$  und aktualisiere die Listen  $\mathcal{L}_O(\theta_i), \mathcal{L}_{O'}(\theta_i), \mathcal{L}_{O''}(\theta_i)$  entsprechend.
    - Falls Suche erfolglos, teste  $\theta_i \in T'$ . Falls ja, entsteht eine konvexe Ecke von  $\mathcal{C}_{\text{verb}}^\theta$ , aber keine neue Komponente und  $\theta_i$  ist gültig, ansonsten wird  $\theta_i$  ignoriert.
  - $T$  besteht aus allen nicht ignorierten Orientierungen aus  $T^*$ .
-

### 2.3.4 Bahnplanung mit dem Kantengraph $E$

Nun stellt sich die Frage, wie uns der Kantengraph bei unserem Bahnplanungsproblem hilft. Wir definieren dazu folgende Abbildung:

**Definition 2.44** Die Abbildung  $\Phi$

$$\begin{aligned} \Phi : \mathcal{C}_{\text{frei}} &\longrightarrow \{ \text{Knoten von } E \} \\ (x, y, \theta) &\longmapsto \text{Knoten } (u, L(u)), \text{ der eine Kante } e_u \text{ von} \\ &\quad \mathcal{C}_{\text{frei}} \text{ repräsentiert, die auf dem Rand der} \\ &\quad \text{Zusammenhangskomponente } Z \text{ von } \mathcal{C}_{\text{frei}} \\ &\quad \text{liegt, die } (x, y, \theta) \text{ enthält.} \end{aligned}$$

sei wie folgt definiert: Verschiebe den Roboter  $R$  nach oben (in Richtung aufsteigender  $Y$ -Werte), bis bei der Plazierung  $(x, y', \theta)$  ein Kontakt mit einem Hindernis erfolgt. Falls bei dieser Plazierung ein Ecke/Ecke-Kontakt oder ein mehrfacher Ecke/Kante-Kontakt entsteht, wurde ein Knoten  $u = (x, y')$  von  $\mathcal{C}_{\text{frei}}^\theta$  gefunden, der auf einer Kante von  $\mathcal{C}_{\text{frei}}$  liegt, welche die Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}$  berandet, die  $(x, y, \theta)$  enthält. Setze  $\Phi(x, y, \theta) := (u, L(u))$ . Ansonsten verschiebe  $R$  unter Aufrechterhaltung des Kontaktes nach links, bis bei einer Ecke  $u = (x'', y'')$  von  $\mathcal{C}_{\text{frei}}$  ein Mehrfachkontakt entsteht und setze  $\Phi(x, y, \theta) := (u, L(u))$ .

Offensichtlich kann  $\Phi(x, y, \theta)$  in Zeit  $O(mn)$  ( $=$  Größe von  $V^\theta$ ) berechnet werden. Für die Bahnplanung ist entscheidend:

**Theorem 2.45** (Kedem, Sharir, 1990)

Seien  $s, t$  zwei (halb-)freie Plazierungen von  $R$ . Dann gilt: Es gibt eine kollisionsfreie Bewegung von  $s$  nach  $t$  genau dann, wenn  $\Phi(s)$  und  $\Phi(t)$  Knoten derselben Zusammenhangskomponente  $Z$  des Kantengraphen  $E$  sind. [?, ?]

Falls  $\Phi(s)$  und  $\Phi(t)$  Knoten derselben Zusammenhangskomponente  $Z$  sind, kann z. B. durch Breitensuche in  $Z$  ein Weg in  $E$  von  $\Phi(s)$  nach  $\Phi(t)$  gefunden und in eine halbfreie Bewegung von  $R$  umgeformt werden. Dies werden wir später noch genauer betrachten. Der Beweis des Theorems stützt sich im wesentlichen auf folgendes Lemma:

**Lemma 2.46** Jeder Zusammenhangskomponente  $W$  von  $\mathcal{C}_{\text{frei}}$  entspricht eindeutig eine Zusammenhangskomponente  $Z$  von  $E$ , deren Knoten genau die Kanten von  $\mathcal{C}_{\text{frei}}$  repräsentieren, die  $W$  begrenzen.

**Beweis.** Wir zeigen zwei Richtungen von

$$\text{Knoten von } Z \longleftrightarrow \text{Kanten in } \partial W.$$

“ $\longrightarrow$ ” Seien  $(u, L(u))$  und  $(v, L(v))$  Knoten derselben Zusammenhangskomponente  $Z$  von  $E$ . Zu zeigen ist, daß  $e_u, e_v$  Kanten derselben Zusammenhangskomponente  $W$  von  $\mathcal{C}_{\text{frei}}$  sind.

Falls  $[(u, L(u)), (v, L(v))]$  sogar eine Kante in  $Z$  ist

$\Rightarrow$  (Def. ??)  $(u, v)$  ist Kante in  $\mathcal{C}_{\text{frei}}^\theta$  für ein  $\theta \in L(u) \cap L(v)$ .

$\Rightarrow$   $e_u, e_v$  — die Kanten, welche durch  $u$  und  $v$  bei wachsendem  $\theta$  entstehen — liegen auf dem Rand derselben Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}$ .

Falls  $(u, L(u)), (v, L(v))$  durch einen Kantenzug in  $Z$  verbunden sind, folgt die Behauptung durch Transitivität.

“ $\leftarrow$ ” Seien  $e_u, e_v$  Kanten auf dem Rand derselben Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}$ . Zu zeigen ist, daß die Knoten  $(u, L(u))$  und  $(v, L(v))$  in derselben Zusammenhangskomponente von  $E$  liegen. Dieser Beweis erfordert etwas mehr Aufwand, da  $e_u, e_v$  bzgl.  $\theta$  “weit auseinander” liegen können. Wir benutzen dabei, daß für jedes feste  $\theta$  Zshkomp. von  $V^\theta \longleftrightarrow$  Zshkomp. von  $\mathcal{C}_{\text{frei}}^\theta$  gilt (Lemma ??, Seite ??).

Sei  $I$  ein offenes Intervall zwischen zwei kritischen Orientierungen — für alle  $\theta \in I$  haben also  $V^\theta$  und  $\mathcal{C}_{\text{frei}}^\theta$  dieselben Strukturen —, sei  $A$  eine Zusammenhangskomponente in  $V^\theta$  und  $Q(A, \theta)$  die entsprechende Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}^\theta$  lt. Lemma ?. Die **Zelle** von  $A$  und  $I$  sei wie folgt definiert:

$$c(A, I) := \{ (x, y, \theta) \mid (x, y) \in Q(A, \theta), \theta \in I \}$$

Die Zellen sind innerhalb des Intervalls  $I$  unabhängig von  $\theta$ , da  $I$  keine kritischen Orientierungen enthält, und liefern eine Zerlegung des  $\mathcal{C}_{\text{frei}}$  in zusammenhängende, paarweise disjunkte Zellen.

Aus Lemma ?? folgt für jedes maximal-unkritische Intervall  $I'$  und jede Zusammenhangskomponente  $A'$  von  $V^\theta$ :<sup>18</sup>

- $A \neq A' \Rightarrow c(A, I) \cap c(A', I) = \emptyset$ .
- $I \neq I' \Rightarrow \overline{c(A, I)} \cap c(A, I') = \emptyset$ .
- $\mathcal{C}_{\text{frei}} = \bigcup c(A, I)$ .

$c(A, I)$  und  $c(A', I')$  heißen **benachbart**  $:\Leftrightarrow I, I'$  haben einen gemeinsamen Endpunkt und es gibt eine (halb-)freie Platzierung  $(x^*, y^*, \theta^*) \in \overline{c(A, I)} \cap \overline{c(A', I')}$ .

Definiere den Zusammenhangsgraph  $G$  als den Graphen, dessen Knoten die Zellen  $c(A, I)$  sind und der Kanten zwischen benachbarten Zellen enthält. In diesem Graphen werden wir uns bewegen, um von  $e_u$  nach  $e_v$  zu gelangen.

**Behauptung 1:**  $c(A, I)$  und  $c(A', I')$  liegen in derselben Zusammenhangskomponente von  $G \Leftrightarrow c(A, I)$  und  $c(A', I')$  sind in derselben Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}$  enthalten.

<sup>18</sup> $\overline{c(A, I)}$  bezeichnet dabei den Abschluß der Zelle.

**Beweis 1:**

“ $\implies$ ” benachbarte Zellen liegen in derselben Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}$ , weil ihre Abschlüsse eine gemeinsame erlaubte Plazierung enthalten.

“ $\impliedby$ ” Seien  $(x, y, \theta) \in c(A, I)$  und  $(x', y', \theta') \in c(A', I')$  zwei freie Plazierungen aus derselben Zusammenhangskomponente  $Q$  von  $\mathcal{C}_{\text{frei}}$ .

$\implies$   $\exists$  stetiger Weg  $\pi$  von  $(x, y, \theta)$  nach  $(x', y', \theta')$ ,  $\pi \in Q$ .

$\implies$  Je zwei von  $\pi$  nacheinander besuchte Zellen sind benachbart.

$\implies$   $c(A, I)$  und  $c(A', I')$  liegen in derselben Zusammenhangskomponente von  $G$ .  $\diamond$

Der Rand einer Zelle  $c(A, I)$  kann Kanten enthalten, die nicht Teil einer Kante von  $\mathcal{C}_{\text{frei}}$  sind, weil sie durch die künstliche Einteilung in Zellen entstanden sind. Es gilt aber:

**Behauptung 2:** Sei  $c(A, I)$  eine Zelle in  $\mathcal{C}_{\text{frei}}$ . Dann gibt es eine Kante  $e \in \partial c(A, I)$  mit  $e \subseteq e_u$  ist Kante von  $\mathcal{C}_{\text{frei}}$  und  $(u, L(u))$  ist Knoten von  $E$ .

**Beweis 2:** Sei  $u$  ein Knoten in der Zusammenhangskomponente  $A$  von  $V^\theta$  für ein  $\theta \in I$ .

$\implies$   $u$  ist Knoten in der zugehörigen Zusammenhangskomponente  $Q(A, \theta)$  von  $\mathcal{C}_{\text{frei}}^\theta$  für jedes  $\theta \in I$ .

$\implies$   $I \subseteq L(u)$ .

$\implies$   $e := e_u \cap c(A, I)$  hat die gewünschte Eigenschaft.  $\diamond$

Eine Verschärfung von Behauptung 2 ist folgende

**Behauptung 3:** Seien  $c(A, I)$  und  $c(A', I')$  benachbarte Zellen. Dann gibt es einen Knoten  $(w, L(w))$  in  $E$ , der eine Kante  $e_w$  in  $\mathcal{C}_{\text{frei}}$  repräsentiert, die die Ränder beider Zellen schneidet.

**Beweis 3:** Sei  $\theta^*$  ein gemeinsamer Endpunkt von  $I$  und  $I'$  — eine kritische Orientierung.

**1. Fall:** die Zusammenhangskomponente  $A$  von  $V^{\theta^*}$ ,  $\theta \in I$ , bleibt bei  $\theta^*$  unverändert.

$\implies$   $A = A'$ .

$\implies$  Jeder Knoten  $w$  von  $Q(A, \theta)$  liefert einen Knoten  $(w, L(w))$  in  $E$  mit der gewünschten Eigenschaft.

**2. Fall:** die Zusammenhangskomponente  $A$  ändert sich bei  $\theta^*$ , d. h.  $A \neq A'$ .

Betrachte die äußeren Ränder der Zusammenhangskomponenten  $Q(A, \theta), \theta \in I$  und  $Q(A', \theta'), \theta' \in I'$  von  $\mathcal{C}_{\text{frei}}^\theta$  bzw.  $\mathcal{C}_{\text{frei}}^{\theta'}$ . Nach Voraussetzung ist  $\mathcal{C}_{\text{frei}}$  beschränkt, also wird jede Zusammenhangskomponente von einem Polygon umschlossen.

$$Q(A, \theta)$$

Jeder der beiden Ränder hat mindestens drei Ecken. Angenommen, in der kritischen Orientierung  $\theta^*$  zwischen  $I$  und  $I'$  sind alle drei Ecken eines Randes betroffen.

- $\Rightarrow$  Bei  $\theta^*$  verschwindet eine Komponente (Abb. ??(ii)).
- $\Rightarrow$  Es gibt keine (halb-)freie Plazierung  $(x^*, y^*, \theta^*)$  in  $c(A, I) \cap c(A', I')$ , da die Plazierung nicht im Abschluß von z. B.  $c(A', I')$  liegen kann, wenn dort keine Komponente vorhanden ist.
- $\Rightarrow$   $c(A, I)$  und  $c(A', I')$  sind nicht benachbart.  $\nexists$

Also gibt es einen Knoten  $w$  auf den äußeren Rändern von  $Q(A, \theta)$  und  $Q(A', \theta')$ , der von  $\theta^*$  nicht betroffen ist.  
 $\Rightarrow e_w \in \partial c(A, I) \cap \partial c(A', I')$ .  $\diamond$

**Behauptung 4:** Seien  $e_u$  und  $e_v$  zwei Kanten einer Zelle  $c(A, I)$  von  $\mathcal{C}_{\text{frei}}$  mit  $e_u \cap \partial c(A, I) \neq \emptyset$  und  $e_v \cap \partial c(A, I) \neq \emptyset$  — nach Behauptung 3 gibt es mindestens eine solche Kante. Dann liegen die zugehörigen Knoten  $(u, L(u))$  und  $(v, L(v))$  in derselben Zusammenhangskomponente von  $E$ .

**Beweis 4:** Weil  $I$  maximal-unkritisch ist, können weder  $e_u$  noch  $e_v$  in  $c(A, I)$  enden.

- $\Rightarrow$  Für jedes  $\theta \in I$  sind  $u$  und  $v$  Knoten derselben Zusammenhangskomponente von  $\mathcal{C}_{\text{frei}}^\theta$ .
- $\Rightarrow$   $u, v$  liegen in derselben Zusammenhangskomponente von  $V^\theta$ .  
Lemma ??
- $\Rightarrow$   $(u, L(u))$  und  $(v, L(v))$  sind in derselben Zusammenhangskomponente von  $E$ .  
Definition ??  $\diamond$

Damit können wir “ $\leftarrow$ ” von Lemma ?? beweisen. Wir wollten zeigen, daß zwei Knoten  $(u, L(u))$  und  $(v, L(v))$  in derselben Zusammenhangskomponente  $Z$  von  $E$  liegen, wenn die Kanten  $e_u, e_v$  auf dem Rand derselben Zusammenhangskomponente  $W$  von  $\mathcal{C}_{\text{frei}}$  liegen.

Die Zusammenhangskomponente  $W$  zerfällt in Zellen.

Sei also  $e_u \cap \partial c(A, I) \neq \emptyset$  und  $e_v \cap \partial c(A', I') \neq \emptyset$ .

$$c(A, I), c(A', I') \in W.$$

$\Rightarrow$  Beh. 1  $c(A, I), c(A', I')$  liegen in derselben Zusammenhangskomponente von  $G$ .

$\Rightarrow$  es gibt eine Folge  $c(A, I) = c(A_0, I_0), c(A_1, I_1), \dots, c(A_m, I_m), c(A_{m+1}, I_{m+1}) = c(A', I')$  von Zellen mit  $c(A_j, I_j)$  und  $c(A_{j+1}, I_{j+1}), j = 0, \dots, m$ , sind benachbart.

$\Rightarrow$  Beh. 3  $\forall j \exists (w_j, L(w_j)) \in E : e_{w_j} \cap \partial c(A_j, I_j) \neq \emptyset$  und  $e_{w_j} \cap \partial c(A_{j+1}, I_{j+1}) \neq \emptyset$ .

$\Rightarrow$  aus Behauptung 4

für  $c(A_0, I_0)$ :  $(u, L(u))$  und  $(w_0, L(w_0))$  liegen in derselben Zusammenhangskomponente von  $E$

für  $c(A_1, I_1)$ :  $(w_0, L(w_0))$  und  $(w_1, L(w_1))$  liegen in derselben Zusammenhangskomponente von  $E$

$\vdots$

für  $c(A_{m+1}, I_{m+1})$ :  $(w_m, L(w_m))$  und  $(v, L(v))$  liegen in derselben Zusammenhangskomponente von  $E$

$\Rightarrow$  Behauptung. □

Damit können wir Theorem ?? beweisen:

**Beweis.** (Theorem ??)

Seien  $\Phi(s) = \Phi(x_s, y_s, \theta_s) = (u, L(u))$ ,  $\Phi(t) = \Phi(x_t, y_t, \theta_t) = (v, L(v))$ , und seien  $e_u, e_v$  die Kanten von  $\mathcal{C}_{\text{frei}}$ , die den Knoten  $(u, L(u))$  und  $(v, L(v))$  von  $E$  entsprechen.

Nach der Definition von  $\Phi$  gibt es kollisionsfreie Translationen von  $s = (x_s, y_s, \theta_s)$  nach  $u$  (Knoten von  $\mathcal{C}_{\text{frei}}^{\theta_s}$ ) und von  $t = (x_t, y_t, \theta_t)$  nach  $v$  (Knoten von  $\mathcal{C}_{\text{frei}}^{\theta_t}$ ), also:

$\exists$  kollisionsfreie Bewegung von  $s$  nach  $t$ .

$\Leftrightarrow \exists$  kollisionsfreie Bewegung von  $u$  nach  $v$ .

$\Leftrightarrow \exists$  kollisionsfreie Bewegung von einem Punkt auf  $e_u$  zu einem Punkt auf  $e_v$ , denn längs der Kanten von  $\mathcal{C}_{\text{frei}}$  ist stets eine Bewegung möglich.

$\Leftrightarrow$  Die Kanten  $e_u$  und  $e_v$  gehören zum Rand derselben Zusammenhangskomponente  $W$  von  $\mathcal{C}_{\text{frei}}$ .

$\Leftrightarrow$  Lemma ?? Die Knoten  $(u, L(u)) = \Phi(s)$  und  $(v, L(v)) = \Phi(t)$  gehören zu derselben Zusammenhangskomponente  $Z$  von  $E$ . □

Wir wissen nun, daß für zwei (halb-)freie Plazierungen ein kollisionsfreier Pfad von  $s$  nach  $t$  genau dann existiert, wenn  $\Phi(s)$  und  $\Phi(t)$  in derselben



Zusammenhangskomponente von  $E$  liegen. Damit bleiben die Konstruktion des Kantengraphen  $E$  und die Berechnung eines Weges als Probleme.

Da  $V^{\theta_0}$  die Komplexität  $O(mn)$  hat und nur  $O(mn \lambda_6(mn))$  neue Knoten und Kanten generiert werden, hat auch der Kantengraph  $E$  eine Komplexität in  $O(mn \lambda_6(mn))$ . Ist die Menge  $T$  gegeben, können wir  $E$  — genauer die Zusammenhangskomponente  $Z_s$  von  $E$ , die  $\Phi(s)$  enthält — in Zeit  $O(mn \lambda_6(mn) \log(mn))$  berechnen. Algorithmus ?? ist ein Sweep im dreidimensionalen Raum. Die Sweepebenen sind Schnitte durch  $\mathcal{C}_{\text{frei}}$  mit  $\theta = \text{konstant}$  und die Ereignisse sind die  $\theta \in T$ .

---

**Algorithmus 2.12** Berechnung des Kantengraphen
 

---

- Berechne die Menge  $T$  aller gültigen kritischen Orientierungen, sortiert nach aufsteigendem  $\theta$ .
  - Wähle  $\theta_0 \notin T$  und berechne  $V^{\theta_0}$ .
  - Initialisiere  $E$  mit den Kanten und Knoten aus  $V^{\theta_0}$ , dabei bleiben die Lebensdauerintervalle  $L(u)$  zunächst undefiniert.
  - Für jede Orientierung  $\theta_i$  in  $T$ :
    - Bestimme die Änderungen in  $V^\theta$  beim Übergang von  $\theta = \theta_i - \varepsilon$  nach  $\theta = \theta_i + \varepsilon$ .
    - Entferne alle verschwindenden Knoten und Kanten  $v_1, \dots, v_r, e_1, \dots, e_s$  aus  $V^\theta$  und setze  $\theta_i$  als obere Grenze der Lebensdauer von  $L(v_j)$  bzw.  $L(e_k)$  ein.
    - Füge alle neu entstandenen Knoten und Kanten  $v_1^*, \dots, v_r^*, e_1^*, \dots, e_s^*$  in  $V^\theta$  und  $E$  ein und setze  $\theta_i$  als untere Grenze der Lebensdauer von  $L(v_j^*)$  bzw.  $L(e_k^*)$  ein.
  - Schließe den Kreis nach einem  $2\pi$  Durchlauf bei  $\theta_0$  und trage die noch fehlenden Lebensdauern nach, d. h. fasse zwei Knoten der Art  $(u, (\theta_1, \text{undef.}))$  und  $(u, (\text{undef.}, \theta_2))$  zu einem Knoten  $(u, (\theta_1, \theta_2))$  zusammen.
- 

Damit erhalten wir folgendes Ergebnis:

**Theorem 2.47** (Kedem, Sharir, (Toledo), 1990/1997)

Sei  $R$  ein konvexer Roboter mit  $m$  Ecken, der sich mit Translation und Rotation zwischen polygonalen Hindernissen  $P_i$  mit insgesamt  $n$  Ecken bewegt. Dann läßt sich in Zeit  $O(mn \lambda_6(mn) \log(mn))$  und Platz  $O(mn \lambda_6(mn))$  feststellen, ob es eine kollisionsfreie Bewegung von  $s$  nach  $t$  gibt und gegebenenfalls eine solche angeben. [?, ?]

**Beweis.**

Wir haben schon gesehen, daß es genau dann eine kollisionsfreie Bewegung von  $s$  nach  $t$  gibt, wenn  $\Phi(s)$  in derselben Zusammenhangskomponente wie  $\Phi(t)$  in  $E$  liegt. Dies kann nach einem Preprocessing in Zeit  $O(mn \lambda_6(mn))$  getestet werden. Ist der Test positiv, wird ein Pfad  $\pi$  in  $E$  gefunden, der wie in Algorithmus ?? beschrieben in eine (halb-)freie Bewegung von  $R$  transformiert werden kann.  $\square$

**Algorithmus 2.13** Bewegungsplanung mit Kantengraph

- Berechne  $\Phi(s)$  und  $\Phi(t)$  nach Definition ?? auf Seite ??.
- Teste, ob  $\Phi(s)$  und  $\Phi(t)$  in derselben Zusammenhangskomponente von  $E$  liegen. Falls nicht, existiert kein Weg von  $s$  nach  $t$ .
- Ansonsten sei

$$\pi = (\xi_1 = \Phi(s), \xi_2, \xi_3, \dots, \xi_\ell = \Phi(t))$$

der Pfad in  $E$ , der  $\Phi(s)$  mit  $\Phi(t)$  verbindet.  $\pi$  hat folgende Eigenschaften:  $\xi_i = (u_i, L(u_i))$  ist Knoten von  $E$ ,  $u_i$  ist Knoten von  $\mathcal{C}_{\text{frei}}^\theta$  und  $e_{u_i}$  ist Kante von  $\mathcal{C}_{\text{frei}}$ .

- Bewege  $R$  von  $s$  nach  $\Phi(s)$  (Translation nach Definition ??).
- Für jedes  $i$ :
  - Wähle eine Orientierung  $\theta_i$  in der Lebensdauer der Kante  $[\xi_i, \xi_{i+1}]$  von  $E$ .
  - Bewege  $R$  entlang der Kante  $e_{u_i}$  von  $\mathcal{C}_{\text{frei}}$ , bis die Orientierung  $\theta_i$  erreicht wird (Kurve mit Grad  $\leq 4$ ).
  - Verschiebe  $R$  in  $\mathcal{C}_{\text{frei}}^{\theta_i}$  vom Knoten  $u_i$  zum Knoten  $u_{i+1}$ , der auf der Kante  $e_{u_{i+1}}$  von  $\mathcal{C}_{\text{frei}}$  liegt.
- Bewege  $R$  von  $\Phi(t)$  nach  $t$  (Translation).

**Bemerkungen:**

- Dieser Algorithmus löst unser wirklich schwieriges Problem, für das eine untere Schranke von  $\Omega(m^2n^2)$  bekannt ist, in Zeit  $O(mn \lambda_6(mn) \log(mn))$ .
- Der Spezialfall  $m = 2$  (Leiter) läßt sich noch etwas effizienter lösen, nämlich in Zeit  $O(n^2 \log n)$ .
- Man kann auch ganz  $\mathcal{C}_{\text{frei}}$  in Zeit  $O(mn \lambda_6(mn) \log(mn))$  konstruieren, nicht nur die Zusammenhangskomponente, die  $s$  enthält (Agarwal, Amenta, Aronov, Sharir, [?]).
- Falls  $R$  nicht konvex ist, hat  $\mathcal{C}_{\text{frei}}$  die Komplexität  $\Theta(m^3n^3)$  (siehe Abbildung ??) und kann in Zeit  $O(m^3n^3 \log(mn))$  konstruiert werden, die Startkomponente in Zeit  $O((mn)^{2+\varepsilon})$  (Halperin, Sharir, [?]).

$R$

Abbildung 2.46: Nicht-konvexer Roboter mit  $\Omega(n^3)$  kritischen Plazierungen.



## Kapitel 3

# Bewegungsplanung für allgemeine Systeme

In den vorherigen Kapiteln haben wir betrachtet, wie wir Bahnen für einen Roboter planen können, der durch ein Polygon beschrieben werden kann. Als Spezialfälle haben wir die Bahnplanung für Punkte, Liniensegmente und Kreise untersucht. Nun stellen wir uns der Frage, ob und wie das Bahnplanungsproblem für allgemeine Systeme gelöst werden kann, also für Roboter, die sich nicht mehr durch Polygone beschreiben lassen, wie z. B. Systeme mit mehreren Armen, die mehrere Teleskop- und Drehgelenke enthalten können.

### 3.1 Eine untere Schranke für das allgemeine Bahnplanungsproblem

Folgendes Theorem zeigt, daß das allgemeine Bahnplanungsproblem bereits mit einer leichten Einschränkung sehr schwierig ist:

**Theorem 3.1** *Ein Roboter  $R$  sei gegeben durch eine Menge von  $m$  unterschiedlich großen, achsenparallelen Rechtecken  $R_i$ , die nicht miteinander verbunden sind. Seine Umgebung sei ein einfaches Polygon  $P$  mit achsenparallelen Kanten. Start- und Zielposition  $s$  und  $t$  seien halbfreie Positionen der  $R_i$  in  $P$ . Zu entscheiden, ob es in diesem Modell eine kollisionsfreie Bewegung von  $s$  nach  $t$  gibt, ist NP-hart.*

**Beweis.** Wir zeigen, daß sich das als NP-vollständig bekannte Problem *Partition* auf die Bahnplanung reduzieren läßt [?]. Bei dem Problem *Partition* ist eine Menge  $X = \{a_1, \dots, a_m\} \subset \mathbb{N}$  gegeben. Zu entscheiden ist, ob es eine Teilmenge  $Y \subset X$  gibt, so daß gilt:

$$\sum_{a_i \in Y} a_i = \sum_{a_j \in X \setminus Y} a_j.$$

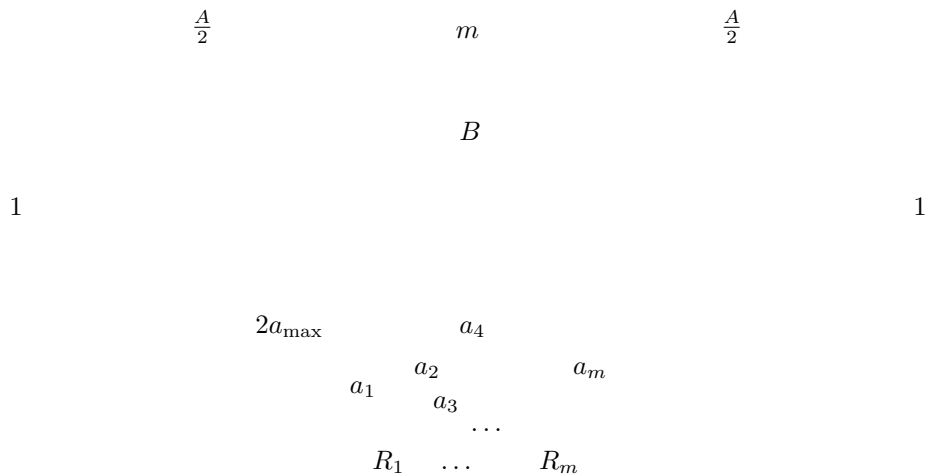


Abbildung 3.1: Umgebung zur Reduktion von Partition auf das allgemeine Bahnplanungsproblem.

Sei  $A := \sum_{i=1}^m a_i$  und  $a_{\max} := \max_{1 \leq i \leq m} a_i$ . Konstruiere nun in polynomieller Zeit zur gegebenen Menge  $X = \{a_1, \dots, a_m\}$  eine Szene für das Bahnplanungsproblem wie folgt (vgl. Abbildung ??):

- Die Umgebung besteht aus einer “Halle” der Breite  $m$  mit zwei Seitenarmen der Breite  $\frac{A}{2}$  und der Höhe  $1$ , die sich in einer Höhe von  $2a_{\max}$  über dem unteren Rand der Halle befinden.
- Für jedes  $a_i$  konstruiere ein Rechteck  $R_i$  des Roboters mit der Höhe  $a_i$  und Breite  $1$ , und positioniere  $R_i$  am unteren Rand der Halle.
- Der Roboter  $R$  bestehe aus allen Rechtecken  $R_i$  und einem weiteren Rechteck  $B$  der Breite  $m$  und Höhe  $> 1$ , das sich am oberen Rand der Halle befindet.
- In der Zielkonfiguration  $t$  sollen sich die  $R_i$  am oberen Rand befinden,  $B$  am unteren Rand.

Eine Bewegung von  $s$  nach  $t$  ist genau dann möglich, wenn alle  $R_i$  in die Seitenarme bewegt werden können (genau dann läßt sich  $B$  vom oberen Rand zum unteren Rand bewegen). Dies wiederum ist genau dann möglich, wenn eine Partition der Menge  $X$  existiert.  $\square$

**Bemerkung 3.2** *Man kann sogar zeigen, daß dieses Problem PSPACE-hart<sup>1</sup> ist (Hopcroft, Schwartz, Sharir, [?]).*

<sup>1</sup>Zur Definition von PSPACE-hart siehe Definition ?? auf Seite ??.

Auch die folgenden Probleme sind NP-hart:

- Gegeben sei ein Roboterarm mit  $m > 3$  Drehgelenken in einer Umgebung mit polygonalen Hindernissen. Gibt es eine kollisionsfreie Bewegung des Roboterkopfes?

Kopf

- Die Bewegung eines Würfels zwischen polyhedralen Hindernissen. (Wir haben gesehen, daß für ein Rechteck in der Ebene das Bahnplanungsproblem effizient lösbar ist, aber bereits eine Dimension höher wird das Problem NP-hart).

$R$

Dies zeigt, daß wir im allgemeinen mit exponentiellen Laufzeiten rechnen müssen.





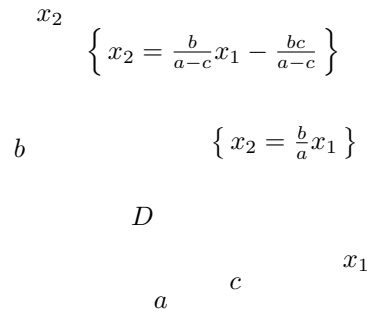
Damit können wir semialgebraische Mengen wie folgt definieren:

**Definition 3.4** Eine Teilmenge  $S \subseteq \mathbb{R}^n$  heißt **semialgebraisch**, falls es einen Tarski–Ausdruck  $T(X_1, \dots, X_n)$  gibt, in dem die Variablen  $X_1, \dots, X_n$  frei vorkommen, so daß gilt:

$$S = \{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid T(x_1, \dots, x_n) \text{ ist erfüllt} \}.$$

Die Menge der semialgebraischen Mengen ist abgeschlossen gegen  $\bigcap^{<\infty}$ ,  $\bigcup^{<\infty}$ , Komplement und Projektion. Bei algebraischen Mengen sind nur Ausdrücke mit “=”, keine Ungleichungen erlaubt. Die Nullstellen von Polynomen sind algebraische Mengen.

Im ersten Schritt beschreiben wir nun  $R$  und die  $P_i$  als semialgebraische Mengen. Da die Oder–Verknüpfung von Tarski–Ausdrücken wieder ein Tarski–Ausdruck ist, genügt es, ein Dreieck  $D$  mit Hilfe eines Tarski–Ausdrucks zu beschreiben. Dazu gibt es mehrere Möglichkeiten. Wir können z. B.  $D$  durch Summen von Vektoren beschreiben (mit Quantoren):



$$\begin{aligned} D &= \{ (x_1, x_2) \mid \exists \lambda \exists \mu (\lambda \geq 0 \wedge \mu \geq 0 \wedge \lambda + \mu \leq 1 \\ &\quad \wedge (x_1, x_2) = \lambda(c, 0) + \mu(a, b) \} \\ &= \{ (x_1, x_2) \mid \exists \lambda \exists \mu (\lambda \geq 0 \wedge \mu \geq 0 \wedge \lambda + \mu \leq 1 \\ &\quad \wedge \underbrace{x_1 = \lambda c + \mu a \wedge x_2 = \mu b}_{T_1(x_1, x_2)} \} . \end{aligned}$$

Eine Alternative ist die Beschreibung von  $D$  als Schnitt von Halbebenen; in diesem Fall kommen wir sogar ohne Quantoren aus:

$$D = \{ (x_1, x_2) \mid \underbrace{x_2 \geq 0 \wedge x_2 \leq \frac{b}{a}x_1 \wedge x_2 \leq \frac{b}{a-c}x_1 - \frac{bc}{a-c}}_{T_2(x_1, x_2)} \} .$$

Triangulieren wir die Hindernisse, so zerfällt ein Hindernis  $P_i$  in Dreiecke  $D_{ij}$  mit  $D_{ij} = \{ (x_1, x_2) \mid T_{ij}(x_1, x_2) \}$  mit adäquatem Tarski–Ausdruck  $T_{ij}(X_1, X_2)$ . Das Hindernis  $P_i$  läßt sich dann als Disjunktion dieser Tarski–Ausdrücke darstellen:

$$P_i(X_1, X_2) := T_{i,1}(X_1, X_2) \vee T_{i,2}(X_1, X_2) \vee \dots \vee T_{i,k}(X_1, X_2)$$

Wir können also die Menge der Hindernisse durch einen Tarski–Ausdruck  $H(X_1, X_2)$  beschreiben. Ebenso kann der Roboter  $R$  in seiner Ausgangslage  $R(0, 0, 90^\circ)$  durch einen Tarski–Ausdruck  $R(X_1, X_2)$  beschrieben werden.

Wie beschreiben wir nun die Bewegungen des Roboters?

**Translation:** Durch Addition eines Translationsvektors  $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ .

**Rotation um  $\theta$ :** Eine Rotation läßt sich durch Anwendung einer Rotationsmatrix  $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$  beschreiben. Das Problem dabei ist, daß die Funktionen  $\sin$  und  $\cos$  nicht algebraisch sind. Da der Winkel  $\theta$  jedoch nicht explizit benötigt wird, reicht es aus, eine Rotationsmatrix  $\begin{pmatrix} S_1 & -S_2 \\ S_2 & S_1 \end{pmatrix}$  mit  $S_1^2 + S_2^2 = 1$  zu benutzen.

Damit können wir Bewegungen als Tarski–Ausdruck

$$B(X_1, X_2, X'_1, X'_2, A_1, A_2, S_1, S_2)$$

beschreiben, mit  $B(x_1, x_2, x'_1, x'_2, a_1, a_2, s_1, s_2) : \iff$  die Zielposition  $\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$  geht aus der Startposition  $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$  durch Rotation des Roboters mit  $\begin{pmatrix} s_1 & -s_2 \\ s_2 & s_1 \end{pmatrix}$  und anschließender Translation um  $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$  hervor, also wenn gilt:

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} s_1 & -s_2 \\ s_2 & s_1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

Der Roboter darf nicht von einem Ausgangspunkt  $(x_1, x_2)$  durch eine Bewegung  $B(\dots)$  in einen Zielpunkt  $(x'_1, x'_2)$  überführt werden, der in einem Hindernis liegt. Der Raum der verbotenen Konfigurationen  $\mathcal{C}_{\text{verb}}$  läßt sich damit beschreiben als:

$$\begin{aligned} \mathcal{C}_{\text{verb}}(A_1, A_2, S_1, S_2) &: \equiv \\ &\exists X_1, X_2, X'_1, X'_2 \left( R(X_1, X_2) \wedge H(X'_1, X'_2) \right. \\ &\quad \left. \wedge B(X_1, X_2, X'_1, X'_2, A_1, A_2, S_1, S_2) \right). \end{aligned} \quad (*)$$

Durch Negation dieses Ausdrucks erhalten wir eine Beschreibung des Raumes der freien Konfigurationen als semialgebraische Menge:

$$\mathcal{C}_{\text{frei}}(A_1, A_2, S_1, S_2) : \equiv \neg \mathcal{C}_{\text{verb}}(A_1, A_2, S_1, S_2)$$

Dies zeigt, daß die Lösung des Bahnplanungsproblems mit den Gültigkeitsbereichen von Tarski–Ausdrücken zusammenhängt. Diese Technik läßt sich auf Roboter mit mehreren Freiheitsgraden verallgemeinern; dabei werden nur die benutzten Ausdrücke komplizierter [?].

### 3.3 Collins' zylindrische algebraische Zerlegung

Wir haben bereits erkannt, daß wir keinen effizienten Algorithmus für das allgemeine Bewegungsplanungsproblem erwarten dürfen. Wir wissen auch — siehe Abschnitt ?? —, daß es Probleme gibt, die so schwierig sind, daß sie überhaupt nicht algorithmisch lösbar sind. In diesem Abschnitt wollen wir uns mit der Entscheidbarkeit zweier Probleme befassen. Die dabei benutzte Beweistechnik können wir danach auf unser Bewegungsplanungsproblem anwenden.

**Problem 1:**

Die *Theorie der reellen Zahlen* ist die Menge aller Tarski-Ausdrücke ohne freie Variablen, die über ganz  $\mathbb{R}$  gelten:

$$\Omega = \{ T \mid T \text{ ist Tarski-Ausdruck ohne freie Variablen} \}$$

$$Th(\mathbb{R}) = \{ T \in \Omega \mid T \text{ gilt über } \mathbb{R} \}$$

Beispiele:

$$\forall x (x^2 \leq 0 \rightarrow x = 0) \quad \in Th(\mathbb{R})$$

$$\forall x \exists y (x \geq 0 \rightarrow x = y^2) \quad \in Th(\mathbb{R})$$

$$\exists x (x^2 = -1) \quad \notin Th(\mathbb{R})$$

**Theorem 3.5** (*Tarski, 1951*)

*Die Theorie der reellen Zahlen ist entscheidbar.*

[?]

**Problem 2:**

Wir haben Beispiele von semialgebraischen Mengen gesehen, die sich durch Tarski-Ausdrücke beschreiben lassen, in denen keine Quantoren auftreten. Das ist kein Zufall! Denn es gilt:

**Theorem 3.6** *Jede semialgebraische Menge läßt sich durch einen quantorenfreien Tarski-Ausdruck definieren.*

Beide Theoreme werden wir im folgenden beweisen, dazu benötigen wir allerdings noch etwas Handwerkszeug:

**Definition 3.7** Eine **Algebraische Zerlegung** des  $\mathbb{R}^n$  ist ein endliches System  $\mathcal{K}_n$  disjunkter Teilmengen (Zellen) des  $\mathbb{R}^n$  für die gilt:

- Jede Zelle  $C_j \in \mathcal{K}_n$  ist semialgebraisch.
- Jede Zelle  $C_j \in \mathcal{K}_n$  ist homöomorph zu einem  $\mathbb{R}^k$ ,  $k \leq n$ , d. h. jede Zelle ist zusammenhängend und hat keine Löcher. Für  $k = 0$  ist  $C_j$  ein algebraischer Punkt.<sup>3</sup>

<sup>3</sup>Ein Punkt, dessen Koordinaten sich durch algebraische Zahlen darstellen lassen.

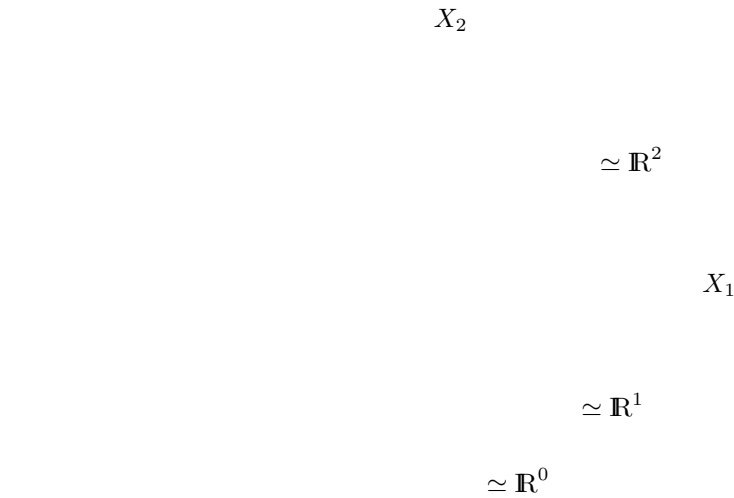


Abbildung 3.3: Beispiel für eine zylindrische algebraische Zerlegungen.

- $\mathcal{K}_n$  ist eine disjunkte Zerlegung des  $\mathbb{R}^n$ :  $\mathbb{R}^n = \bigcup_{j=1}^m C_j$ .

Eine **Probefunktion**  $\sigma_n$  ist eine Funktion  $\sigma_n : \mathcal{K}_n \rightarrow \mathbb{R}^n$ , die jeder Zelle  $C_j$  einen Probepunkt  $\sigma_n(C_j) \in C_j$  zuordnet.

Die **Projektion**  $\pi_n$  sei die Abbildung

$$\begin{aligned} \pi_n : \mathbb{R}^n &\longrightarrow \mathbb{R}^{n-1} \\ (X_1, \dots, X_n) &\longmapsto (X_1, \dots, X_{n-1}) \end{aligned}$$

Die Projektion einer Menge  $\mathcal{M}$  ist die Projektion ihrer Elemente:  $\pi_n(\mathcal{M}) := \{y \mid y = \pi_n(x), x \in \mathcal{M}\}$

**Definition 3.8** Ein System  $(\mathcal{K}_n, \sigma_n)$  heißt **zylindrische algebraische Zerlegung** (cylindrical algebraic decomposition, CAD, hier: ZAZ) des  $\mathbb{R}^n$  falls gilt:

- $\mathcal{K}_n$  ist eine algebraische Zerlegung mit Probefunktion  $\sigma_n$ .
- Für  $n > 1$  läßt sich der  $\mathbb{R}^{n-1}$  rekursiv zerlegen, d. h. es existiert eine ZAZ  $(\mathcal{K}_{n-1}, \sigma_{n-1})$  des  $\mathbb{R}^{n-1}$  mit:

- Die Probefunktion zu  $\mathcal{K}_{n-1}$  ist die Projektion der Probefkt. zu  $\mathcal{K}_n$ :

$$\begin{aligned} \sigma_{n-1} : \mathcal{K}_{n-1} &\longrightarrow \mathbb{R}^{n-1} \\ \pi_n(C) &\longmapsto \pi_n(\sigma_n(C)) \end{aligned}$$

- Zu jeder Zelle  $C_j \in \mathcal{K}_n$  existiert eine Zelle  $C'_j \in \mathcal{K}_{n-1}$  mit  $C'_j = \pi_n(C_j)$ , die Projektion einer Zelle in  $\mathcal{K}_n$  ist eine Zelle in  $\mathcal{K}_{n-1}$ .

– Für Zellen  $C_i, C_j \in \mathcal{K}_n$  mit  $\pi_n(C_i) = \pi_n(C_j)$  gilt  $\pi_n(\sigma_n(C_i)) = \pi_n(\sigma_n(C_j))$ .

- Für  $n = 1$  ist  $\mathcal{K}_1$  eine Zerlegung von  $\mathbb{R}$  in eine endliche Menge algebraischer Zahlen, den (endlichen und unendlichen) offenen Intervallen zwischen diesen Zahlen.
- Der **Zylinder** über  $C'$  sind alle Zellen  $C \in \mathcal{K}_n$  mit  $\pi_n(C) = C'$ .

Abbildung ?? zeigt ein Beispiel für eine zylindrische algebraische Zerlegungen. Die Zerlegung besteht aus vier Viertelebenen (homöomorph zum  $\mathbb{R}^2$ ), vier Halbgraden ( $\simeq \mathbb{R}^1$ ) und einem Punkt ( $\simeq \mathbb{R}^0$ ). Die Probepunkte sind mit  $\times$  gekennzeichnet.

**Definition 3.9** Sei  $\mathcal{P} \subseteq \mathbb{Q}[X_1, \dots, X_n]$  eine endliche Menge von Polynomen über  $\mathbb{Q}[X_1, \dots, X_n]$ . Eine zylindrische algebraische Zerlegung  $(\mathcal{K}_n, \sigma_n)$  heißt  **$\mathcal{P}$ -invariant**, wenn gilt:

$$\forall f(X_1, \dots, X_n) \in \mathcal{P} : \forall C \in \mathcal{K}_n : \text{sign}\left(f\Big|_C\right) = \text{const},$$

d.h. für alle Punkte  $(x_1, \dots, x_n)$  derselben Zelle  $C$  hat  $f(x_1, \dots, x_n)$  das gleiche Vorzeichen.

Für den Beweis von Theorem ??, Theorem ?? und für das Bewegungsplanungsproblem ist folgendes Theorem entscheidend. In der Bewegungsplanung entspricht die Dimension der Anzahl der Freiheitsgrade des Robotersystems.

**Theorem 3.10** (Collins, 1975)

*Man kann eine  $\mathcal{P}$ -invariante zylindrische algebraische Zerlegung effektiv konstruieren. Die Laufzeit ist polynomiell in Anzahl und Grad der Polynome in  $\mathcal{P}$ , doppelt exponentiell in der Anzahl der Dimensionen. [?]*

Die Idee ist, die Polynome in  $\mathcal{P}$  zu einem Polynom  $P$  zusammenzufassen und die Zerlegung schrittweise mit wachsender Dimension zu konstruieren. Für eine Dimension  $n$  fassen wir  $P$  als Polynom in einer Variablen auf und bestimmen dessen Nullstellen, die als Funktionen über  $n - 1$  Variablen dargestellt werden. Diese Nullstellen dienen als Intervallgrenzen bei der Konstruktion des Zylinders. Genauer:

**Beweis.** Durch Induktion über  $n$ . Sei  $P(X_1, \dots, X_n) := \prod_{P_i \in \mathcal{P}} P_i$ .

$n = 1$ :  $P(X_1)$  habe den Grad  $m \Rightarrow P(X_1)$  hat  $r \leq m$  viele verschiedene reelle Nullstellen. Die ZAZ in Abbildung ?? ist nach dem Zwischenwertsatz  $\mathcal{P}$ -invariant.

Wegen Stetigkeit kein Vorzeichenwechsel in den Intervallen

$$X_1 \quad X_2 \quad \dots \quad X_{r-1} \quad X_r$$

Abbildung 3.4: Zylindrische algebraische Zerlegung für  $n = 1$ .

$n > 1$ : Fasse  $P(X_1, \dots, X_n) := \prod_{P \in \mathcal{P}} P(X_1, \dots, X_n)$  als Polynom in  $X_n$  mit Koeffizienten in  $\mathbb{Q}[X_1, \dots, X_{n-1}]$  auf.

Sei  $(x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$  ein festes Tupel,  $\hat{P}(X_n) := P_{x_1, \dots, x_{n-1}}(X_n)$  und sei  $r := \text{grad}_{X_n}(\hat{P}(X_n))$  der Grad der Variablen  $X_n$ . Dann ist die Anzahl der verschiedenen reellen und komplexen Nullstellen von  $\hat{P}(X_n)$  gleich

$$r - \text{grad}_{X_n} \left( \text{ggT} \left( \hat{P}(X_n), \underbrace{\hat{P}'(X_n)}_{\text{Ableitung nach } X_n} \right) \right).$$

Begründung:

$$\begin{aligned} & \alpha \text{ ist } k\text{-fache Nullstelle von } f(x) \\ \Rightarrow & f(x) = (x - \alpha)^k \cdot g(x) \text{ mit } g(\alpha) \neq 0 \\ \Rightarrow & f'(x) = k(x - \alpha)^{k-1} \cdot g(x) + (x - \alpha)^k \cdot g'(x) \text{ hat } \alpha \text{ als} \\ & (k - 1)\text{-fache Nullstelle.} \end{aligned}$$

Ist also  $f(x) = \prod_{i=1}^j (x - \alpha_i)^{k_i}$ , so ist  $\text{ggT}(f(x), f'(x)) = \prod_{i=1}^j (x - \alpha_i)^{k_i - 1}$ , und der Grad von  $f$  ist um  $j$  höher als der des  $\text{ggT}$ .

Man kann nun polynomielle Bedingungen dafür aufstellen, daß sich  $\text{grad}_{X_n}(\hat{P}(X_n))$  und  $\text{grad}_{X_n} \left( \text{ggT}(\hat{P}(X_n), \hat{P}'(X_n)) \right)$  nicht ändern.<sup>4</sup> Wendet man die Induktionsvoraussetzung auf diese polynomiellen Bedingungen an, dann existiert eine ZAZ  $(\mathcal{K}_{n-1}, \sigma_{n-1})$  des  $\mathbb{R}^{n-1}$ , bei der  $\text{grad}_{X_n}(\hat{P}(X_n))$  und  $\text{grad}_{X_n}(\text{ggT}(\hat{P}(X_n), \hat{P}'(X_n)))$  auf jeder Zelle  $C$  konstant sind.

$$\begin{aligned} \Rightarrow & \text{Die Anzahl der verschiedenen reellen Nullstellen von} \\ & \hat{P}(X_n) \text{ ist auf jeder Zelle konstant.} \\ \Rightarrow & \text{Die verschiedenen reellen Nullstellen lassen sich auf jeder} \\ & \text{Zelle } C \text{ durch stetige Funktionen} \\ & f_1(X_1, \dots, X_{n-1}) < f_2(X_1, \dots, X_{n-1}) < \dots < f_j(X_1, \dots, X_{n-1}) \\ & \text{darstellen.} \end{aligned}$$

<sup>4</sup>Dazu braucht man Kenntnisse aus der Resultantentheorie, die wir später erläutern.

Man benutzt diese  $f_j$  auf folgende Weise, um den Zylinder  $C \times \mathbb{R}$  in Zellen zu zerlegen, auf denen  $P(X_1, \dots, X_n)$  ein konstantes Vorzeichen hat (siehe Abbildung ??):

$$C_{0,1} = \{ (x_1, \dots, x_{n-1}, y) \mid (x_1, \dots, x_{n-1}) \in C, y < f_1(x_1, \dots, x_{n-1}) \}$$

...

$$C_i = \{ (x_1, \dots, x_{n-1}, y) \mid (x_1, \dots, x_{n-1}) \in C, y = f_i(x_1, \dots, x_{n-1}) \}$$

$$C_{i,i+1} = \{ (x_1, \dots, x_{n-1}, y) \mid (x_1, \dots, x_{n-1}) \in C, f_i(x_1, \dots, x_{n-1}) < y < f_{i+1}(x_1, \dots, x_{n-1}) \}$$

...

$$C_{j,\infty} = \{ (x_1, \dots, x_{n-1}, y) \mid (x_1, \dots, x_{n-1}) \in C, f_j(x_1, \dots, x_{n-1}) < y \}$$

Die Probepunkte werden entsprechend definiert. So entsteht  $(\mathcal{K}_n, \sigma_n)$ .

□

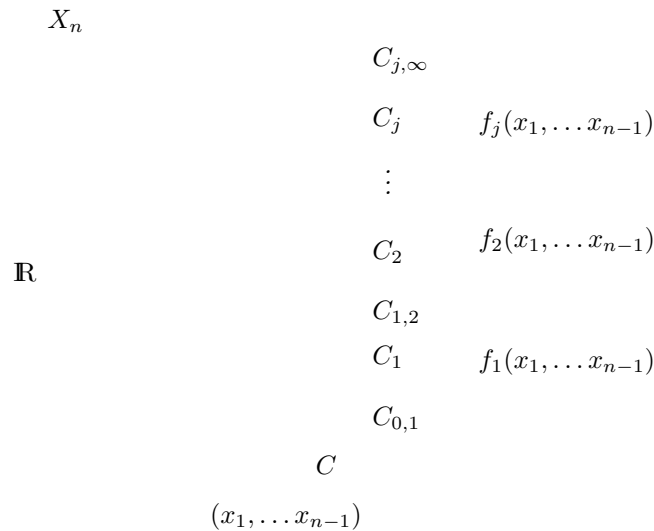


Abbildung 3.5: Zerlegung des Zylinders  $C \times \mathbb{R}$ .

Damit kann man die Entscheidbarkeit der Theorie der reellen Zahlen (Theorem ??) beweisen. Wir beschränken uns hier auf ein signifikantes

**Beispiel:** Ist die Tarski-Formel  $\forall X \exists Y : T(X, Y)$  gültig?

Berechne eine  $T$ -invariante ZAZ  $(\mathcal{K}_2, \sigma_2)$  des  $\mathbb{R}^2$ , d. h. die Polynome in  $X$  und  $Y$ , die in  $T$  vorkommen, haben auf jeder Zelle ein konstantes Vorzeichen. Sei  $(\mathcal{K}_1, \sigma_1)$  die projizierte eindimensionale ZAZ, dann gilt:

$\forall X \exists Y : T(X, Y)$  ist gültig  $\iff$  für jede Zelle  $C' \in \mathcal{K}_1$  gibt es eine Zelle  $C \in \mathcal{K}_2$  im Zylinder über  $C'$  mit  $T(x, y)$  ist wahr für  $(x, y) = \sigma_2(C)$ .

“ $\implies$ ” Sei  $C' \in \mathcal{K}_1$  beliebig. Betrachte  $x = \sigma_1(C') \in C'$ . Nach Voraussetzung gibt es ein  $\tilde{y}$  mit  $T(x, \tilde{y})$ . Sei  $C$  die Zelle über  $C'$  mit  $(x, \tilde{y}) \in C$ . Wegen der  $T$ -Invarianz gilt dann auch  $T(x, y)$  für  $(x, y) = \sigma_2(C)$ .

“ $\impliedby$ ” Sei  $\tilde{x}$  beliebig. Sei  $C'$  die Zelle in  $\mathcal{K}_1$  mit  $\tilde{x} \in C'$ . Nach Voraussetzung gibt es ein  $C \in \mathcal{K}_2$  über  $C'$  mit  $T(x, y)$  für  $(x, y) = \sigma_2(C)$ . Es existiert ein  $\tilde{y}$  mit  $(\tilde{x}, \tilde{y}) \in C$ , da  $C$  im Zylinder von  $C'$  liegt, wegen der  $T$ -Invarianz gilt dann auch  $T(\tilde{x}, \tilde{y})$ .

Dies ist ein endliches Problem und kann in endlich vielen Schritten getestet werden. Der Beweis für andere Quantoren und Formeln mit längerem Präfix ist analog.  $\square$

Auch die Eliminierung von Quantoren (Theorem ??) kann man nun beweisen:

**Beispiel:** Sei  $S = \{x \mid \exists Y : H(x, Y)\}$  eine semialgebraische Menge, wobei  $H$  ein Tarski-Ausdruck ist. Darin ist  $x$  eine freie,  $Y$  eine gebundene Variable.

Wie oben sei  $(\mathcal{K}_2, \sigma_2)$  eine ZAZ für die Polynome in  $H$ . Für eine Zelle  $C$  sei  $\alpha_C$  der quantorenfreie Tarski-Ausdruck, welcher  $C$  definiert (nach Collins existiert ein solcher).

Betrachte

$$\mathcal{C} := \{C' \in \mathcal{K}_1 \mid \exists C \in \mathcal{K}_2 \text{ über } C' \text{ mit } H(x, y) \text{ für } (x, y) = \sigma_2(C)\}$$

und setze  $\alpha(x) := \bigvee_{C' \in \mathcal{C}} \alpha_{C'}(x)$ . Beachte:  $\alpha(x)$  ist quantorenfrei!

**Behauptung:**  $S = \{x \mid \exists Y H(x, Y)\} \stackrel{!}{=} \{x \mid \alpha(x)\}$ .

“ $\subseteq$ ” Gelte  $H(\tilde{x}, \tilde{y})$ . Dann sei  $\tilde{x} \in \tilde{C}$ , und  $C$  sei die Zelle über  $\tilde{C}$  mit  $(\tilde{x}, \tilde{y}) \in C$ . Dann gilt auch  $H(x, y)$  für  $(x, y) = \sigma_2(C)$ . Also ist  $\tilde{C} = C' \in \mathcal{C}$  und wegen  $x, \tilde{x} \in C'$  folgt  $\alpha(\tilde{x})$ .

“ $\supseteq$ ” Gelte  $\alpha(\tilde{x})$ . Dann gilt auch  $\alpha_{C'}(\tilde{x})$  für ein  $C' \in \mathcal{C}$ , also auch  $\tilde{x} \in C'$ . Sei  $C$  die Zelle über  $C'$  mit  $H(x, y)$  für  $(x, y) = \sigma_2(C)$ . Dann existiert ein  $\tilde{y}$  mit  $(\tilde{x}, \tilde{y}) \in C$ , da  $C$  im Zylinder über  $C'$  liegt. Dann gilt auch  $H(\tilde{x}, \tilde{y})$ .  $\square$

Im Folgenden wollen wir zeigen, welche polynomiellen Bedingungen gestellt werden müssen um die zylindrische algebraische Zerlegung rekursiv durchzuführen. Eine Bedingung war unter anderem, dass für das Polynom  $\hat{P}(X_n)$  gilt, dass  $\text{grad}_{X_n}(\text{ggT}(\hat{P}(X_n), \hat{P}'(X_n)))$  konstant ist.

Allgemeiner bestimmen wir die Anzahl gemeinsamer Nullstellen zweier Polynome  $A(X)$  und  $B(X)$  mit Grad  $a$  respektive Grad  $b$ . Wir betrachten dazu für  $(0 \leq j \leq \min\{a, b\} - 1)$  Polynome  $U_j(X)$  mit Grad  $b - j - 1$ , und  $V_j(X)$  mit Grad  $a - j - 1$ . Diese Polynome sollen die Gleichung



$$A(X)U_j(X) - B(X)V_j(X) = 0$$

erfüllen. Das Polynom  $U_j$  hat  $b-j$  Koeffizienten  $u_i$  und das Polynom  $V_j$  hat  $a-j$  Koeffizienten  $v_i$ . Das Polynom auf der linken Seite obiger Gleichung hat somit Grad  $a + b - j - 1$ .

**Lemma 3.11** *Genau dann wenn  $A(X)U_j(X) - B(X)V_j(X) = 0$  eine nicht-triviale Lösung hat, haben die Polynome  $A(X)$  und  $B(X)$   $j+1$  gemeinsame Nullstellen.*

**Beweis.** O.B.d.A. betrachten wir  $A, B, U$  und  $V$  als normierte Polynome. Also sei  $A(X) = (X - \alpha_1) \cdots (X - \alpha_a)$ ,  $B(X) = (X - \beta_1) \cdots (X - \beta_b)$ ,  $U_j(X) = (X - \mu_1) \cdots (X - \mu_{b-j-1})$  und  $V_j(X) = (X - \nu_1) \cdots (X - \nu_{a-j-1})$ . Sei  $j$  die kleinste Zahl, so dass  $A(X)U_j(X) - B(X)V_j(X) = 0$  eine nicht-triviale Lösung hat. Dann gilt  $\mu_r \neq \nu_s$ ,  $r \in [1, b-j-1]$ ,  $s \in [1, a-j-1]$ . Sonst liesse sich ein Faktor entfernen und die Gleichung gelte immer noch. Ein Widerspruch zum kleinsten  $j$ ! Es existiert eine eindeutige Faktorisierung von  $A(X)U_j(X)$  und  $B(X)V_j(X)$ . Jede Nullstelle von  $U_j$  muss nun eine Nullstelle von  $B$  und jede Nullstelle von  $V_j$  eine Nullstelle von  $A$  sein. Dann haben  $A$  und  $B$   $j+1$  gemeinsame Nullstellen!

Umgekehrt, falls  $A(X)$  und  $B(X)$   $j+1$  gemeinsame Nullstellen haben, dann können wir die restlichen Nullstellen von  $A$  auf  $V_j$  und die restlichen Nullstellen von  $B$  auf  $U_j$  verteilen und  $A(X)U_j(X) - B(X)V_j(X) = 0$  hat eine nicht-triviale Lösung!  $\square$

Mit diesen Aussagen können wir eine Bedingung für die Anzahl der gemeinsamen Nullstellen aufstellen. Das lineare Gleichungssystem der Koeffizienten von  $A(X)U_j(X) - B(X)V_j(X) = 0$  besteht aus  $a + b - j$  Gleichungen aber nur  $a + b - 2j$  Unbekannten (Koeffizienten aus  $U_j(X)$  und  $V_j(X)$ ). Wir betrachten lediglich die ersten  $a + b - 2j$  Koeffizienten des Polynoms  $A(X)U_j(X) - B(X)V_j(X)$ . Dann ergibt sich ein lineares Gleichungssystem mit  $a + b - 2j$  Unbekannten und  $a + b - 2j$  Gleichungen.

Sei  $\psi_j(A, B)$  die Determinante der Matrix dieses Gleichungssystems. Wir werden zeigen, dass es ausreicht, das kleinste  $j$  zu bestimmen, so dass  $\psi(A, B) \neq 0$  ist.

**Theorem 3.12** *Die Anzahl der gemeinsamen Nullstellen von  $A$  und  $B$  ist die kleinste Zahl  $j$ , so dass  $\psi_j(A, B) \neq 0$  gilt.*

**Beweis.** Sei  $j = 0$  und  $\psi_j(A, B)$  die Determinante der ersten  $a + b - 2j$  Gleichungen für die Koeffizienten mit höchstem Grad in  $X$  aus  $A(X)U_j(X) - B(X)V_j(X) = 0$ . Also haben wir hier ein lineares Gleichungssystem mit  $a + b$  Gleichungen und  $a + b$  Unbekannten. Falls  $\psi_j(A, B) = 0$  ist, dann hat  $A(X)U_j(X) - B(X)V_j(X) = 0$  gemäß Lemma ?? eine nicht-triviale Lösung und somit eine gemeinsame Nullstelle. Ansonsten gibt es keine gemeinsame Nullstelle.

Nehmen wir nun an es gäbe nicht-triviale Lösungen für  $A(X)U_i(X) - B(X)V_i(X) = 0$  für  $i = 0, \dots, j-1$ , so dass  $A$  und  $B$  mindestens  $j$  gemeinsame Nullstellen haben. Wir zeigen: Das Vorzeichen von  $\psi_j(A, B)$  entscheidet, ob  $j+1$  oder nur  $j$  gemeinsame Nullstellen existieren.

Betrachte dazu  $A(X)U_j(X) - B(X)V_j(X) = C_j(X)$ , wobei  $C_j$  ein Polynom vom Grad  $\leq j-1$  sei. Sei  $\Sigma$  das lineare Gleichungssystem mit Koeffizienten aus  $U_j, V_j$  und  $C_j$ . Zu den  $a+b-2j$  Unbekannten kommen also  $j$  Unbekannte dazu und  $\Sigma$  ist ein quadratisches System. Das Subsystem  $\Sigma'$  mit den ersten  $a+b-2j$  Unbekannten und  $a+b-2j$  Gleichungen für  $X^i$  für  $i \in [j, a+b-j-1]$  benutzt keine Koeffizienten aus  $C_j$ .

Es ist leicht zu sehen, dass  $\Sigma'$  genau dann nicht-triviale Lösung ( $\psi_j(A, B) = 0$ ) hat, wenn  $\Sigma$  eine nicht-triviale Lösung besitzt. Die Matrix von  $\Sigma'$  ist eine echte Submatrix von  $\Sigma$ , die Determinante von  $\Sigma'$  ist genau dann gleich null, wenn die Determinante von  $\Sigma$  null ist.

Falls nun  $\psi_j(A, B) \neq 0$  gilt, dann hat  $A(X)U_j(X) - B(X)V_j(X) = C_j(X)$  nur die triviale Lösung. Insbesondere gilt das dann auch für  $A(X)U_j(X) - B(X)V_j(X) = 0$ , nach Lemma ?? haben  $A$  und  $B$  genau  $j$  gemeinsame Nullstellen.

Falls aber  $\psi_j(A, B) = 0$  gilt, dann existieren  $U_j, V_j$  und  $C_j$  die die Gleichung  $A(X)U_j(X) - B(X)V_j(X) = C_j(X)$  erfüllen. Nach Annahme haben  $A$  und  $B$  mindestens  $j$  gemeinsame Nullstellen. Dann teilt das Produkt der Nullstellen das Polygon  $C_j(X)$ . Da der Grad von  $C_j$  aber  $j-1$  ist, kann nur  $C_j \equiv 0$  gelten und die Gleichung  $A(X)U_j(X) - B(X)V_j(X) = 0$  hat ebenfalls eine nicht-triviale Lösung. Nach Lemma ?? haben  $A$  und  $B$  dann mind  $j+1$  gemeinsame Nullstellen.  $\square$

Wir betrachten nun zwei Beispiele. Zunächst ein Beispiel, dass zwei Polynome mit festen Koeffizienten verwendet. Danach ein Beispiel, dass gemäß der Aufgabenstellung Koeffizienten verwendet, die selbst wieder Polynome sind.

**Beispiel I:** Betrachte  $A(X) = X^3 - 6X^2 + 11X - 6$  und  $B(X) = 2X^2 - 6X + 4$ , dann ist ( $0 \leq j \leq \min\{a, b\} - 1 = 1$ ). Als Übungsaufgabe für den Leser ist zu zeigen, dass  $\psi_0(A, B) = 0$  gilt.

Sei nun  $j = 1$ , dann wähle  $U_1(X) = u_0$  mit Grad  $b - j - 1 = 0$ , und  $V_1(X) = v_1X + v_0$  mit Grad  $a - j - 1 = 1$ . Aus  $A(X)U_j(X) - B(X)V_j(X) = 0$  folgt  $u_0(X^3 - 6X^2 + 11X - 6) = (2X^2 - 6X + 4)(v_1X + v_0)$ . Wir betrachten die ersten  $a + b - 2j = 3$  Gleichungen von  $a + b - j = 4$  Gleichungen mit  $a + b - 2j = 3$  Unbekannten. Aus  $u_0(X^3 - 6X^2 + 11X - 6) = 2v_1X^3 + (-6v_1 + 2v_0)X^2 + (4v_1 - 6v_0)X + 4v_0$  erhalten wir folgende Rechnung:

$$\psi_1(A, B) = \begin{vmatrix} 2 & 0 & -1 \\ -6 & 2 & 6 \\ 4 & -6 & -11 \end{vmatrix} = \begin{vmatrix} 2 & 0 & -1 \\ 6 & 2 & 0 \\ 18 & -6 & 0 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 1 \\ 6 & -2 & 0 \\ 0 & 0 & 0 \end{vmatrix} = 0.$$

Daraus folgt, dass  $A(X) = X^3 - 6X^2 + 11X - 6$  und  $B(X) = 2X^2 - 6X + 4$  mindestens zwei gemeinsame Nullstellen haben. Das stimmt auch,

denn  $(X-3)(X-2)(X-1) = A(X)$  und  $2(X-2)(X-1) = A(X)$  lässt sich leicht verifizieren.

**Beispiel II:** Sei  $P(X, Y, Z) = X^2 + Y^2 + Z^2 - 1$ . Gemäß der rekursiven Konstruktion einer zylindrischen Zerlegung wollen wir den Grad von  $\hat{P}(X_n) := P_{x_1, \dots, x_{n-1}}(X_n)$  und die Anzahl der gemeinsamen Nullstellen von  $\hat{P}(X_n)$  und  $\hat{P}'(X_n)$  konstant halten! Nun sei  $P_{X,Y}(Z) = Z^2 + (X^2 + Y^2 - 1) =: A(Z)$  und  $P'_{X,Y}(Z) = 2Z =: B(Z)$ . Offensichtlich hat  $P_{X,Y}(Z)$  konstant Grad 2 unabhängig von  $X$  und  $Y$ .

Für die Festlegung der Anzahl gemeinsamer Nullstellen betrachten wir  $(0 \leq j \leq \min\{a, b\} - 1 = 0)$ . Für  $j = 0$  haben wir  $U_0(Z) = u_0$  mit Grad  $b - j - 1 = 0$  und  $V_0(Z) = v_1Z + v_0$  mit Grad  $a - j - 1 = 1$ . Dann ist  $A(Z)U_j(Z) - B(Z)V_j(Z) = u_0(Z^2 + (X^2 + Y^2 - 1)) - (v_1Z + v_0)2Z = (u_0 - 2v_1)Z^2 - 2v_0Z + u_0(X^2 + Y^2 - 1) = 0$ . Die ersten  $a + b - 2j = 3$  Gleichungen mit  $a + b - j = 3$  Unbekannten ergeben:

$$\begin{vmatrix} -2 & 0 & 1 \\ 0 & -2 & 0 \\ 0 & 0 & (X^2 + Y^2 - 1) \end{vmatrix} = 4(X^2 + Y^2 - 1)$$

und das Vorzeichen von  $Q(X, Y) = 4(X^2 + Y^2 - 1)$  bestimmt die Anzahl der Nullstellen von  $P$ .

Aus den bisherigen Überlegungen lässt sich der folgende Vorschlag für die allgemeine Aufstellung der polynomiellen Bedingungen extrahieren:

**Theorem 3.13** Sei  $P_{(X_1, \dots, X_{n-1})}(X_n)$  gegeben und der Grad  $X_n$  in  $P$  sei  $k$ . Für  $j = 0, \dots, k$  sei  $Q_j(X_1, \dots, X_{n-1})$  der Koeffizient von  $P_{(X_1, \dots, X_{n-1})}(X_n)$  mit Grad  $j$ . Sei  $P_{(j, (X_1, \dots, X_{n-1}))}(X_n)$  die Summe aller Terme des Polynoms  $P_{(X_1, \dots, X_{n-1})}(X_n)$  mit Grad  $\leq j$  in  $X_n$ . Wir definieren weiterhin  $R_{(j,l)} := \psi_l \left( P_{(j, (X_1, \dots, X_{n-1}))}(X_n), P'_{(j, (X_1, \dots, X_{n-1}))}(X_n) \right)$  für  $l = 0, \dots, j - 2$ . Sei  $\mathcal{Q}$  die Menge der Polynome aus  $Q_j$  und  $R_{(j,l)}$ . Dann gilt: Die Anzahl der verschiedenen Nullstellen von  $P_{(X_1, \dots, X_{n-1})}(X_n)$  ist konstant falls alle Polynome aus  $\mathcal{Q}$  ein konstantes Vorzeichen haben.

Die Menge der betrachteten Polynome erscheint auf den ersten Blick etwas hochgegriffen. Es kann aber tatsächlich sein, dass einige Koeffizienten entfallen und die Betrachtungen dann für die entstehenden Polynome mit kleinerem Grad ebenfalls durchgeführt werden müssen. Insofern ist dieser konservative Ansatz gerechtfertigt.

### 3.4 Anwendung auf die Bewegungsplanung

Mit der Technik des vorherigen Abschnittes können wir nun folgendes zeigen:

**Theorem 3.14** (Schwartz, Sharir, 1983)

Das Problem der Bewegungsplanung für allgemeine Systeme ist effektiv berechenbar. [?]

**Beweis.**

$\mathcal{C}_{\text{frei}}$  sei als semialgebraische Menge gegeben (vgl. Formel (\*) auf Seite ??):

$$\mathcal{C}_{\text{frei}} = \left\{ (a_1, a_2, s_1, s_2) \in \mathbb{R}^4 \mid \forall X_1, X_2, X'_1, X'_2 : \underbrace{(\neg R(X_1, X_2) \vee \neg H(X'_1, X'_2) \vee \neg B(X_1, X_2, X'_1, X'_2, a_1, a_2, s_1, s_2))}_T \right\}$$

- Sei  $(\mathcal{K}_4, \sigma_4)$  eine  $T$ -invariante zylindrische algebraische Zerlegung des  $\mathbb{R}^4$  für die Polynome in  $R, H$  und  $B$  (diese ist effektiv, aber nicht effizient berechenbar). Sei  $\mathcal{K}_{\text{frei}}$  die Menge aller Zellen  $C$  in  $\mathcal{K}_4, \mathcal{K}_3, \mathcal{K}_2, \mathcal{K}_1$  mit  $C \subseteq \mathcal{C}_{\text{frei}}$ .

Wir nehmen zusätzlich an, daß  $\mathcal{C}_{\text{frei}}$  offen ist, also keine Hinderniskanten aneinanderstoßen, wie in der Abbildung skizziert. Damit genügt es für die Bahnplanung, alle drei- und vierdimensionalen freien Zellen zu betrachten.

4D

- Wir bauen den Zusammenhangsgraph aller drei- und vierdimensionalen Zellen aus  $\mathcal{K}_{\text{frei}}$ ; dabei ist stets 4D    3D eine vierdimensionale Zelle zu einer dreidimensionalen benachbart.

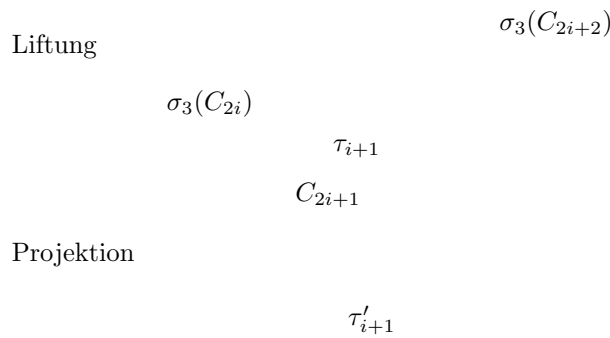
Zum Test einer dreidimensionalen Zelle  $C$  und einer vierdimensionalen Zelle  $C'$  auf Nachbarschaft benutzen wir folgende Eigenschaft:

$$\forall \varepsilon \forall q \exists q' : (\alpha_C(q) \wedge \varepsilon > 0 \rightarrow \alpha_{C'}(q') \wedge |q - q'| < \varepsilon).$$

Dabei bezeichnen  $\alpha_C, \alpha_{C'}$  die definierenden Tarski-Ausdrücke (quantorenfrei) für die Zellen  $C, C'$ . Nach Theorem ?? auf Seite ?? kann dieser Test effektiv entschieden werden.

Zu gegebener Start- und Zielkonfiguration  $s, t$ :

- Bestimme die Zellen  $C_s \ni s$  und  $C_t \ni t$  in  $\mathcal{K}_{\text{frei}}$  (teste durch Einsetzen von  $s, t$  in die definierenden Polynome sämtlicher Zellen  $C \in \mathcal{K}_{\text{frei}}$ , ob  $\alpha_C(s)$  bzw.  $\alpha_C(t)$  gilt).
- Wende Breitensuche an, um einen Pfad von  $C_s$  nach  $C_t$  im Zusammenhangsgraphen zu finden. Falls erfolgreich, liefert die Breitensuche eine Folge von Zellen  $C_s = C_1, C_2, C_3, \dots, C_{2p+1} = C_t$  mit  
 $C_{2i+1}$ ,  $0 \leq i \leq p$  hat Dimension 4,  
 $C_{2i}$ ,  $0 \leq i \leq p$  hat Dimension  $\leq 3$ ,  
 $C_{2i+1}$  ist zu  $C_{2i+2}$  benachbart ( $0 \leq i \leq p-1$ ),  $C_{2p}$  zu  $C_{2p+1}$ .

 $X_4$ Abbildung 3.6: Induktive Konstruktion von  $\pi$ : Projektion und Liftung.

- Konstruiere einen Pfad  $\pi$  von  $s$  nach  $t$  durch  $C_1, C_2, \dots, C_{2p+1}$ :
  - Von  $s$  nach  $\sigma_3(C_2)$ ,
  - von  $\sigma_3(C_{2i})$  zu  $\sigma_3(C_{2i+2})$ ,
  - von  $\sigma_3(C_{2p})$  nach  $t$ .

Die Konstruktion erfolgt induktiv über die Dimensionen: Angenommen, in  $\pi_4(C_{2i+1})$  ist Pfad  $\tau'_{i+1}$  von  $\pi_3(\sigma_3(C_{2i}))$  nach  $\pi_3(\sigma_3(C_{2i+2}))$  bereits konstruiert, dann lifte  $\tau'_{i+1}$  zu Pfad  $\tau_{i+1} \subset C_{2i+1}$  unter Beachtung der Abstände zu den Zellenwänden von  $C_{2i+1}$  in Richtung  $-X_4$  und  $+X_4$ , vgl. Abbildung ??.

Damit ist — im wesentlichen — gezeigt, daß das allgemeine Bewegungsplanungsproblem effektiv gelöst werden kann.  $\square$

Da das allgemeine Bewegungsplanungsproblem zwar effektiv, aber nicht effizient gelöst werden kann, setzt man in der Praxis approximative Verfahren ein oder verfolgt andere Ansätze zur Zellerlegung, die weniger Zellen produzieren.



## Kapitel 4

# Roboter in der Fertigungstechnik

Abschließend wollen wir ein Thema behandeln, das im Zusammenhang mit Robotern in der industriellen Fertigung interessant ist.

### 4.1 Orientierung polygonaler Werkstücke ohne Sensoren

Parts-  
Feeder

Abbildung 4.1: Parts-Feeder zur Orientierung von Werkstücken.

Stellen wir uns folgendes Problem vor: In einer Fertigungsanlage wird ein Fließband mit planaren Werkstücken — z. B. Rohlingen für Schlüssel — beschickt. Diese werden einer Maschine zur Verarbeitung zugeführt, die die Werkstücke nur verarbeiten kann, wenn diese in einer bestimmten Lage auf dem Fließband liegen. Die Lage der Werkstücke auf dem Fließband ist jedoch zunächst einmal zufällig; wenn etwa die Werkstücke aus einem Behälter auf das Band fallen. Zu konstruieren ist also eine Einheit, die die Werkstücke in die richtige Lage bringt, ein sogenannter Parts-Feeder. Die Eingabe in den Parts-Feeder ist ein Strom von Werkstücken in beliebiger Orientierung, die Ausgabe ein Strom von Werkstücken in eindeutiger Orientierung, siehe Abbildung ??.

Häufig sind Parts-Feeder direkt auf die Rohlinge abgestimmt, was zur Folge hat, daß die Hardware des Parts-Feeders auszutauschen ist, wenn sich die Werkstücke ändern. Interessanter wären reprogrammierbare Parts-Feeder, die aus Kostengründen ohne Sensoren auskommen sollten.

Abbildung 4.2: Parallel-Jaw Gripper.

Wir betrachten hier den Parallel-Jaw Gripper, der aus einem Greifer mit zwei parallelen Backen besteht und ein Werkstück greifen und drehen kann, siehe Abbildung ???. Dieser Greifer ermöglicht die Ausführung der folgenden Aktion (Greifaktion)  $A$ :

- Drehe den Greifer auf einen Winkel  $\alpha$  bzgl. einer Normalstellung.
- Schließe beide Backen gleichzeitig, bis das Werkstück eingeklemmt ist.
- Öffne beide Backen.

Der einzige Parameter einer Greifaktion ist der Winkel  $\alpha$ ; wir können also eine Folge von Aktionen  $A_i$  durch eine Folge von Winkeln  $\alpha_i$  beschreiben. Das Problem läßt sich damit folgendermaßen spezifizieren: Gegeben ist die Beschreibung eines polygonalen Werkstücks<sup>1</sup> mit  $n$  Kanten und eine Zielorientierung des Werkstücks. Finde einen Plan  $\mathcal{A} = (A_1, \dots, A_k) = (\alpha_1, \dots, \alpha_k)$  zur Ansteuerung des Greifers, der ein Werkstück in beliebiger Ausgangslage in die Zielorientierung bringt. Abbildung ??? zeigt vier Beispiele für die Ausführung eines Plans  $\mathcal{A} = (0^\circ, 45^\circ)$  mit unterschiedlichen Ausgangslagen.

Die Zielorientierung kann dabei nur bis auf Symmetrien in der konvexen Hülle eindeutig sein, in nebenstehender Abbildung hat der Greifer keine Möglichkeit, die Orientierungen (a) und (b) voneinander zu unterscheiden.

1	4
(a)	
2	3
3	2
(b)	
4	1

Fassen wir die Annahmen zusammen:

1. Der Greifer besteht aus zwei parallelen Backen.

---

<sup>1</sup>D.h. die konvexe Hülle des Werkstücks ist ein Polygon.



Abbildung 4.3: Vier Beispiele für die Ausführung eines Plans  $\mathcal{A} = (0^\circ, 45^\circ)$ .

2. Die Richtung des Greifers ist orthogonal zu den Backen.
3. Die konvexe Hülle des Werkstücks wird als festes, planares Polygon angesehen, alle Aktionen beziehen sich auf die konvexe Hülle des Werkstücks.
4. Das Werkstück liegt isoliert vor, d. h. es muß nicht aus einem Haufen anderer Werkstücke gezogen werden und der Abstand zwischen zwei Werkstücken auf dem Band ist groß genug.
5. Die Ausgangsposition des Werkstücks ist unbekannt; bekannt ist nur, daß es zwischen den Backen des Greifers liegt.
6. Beide Backen treffen gleichzeitig auf das Werkstück (Pure Squeezing). Wir werden später betrachten, wie man mit einer kombinierten Schiebe-/Greifaktion ohne diese Annahme auskommt.
7. Es gibt keine Reibung zwischen den Backen und dem Werkstück.

Die letzte Annahme soll verhindern, daß sich das Werkstück zwischen den Backen einklemmt, wir nennen so eine Lage ein **instabiles Gleichgewicht**. Man kann diese Annahme durch mechanische Abhilfen wie Rollbänder am Greifer oder Vibrationen simulieren und algorithmisch verhindern, daß das Werkstück nach dem ersten Schritt in eine Lage gebracht wird, die zu einem instabilen Gleichgewicht führen könnte.

Eine Lage, in der mindestens eine Kante der konvexen Hülle des Werkstücks mit einer Backe Kontakt hat, nennen wir ein **stabiles Gleichgewicht**.

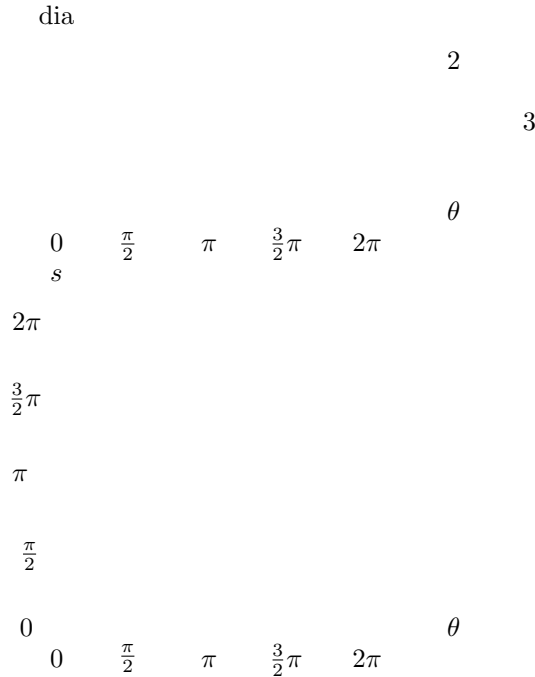


Abbildung 4.4: Durchmesser- und Greiffunktion.

Der Algorithmus zur Berechnung eines Plans für den Greifer stützt sich auf die beiden im folgenden definierten Funktionen:

**Definition 4.1**

- (i) Die **Durchmesserfunktion**  $\text{dia} : [0, 2\pi) \rightarrow \mathbb{R}$ , gibt für eine feste Orientierung des Werkstücks den Abstand  $\text{dia}(\theta)$  zweier paralleler Linien mit Winkel  $\theta$  an, die das Werkstück berühren.
- (ii) Die **Greiffunktion**  $s : [0, 2\pi) \rightarrow [0, 2\pi)$  bildet eine Orientierung  $\theta$  des Werkstücks *bezüglich des Greifers* auf eine Orientierung  $s(\theta)$  bezüglich des Greifers nach der Greifaktion ab. Sie ist stückweise konstant und berechnet sich aus der Durchmesserfunktion, indem alle Orientierungen, die zwischen zwei lokalen Maxima von  $\text{dia}(\theta)$  liegen, auf die Orientierung  $\theta$  abgebildet werden, an der  $\text{dia}(\theta)$  das lokale Minimum zwischen den beiden Maxima annimmt, vgl. Abbildung ???. Wir legen fest, daß die Intervalle nach rechts offen und nach links abgeschlossen sind.
- (iii) Für eine Sequenz von Aktionen  $\mathcal{A} = (\alpha_1, \alpha_2, \dots, \alpha_{k-1}, \alpha_k)$  und eine Ausgangsposition  $\theta$  bezeichne

$$S(\mathcal{A}, \theta) := s(s(\dots s(s(\theta - \alpha_1) - \alpha_2) \dots) - \alpha_{k-1}) - \alpha_k)$$

die Position des Werkstücks bzgl. des Greifers nach Ausführung der Sequenz  $\mathcal{A}$ .

Bei der Greiffunktion ist zu beachten, daß sie bezüglich des Greifers definiert ist. Befindet sich das Werkstück in einer Orientierung  $\theta$  bzgl. des Greifers und der Greifer führt eine Greifaktion bei Winkel  $\alpha$  aus, so ist die Orientierung des Werkstücks bzgl. des Greifers nach der Aktion  $s(\theta - \alpha)$ .

Die Durchmesserfunktion ist uns schon bei der Berechnung von kürzesten Pfaden für Liniensegmenten, in Theorem ?? auf Seite ?? begegnet. Sie beschreibt, wie sich der Abstand zweier paralleler Linien — in unserem Fall die Backen des Greifers — ändert, wenn die Linien um das Werkstück rotieren. Ein stabiles Gleichgewicht wird offensichtlich in den lokalen Minima der Durchmesserfunktion erreicht.

dia

dreifache	vierfache	fünffache	$\pi$	$\theta$
	Rotationssymmetrie			
$T = \frac{\pi}{3}$	$T = \frac{\pi}{2}$	$T = \frac{\pi}{5}$	$r = 3$	

Abbildung 4.5: Rotationssymmetrie und Periodizität von dia; dia für  $r = 3$ .

Die Durchmesserfunktion hat aufgrund ihrer Definition stets eine Periode von  $\pi$ : eine Rotation der parallelen Linien (bzw. des Greifers) um  $180^\circ$  liefert denselben Durchmesser. Weitere Perioden entstehen durch Rotationssymmetrie in der konvexen Hülle des Werkstücks, siehe Abbildung ?. Perioden in der Durchmesserfunktion induzieren eine Struktur in die Greiffunktion:

**Definition 4.2** Eine Greiffunktion  $s$  ist **T-periodisch**, falls für alle  $\theta \in [0, 2\pi)$  gilt:

$$s(\theta + T) = s(\theta) + T.$$

**Lemma 4.3**

(i) Für Polygone mit  $r$ -facher Rotationssymmetrie ist die Greiffunktion  $T$ -periodisch mit

$$T = \frac{2\pi}{r(1 + (r \bmod 2))}.$$

(ii) Für eine  $T$ -periodische Greiffunktion  $s$  gilt:

$$S(\mathcal{A}, \theta + T) = S(\mathcal{A}, \theta) + T.$$

**Beweis.** Übungsaufgabe. □

Aus Lemma ??(ii) folgt, daß es keinen Plan gibt, der die Orientierungen  $\theta$  und  $\theta + T$  in eine gemeinsame Zielorientierung überführen kann.

**Definition 4.4** Gegeben sei ein Werkstück  $W$ . Sei  $T$  die kleinste Periode der zu  $W$  gehörenden Greiffunktion  $s$ . Ein Plan  $\mathcal{A}$  orientiert die konvexe Hülle von  $W$  **bis auf Symmetrie**, falls in der Menge der möglichen letzten Orientierungen  $S(\mathcal{A}, \theta)$  genau  $\frac{2\pi}{T}$  Orientierungen sind, die gleichverteilt auf  $[0, 2\pi)$  liegen.

Für das Werkstück in Abbildung ?? z. B. gilt  $r = 2, T = \pi$ . Für den Plan  $\mathcal{A} = (0^\circ, 45^\circ)$  gibt es die möglichen Zielorientierungen  $\frac{\pi}{4}$  und  $\frac{5}{4}\pi$ .

Zur Beschreibung des Algorithmus müssen wir noch einige Begriffe festlegen:

- (i) Ein **Intervall**  $\Theta$  ist eine zusammenhängende Teilmenge von  $[0, 2\pi)$ ,  $|\Theta|$  bezeichne die Länge des Intervalls.
- (ii) Ein **s-Intervall** ist ein halboffenes Intervall  $[\xi_i, \nu_i)$ , wobei  $\xi_i$  und  $\nu_i$  Unstetigkeitsstellen der Greiffunktion sind. Da die Greiffunktion  $O(n)$  Unstetigkeitsstellen hat, gibt es  $O(n^2)$  eindeutige s-Intervalle.
- (iii) Zu einem s-Intervall  $\Theta$  sei das **s-Image**  $s(\Theta)$  das kleinste Intervall, das die Menge

$$\{ s(\theta) \mid \theta \in \Theta \}$$

enthält. Ein s-Image ist stets ein abgeschlossenes Intervall.

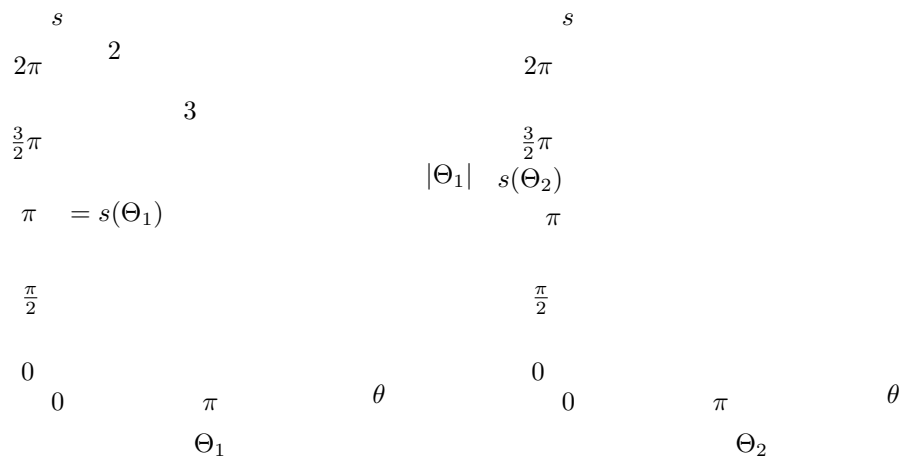


Abbildung 4.6: Finden von s-Intervallen.

Algorithmus ?? findet einen Plan zur Orientierung des Werkstücks. Zur Illustration betrachten wir ein Rechteck mit den Kantenlängen 2 und 3 (Abbildung ??). Da das Seitenverhältnis  $\frac{3}{2}$  beträgt, sei  $a = \arctan \frac{3}{2}$ . In Schritt 2 wird der größte Schritt in der Greiffunktion gefunden, in unserem

**Algorithmus 4.1** Orientierung eines Werkstücks (Pure Squeezing)

- Berechne die Durchmesserfunktion und die Greiffunktion.
- Bestimme das längste  $s$ -Intervall  $\Theta_1$ , über dem die Greiffunktion stetig ist. Wir legen fest, daß dieses Intervall zur Periode gehört. Setze  $i := 1$ .
- Solange ein  $s$ -Intervall  $\Theta$  mit  $|s(\Theta)| < |\Theta_i|$  existiert:
  - Setze  $\Theta_{i+1}$  auf das Größte dieser  $s$ -Intervalle.
  - Inkrementiere  $i$ .

$\Rightarrow$  Liste  $\mathcal{L} = (\Theta_1, \Theta_2, \dots, \Theta_i)$  mit  $|\Theta_i| = T$ .

- Berechne aus  $\mathcal{L}$  einen Plan  $\mathcal{A} = (\alpha_i, \alpha_{i-1}, \dots, \alpha_1)$ :

- $\alpha_i := 0$ .
- Für  $j = i - 1, i - 2, \dots, 1$ :

$$\alpha_j := s(\xi_{j+1}) - \xi_j - \varepsilon_j + \alpha_{j+1}.$$

Dabei ist  $\varepsilon_j = \frac{1}{2} (|\Theta_j| - |s(\Theta_{j+1})|)$  eine Fehlertoleranz, u.a. zur Vermeidung instabiler Gleichgewichte.

Fall  $\Theta_1 = [\pi - a, \pi + a)$  mit  $|\Theta_1| = 2a$ . Man beachte, daß  $s(\Theta_1)$  eine einzige Orientierung mit Winkel  $\pi$  ist (Abbildung ??(i)). Im dritten Schritt suchen wir das größte  $s$ -Intervall mit einem  $s$ -Image kleiner als  $\Theta_1$ . Wie in Abbildung ??(ii) gezeigt, kann dies durch Anlegen eines Quadrats mit Seitenlängen  $|\Theta_i|$  linksbündig an alle Stufen der Greiffunktion gesucht werden. Berührt das nächste Intervall der Greiffunktion das Quadrat, hat man ein solches  $s$ -Intervall gefunden, hier bei  $\Theta_2 = [\pi - a, 2\pi - a)$  mit  $s(\Theta_2) = [\pi, \frac{3}{2}\pi]$  und  $|s(\Theta_2)| = \frac{\pi}{2} < |\Theta_1|$ . So fahren wir fort, bis der Algorithmus mit  $|\Theta_i| = T$  terminiert, was in unserem Beispiel wegen  $|\Theta_2| = \pi$  bereits der Fall ist.

Unser Werkstück liegt zu Beginn in einer Orientierung  $\theta \in \Theta_2$ . Nach der ersten Aktion bei Winkel  $0^\circ$  liegt die Orientierung des Werkstücks in  $s(\Theta_2)$ . Wegen  $|s(\Theta_2)| < |\Theta_1|$  können alle Orientierungen aus  $s(\Theta_2)$  mit einer Aktion auf einen Punkt abgebildet werden, wenn  $s(\Theta_2)$  in  $\Theta_1$  ausgerichtet werden kann. Dies kann durch Drehung des Greifers auf den Winkel  $s(\xi_2) - \xi_1$  erreicht werden. Da die Orientierung des Werkstücks dann in  $\Theta_1$  bzgl. der Greiferstellung liegt, kann mit einer zweiten Aktion das Werkstück in die Orientierung  $s(\Theta_1)$  gebracht werden.

Wenn die Orientierung des Werkstücks zu Beginn im Bereich  $\Theta_i$  liegt, wird es auf die Orientierung  $s(\Theta_i) + \alpha_i$  gedreht. Wegen  $|\Theta_i| = T$  orientiert der Plan  $\mathcal{A}$  das Werkstück bis auf Symmetrie. Unser Rechteck wird also mit dem Plan  $\mathcal{A} = (0, \frac{\pi}{4})$  auf den Winkel  $\frac{\pi}{4}$  gedreht, wenn seine Orientierung

zu Beginn im Bereich  $[-a, \pi - a)$  lag, und auf den Winkel  $\frac{5}{4}\pi$ , wenn es im Bereich  $[\pi - a, -a)$  lag.

**Theorem 4.5** (Goldberg, 1993) *Gegeben sei eine Liste von  $n$  Kanten, die die konvexe Hülle eines gegebenen Werkstücks repräsentieren. Dann läßt sich in Zeit  $O(n^2 \log n)$  die kürzeste Sequenz von Greifaktionen finden, die eine Orientierung des Werkstücks bis auf Symmetrie garantiert. Der gefundene Plan hat eine Länge von  $O(n^2)$ . [?]*

**Beweis.**

Wir zeigen die Korrektheit und die Vollständigkeit von Algorithmus ?? und schätzen seine Laufzeit ab:

**Korrektheit:**

Zu zeigen ist: (i) der berechnete Plan orientiert das Werkstück eindeutig bis auf Symmetrie, und (ii) es gibt keinen kürzeren.

- (i) Der Algorithmus findet einen Plan, der ein  $s$ -Intervall  $\Theta$  der Länge  $T$  auf einen Punkt  $\theta'$  abbildet, wobei  $T$  die kleinste Periode der Greiffunktion des Werkstücks ist. Wegen  $s(\theta + T) = s(\theta) + T$  gilt für jeden Plan  $\mathcal{A}$ :  $\mathcal{A}(\theta + T) = \mathcal{A}(\theta) + T$ . Z.B. gilt für einen zweistufigen Plan  $\mathcal{A} = (\alpha_1, \alpha_2)$ :

$$\begin{aligned} S(\mathcal{A}, \theta + T) &= s(s(\theta + T - \alpha_1) - \alpha_2) \\ &= s(s(\theta - \alpha_1) + T - \alpha_2) \\ &= s(s(\theta - \alpha_1) - \alpha_2) + T \\ &= S(\mathcal{A}, \theta) + T \\ &= \theta' + T \end{aligned}$$

Wenn der Plan alle Orientierungen eines  $s$ -Intervalls  $\Theta$  auf  $\theta'$  abbildet, bildet er alle Orientierungen eines  $s$ -Intervalls  $\Theta'$  mit gleicher Länge wie  $\Theta$  und einem Offset von  $kT$  auf  $\theta' + kT$  ab. Die Zielorientierung wird also eine der  $\frac{2\pi}{T}$  Orientierungen sein, die gleichverteilt auf  $[0, 2\pi)$  liegen. Der Plan wird also das Werkstück bis auf Symmetrie orientieren.

- (ii) Zum Beweis, daß es keinen kürzeren Plan gibt, benutzen wir folgende

**Behauptung:** Jeder Plan, der eine Menge  $\Theta \subseteq [0, \pi)$  auf einen Punkt  $\theta$  abbildet, bildet auch das kleinste zusammenhängende Intervall, das  $\Theta$  enthält, auf  $\theta$  ab.

**Beweis:** Sei  $\Theta'$  das kleinste zusammenhängende Intervall, das  $\Theta$  enthält. Wegen der Monotonie der Greiffunktion gilt  $s(\Theta') = s(\Theta)$ . Die erste Greifaktion wird also in beiden Fällen das gleiche  $s$ -Image erzeugen und mit dem Rest des Plans kann  $\Theta'$  auf einen Punkt abgebildet werden.  $\diamond$

Nehmen wir nun an, es existiert ein Plan  $\mathcal{A}'$  mit weniger Schritten als  $\mathcal{A}$ . Sei  $(\Theta_1, \Theta_2, \dots, \Theta_i)$  die Liste der  $s$ -Intervalle, die der Algorithmus gefunden hat,  $(\Theta'_1, \Theta'_2, \dots, \Theta'_j)$  seien die zum Plan  $\mathcal{A}'$  gehörenden Intervalle erweitert auf zusammenhängende Intervalle. Aufgrund des Algorithmus gilt  $|\Theta_1| \geq |\Theta'_1|$ . Da  $\mathcal{A}'$  nach  $j$  Schritten terminiert,  $\mathcal{A}$  jedoch nicht, muß  $|\Theta_j| < |\Theta'_j|$  gelten. Also existiert ein  $k$  mit  $|\Theta_k| \geq |\Theta'_k|$  und  $|\Theta_{k+1}| < |\Theta'_{k+1}|$ .  
 $\zeta$ , der Algorithmus hätte das größere Intervall  $\Theta'_{k+1}$  gewählt.

### Vollständigkeit:

Für jedes Werkstück finden wir einen Plan. Aus technischen Gründen erweitern wir hier die Greiffunktion mit Periode  $T$  auf  $s : \mathbb{R} \rightarrow \mathbb{R}$  — wegen  $s(\theta + T) = s(\theta) + T$  ist dies möglich. Wir beweisen die Vollständigkeit, indem wir zeigen, daß wir für jede Greiffunktion stets eine Sequenz von  $s$ -Intervallen  $(\Theta_1, \Theta_2, \dots, \Theta_n)$  finden, bei der  $|\Theta_j|$  größer ist als  $|s(\Theta_{j+1})|$ . Wir zeigen, daß wir für jede stückweise konstante, monotone Greiffunktion und jedes  $s$ -Intervall stets ein größeres  $s$ -Intervall finden, bis wir auf die Periode  $T$  in der Funktion stoßen.

Sei  $h$  die Länge des bereits gefundenen  $s$ -Intervalls, dann ist zu zeigen:

- (i) Entweder finden wir ein  $s$ -Intervall, dessen  $s$ -Image kleiner ist:

$$\exists \theta : s(\theta + h) - s(\theta) < h,$$

- (ii) oder  $h$  ist die Periode der Greiffunktion:

$$\forall \theta : s(\theta + h) = s(\theta) + h.$$

(i) sagt aus, daß zum aktuellen  $s$ -Intervall  $\Theta_j = [\theta_j, \theta_j + h)$  ein Intervall  $\Theta = [\theta, \theta + h]$  existiert, dessen  $s$ -Image kleiner ist als  $\Theta_j$ . Wir können  $\Theta$  nach rechts bis zur nächsten Unstetigkeitsstelle erweitern, ohne das  $s$ -Image zu vergrößern. Dies liefert ein  $s$ -Intervall  $\Theta'$  mit kleinerem  $s$ -Image als  $\Theta_j$ , dessen Differenz zu  $\Theta$  ein offenes Intervall ist, dessen Länge somit ungleich Null ist. Also gilt  $|\Theta'| > |\Theta_j|$ .

Um zu zeigen, daß entweder (i) oder (ii) gilt, betrachte das Integral von  $s(\theta + h) - s(\theta) - h$  über  $[0, T)$ :

$$\begin{aligned} \int_0^T s(\theta + h) - s(\theta) - h \, d\theta &= \int_h^{T+h} s(\theta) \, d\theta - \int_0^T s(\theta) \, d\theta - hT \\ &=^2 - \int_0^h s(\theta) \, d\theta + \int_T^{T+h} s(\theta) \, d\theta - hT \\ &= - \int_0^h s(\theta) \, d\theta + \int_0^h s(\theta) + T \, d\theta - hT \\ &= - \int_0^h s(\theta) \, d\theta + \int_0^h s(\theta) \, d\theta + hT - hT \\ &= 0 \end{aligned}$$

Also existiert ein Punkt, an dem die Funktion einen negativen Wert annimmt (i), oder die Funktion ist konstant Null (ii).

**Laufzeit:**

Durchmesser- und Greiffunktion lassen sich in Zeit  $O(n)$  berechnen. Auch der größte Teilschritt in der Greiffunktion (Schritt 2) läßt sich in Zeit  $O(n)$  berechnen, da die Funktion maximal  $O(n)$  Unstetigkeitsstellen hat. In der Schleife (Schritt 3) werden  $O(n^2)$  s-Intervalle durchlaufen, die in einer nach  $|s(\Theta)|$  sortierten Liste gespeichert sind; Schritt 3 benötigt also Zeit  $O(n^2)$ . Die Konstruktion des Plans aus der Liste der Intervalle  $(\Theta_1, \Theta_2, \dots, \Theta_i)$  ist in Zeit  $O(i)$  möglich. Es dominiert die Zeit für das Sortieren der Intervalle; insgesamt erhalten wir eine Laufzeit von  $O(n^2 \log n)$ . □

Bisher sind wir von der unrealistischen Annahme ausgegangen, daß beide Backen des Greifers das Werkstück gleichzeitig berühren. Von dieser Annahme wollen wir uns im folgenden lösen und Aktionen betrachten, bei denen das Werkstück erst ein Stück geschoben und dann gegriffen wird (Push-and-Grasp). Eine Aktion des Greifers besteht jetzt aus folgenden Schritten:

- Drehe den Greifer auf einen Winkel  $\alpha$  bzgl. einer Normalstellung.
- Verschiebe den Greifer in Richtung  $\alpha + \frac{\pi}{2}$  um eine Distanz  $\delta$ .
- Schließe beide Backen gleichzeitig, bis das Werkstück eingeklemmt ist.
- Verschiebe den Greifer in Richtung  $-(\alpha + \frac{\pi}{2})$  um  $-\delta$ .
- Öffne beide Backen.

Wir nehmen jetzt an, daß der Schwerpunkt des Werkstücks bekannt ist und die Distanz  $\delta$  so groß ist, daß das Werkstück durch Verschieben in eine stabile Orientierung gebracht werden kann.

Analog zur Durchmesser- und Greiffunktion benutzen wir hier folgende Funktionen:

**Definition 4.6**

- (i) Die **Radiusfunktion**  $\text{rad} : [0, 2\pi) \rightarrow \mathbb{R}$  gibt für eine feste Orientierung des Werkstücks den Abstand  $\text{rad}(\theta)$  zwischen dem Schwerpunkt und einer Linie mit Winkel  $\theta$  an, die das Werkstück berührt.
- (ii) Die **Schiebefunktion**  $p : [0, 2\pi) \rightarrow [0, 2\pi)$  bildet eine Orientierung  $\theta$  des Werkstücks *bezüglich des Greifers* auf eine Orientierung  $p(\theta)$  bezüglich des Greifers nach der Schiebeaktion ab. Sie ist stückweise konstant und berechnet sich aus der Radiusfunktion, indem alle Orientierungen, die zwischen zwei lokalen Maxima von  $\text{rad}(\theta)$  liegen, auf die Orientierung  $\theta$  abgebildet werden, an der  $\text{rad}(\theta)$  das lokale Minimum zwischen den beiden Maxima annimmt.

---

<sup>2</sup>Wegen der Periodizität gilt:  $\int_h^{T+h} \dots = \int_0^T \dots - \int_0^h \dots + \int_T^{T+h} \dots$



- (iii) Die **Transferfunktion**  $g : [0, 2\pi) \rightarrow [0, 2\pi)$  beschreibt eine Aktion, bei der zunächst geschoben, danach gegriffen wird:

$$g := s \circ p.$$

Damit können wir mit Algorithmus ?? einen Plan berechnen; im ersten Schritt ist natürlich die Transferfunktion statt der Greiffunktion zu berechnen.



# Anhang A

## Grundlagen

In diesem Kapitel sollen einige Themen behandelt werden, die nicht unmittelbar zum Thema Bewegungsplanung für Roboter gehören, wohl aber Grundlagen zu diesem Thema sind und zum Teil aus anderen Vorlesungen (Theoretische Informatik, Algorithmische Geometrie etc.) bekannt sind. Aus diesem Grunde hat die Darstellung hier auch eher den Charakter eines kurzen Einblicks zur Wiederholung des Stoffs.

### A.1 Komplexitätsklassen

Die Klasse  $\mathcal{NP}$  ist die Menge der Probleme, die von einer nichtdeterministischen Turingmaschine in polynomieller Zeit gelöst werden können. Ein Problem  $S$  ist eine Menge  $S \subset \Omega$ . Zu einer Eingabe  $x \in \Omega$  ist zu entscheiden, ob  $x \in S$  gilt.

Eine nichtdeterministische Turingmaschine  $M$  löst ein Problem in polynomieller Zeit, wenn es ein Polynom  $p(n)$  gibt, so daß für alle  $x \in \Omega$  genau dann  $x \in S$  gilt, wenn ein Rechengang von  $M$  existiert, so daß  $M$ , angesetzt auf  $x$ , nach  $p(|x|)$  Schritten mit der Antwort "ja" stoppt.

**Definition A.1** Ein Problem  $S$  über  $\Omega$  heißt **NP-vollständig**, wenn gilt:

1.  $S \in \mathcal{NP}$ .
2. Jedes andere Problem  $S' \in \mathcal{NP}$  läßt sich in polynomieller Zeit auf  $S$  reduzieren, d. h. für alle  $S'$  über  $\Omega'$  in  $\mathcal{NP}$  gibt es eine Funktion  $f : \Omega' \rightarrow \Omega$  und ein Polynom  $p(n)$ , so daß:
  - (i)  $\forall x' \in \Omega' : f(x')$  ist in Zeit  $p(|x'|)$  berechenbar, und
  - (ii)  $\forall x' \in \Omega' : f(x') \in S \Leftrightarrow x' \in S'$ .

Anders ausgedrückt: gäbe es eine schnelle Lösung für  $S$ , so auch für jedes andere Problem  $S' \in \mathcal{NP}$ .

Ein Problem  $S$  heißt **NP-hart**, wenn es die zweite Forderung erfüllt.

Ein prominentes Beispiel für NP-vollständige Probleme ist das Problem SAT: gegeben seien  $m$  Klauseln über  $n$  booleschen Variablen. Eine Klausel  $C_i$  ist eine Disjunktion von  $k_i$  Literalen, also negierten oder nicht negierten Variablen:

$$C_i = \bigvee_{j=1}^{k_i} \pi_{ij}, \quad \pi_{ij} \in \{X_\ell, \overline{X}_\ell, 1 \leq \ell \leq n\}, \quad 1 \leq i \leq m.$$

Es soll entschieden werden, ob eine Belegung  $w = (w_1, \dots, w_n) \in \{0, 1\}^n$  existiert, so daß alle Klauseln erfüllt sind.

SAT liegt in  $\mathcal{NP}$ , da eine nichtdeterministische Turingmaschine eine Variablenbelegung raten und in polynomieller Zeit überprüfen kann, ob der Ausdruck damit “wahr” wird. Weiterhin kann SAT zur Simulation der Rechengänge einer nichtdeterministischen Turingmaschine — die ein beliebiges Problem  $S \in \mathcal{NP}$  lösen kann — benutzt werden. Also gilt:

**Theorem A.2** (Cook, 1971)

*SAT ist NP-vollständig.*

[?]

Bei dem Problem 3-SAT sind  $m$  Klauseln mit je drei Literalen gegeben, und es soll entschieden werden, ob eine Wahrheitswertbelegung existiert, die den Ausdruck erfüllt:

$$\Omega = \left\{ \alpha = \bigwedge_{i=1}^m (\pi_{i1} \vee \pi_{i2} \vee \pi_{i3}), \quad \pi_{ij} \in \{X_\ell, \overline{X}_\ell, 1 \leq \ell \leq n\} \right\},$$

$$S = \left\{ \alpha \in \Omega \mid \alpha \text{ erfüllbar} \right\}.$$

Da SAT sich auf 3-SAT reduzieren läßt, ist bereits 3-SAT NP-vollständig.<sup>1</sup>

Um ein anderes Problem als NP-hart nachzuweisen, genügt es nun zu zeigen, daß SAT oder 3-SAT sich auf dieses Problem reduzieren läßt. Für die NP-Vollständigkeit ist zusätzlich zu zeigen, daß das Problem in  $\mathcal{NP}$  liegt.

Eine andere Klasse von Problemen ist die Klasse  $\mathcal{PSPACE}$ :

**Definition A.3** Ein Problem  $S$  gehört zu  $\mathcal{PSPACE}$ , wenn  $S$  von einer deterministischen Turingmaschine auf polynomiell beschränktem Band berechnet werden kann. Ein Problem  $S$  heißt **PSPACE-hart**, wenn sich jedes andere Problem aus  $\mathcal{PSPACE}$  auf  $S$  reduzieren läßt, analog zu Definition ??

<sup>1</sup>Das Problem 2-SAT dagegen, bei dem die Klauseln aus zwei Literalen bestehen, läßt sich in polynomieller Zeit lösen.

Da bei den Problemen aus  $\mathcal{PSPACE}$  keine Beschränkung in der Laufzeit gefordert ist, die Laufzeit bei Problemen der Klasse  $\mathcal{NP}$  jedoch polynomiell beschränkt ist — und damit erst recht der verwendete Speicherplatz —, kann eine deterministische Turingmaschine alle möglichen Rechengänge einer nichtdeterministischen Turingmaschine nacheinander berechnen und prüfen. Der dabei benutzte Speicherplatz ist polynomiell beschränkt. Es gilt somit die Hierarchie:<sup>2</sup>

$$\mathcal{P} \subseteq \mathcal{NP} \subseteq \mathcal{PSPACE}.$$

Neben der berühmten Frage  $\mathcal{P} \stackrel{?}{=} \mathcal{NP}$  ist auch die Frage  $\mathcal{P} \stackrel{?}{=} \mathcal{PSPACE}$  offen.

Ein  $\mathcal{PSPACE}$ -vollständiges Problem ist die Erfüllbarkeit boolescher Ausdrücke mit Quantoren (QSAT), also Ausdrücke wie z. B.

$$\forall x \exists y \forall z : (\bar{x} \wedge z \longrightarrow y \vee \bar{z}).$$

Als Vertreter für NP-vollständige Probleme haben wir also die Erfüllbarkeit boolescher Ausdrücke *ohne* Quantoren (SAT), als Vertreter für  $\mathcal{PSPACE}$ -vollständige Probleme das gleiche *mit* Quantoren.

Mehr zum Thema Komplexitätsklassen und NP-vollständige Probleme findet sich z. B. im Buch von Garey und Johnson [?].

---

<sup>2</sup> $\mathcal{P}$  bezeichne die Menge aller in polynomieller Zeit lösbarer Probleme.

## A.2 Entscheidbarkeit

In der Regel denken wir über Fragen der Laufzeit nach. Manche Probleme sind allerdings so schwierig, daß sich die Frage stellt, ob sie überhaupt mit Hilfe von Computern lösbar sind. Doch was heißt überhaupt entscheidbar?

**Definition A.4** Ein Problem  $S \subset \Omega$  heißt **entscheidbar**, wenn es eine Turingmaschine  $M$  gibt, so daß

$$\forall x \in \Omega : \quad M \text{ angesetzt auf } x$$

- 1) stoppt und
- 2) liefert die Antwort "ja", falls  $x \in S$   
liefert die Antwort "nein", falls  $x \notin S$

Tatsächlich gibt es unentscheidbare Probleme — und das ist eine sehr wichtige Erkenntnis. Daß es unentscheidbare Probleme geben muß, kann man mit einem einfachen Abzählbarkeitsargument begründen: es gibt nur abzählbar unendlich viele Turingmaschinen, aber überabzählbar viele Teilmengen  $S$  der Menge  $\Omega = \mathbb{N}$ .

Beispiele für unentscheidbare Probleme sind:

- Die Überprüfung der semantischen Korrektheit für (Turing-)Programme. Bereits das Halteproblem, also die Frage, ob ein gegebenes Programm nach endlicher Zeit anhält, ist unentscheidbar.
- Gegeben seien  $n$  Typen von Dominosteinen mit je 4 Farben
 

blau		gelb	
blau	rot	rot	gelb
gelb		grün	

 (siehe Abbildung). Die Frage, ob man mit diesen Steinen die Ebene so kacheln kann, so daß nur gleiche Farben benachbart sind, ist unentscheidbar.
- Das "Game of Life" von John Conway — vorgestellt z. B. von Gardner [?]<sup>3</sup> — wird in einem zweidimensionalen Gitter gespielt. Jede Zelle des Gitters hat entweder den Status "lebend" oder "tot". Eine Verteilung von lebenden und toten Zellen (Generation) erzeugt eine Nachfolgeneration, indem für jede Zelle gezählt wird, wie viele ihrer acht Nachbarzellen leben. Der Status einer Zelle in der Nachfolgeneration ergibt sich durch folgende Regeln:

<sup>3</sup><http://acf5.nyu.edu/~mm64/x52.9264/october1970.html>  
 Siehe auch <http://hensel.lifepatterns.net/> (Java-Applet)

**Tod:** Eine Zelle ist in der folgenden Generation tot, wenn sie in der aktuellen Generation weniger als 2 oder mehr als 3 lebende Nachbarn hat (Isolation bzw. Überbevölkerung).

**Überleben:** Eine lebende Zelle überlebt, wenn sie genau 2 oder 3 lebende Nachbarn hat.

**Geburt:** Eine tote Zelle wird zu einer lebenden, wenn sie genau 3 lebende Nachbarn hat.

Das Schicksal einer Population im Game of Life ist unentscheidbar.

- Die Lösbarkeit Diophantischer Gleichungen (Hilberts 10. Problem): gegeben ein Polynom  $P(X_1, \dots, X_n) \in \mathbb{Z}[X_1, \dots, X_n]$ . Gibt es eine Belegung  $(x_1, \dots, x_n) \in \mathbb{Z}^n$ , so daß  $P(x_1, \dots, x_n) = 0$  gilt?

Ein Beispiel für eine Diophantische Gleichung ist Fermats Problem:

$$X^n + Y^n = Z^n.$$

Entscheidbar dagegen ist die Theorie der reellen Zahlen — die Menge aller Tarski–Ausdrücke<sup>4</sup> ohne freie Variablen, die über ganz  $\mathbb{R}$  gelten:

$$\Omega = \{ T \mid T \text{ ist Tarski–Ausdruck ohne freie Variablen} \}$$

$$Th(\mathbb{R}) = \{ T \in \Omega \mid T \text{ gilt über } \mathbb{R} \}.$$

Weitere Ergebnisse zum Thema Entscheidbarkeit finden sich z. B. in dem Buch von Börger et. al. [?].

---

<sup>4</sup>Siehe Definition ?? auf Seite ??.

## A.3 Voronoi–Diagramme

### A.3.1 Definition und Eigenschaften

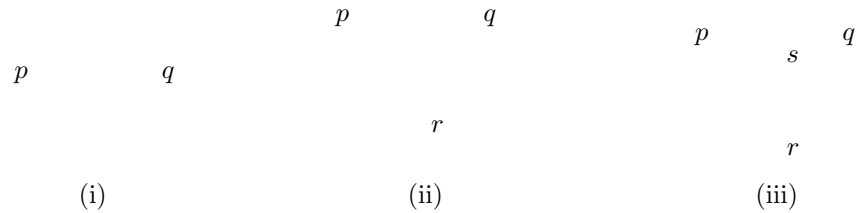


Abbildung A.1: (i) Bisektor zwischen  $p$  und  $q$ , (ii) Regionen von  $p, q, r$ , (iii) Regionen von  $p, q, r, s$ .

Es waren einmal zwei mächtige Fürsten — da ihre Namen nicht überliefert sind, nennen wir sie  $p$  und  $q$  —, die das weite Land Errquadrat unter sich aufteilen wollten. Da beide Fürsten gleich mächtig waren, beschlossen ihre Weisen — Delaunay und Voronoi — jeden Punkt im Lande dem Fürsten zuzuordnen, zu dessen Schloß er den kleinsten Abstand hat. Die dabei entstehende Grenze — die Weisen nannten sie den Bisektor — zwischen den Fürstentümern war eine Gerade (Abbildung ??(i)), auf der jeder Punkt von beiden Schlössern gleich entfernt war:

$$B(p, q) = \{x \in \mathbb{R}^2 \mid |xp| = |xq|\}.$$

Dabei bezeichnet  $|xq|$  den euklidischen Abstand,  $x = (x_1, x_2)$ ,  $q = (q_1, q_2)$ :

$$|xq| = \sqrt{(x_1 - q_1)^2 + (x_2 - q_2)^2}.$$

So lebten denn  $p$  und  $q$  glücklich, bis ein dritter Fürst  $r$  hinzukam, der ebenso mächtig war wie  $p$  und  $q$  und das Land jetzt unter den dreien aufgeteilt werden mußte. Die Weisen beschlossen, ihr Kriterium beizubehalten und die Grenzen — die Bisektoren — nach diesem Kriterium neu zu berechnen. Zu der Geraden zwischen  $p$  und  $q$  fügten sie die Geraden  $p/r$  und  $q/r$  und schnitten sie dort ab, wo die Distanz zum dritten Schloß geringer wurde als die Distanz zu den beiden Schlössern, zwischen denen der jeweilige Bisektor verläuft (Bild (ii)).

Auch als schließlich ein vierter Fürst  $s$  sein Schloß genau in der Mitte des Dreiecks der drei anderen Schösser baute, wurde das Distanzkriterium beibehalten, obwohl das Gebiet von  $s$  dadurch recht klein ausfiel (Bild (iii)).<sup>5</sup>

<sup>5</sup>Unter <http://web.informatik.uni-bonn.de/I/GeomLab/VoroGlide/> findet sich ein Java-Applet zur Visualisierung von Voronoi–Diagrammen, Delaunay–Triangulationen und konvexen Hüllen von Punkten im  $\mathbb{R}^2$ . [?]



**Definition A.5** Sei  $S \subset \mathbb{R}^2$  eine Menge von Punkten (Sites),  $p \in S$ . Die **Voronoi-Region** des Punktes  $p$  bzgl. der Punktmenge  $S$  ist die Menge aller Punkte aus  $\mathbb{R}^2$ , denen  $p$  näher ist als jeder andere Punkt aus  $S$ :

$$\text{VR}(p, S) := \left\{ x \in \mathbb{R}^2 \mid |px| < \min_{q \in S \setminus \{p\}} \{|qx|\} \right\}.$$

Das **Voronoi-Diagramm** der Menge  $S$  besteht aus den Punkten, die auf einer Grenze zwischen zwei Voronoi-Regionen liegen:

$$V(S) := \bigcup_{p \in S} \partial \text{VR}(p, S).$$

Die Stücke der Bisektoren, die die Voronoi-Regionen beranden, heißen **Voronoi-Kanten**, die Punkte, an denen mehrere Voronoi-Kanten zusammentreffen, heißen **Voronoi-Knoten**.

Eine andere Ansicht von Voronoi-Diagrammen ist, von jedem Punkt aus  $S$  einen Kreis wachsen zu lassen. Alle Kreise starten zum gleichen Zeitpunkt und wachsen gleich schnell. Ein Punkt  $x$  wird der Region zugeordnet, dessen Kreis ihn als erster erreicht. Erreichen  $x$  mehrere Kreise zur gleichen Zeit, so liegt  $x$  auf einem Bisektor. Auch umgekehrt läßt sich  $x$  dem Punkt  $p \in S$  zuordnen, den ein um  $x$  wachsender Kreis als erster erreichen würde (Abbildung ??(ii)). Liegt  $x$  auf einer Voronoi-Kante, so trifft der Kreis um  $x$  zwei Punkte aus  $S$  gleichzeitig (Abbildung ??(iii)), liegt er auf einem Voronoi-Knoten, so trifft er drei Punkte aus  $S$  (Abbildung ??(iv)).

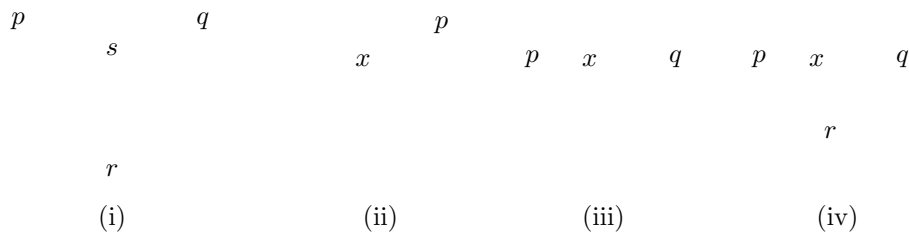


Abbildung A.2: (i) Delaunay-Triangulation und konvexe Hülle,  $x$  ist (ii) in Region von  $p$ , (iii) auf Voronoi-Kante zwischen  $p$  und  $q$ , (iv) ein Voronoi-Knoten.

**Lemma A.6** Das Voronoi-Diagramm  $V(S)$  einer Menge  $S$  von  $n$  Punkten hat die folgenden Eigenschaften:

- (i)  $V(S)$  hat  $O(n)$  Kanten, Knoten und Flächen (folgt aus der Euler-Formel).
- (ii)  $V(S)$  läßt sich in Zeit  $O(n \log n)$  mittels Divide-and-Conquer, Sweep oder inkrementell berechnen.

- (iii) Das Problem der nächsten Nachbarn (gegeben:  $n$  Punkte  $p_1, \dots, p_n$ , Query-Punkt  $x$ , gesucht:  $p_i$  mit  $|xp_i| \leq \min_{j \neq i} |xp_j|$ ) reduziert sich auf das Lokalisierungsproblem im Voronoi-Diagramm: In welcher Voronoi-Region  $\text{VR}(p_i, \{p_1, \dots, p_n\})$  liegt  $x$ ? Dies läßt sich bei gegebenem  $V(S)$  nach Vorbereitungszeit  $O(n)$  in Zeit  $O(\log n)$  pro Query beantworten.
- (iv) Die Punkte mit unbeschränkten Voronoi-Regionen bilden die konvexe Hülle von  $S$  (in Abbildung ??(i) fett gezeichnet).
- (v) Verbindet man in  $S$  alle Punkte, deren Voronoi-Regionen gemeinsame Kanten haben, so entsteht ein zum Voronoi-Diagramm dualer Graph. Sind die Punkte aus  $S$  in allgemeiner Lage, ist dieser Graph eine Triangulation der Punktmenge  $S$ , die **Delaunay-Triangulation**, bei der der kleinste auftretende Winkel maximal unter allen möglichen Triangulationen von  $S$  ist. Die Delaunay-Triangulation ist in Abbildung ??(i) gestrichelt, die konvexe Hülle ist Teil der Delaunay-Triangulation.
- (vi) Zu jedem Dreieck der Delaunay-Triangulation enthält der Umkreis keinen Punkt aus  $S$ . Dies ist eine äquivalente Charakterisierung der Delaunay-Triangulation. Algorithmus ?? läßt sich aus dieser Eigenschaft ableiten.

---

**Algorithmus A.1** Inkrementelle Berechnung der Delaunay-Triangulation
 

---

- Beginne mit 3 beliebigen Punkten.
- Füge sukzessive die übrigen  $n - 3$  Punkte ein:
  - Falls  $p_{i+1}$  außerhalb der konvexen Hülle der bisher eingefügten Punkte liegt: verbinde  $p_{i+1}$  mit allen Punkten der konvexen Hülle, die von  $p_{i+1}$  aus sichtbar sind.
  - Andernfalls lokalisiere das Dreieck  $D$ , in dem  $p_{i+1}$  liegt, und verbinde  $p_{i+1}$  mit allen Ecken von  $D$ . (Dabei entsteht eine Triangulation, jedoch nicht notwendigerweise eine Delaunay-Triangulation, da der Umkreis vorhandener Delaunay-Dreiecke jetzt  $p_{i+1}$  enthalten kann).
  - Führe solange Edge-Flips durch, bis alle Dreiecke die Umkreisbedingung erfüllen, vgl. Abbildung ??(ii).

---

Definition ?? läßt sich verallgemeinern, indem man statt Punkten eine Menge beliebiger Objekte  $\mathcal{M} = \{M_1, \dots, M_n\}$  in  $\Omega$  betrachtet und für jedes Objekt  $M_i$  eine Abstandsfunktion  $d_i : \Omega \rightarrow \mathbb{R}^{\geq 0}$  angibt; die Distanzfunktion kann jetzt abhängig vom Objekt sein! Die Voronoi-Region eines Objekts  $M_i$  ist dann:

$$\text{VR}(M_i, \mathcal{M}) := \{w \in \Omega \mid d_i(w) < \min_{j \neq i} d_j(w)\}.$$

Für die Bewegungsplanung spielen die folgenden Verallgemeinerungen eine Rolle.

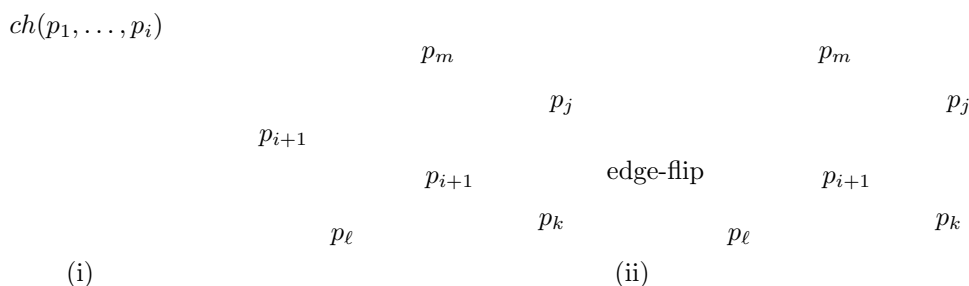


Abbildung A.3: (i) Der neue Punkt  $p_{i+1}$  liegt außerhalb der konvexen Hülle, (ii)  $p_{i+1}$  liegt innerhalb des Umkreises des Dreiecks  $(p_j, p_k, p_m)$ , die Kante  $\overline{p_k p_m}$  wird umgeklappt.

### A.3.2 Voronoi–Diagramme bzgl. konvexer Distanzfunktionen

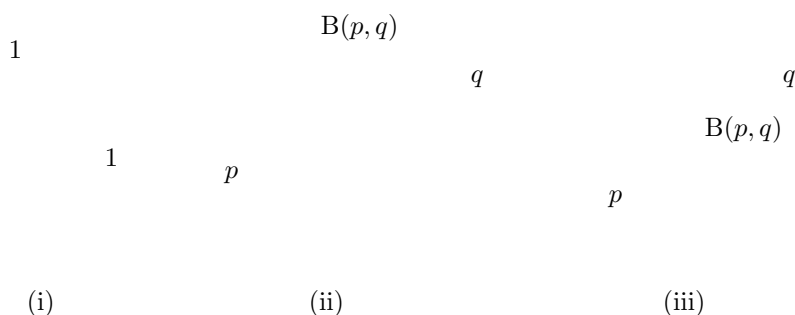


Abbildung A.4: (i) Einheitskreis der  $L_1$ –Metrik, (ii) Bisektor zwischen  $p$  und  $q$ , (iii) der Bisektor zwischen  $p$  und  $q$  kann zwei Viertelebenen enthalten.

Bleiben wir zunächst bei Punkten im  $\mathbb{R}^2$  als Objekte, verwenden aber eine Metrik, die von der euklidischen Metrik verschieden ist, z. B. die  $L_1$ –Metrik — auch als Manhattan–Metrik bekannt — die anschaulich die Distanz zwischen zwei Orten in einer Stadt mit orthogonalem Straßengitter angibt. Die Distanz zwischen zwei Punkten  $p = (p_1, p_2)$  und  $q = (q_1, q_2)$  beträgt

$$L_1(p, q) = |p_1 - q_1| + |p_2 - q_2|.$$

Voronoi–Diagramme bzgl. der  $L_1$ –Metrik enthält man durch Anwendung von Definition ?? mit der  $L_1$ –Metrik statt der euklidischen Metrik oder mit dem Modell der Kreise, die sich um die Sites ausbreiten, wobei die Kreise jetzt die Gestalt der Einheitskreise der  $L_1$ –Metrik haben müssen.

Die Punkte  $p = (x, y)$  auf dem Einheitskreis bzgl. der  $L_1$ –Metrik müssen der Gleichung  $|x| + |y| = 1$  genügen, der Einheitskreis ist daher ein um  $45^\circ$  gedrehtes Quadrat der Seitenlänge  $\sqrt{2}$  um den Ursprung, siehe Abbildung ??(i).

Daraus ergeben sich für den Bisektor zwischen zwei Punkten  $p$  und  $q$  drei Fälle: liegen beide Punkte auf einer gemeinsamen  $X$ - oder  $Y$ -Koordinate, so ist der Bisektor eine achsenparallele Gerade zwischen  $p$  und  $q$ . Spannen  $p$  und  $q$  ein echtes Rechteck auf, besteht der Bisektor aus zwei Halbgeraden außerhalb des Rechtecks und einem Liniensegment mit Winkel  $\pm\frac{\pi}{4}$ , das beide Halbgeraden innerhalb des Rechtecks verbindet, siehe Abbildung ??(ii). Eine Besonderheit tritt auf, wenn  $p$  und  $q$  ein Quadrat aufspannen, dann enthält der Bisektor zwei Viertelkreise (Abbildung ??(iii)), alle diese Punkte haben ja bzgl. der  $L_1$ -Metrik zu  $p$  und  $q$  den gleichen Abstand!

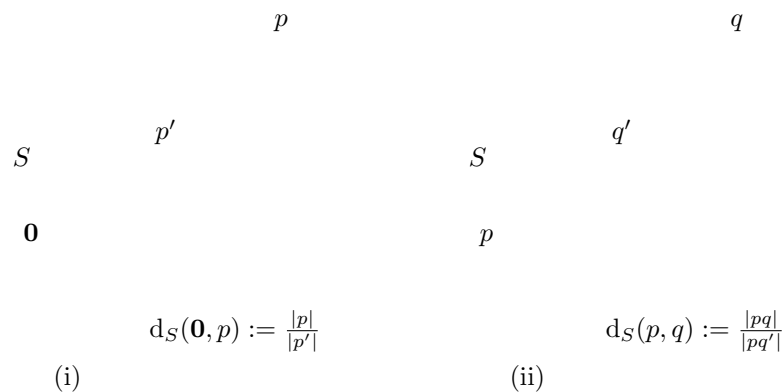


Abbildung A.5: (i) Abstand von  $p$  zum Ursprung bzgl. konvexer Distanzfunktion, (ii) Abstand von  $p$  und  $q$ .

Allgemeiner sind konvexe Distanzfunktionen, die durch eine kompakte, konvexe Menge  $S \subset \mathbb{R}^2$ , die den Ursprung  $\mathbf{0}$  enthält, definiert werden. Die Einheitskreise der  $L_1$ - und  $L_2$ -Metrik sind Spezialfälle einer solchen Menge,  $S$  ist der Einheitskreis der definierten Metrik. Wegen der Konvexität gibt es zu jedem Punkt  $p \in \mathbb{R}^2$  genau einen Schnittpunkt von  $\overline{\mathbf{0}p}$  mit  $S$ , also auch genau einen Punkt  $p'$  auf dem Rand von  $S$ , siehe Abbildung ??(i). Damit läßt sich der Abstand von  $p$  zum Ursprung definieren als  $d_S(\mathbf{0}, p) := \frac{|p|}{|p'|}$ . Die Distanz zwischen zwei Punkten  $p$  und  $q$  ergibt sich dann durch Verschiebung der Menge  $S$  vom Ursprung zum Punkt  $p$  und Anwendung obiger Definition, vgl. Abbildung ??(ii):

$$d_S(p, q) = \frac{|pq|}{|pq'|}.$$

Insbesondere konvexe Polygone als Einheitskreis  $S$  sind für die Bewegungsplanung interessant.

### A.3.3 Voronoi-Diagramme von Liniensegmenten

Im zweiten Fall wollen wir bei der euklidischen Metrik als Abstandsfunktion bleiben, jedoch von Punkten verschiedene Objekte, nämlich  $n$  Liniensegmente

im  $\mathbb{R}^2$  betrachten. Auch hier soll entsprechend Definition ?? ein Punkt  $x$  der Ebene in der Voronoi-Region des Liniensegmentes liegen, das zu  $x$  einen geringeren Abstand hat als alle anderen Liniensegmente. Dazu müssen wir zunächst festlegen, was wir unter dem Abstand eines Punktes  $p$  zu einem Liniensegment  $\ell$  verstehen. Wir definieren dies als den Abstand zwischen  $p$  und dem zu  $p$  nächstgelegenen Punkt auf  $p_\ell \in \ell$ :

$$d(p, \ell) := \min_{q \in \ell} |pq|.$$

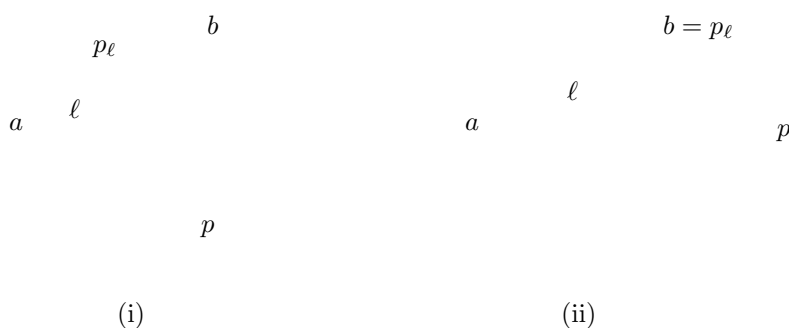


Abbildung A.6: Der zu  $p$  nächstgelegene Punkt auf einem Liniensegment.

Hier können zwei Fälle auftreten: liegt  $p$  in dem Streifen der beiden senkrecht zu  $\ell$  durch die Endpunkte von  $\ell$  verlaufenden Geraden, so ist  $p_\ell$  der Fußpunkt des Lots von  $p$  auf  $\ell$ . Liegt  $p$  außerhalb dieses Streifens, so ist  $p_\ell$  einer der Endpunkte von  $\ell$ , vgl. Abbildung ??.

Wie sieht nun der Bisektor zwischen einem Punkt und einem Liniensegment aus? Betrachten wir zunächst den Bisektor zwischen einem Punkt  $p = (0, a)$ ,  $a \neq 0$  und der X-Achse. Für einen Punkt  $(x, y)$  auf dem Bisektor müssen die Abstände zu  $p$  und zur X-Achse gleich sein:

$$\begin{aligned} (x, y) &\in B(p, X) \\ \Leftrightarrow y &= \sqrt{x^2 + (y - a)^2} \\ \Leftrightarrow y^2 &= x^2 + y^2 - 2ay + a^2 \\ \Leftrightarrow y &= \frac{x^2}{2a} + \frac{a}{2}. \end{aligned}$$

Dies ist die Gleichung einer Parabel. Für Liniensegmente bedeutet dies, daß der Bisektor innerhalb des Streifens aus einem Parabelstück besteht. Außerhalb des Streifens ist der Bisektor gerade der Bisektor zwischen  $p$  und dem jeweiligen Endpunkt von  $\ell$ , der Bisektor zwischen zwei Punkten ist eine Gerade. Also besteht  $B(p, \ell)$  außerhalb des Streifens aus zwei Halbgeraden,

vgl. Abbildung ?? . Der Bisektor zweier Liniensegmente besteht aus max. sieben Teilstücken, die äußeren sind Halbgeraden, die übrigen Parabelstücke oder Liniensegmente. Abbildung ?? zeigt ein Beispiel eines solchen Bisektors, an den einzelnen Teilstücken ist dabei vermerkt, zwischen welchen Objekten sie verlaufen. Abbildung ?? zeigt ein Voronoi-Diagramm von Liniensegmenten.

Abbildung A.7: Der Bisektor eines Punktes und eines Liniensegmentes.

Abbildung A.8: Der Bisektor zwischen zwei Liniensegmenten.

Abbildung A.9: Voronoi-Diagramm von sieben Liniensegmenten.

### A.3.4 Voronoi-Diagramme mit additiven Gewichten

Betrachten wir nun wieder eine Menge  $S$  von Punkten im  $\mathbb{R}^2$ , hier jedoch mit einer vom Objekt abhängigen Abstandsfunktion, bei der zur euklidischen Distanz zwischen  $p_i \in S$  und  $x \in \mathbb{R}^2$  eine Konstante  $a_i$  addiert wird:

$$d_i(x) := |p_i x| + a_i.$$

Ein Bisektor zwischen zwei Punkten  $p_i, p_j \in S$  ist, wie üblich, die Menge aller Punkte  $x \in \mathbb{R}^2$ , die zu  $p_i$  und  $p_j$  den gleichen Abstand — gemessen mit unserer neuen Abstandsfunktion — haben:

$$\begin{aligned} x &\in B(p_i, p_j) && B(p_i, p_j) \\ \Leftrightarrow d_i(x) &= d_j(x) \\ \Leftrightarrow |p_i x| + a_i &= |p_j x| + a_j && p_i \quad p_j \\ \Leftrightarrow |p_i x| - |p_j x| &= a_j - a_i. && a_j > a_i \end{aligned}$$

Dies ist die Gleichung einer Hyperbel. Nebenstehende Abbildung zeigt einen Bisektor für  $a_j > a_i$ .

Im Modell der um die  $p_i$  wachsenden Kreise bedeutet ein additives Gewicht, daß nicht alle Kreise zum gleichen Zeitpunkt entstehen. Das Gewicht  $a_i$  gibt die Zeit an, die verstreichen muß, bis ein Kreis um  $p_i$  mit seinem Wachstum beginnen darf.

Weiterführendes zum Thema Voronoi-Diagramme findet sich z. B. im Buch von Klein [?] oder dem Übersichtsartikel von Aurenhammer und Klein [?].

## A.4 Davenport–Schinzel–Sequenzen

### A.4.1 Definition und Eigenschaften

**Definition A.7** Gegeben sei ein Alphabet  $\Sigma = \{A, B, C, \dots\}$  über  $n$  Buchstaben. Ein Wort  $w$  über  $\Sigma$  heißt **Davenport–Schinzel–Sequenz** der Ordnung  $s$ , wenn

- (i) in  $w$  keine benachbarten Buchstaben gleich sind und
- (ii) keine zwei verschiedenen Buchstaben einander mehr als  $s$  mal abwechseln. [?]

Das Wort  $ABBA$  ist also wegen  $BB$  keine Davenport–Schinzel–Sequenz, und das Wort

ist eine Davenport–Schinzel–Sequenz der Ordnung 4.

Interessanter als die Davenport–Schinzel–Sequenzen selbst ist die Frage, wie lang eine Davenport–Schinzel–Sequenz zu gegebenem  $n$  und  $s$  maximal werden kann.

### Definition A.8

$\lambda_s(n) :=$  die maximale Länge einer Davenport–Schinzel–Sequenz der Ordnung  $s$  über einem Alphabet  $\Sigma$  mit  $n$  Buchstaben.

Die Berechnung der Funktion  $\lambda_s(n)$  ist ein kompliziertes kombinatorisches Problem. Bekannt darüber ist:

**Theorem A.9** (*Hart, Sharir, 1986*)

$$\begin{aligned} \lambda_1(n) &= n \\ \lambda_2(n) &= 2n - 1 \\ \lambda_3(n) &\in \Theta(n \alpha(n)) \\ \lambda_4(n) &\in \Theta(n \cdot 2^{\alpha(n)}) \\ \lambda_s(n) &\in O(n \log^*(n)) \quad [?] \end{aligned}$$

Dabei ist  $\alpha(n)$  das Inverse der Ackermann–Funktion [?]:

$$\begin{aligned} \alpha(m) &:= \min \{ n \mid A(n, n) \geq m \} \\ A(0, n) &:= n + 1 \\ A(k, 0) &:= A(k - 1, 1) \\ A(k, n) &:= A(k - 1, A(k, n - 1)) \end{aligned}$$

Die Ackermann–Funktion wächst extrem schnell, z. B. ist bereits

$$A(4, m) = 2^{2^{\cdot^{\cdot^2}}} \}^{m\text{-mal}}$$



Dementsprechend langsam wächst  $\alpha(n)$  und kann als nahezu konstant angesehen werden. Auch

$$\log^*(n) := \text{kleinstes } m \text{ mit } \underbrace{\log(\log(\dots(\log n)))}_{m \text{ mal}} \leq 1$$

wächst extrem langsam und ist “fast konstant”. Die  $\lambda_s(n)$  sind also fast linear.

### A.4.2 Anwendung: die untere Kontur von Funktionen



Abbildung A.10: Die untere Kontur von Funktionen (i) über einem gemeinsamen Intervall  $I$  und (ii) über Intervallen  $I_j$ .

**Definition A.10** Seien  $f_1, f_2, \dots, f_n$  reellwertige Funktionen, entweder

- (i) über einem gemeinsamen Intervall  $I$  oder
- (ii) über je einem Intervall  $I_j \subseteq I, 1 \leq j \leq n$ .

Je zwei Funktionen mögen höchstens  $s$  gemeinsame Schnittpunkte haben. Dann ist  $L(n) := \min \{ f_i(x) \mid 1 \leq i \leq n \}$  die **untere Kontur** (lower envelope) der  $n$  Funktionsgraphen.

**Theorem A.11**

In den oben definierten Fällen besteht  $L(x)$  aus maximal

- (i)  $\lambda_s(n)$
- (ii)  $\lambda_{s+2}(n)$

Segmenten.

[?]

Der Beweis dazu findet sich z. B. in [?].

Abbildung ?? zeigt ein Beispiel für untere Konturen, Abbildung ?? zeigt als Spezialfall die untere Kontur eines Arrangements von Liniensegmenten.

**Bemerkung A.12** Die untere Kontur von  $n$  Liniensegmenten hat die Komplexität  $O(\lambda_3(n)) \in O(n \alpha(n))$ . Diese Schranke ist scharf.

Abbildung A.11: Die untere Kontur von Liniensegmenten.

Mehr über Davenport–Schinzel–Sequenzen und ihre Anwendungen findet sich im Buch [?] oder im Übersichtsartikel von Agarwal und Sharir [?].

# Anhang B

## Roboter

Abbildung ??

Abbildung ??(ii)

$-R$

Abbildung ??

Abbildung ??(i)

Abbildung ??

Abbildung ??

Seite ?? unten

Abbildung ??

Abbildung ??

Abbildung ??

Abbildung ??

Abbildung ??

Abbildung B.1: Roboter zum Ausschneiden



INHALTSVERZEICHNIS

# Inhaltsverzeichnis



ABBILDUNGSVERZEICHNIS

# Abbildungsverzeichnis





VERZEICHNIS DER ALGORITHMEN

# Verzeichnis der Algorithmen

