Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology

DIPLOMARBEIT

Approximation mit abwickelbaren Flächen

Ausgeführt am Institut für

Diskrete Mathematik und Geometrie
der Technischen Universität Wien

unter der Anleitung von
Ao.Univ.Prof. Mag.rer.nat. Dr.techn. Martin PETERNELL

durch

Lukas GOTTHART
Neugebäudestraße 24/2/7, 1110 Wien

12. September 2011

Kurzfassung

Die vorliegende Arbeit behandelt die Approximation allgemeiner Flächen durch abwickelbare Flächen. Es sollen im Folgenden unterschiedliche Herangehensweisen an dieses Problem erläutert werden. Zum einen sind dies Ansätze, die ein gegebenes Netz unterteilen, um somit annähernd abwickelbare Bereiche zu erhalten und andererseits Zugänge, die eine gegebene Punktmenge durch abwickelbare Flächen approximieren. Da die gegebenen Flächen bzw. Punktmengen im Allgemeinen nicht abwickelbar sein werden bzw. nicht von einer abwickelbaren Fläche stammen, ist offensichtlich, dass ein Kompromiss eingegangen werden muss. Entweder beschreibt das Ergebnis exakt das gegebene Netz, aber die Segmentierung ist nur annähernd abwickelbar oder die erhaltene Fläche ist abwickelbar, aber beschreibt nicht exakt die gegebene Fläche bzw. Punktmenge, sondern nähert sie nur an. Eine weitere Möglichkeit wäre mit Dreiecksstreifen zu arbeiten, wodurch mitunter eine exakte Beschreibung der Fläche und die Abwickelbarkeit des Ergebnisses gewährleistet ist, allerdings liefern diese Lösungen zumeist eine große Anzahl an Streifen und lange Grenzen, was häufig unerwünscht ist. In den Beispielen, die die Netzsegmentierung behandeln, sind die entstehenden Bereiche annähernd abwickelbar und dies liefert für Anwendungen, die mit flexiblen Materialien arbeiten, etwa Stoff, eine hinreichend gute Genauigkeit, deren Abweichung das Material ausgleicht. Die Anwendungen für abwickelbare Flächen und somit auch für diese Algorithmen, sind weitläufig. Sie reichen vom Flugzeug- und Schiffbau über Texturen in der Computergrafik, bis zu Modedesign, um nur einige zu nennen.

Die weitere Arbeit ist wie folgt gegliedert: Zu Beginn werden die notwendigen Grundlagen aus der Differentialgeometrie und der Liniengeometrie erläutert. Anschließend wird beleuchtet, wie man, bei gegebener Punktwolke bzw. Triangulierung, die Normalvektoren in den Datenpunkten schätzen kann, was für viele Algorithmen aus dem untersuchten Gebiet eine essentielle Rolle spielt. In den Kapiteln 3 und 4 werden dann zwei Methoden zur Netzsegmentierung vorgestellt, bevor abschließend drei Algorithmen zur Approximation mit abwickelbaren Flächen, betrachtet werden.

Inhaltsverzeichnis

1 Grundlagen **1**

1.1 Differentialgeometrische Grundlagen 1

1.1.1 Flächen im $\mathbb{R}^n$ 1

1.1.2 Krümmungstheorie der Hyperflächen 4

1.1.3 Regelflächen 10

1.2 Grundlagen der Liniengeometrie 13

1.2.1 Plücker-Koordinaten 13

1.2.2 Lineare Komplexe 15

1.2.3 Der Zusammenhang zwischen linearen Komplexen und Schraubungen 16

2 Das Schätzen von Normalvektoren **19**

2.1 Normalvektorschätzung mit quadratischen Kurven 19

2.2 Gewichte für die Berechnung von Normalvektoren 23

3 D-charts: Quasi-Developable Mesh Segmentation **25**

3.1 Der Lloyd-Algorithmus 25

3.2 Der D-charts-Algorithmus 26

4 Mesh Segmentation Driven by Gaussian Curvature **31**

5 Flächenapproximation mittels abwickelbarer Flächen **39**

6 Approximation von Punktwolken durch abwickelbare Flächen **47**

6.1 Grundidee 47

6.2 Blaschke-Bilder spezieller Flächen 49

6.3 Der Algorithmus 51

7 Curved Folding **59**

1 Grundlagen

1.1 Differentialgeometrische Grundlagen

Da die vorliegende Arbeit von abwickelbaren Flächen handelt, soll in diesem Kapitel erläutert werden, was man unter diesem Begriff versteht. Dazu ist es zunächst nötig einige grundlegende Definitionen und Begriffe der Differentialgeometrie einzuführen. Nach einer ersten Definition des Begriffs *abwickelbar*, sollen noch weitere Charakterisierungen folgen, die diesen Begriff von anderen Seiten beleuchten. Die hier ausgeführten Definitionen und Sätze sind in jedem Lehrbuch der Differentialgeometrie nachzulesen, etwa [3].

1.1.1 Flächen im $\mathbb{R}^n$

In diesem Kapitel wird die Einsteinsche Summenkonvention verwendet, also für einen Vektor $v \in \mathbb{V}$ mit Koordinaten $v^j \in \mathbb{R}$ bezüglich der Basis $A = \{a_j | j = 1 \dots m\}$ gilt

$$v = \sum_{j=1}^{m} v^j a_j = v^j a_j.$$

Wir bezeichnen im Folgenden die Koordinatenfunktionen des m-dimensionalen Zahlenraumes $\mathbb{R}^m$ $(m \geq 2)$ mit u^j $(j = 1, \dots, m)$ und fassen die Menge $\mathbb{R}^n$ als euklidischen Raum mit den Koordinatenfunktionen x^α $(\alpha = 1, \dots, n)$ auf.

Sei $U \subset \mathbb{R}^m$ eine offene, nichtleere Teilmenge von $\mathbb{R}^m$ und $f : U \to \mathbb{R}^n$, also

$$\mathbf{a} = (a^1, \dots, a^m) \in U \mapsto f(\mathbf{a}) = (f^1(\mathbf{a}), \dots, f^n(\mathbf{a})) \in \mathbb{R}^n$$

dann schreiben wir für die partiellen Ableitungen

$$\frac{\partial f^\alpha}{\partial u^j}(\mathbf{a}) =: f^\alpha{}_{,j}(\mathbf{a}) \quad (1 \leq \alpha \leq n, 1 \leq j \leq m)$$

und weiters für den Ableitungsvektor von f nach u^j in $\mathbf{a}$

$$f_{,j}(\mathbf{a}) := \frac{\partial f}{\partial u^j}(\mathbf{a}) = \left(f^1{}_{,j}\,(\mathbf{a}), \ldots, f^n{}_{,j}\,(\mathbf{a})\right) \in \mathbb{R}^n.$$

Man nennt die $(m \times n)$-Matrix

$$\begin{pmatrix} f_{,1}\,(\mathbf{a}) \\ \vdots \\ f_{,m}\,(\mathbf{a}) \end{pmatrix} = \begin{pmatrix} f^1{}_{,1}\,(\mathbf{a}) & \ldots & f^n{}_{,1}\,(\mathbf{a}) \\ \vdots & & \vdots \\ f^1{}_{,m}\,(\mathbf{a}) & \ldots & f^n{}_{,m}\,(\mathbf{a}) \end{pmatrix} = J_f(\mathbf{a})$$

die *C. Jacobi*[1]*-Matrix* von f in $\mathbf{a}$.

Definition 1.1. Eine r-mal stetig differenzierbare Abbildung $f : U \to \mathbb{R}^n$ $(r \geq 1)$ einer offenen Menge $U \subset \mathbb{R}^m$ in $\mathbb{R}^n$ $(2 \leq m \leq n-1)$ heißt ein m-dimensionaler $C^r(U)$-Weg ($C^r(U)$–m-Weg) in $\mathbb{R}^n$. Die Punktmenge $f(U) \subset \mathbb{R}^n$ heißt C^r-Fläche (C^r–m-Fläche) und die als nicht leer vorausgesetzte offene Definitionsmenge U heißt Parametermenge des m-Weges f. Für $n = m+1$ heißt f ein Hyperflächenweg und $f(U)$ eine Hyperfläche.

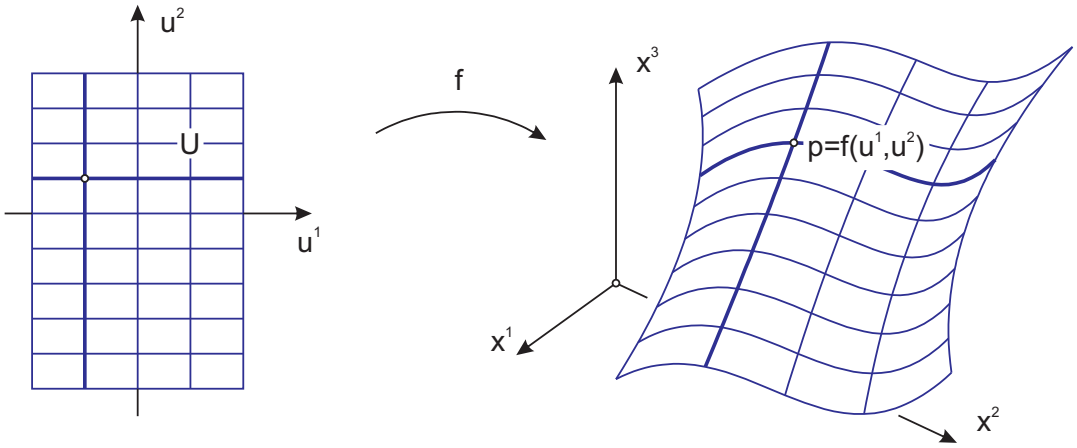

Abb. 1.1: Parametermenge und Punktmenge eines Weges

Definition 1.2. Ist $RgJ_f(\mathbf{a}) = m$ so heißt $\mathbf{a} \in U$ bzw. $f(\mathbf{a}) \in f(U)$ ein regulärer Punkt von f bzw. $f(U)$. Falls $RgJ_f(\mathbf{a}) = m$ für alle $\mathbf{a} \in U$ gilt, so heißt f eine $C^r(U)$-Immersion und $f(U)$ eine reguläre $C^r(U)$-m-Fläche. Ist f injektiv, so heißt $f(U)$ eine einfache Fläche.

[1]Carl Gustav Jacob Jacobi, geboren am 10.12.1804 in Potsdam, gestorben am 18.02.1851 in Berlin

Definition 1.3. Sei $f : U \to \mathbb{R}^n$ und $c : I \subset \mathbb{R} \to U$, $c \in C^s(I)$ $(s \geq 1)$. Dann heißt

$$c_f := f \circ c : I \to f(U) \subset \mathbb{R}^n$$

ein $C^s(I)$-Flächenweg und die Punktmenge $c_f(I) \subset f(U) \subset \mathbb{R}^n$ eine $C^s(I)$-Flächenkurve.

Definition 1.4. Sei $\mathbb{A}(\mathbb{V})$ der m-dimensionale reelle affine Raum mit der Punktmenge $a+\mathbb{V} := \{a+v | v \in \mathbb{V}\}$. Dann heißt $V_a := (a; v)$ $(v \in \mathbb{V})$ ein Tangentialvektor von $\mathbb{A}(\mathbb{V})$ in a und $\{V_a = (a; v) | v \in \mathbb{V}\} =: T_a\mathbb{A}(\mathbb{V})$ der Tangentialvektorraum von $\mathbb{A}(\mathbb{V})$ in $a \in \mathbb{A}(\mathbb{V})$.

Somit verstehen wir also in diesem Sinne einen Tangentialvektor als Aufhängepunkt a und zugehörigem Vektorbestandteil v.

Definition 1.5. Sei $f : U \to \mathbb{R}^n$, $a \in U$. Dann heißt

$$T_a f := \{(f(a); v) | v \in \mathbb{H}(f_{,1}(a), \ldots, f_{,m}(a))\}$$

der Tangentialvektorraum der Fläche $f(U)$ im Punkt $f(a)$. $V_{f(a)} \in T_a f$ heißt ein Tangentialvektor der Fläche $f(U)$ im Punkt $f(a)$.

Definition 1.6. Seien $f(U), \overline{f}(\overline{U})$ reguläre, einfache Flächen im $\mathbb{R}^n$ und $\alpha : f(U) \to \overline{f}(\overline{U})$ differenzierbar. Für $X_p \in T_a f$ $(p = f(a))$ sei $c_f = f \circ c : I \to f(U) \subset \mathbb{R}^n$ ein regulärer $C^1(I)$-Flächenweg mit

$$X_p = (c_f(0); \dot{c}_f(0)).$$

Dann heißt mit $A := \overline{f}^{-1} \circ \alpha \circ f : U \to \overline{U}$ die Abbildung

$$\alpha_{*p} : T_a f \to T_{A(a)}\overline{f} \text{ mit } X_p \mapsto \alpha_{*p}(X_p) := \left(\alpha(p); \frac{d}{dt}(\alpha \circ f \circ c)(0)\right)$$

die derivierte Abbildung von α in p.

Definition 1.7. Es seien $f(U), \overline{f}(\overline{U})$ reguläre, einfache Flächen im $\mathbb{R}^n$. Dann heißt die differenzierbare Abbildung $\alpha : f(U) \to \overline{f}(\overline{U})$ eine *lokale Isometrie*, wenn die derivierte Abbildung α_{*p} für alle $p \in f(U)$ ein orthogonaler Vektorraumisomorphismus ist.

Da nun alle nötigen Begriffe definiert wurden, können wir uns nun unserer eigentlichen Intention, nämlich der Erläuterung von *abwickelbar*, zuwenden.

Definition 1.8. Eine reguläre C^1-m-Fläche $f(U) \subset \mathbb{R}^n$ heißt *abwickelbar*, wenn $f(U)$ lokal isometrisch in den m-dimensionalen euklidischen Raum abgebildet werden kann.

Beispiele für abwickelbare Flächen im dreidimensionalen Raum sind (und dies sind im wesentlichen alle Flächen dieser Art):

- Zylinderflächen mit regulärem C^1-Normalschnitt
- in der Spitze punktierte Halbkegelflächen mit regulärer C^1-Richtkurve
- längs $c(I)$ aufgeschnittene Tangentenflächen regulärer, wendepunktfreier C^2-Kurven $c(I) \subset \mathbb{R}^n$

1.1.2 Krümmungstheorie der Hyperflächen

Nun wollen wir uns noch einer anderen Charakterisierung von abwickelbaren Flächen widmen, nämlich jener mit Hilfe der sogenannten Gaußschen[2] Krümmung. Auf dem Weg zur Definition selbiger, begegnen wir zunächst der Gauß-Abbildung, auch sphärische Abbildung genannt, die uns im späteren Verlauf dieser Arbeit wieder unterkommen wird. In diesem Abschnitt gehen wir von Hyperflächen $f(U) \subset \mathbb{R}^{m+1}$ aus, also $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ mit $n = m + 1$ ist ein Hyperflächenweg. Auf Ausnahmen dieser Regel wird explizit hingewiesen. Wir bezeichnen weiters die orientierte Fläche mit $\overrightarrow{f(U)}$ und den normierten Normalvektor im Punkt $f(a)$ mit $n_f(a)$. Damit erhalten wir die folgende

Definition 1.9. Sei $f \in C^2(U)$ regulär und $S^m \subset \mathbb{R}^{m+1}$ die m-dimensionale Einheitshypersphäre, so heißt die Abbildung $\sigma : \overrightarrow{f(U)} \to S^m$ mit

$$f(a) = p \in f(U) \mapsto \sigma(p) := n_f(a) \in S^m$$

die Gauß-Abbildung (sphärische Abbildung) der Hyperfläche f(U).

Ein wichtiger Schritt auf dem Weg zur Gauß Krümmung ist die Weingarten[3]-Abbildung, die uns direkt zu den Hauptkrümmungen und der Gauß Krümmung führt.

Definition 1.10. Ist f regulär und wird für $a \in U$ der Tangentialvektorraum $T_a n_f$ mit $T_a f$ identifiziert, so heißt mit $p = f(a)$ der Vektorraumendomorphismus

$$\omega_p : T_a f \to T_a f \text{ mit } \omega_p := -\sigma_{*p}$$

die Weingarten-Abbildung der orientierten Hyperfläche $\overrightarrow{f(U)}$.

[2] Johann Carl Friedrich Gauß, geboren am 30.04.1777 in Braunschweig, gestorben am 23.02.1855 in Göttingen

[3] Julius Weingarten, geboren am 02.03.1836 in Berlin, gestorben am 16.06.1910 in Freiburg im Breisgau

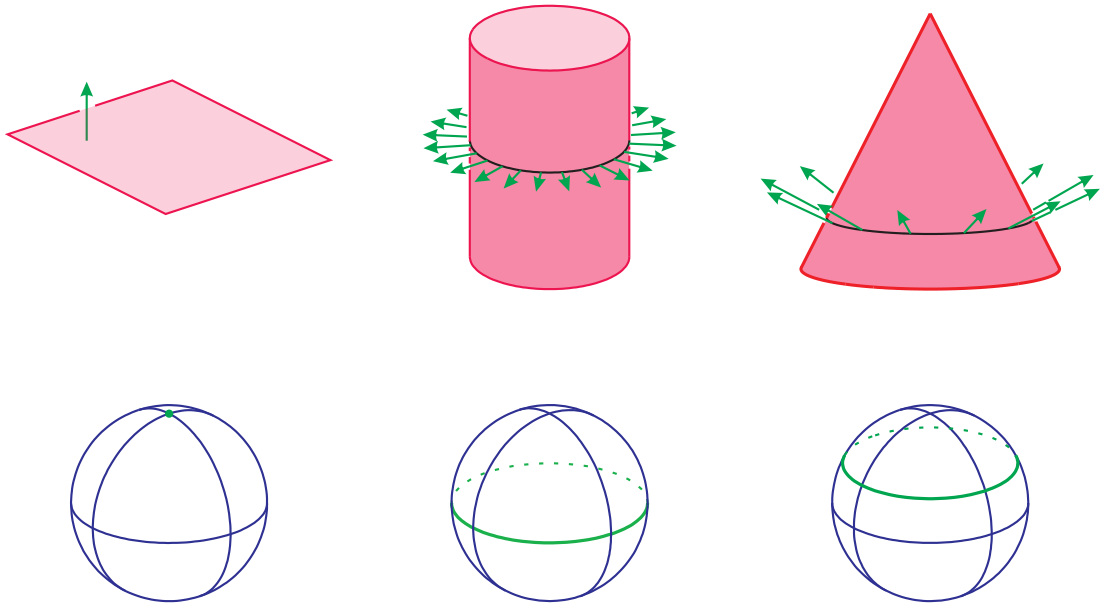

Abb. 1.2: oben v.l.n.r.: Ebene, Zylinder, Kegel; unten: die zugehörigen Gaußbilder

Sei V_p ein Tangentialvektor mit $p = f(a)$, $F_j(a) := (f(a); f_{,j}(a)), (j = 1 \dots m)$ eine Basis des zugehörigen Tangentialvektorraumes und $N(a) = (p; n(a))$ der Normalvektor im Punkt $f(a)$, dann wirkt die Weingarten-Abbildung also in der Art

$$V_p = v^j F_j(a) \mapsto \omega_p(V_p) = -v^j N_{,j}(a) \tag{1.1}$$

und insbesondere

$$\omega_p(F_j(a)) = -N_{,j}(a). \tag{1.2}$$

Satz 1.1. *Ist $f \in C^2(U)$, so ist die Weingarten-Abbildung ω_p selbstadjungiert für alle $p = f(a) \in f(U)$*

Beweis. Aufgrund der Linearität von ω_p genügt es zu zeigen, dass für die Basisvektoren die Beziehung

$$F_j(a) \cdot \omega_p(F_k(a)) = \omega_p(F_j(a)) \cdot F_k(a) \text{ bzw. } f_{,j} \cdot n_{,k} = n_{,j} \cdot f_k$$

gilt. Wir wissen, dass für alle j die Beziehung $f_{,j} \cdot n = 0$ in U gilt und somit folgt durch Differenziation

$$f_{,jk} \cdot n + f_{,j} \cdot n_{,k} = 0 \text{ und analog } f_{,kj} \cdot n + f_{,k} \cdot n_{,j} = 0.$$

Wegen $f \in C^2$ folgt durch Differenzbildung die Behauptung. $\square$

Aus der linearen Algebra ist bekannt, dass bei einer selbstadjungierten Abbildung alle Eigenwerte reel sind und dass eine Orthonormalbasis aus Eigenvektoren existiert. Somit ist nachfolgende Definition sinnvoll:

Definition 1.11. Es sei $f \in C^2(U)$ regulär. Die m reellen Eigenwerte ${}_{(1)}\kappa_p, \ldots, {}_{(m)}\kappa_p$ von ω_p heißen Hauptkrümmungen von $\overrightarrow{f(U)}$ in p. Ein zu einer Hauptkrümmung gehöriger Eigenvektor heißt Krümmungsvektor, die zugehörige Flächentangente Krümmungstangente.
$K(a) := \det \omega_p = {}_{(1)}\kappa_p \cdots {}_{(m)}\kappa_p$ heißt Gaußsche Krümmung von $\overrightarrow{f(U)}$ in $p = f(a)$.

Bevor wir uns nun damit beschäftigen, wie die Gaußsche Krümmung mit abwickelbaren Flächen in Verbindung steht, wollen wir uns noch eine weitere Betrachtungsweise der Gaußschen Krümmung zuwenden, die uns später in Kapitel 4 wieder begegnen wird. Dazu benötigen wir allerdings noch einige Vorkenntnisse, die zunächst erläutert werden sollen.

Das kanonische innere Produkt im $\mathbb{R}^n$ induziert in jedem Tangentialvektorraum $T_p\mathbb{R}^n$ ein inneres Produkt. Dabei gilt mit $X_p = (p; v)$ und $Y_p = (p; w) \in T_p\mathbb{R}^n$, dass $X_p \cdot Y_p = v \cdot w$. Die folgenden zwei Definitionen, der metrischen Grundform und der Oberfläche, beschränken sich nicht nur auf Hyperflächen.

Definition 1.12. Unter der metrischen (ersten) Grundform der Fläche $f(U)$ im Punkt $f(a)$ versteht man mit $X_p = v^j F_j(a), Y_p = w^j F_j(a) \in T_a f$ die Abbildung $G_{f(a)} : T_a f \times T_a f \to \mathbb{R}$ mit

$$G_{f(a)}(X_p, Y_p) := X_p \cdot Y_p = v^j w^k F_j(a) \cdot F_k(a) = v^j w^k f_{,j}(a) \cdot f_{,k}(a) = v^j w^k g_{jk}(a) \quad (1.3)$$

mit

$$g_{jk}(a) := G_{f(a)}(F_j(a), F_k(a)) = f_{,j}(a) \cdot f_{,k}(a) \text{ und } \Delta(a) := \det(g_{jk}(a)). \quad (1.4)$$

Als nächstes soll die Oberfläche definiert werden, was wir anhand einer Überlegung im dreidimensionalen Raum motivieren wollen. Wir gehen von einem Weg $f : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ mit $f(u^1, u^2) = (u^1, u^2, z(u^1, u^2))$ aus. Der Grundgedanke ist, dass ein achsenparalleles Rechteck mit Flächeninhalt A im Parametergebiet U mit Mittelpunkt a durch Projektion parallel zur x^3-Achse in die Tangentialebene von $f(U)$ in $f(a)$ auf ein Parallelogramm abgebildet wird, welches den Flächeninhalt $A' = A\sqrt{\Delta}$ hat. Stellt man sich nun eine

kompakte Menge $T \subset U$ vor, so kann man diese durch infinitesimale Rechtecke annähern und die projizierten Rechtecke nähern die Oberfläche beliebig gut an. Ganz allgemein können wir die Oberfläche folgendermaßen definieren:

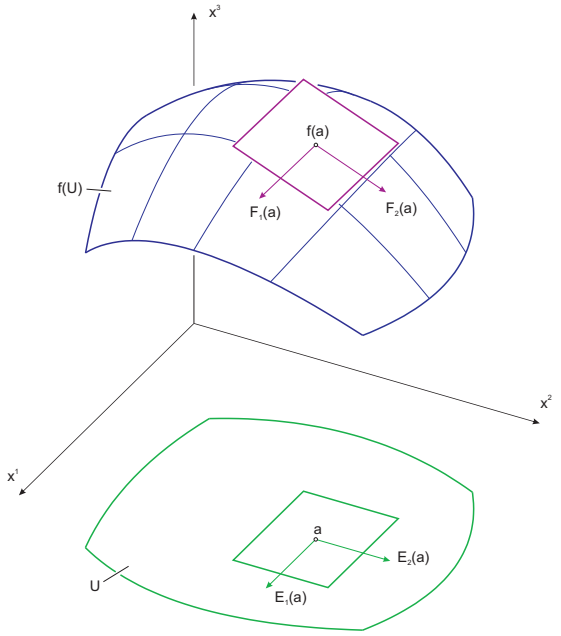

Abb. 1.3: Geometrische Motivation der Oberflächendefinition

Definition 1.13. Sei $f(U)$ eine reguläre Fläche. Dann heißt für jede beschränkte, Jordan-messbare Teilmenge $T \subset U$

$$O(f(T)) := \int_T \sqrt{\Delta} du^1 \dots du^m$$

die Oberfläche von $f(T)$.

Bevor wir unsere Ausführungen zur Gaußschen Krümmung abschließen können, brauchen wir noch ein letztes Hilfsmittel, nämlich die sogenannte Krümmungsform:

Definition 1.14. Sei $f \in C^2(U)$ regulär, $p = f(a)$. Dann versteht man unter der Krümmungsform (2. Grundform) der orientierten Fläche $\overrightarrow{f(U)}$ im Punkt $f(a)$ die Abbildung $H_{f(a)} : T_a f \times T_a f \to \mathbb{R}$ mit

$$H_{f(a)}(X_p, Y_p) := G_{f(a)}(X_p, \omega_p(Y_p)) = X_p \cdot \omega_p(Y_p).$$

Die Funktionen $h_{jk} : U \to \mathbb{R}$ mit

$$h_{jk}(a) := H_{f(a)}(F_j(a), F_k(a))$$

heißen Koordinatenfunktionen von H_f.

Fassen wir nun noch einige Eigenschaften zusammen:
Aus $f_{,j} \cdot n = 0$ folgt durch Differenziation $f_{,jk} \cdot n + f_{,j} \cdot n_{,k} = 0$. Somit gilt

$$h_{jk}(a) = F_j(a) \cdot \omega_p(F_k(a)) = -f_{,j}(a) \cdot n_{,k}(a) = n(a) \cdot f_{,jk}(a). \tag{1.5}$$

Weiters gilt mit $\omega_p(F_j(a)) = h_j^l(a) F_l(a)$

$$h_{jk}(a) = F_j(a) \cdot \omega_p(F_k(a)) = f_{,j}(a) h_k^l(a) f_{,l}(a) = h_k^l(a) g_{jl}(a), \tag{1.6}$$

also gilt, wenn $(g^{jk}(a))$ die zu $(g_{jk}(a))$ inverse Matrix bezeichnet

$$h_{jk}(a) = h_k^l(a) g_{jl}(a) \Leftrightarrow h_k^l(a) = h_{ks} g^{sl}(a). \tag{1.7}$$

Somit lässt sich nun auch die Gaußsche Krümmung über die Determinanten der ersten und zweiten Grundform berechnen, nämlich durch

$$K(a) = \det \omega_p = \det \left(h_k^j(a) \right) = \frac{\det \left(h_{jk}(a) \right)}{\det \left(g_{jk}(a) \right)}. \tag{1.8}$$

Kommen wir nun zur zuvor erwähnten alternativen Darstellung der Gaußschen Krümmung.

Satz 1.2. *Sei $f \in C^2(U)$ regulär, $a \in U$ und $B_\varepsilon := \{u \in U | \, \|u - a\| \leq \varepsilon\}$. Dann gilt mit $\Omega := O(f(B_\varepsilon))$ und $\overline{\Omega} := O(n_f(B_\varepsilon))$*

$$|K(a)| = \lim_{\varepsilon \to 0} \frac{\overline{\Omega}}{\Omega}. \tag{1.9}$$

Beweis. Es ist $\Omega = O(f(B_\varepsilon)) = \int_{B_\varepsilon} \sqrt{\Delta} du^1 du^2$ und $\overline{\Omega} = O(n(B_\varepsilon)) = \int_{B_\varepsilon} \sqrt{\overline{\Delta}} du^1 du^2$ mit $\overline{\Delta} := \det(\overline{g}_{jk})$ und $\overline{g}_{jk} := n_{,j} \cdot n_{,k}$. Weiters gilt

$$\overline{g}_{jk} = n_{,j} \cdot n_{,k} = (-h_j^l f_{,l}) \cdot (-h_k^s f_{,s}) = h_j^l h_k^s g_{ls} \Rightarrow \overline{\Delta} = \det(\overline{g}_{jk}) = K^2 \Delta$$

und damit ergibt sich mit $b \in B_\varepsilon$ für $\overline{\Omega}$

$$\overline{\Omega} = \int\limits_{B_\varepsilon} |K| \sqrt{\Delta} du^1 du^2 = |K(b)| \int\limits_{B_\varepsilon} \sqrt{\Delta} du^1 du^2 = |K(b)| \, \Omega$$

Das zweite Gleichheitszeichen folgt dabei aus dem verallgemeinerten Mittelwertsatz der Integralrechnung. Für $\varepsilon \to 0$ gilt somit

$$\lim_{\varepsilon \to 0} \frac{\overline{\Omega}}{\Omega} = \lim_{\varepsilon \to 0} \frac{K(b)\Omega}{\Omega} = K(a). \qquad \square$$

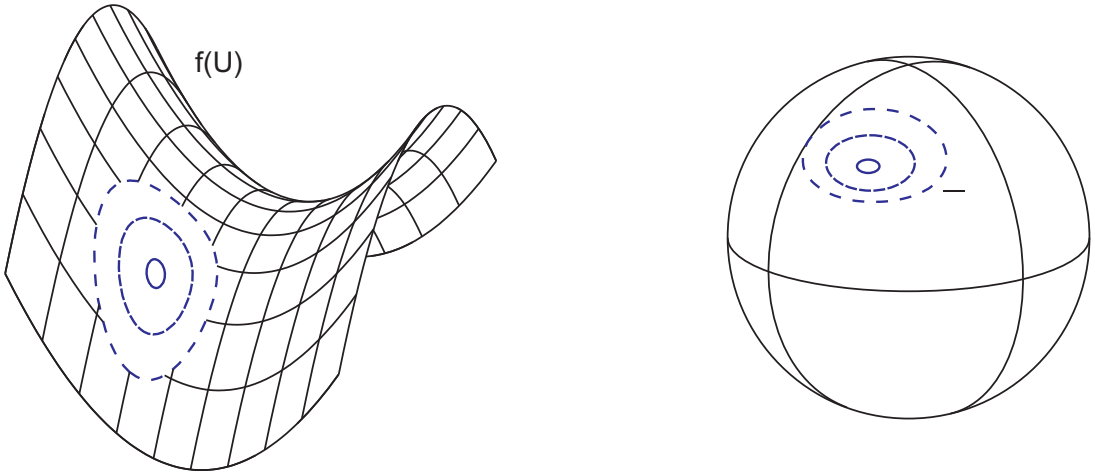

Abb. 1.4: Gaußsche Krümmung als Grenzwert eines Verhältnisses von Flächen

Wir wollen im Folgenden betrachten, wie die Gaußsche Krümmung mit abwickelbaren Flächen in Zusammenhang steht. Dazu werfen wir einen Blick auf Gauß berühmtes Theorema Egregium, welches wir aber nicht beweisen.

Satz 1.3. *Die Gaußsche Krümmung K einer regulären C^3-2-Fläche im $\mathbb{R}^3$ ist innergeometrisch.*

Für einen Beweis siehe z.B. [3]. Daraus ergibt sich der Umstand, dass für $f(U) \subset \mathbb{R}^3$ die Gaußsche Krümmung invariant bei Isometrien ist.

Definition 1.15. Eine reguläre C^2-Fläche im $\mathbb{R}^3$ mit $K = 0$ heißt Torse.

Somit können wir nun den Zusammenhang zwischen Gaußscher Krümmung und Abwickelbarkeit herstellen, denn es gilt:

Satz 1.4. *Eine reguläre C^3-2-Fläche im $\mathbb{R}^3$ ist genau dann abwickelbar, wenn sie eine C^3-Torse ist.*

1.1.3 Regelflächen

Wir wollen diesen Abschnitt mit einigen Ausführungen über Regelflächen abschließen, die uns im späteren Verlauf noch einige Male begegnen werden.

Definition 1.16. Eine Fläche in $\mathbb{R}^3$ heißt eine C^r-Regelfläche ($r \geq 1$), wenn sie mit $U := I \times \mathbb{R} \subset \mathbb{R}^2$ eine C^r-Parameterdarstellung $f : U \to \mathbb{R}^3$ der Gestalt

$$(u^1, u^2) \in U \mapsto f(u^1, u^2) = l(u^1) + u^2 r(u^1) \in \mathbb{R}^3 \text{ mit } l, r : U \to \mathbb{R}^3 \text{ und } r \neq \mathbf{0} \text{ in } U$$

gestattet. Die geradlinigen Flächenkurven $u^1 =$ konstant heißen Erzeugende der Regelfläche. Die Kurve $l(I) \subset f(U) \subset \mathbb{R}^3$ heißt eine Leitkurve von $f(U)$, die Kurve $r(I)$ eine Richtkurve.

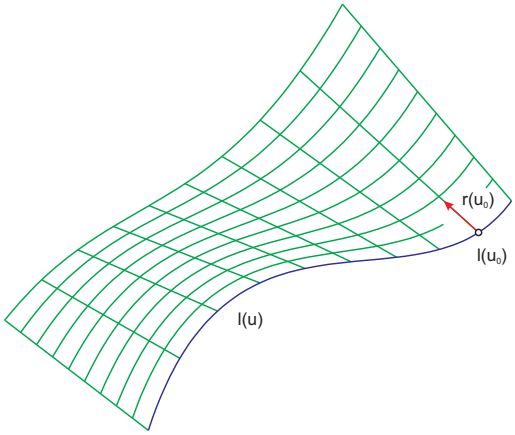

Abb. 1.5: Regelfläche

Bemerkung: Wegen $r \neq \mathbf{0}$ kann ohne Einschränkung $\|r\| = 1$ in I vorausgesetzt werden.

Definition 1.17. Trägt eine Erzeugende nur singuläre Punkte, so heißt sie singuläre Erzeugende, sonst nichtsinguläre Erzeugende.

Satz 1.5. *(1) Eine Erzeugende $e(u_0^1)$ ist singulär* $\Leftrightarrow$

$$(\dot{l} \times r)(u_0^1) = (\dot{r} \times r)(u_0^1) = \mathbf{0}.$$

(2) Trägt eine Erzeugende $e(u_0^1)$ zwei verschiedene singuläre Punkte, so ist sie singulär.

Beweis. (1) $f_{,u^1} = \dot{l}(u^1) + u^2 \dot{r}(u^1)$ und $f_{,u^2} = r(u^1)$ und somit

$$(f_{,u^1} \times f_{,u^2})(u_0^1, u^2) = \mathbf{0} \Leftrightarrow (\dot{l} \times r)(u_0^1) + u^2 (\dot{r} \times r)(u_0^1) = \mathbf{0}.$$

(2) Gilt die soeben genannte Beziehung für zwei verschiedene u^2-Werte, so gilt $(\dot{l} \times r)(u_0^1) = (\dot{r} \times r)(u_0^1) = \mathbf{0}$ also $(f_{,u^1} \times f_{,u^2})(u_0^1, u^2) = \mathbf{0}$ $\forall u^2$, also ist jeder Punkt von $e(u_0^1)$ singulär. □

Definition 1.18. Eine nichtsinguläre Erzeugende $e(u_0^1)$ heißt torsale Erzeugende, wenn in zwei verschiedenen regulären Punkten von $e(u_0^1)$ die Tangentialebenen übereinstimmen. Besitzt eine Regelfläche nur torsale Erzeugende, so heißt sie torsale Regelfläche.

Satz 1.6. *Für eine nichtsinguläre Erzeugende $e(u_0^1)$ gilt:*

$$e(u_0^1) \text{ ist torsal } \Leftrightarrow \det(\dot{l}, r, \dot{r}) = 0$$

Beweis. Sei $e(u_0^1)$ torsal. Dann gibt es zwei verschiedene Zahlen $\alpha, \beta \in \mathbb{R}$, sodass gilt

$$[(\dot{l} \times r) + \alpha(\dot{r} \times r)] \times [(\dot{l} \times r) + \beta(\dot{r} \times r)] = 0 \Leftrightarrow (\alpha - \beta)[(\dot{r} \times r) \times (\dot{l} \times r)] = 0$$

Dies ist aber wegen

$$(a \times b) \times (c \times d) = c \det(a, b, d) - d \det(a, b, c)$$

äquivalent mit

$$\det(\dot{l}, r, \dot{r})(u_0^1) = 0.$$ □

Bemerkung: Eine torsale Erzeugende besitzt somit in allen regulären Punkten dieselbe Tangentialebene.

Definition 1.19. Sei $e(u_0^1)$ eine torsale Erzeugende.
(1) Trägt $e(u_0^1)$ einen singulären Punkt, so heißt dieser Kuspidalpunkt.
(2) Trägt $e(u_0^1)$ keinen singulären Punkt, so heißt $e(u_0^1)$ zylindrisch.

Der Kuspidalpunkt einer nicht singulären Erzeugenden eines Kegels bzw. einer Tangentenfläche ist die Kegelspitze bzw. der Gratpunkt und jede nicht singuläre Erzeugende eines Zylinders ist zylindrisch.

Satz 1.7. *Jede von singulären Punkten freie C^2-Fläche, die einer torsalen Regelfläche angehört, ist eine Torse.*

Der abschließende Satz zeigt nun, dass Zylinder, Kegel und Tangentenfläche im wesentlichen die einzigen torsalen Regelflächen sind.

Satz 1.8. *Ist $f : I \times \mathbb{R} \to \mathbb{R}^3$ ein torsaler C^2-Regelflächenweg, so existiert zu jedem $u_0^1 \in I$ in jeder offenen Umgebung $I_0 \subset I$ von u_0^1 ein Punkt $\overline{u}^1 \in I_0$ derart, dass für eine offene Umgebung $\overline{I}$ von $\overline{u}^1$ dann $f(\overline{I} \times \mathbb{R})$ zylindrisch, kegelig oder tangentenflächig ist.*

Beweis. Da alle Erzeugenden torsal sind gilt:

- keine singuläre Erzeugende $\Leftrightarrow (\dot{l} \times r), (\dot{r} \times r) \neq (0,0)$ in I
- jede Erzeugende ist torsal $\Leftrightarrow \det(\dot{l}, r, \dot{r}) = 0$ in I.

(1) Sei $\dot{r} \times r = \mathbf{0}$ in I
Für $\|r\| = 1$ in I (o.B.d.A.) folgt $\|\dot{r} \times r\| = \|\dot{r}\|$, also $\dot{r} = 0 \Rightarrow r = r_0 \in \mathbb{R}^3$ konst.

$$\Rightarrow f(u,v) = l(u) + vr_0 \Rightarrow f(U)\text{zylindrisch.}$$

(2) Sei $\dot{r} \times r \neq 0$ in I
Wegen $\det(\dot{l}, r, \dot{r}) = 0$ folgt

$$\dot{l} = Ar + B\dot{r},\ A, B : I \to \mathbb{R}.$$

Wir betrachten die Flächenkurve $s(I)$ mit $s(u) := l(u) - B(u)r(u)$. Dann gilt:

$$\dot{s} = \dot{l} - \dot{B}r - B\dot{r} = Ar - \dot{B}r = (A - \dot{B})r.$$

(a) $A - \dot{B} = 0$ in $I \Rightarrow \dot{s} = 0$ in $I \Rightarrow s(I)$ ist einpunktig $\Rightarrow f(U)$ kegelig.
(b) $A - \dot{B} \neq 0$ in $I \Rightarrow \dot{s} \neq 0$ in $I \Rightarrow s(I)$ ist eine reguläre Flächenkurve $\Rightarrow f(U)$ tangentenflächig.

Weiters gilt:

- Ist eine stetige Funktion an einer Stelle ungleich Null, so gilt dies lokal.
- Ist eine stetige Funktion an einer Stelle gleich Null, so gilt dies entweder lokal oder die Stelle ist eine isolierte Nullstelle.

Daraus folgt die Behauptung. □

1.2 Grundlagen der Liniengeometrie

Die nachfolgenden Grundlagen können in jedem Lehrbuch über Liniengeometrie, wie etwa [4], nachgelesen werden.

1.2.1 Plücker-Koordinaten

Für die Menge $\mathcal{L}$ der Geraden im dreidimensionalen projektiven Raum $\mathcal{P}^3$ können nach J. Plücker[4] Koordinaten eingeführt werden. Eine Gerade L kann mittels zwei ihrer Punkte X und Y geschrieben werden in der Form $L = X \vee Y$. Wir definieren ein äußeres Produkt $\wedge$, das zwei Vektoren $\mathbf{x}$ und $\mathbf{y}$ des $\mathbb{R}^4$ einen Vektor zuweist, der in einem neuen Vektorraum enthalten ist, welchen wir mit $\Lambda^2\mathbb{R}^4$ bezeichnen. Die Operation $\wedge$ soll bilinear in beiden Argumenten, also

$$\mathbf{x} \wedge (\lambda\mathbf{y} + \mu\mathbf{z}) = \lambda\mathbf{x} \wedge \mathbf{y} + \mu\mathbf{x} \wedge \mathbf{z}, \quad (\lambda\mathbf{x} + \mu\mathbf{y}) \wedge \mathbf{z} = \lambda\mathbf{x} \wedge \mathbf{z} + \mu\mathbf{y} \wedge \mathbf{z}$$

und antikommutativ sein, also

$$\mathbf{x} \wedge \mathbf{y} = -\mathbf{y} \wedge \mathbf{x}.$$

Seien $\mathbf{e}_0, \ldots, \mathbf{e}_3$ eine Basis in $\mathbb{R}^4$ und seien $\mathbf{x} = \sum x_i\mathbf{e}_i$ und $\mathbf{y} = \sum y_i\mathbf{e}_i$, dann folgt aus der Bilinearität von $\wedge$, dass

$$\mathbf{x} \wedge \mathbf{y} = \sum_{i,j=0}^{3} x_i y_j \mathbf{e}_i \wedge \mathbf{e}_j.$$

Aus der Antikommutativität von $\wedge$ folgt, dass $\mathbf{e}_i \wedge \mathbf{e}_i = \mathbf{0},\ \forall i$ und daher bleiben nur sechs nichttriviale Produkte, nämlich

$$\mathbf{e}_0 \wedge \mathbf{e}_1,\ \mathbf{e}_0 \wedge \mathbf{e}_2,\ \mathbf{e}_0 \wedge \mathbf{e}_3,\ \mathbf{e}_2 \wedge \mathbf{e}_3,\ \mathbf{e}_3 \wedge \mathbf{e}_1,\ \mathbf{e}_1 \wedge \mathbf{e}_2. \tag{1.10}$$

Somit gilt für das Produkt $\mathbf{x} \wedge \mathbf{y}$

$$\begin{aligned} &\mathbf{x} \wedge \mathbf{y} = \sum_{(i,j)\in I} (x_i y_j - x_j y_i)\mathbf{e}_i \wedge \mathbf{e}_j, \\ &\text{mit} \quad I = \{(0,1),(0,2),(0,3),(2,3),(3,1),(1,2)\}. \end{aligned} \tag{1.11}$$

Definition 1.20. Sei der Vektorraum $\Lambda^2\mathbb{R}^4$ sechsdimensional mit der Basis (1.10). Das äußere Produkt der Vektoren $\mathbf{x} = \sum x_i\mathbf{e}_i$ und $\mathbf{y} = \sum y_i\mathbf{e}_i$ ist ein Element von $\Lambda^2\mathbb{R}^4$,

[4] Julius Plücker, geboren am 16.6.1801 in Elberfeld, gestorben 22.5.1868 in Bonn

und ist gegeben durch (1.11).

Die Koordinaten von $\mathbf{x} \wedge \mathbf{y}$ werden *Plücker-Koordinaten* der Geraden L, die durch die Punkte $\mathbf{x}\mathbb{R}$ und $\mathbf{y}\mathbb{R}$ bestimmt ist, genannt. Wir schreiben

$$(x_0, \ldots, x_3) \wedge (y_0, \ldots, y_3) = (l_{01}, l_{02}, l_{03}, l_{23}, l_{31}, l_{12}),\ l_{ij} = x_i y_j - x_j y_i$$

l_{ij} ist somit die Determinante der i-ten und j-ten Spalte der Matrix

$$\begin{pmatrix} x_0 & x_1 & x_2 & x_3 \\ y_0 & y_1 & y_2 & y_3 \end{pmatrix}$$

Lemma 1.9. *Die Plücker-Koordinaten sind, bis auf einen skalaren Faktor, unabhängig von der Wahl der Punkte der Geraden.*
Ein 6-Tupel $L = (l_{01}, \ldots, l_{12}) \neq (0, \ldots, 0)$ von reellen Zahlen ist genau dann das Ergebnis eines äußeren Produkts, wenn es der Bedingung

$$\Omega_q(L) = l_{01} l_{23} + l_{02} l_{31} + l_{03} l_{12} = 0 \tag{1.12}$$

genügt.

Beweis. $(a_1\mathbf{x} + b_1\mathbf{y}) \wedge (a_2\mathbf{x} + b_2\mathbf{y}) = a_1 a_2 \mathbf{x} \wedge \mathbf{x} + a_1 b_2 \mathbf{x} \wedge \mathbf{y} + b_1 a_2 \mathbf{y} \wedge \mathbf{x} + b_1 b_2 \mathbf{y} \wedge \mathbf{y} = (a_1 b_2 - b_1 a_2)\mathbf{x} \wedge \mathbf{y}$, was den ersten Teil der Aussage beweist. Dass L (1.12), die sogenannte Plücker-Identität, erfüllt, wenn $l_{ij} = x_i y_j - x_j y_i$ gilt, folgt durch elementares Ausrechnen. Um die Umkehrung zu zeigen, betrachten wir die Vektoren

$$\begin{aligned} \mathbf{s}_0 &= (0, l_{01}, l_{02}, l_{03}), \\ \mathbf{s}_1 &= (-l_{01}, 0, l_{12}, -l_{31}), \\ \mathbf{s}_2 &= (-l_{02}, -l_{12}, 0, l_{23}), \\ \mathbf{s}_3 &= (-l_{03}, l_{31}, -l_{23}, 0). \end{aligned}$$

Das äußere Produkt zweier dieser Vektoren ist ein skalares Vielfaches von L, und nicht alle sechs möglichen äußeren Produkte sind null, außer L ist null. □

Bemerkung. Wir sehen somit, dass jeder Geraden im $\mathcal{P}^3$ ein Punkt der sogenannten Kleinschen[5] Quadrik $\Omega_q = 0$ im $\mathcal{P}^5$ zugewiesen werden kann. Diese Zuweisung ist, wie wir gesehen haben bijektiv. Die Kleinsche Quadrik wird meist mit M_2^4 bezeichnet.

[5] Felix Christian Klein, geboren am 25.4.1849 in Düsseldorf, gestorben am 22.6.1925 in Göttingen

Die Plücker-Koordinaten $(l_{01}, l_{02}, l_{03}, l_{23}, l_{31}, l_{12})$ einer Geraden L werden im Folgenden häufig in der komprimierten Schreibweise

$$L = (\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}, \text{ mit } \quad \mathbf{l} = (l_{01}, l_{02}, l_{03}), \ \bar{\mathbf{l}} = (l_{23}, l_{31}, l_{12}),$$

wobei $\mathbf{l}$ der Richtungsvektor der Geraden L ist und $\bar{\mathbf{l}}$ der sogennante Momentenvektor, der orthogonal auf die Verbindungsebene von L mit dem Ursprung steht. Schreibt man

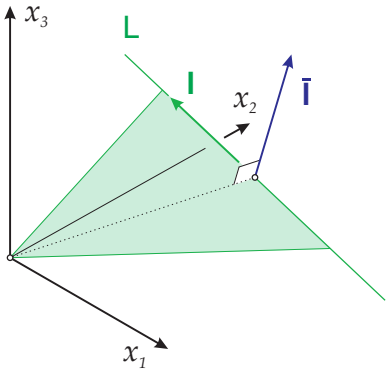

Abb. 1.6: Richtungs- und Momentenvektor

weiters die homogenen Koordinaten eines Punktes in der Form

$$(x_0 : x_1 : x_2 : x_3) = (x_0, \mathbf{x})\mathbb{R} \quad \text{mit} \quad \mathbf{x} = (x_1, x_2, x_3),$$

so kann man die Plücker-Koordinaten der Geraden L, die von $(x_0, \mathbf{x})\mathbb{R}$ und $(y_0, \mathbf{y})\mathbb{R}$ aufgespannt wird, mit

$$L = (x_0, \mathbf{x}) \wedge (y_0, \mathbf{y}) = (x_0\mathbf{y} - y_0\mathbf{x}, \mathbf{x} \times \mathbf{y})\mathbb{R} \tag{1.13}$$

berechnen.

1.2.2 Lineare Komplexe

Definition 1.21. Sei π eine Nullpolarität, dann nennen wir $P\pi$ die Nullebene von P und P heißt der Nullpunkt von $P\pi$. Eine Fixgerade $L = L\pi$ heißt Nullgerade.

Lemma 1.10. *Sind P, Q konjugierte Punkte von π, d.h. $P \in Q\pi$ bzw. $Q \in P\pi$, dann ist $P \vee Q$ eine Nullgerade, und alle Nullgeraden können derart dargestellt werden.*

Beweis. Es gilt $P \in P\pi$ und $Q \in Q\pi$, weil π eine Nullpolarität ist. Weiters gilt $P \in Q\pi$ und $Q \in P\pi$, weil P und Q konjugiert sind. $(P \vee Q)\pi = P\pi \cap Q\pi$, gilt für alle Korrela-

tionen. Daher enthält $(P \vee Q)\pi$ die Punkte P und Q, und $(P \vee Q)\pi = P \vee Q$. Sei nun L eine Nullgerade. Wähle $P, Q \in L$. Dann gilt $L = L\pi$, $L \subset P\pi$ und $L \subset Q\pi$, und somit sind P und Q konjugiert. □

Lemma 1.11. *Sei π eine Nullpolarität und sei C die zugehörige — notwendigerweise schiefsymmetrische — Matrix mit den Einträgen c_{ij}. Weiters setzen wir*

$$\overline{\mathbf{c}} = (c_{01}, c_{02}, c_{03}), \quad \mathbf{c} = (c_{23}, c_{31}, c_{12})$$

Nehmen wir an die Gerade L habe die Plücker-Koordinaten $(\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$. Dann gilt

$$L = L\pi \Leftrightarrow \overline{\mathbf{c}} \cdot \mathbf{l} + \mathbf{c} \cdot \bar{\mathbf{l}} = 0. \tag{1.14}$$

Beweis. Nehmen wir an, dass L durch die Punkte $\mathbf{x}\mathbb{R}$ und $\mathbf{y}\mathbb{R}$ aufgespannt wird und dass $(\mathbf{l}, \bar{\mathbf{l}}) = \mathbf{x} \wedge \mathbf{y}$. Konjugiertheit von $\mathbf{x}\mathbb{R}$ und $\mathbf{y}\mathbb{R}$ bedeutet

$$\mathbf{x}^T \cdot C \cdot \mathbf{y} = \sum_{i<k} c_{ik}(x_i y_k - x_k y_i) = 0.$$

Per Definition des äußeren Produkts ist das äquivalent zu $\overline{\mathbf{c}} \cdot \mathbf{l} + \mathbf{c} \cdot \bar{\mathbf{l}} = 0$. □

Definition 1.22. Eine Menge $\mathcal{C}$ von Geraden, die durch eine lineare homogene Gleichung in Plücker-Koordinaten gegeben ist, heißt linearer Linienkomplex oder kurz linearer Komplex.

Definition 1.23. Ein linearer Komplex $\mathcal{C}$, der durch $\overline{\mathbf{c}} \cdot \mathbf{x} + \mathbf{c} \cdot \overline{\mathbf{x}} = 0$ definiert ist, heißt singulär, wenn $C = (\mathbf{c}, \overline{\mathbf{c}})$ eine Gerade beschreibt, also $\Omega_q(C) = 0$; sonst heißt er regulär.

1.2.3 Der Zusammenhang zwischen linearen Komplexen und Schraubungen

Definition 1.24. Die Familie der euklidischen Kongruenztransformationen, die in einem geeigneten kartesischen Koordinatensystem Punkte nach der Vorschrift

$$\mathbf{x} \mapsto \mathbf{x}(t) = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 0 \\ 0 \\ pt \end{pmatrix} \tag{1.15}$$

abbilden, nennt man, falls $p \neq 0$, Schraubungen, andernfalls heißen sie Drehungen. Eine Schraubung ist die Überlagerung einer Drehung mit einer Translation in Richtung $(0, 0, pt)$. Die Gerade $x_1 = x_2 = 0$ wird Achse genannt und p heißt Schraubparameter.

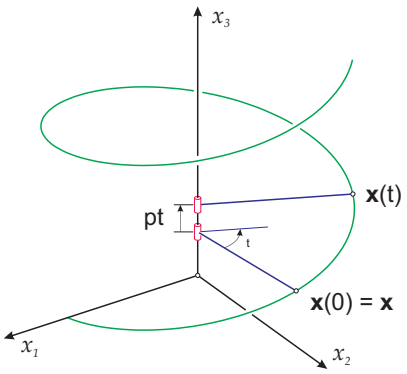

Abb. 1.7: Schraubung

Das Geschwindigkeitsvektorfeld einer Schraubung ist unabhängig von t: Sei $\mathbf{x}(t) = (x, y, z)$, dann gilt

$$\mathbf{v} = \frac{d\mathbf{x}(t)}{dt} = \begin{pmatrix} -y \\ x \\ p \end{pmatrix}. \tag{1.16}$$

Satz 1.12. *Die Bahnnormalen einer Schraubung im euklidischen dreidimensionalen Raum sind die eigentlichen Geraden eines regulären linearen Komplexes $\mathcal{C}$ von $\mathcal{P}^3$.*

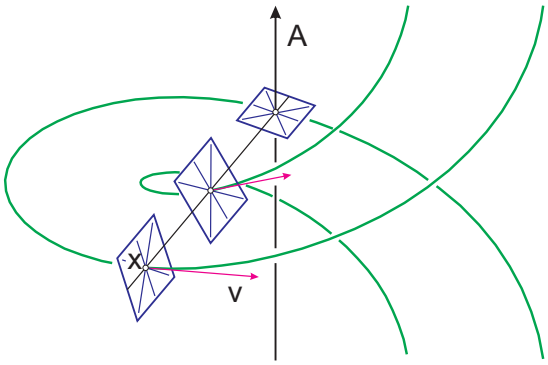

Abb. 1.8: Bahnnormalen einer Schraubung

Beweis. Die Plücker-Koordinaten $(\mathbf{l}, \bar{\mathbf{l}})$ einer Bahnnormale einer Schraubung berechnen sich aus (1.15) und (1.16): Die Normale wird durch $(1, \mathbf{x})$ und $(0, \mathbf{n})$, wobei $\mathbf{n}$ orthogonal auf $\mathbf{v}$ steht, aufgespannt. Somit gilt $(\mathbf{l}, \bar{\mathbf{l}}) = (\mathbf{n}, \mathbf{x} \times \mathbf{n})$ und $\mathbf{v} \cdot \mathbf{l} = -yl_{01} + xl_{02} + pl_{03} = 0$. Außerdem erhält man $l_{12} = -yl_{01} + xl_{02}$ aus $\bar{\mathbf{l}} = \mathbf{x} \times \mathbf{l}$. Das zeigt nun, dass $pl_{03} + l_{12} = 0$

und daher alle Normalen im linearen Komplex mit der Gleichung

$$pl_{03} + l_{12} = 0 \tag{1.17}$$

enthalten sind. □

Nun wollen wir noch überlegen, wie wir die Achse und den Schraubparameter einer Schraubung berechnen können, wenn wir den linearen Komplex gegeben haben.

Satz 1.13. *Der Schraubparameter p und die Plücker-Koordinaten $(\mathbf{a}, \overline{\mathbf{a}})\mathbb{R}$ der Achse A eines regulären linearen Komplexes $\mathcal{C}$ mit der Gleichung $\overline{\mathbf{c}} \cdot \mathbf{l} + \mathbf{c} \cdot \overline{\mathbf{l}} = 0$ sind gegeben durch*

$$p = \frac{\mathbf{c} \cdot \overline{\mathbf{c}}}{\mathbf{c}^2}, \quad (\mathbf{a}, \overline{\mathbf{a}}) = (\mathbf{c}, \overline{\mathbf{c}} - p\mathbf{c}). \tag{1.18}$$

2 Das Schätzen von Normalvektoren

Da im Folgenden die Flächennormalen eine essenzielle Rolle spielen, soll in diesem Kapitel exemplarisch beleuchtet werden, wie man die Normalvektoren, bei gegebener Punktwolke, schätzen kann. Dazu werfen wir zum einen einen Blick auf die Arbeit [8] von Daoshan OuYang und Hsi-Yung Feng aus dem Jahr 2004 und zum anderen betrachten wir einen Ansatz von Nelson Max.

2.1 Normalvektorschätzung mit quadratischen Kurven

Der erste Schritt zu einem vernünftigen Schätzer für den Normalvektor $\mathbf{n}$ in einem Datenpunkt P einer Punktwolke ist, die Nachbarn von P zu identifizieren. Eine beliebte Methode hierfür ist die K Punkte mit kleinstem Abstand zu P zu ermitteln und darauf aufbauend den Normalvektor in P zu schätzen. Der Ansatz, in der nun beschriebenen Arbeit basiert auf der Idee des globalen Delaunay-Triangulations-Netzes, welches aus dem 3D Voronoi Diagramm ermittelt wird. Hierzu soll zunächst der Begriff Voronoi Diagramm erläutert werden:

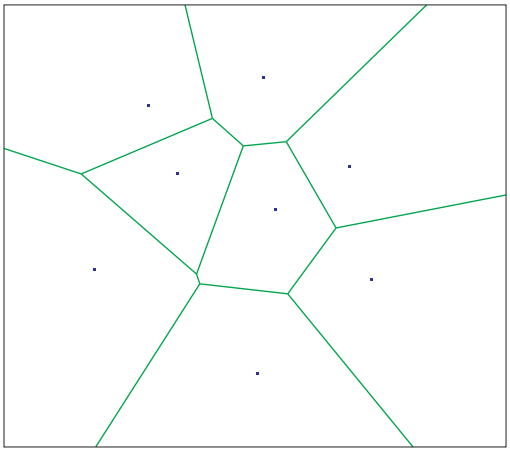

Abb. 2.1: 2D Voronoi-Diagramm

Definition 2.1. Sei P eine Menge von n Punkten im $\mathbb{E}^d$, dann ist das Voronoi[1]-Diagramm $V(P)$ ein Zellkomplex in $\mathbb{E}^d$, welcher Lagebeziehungen zwischen den Punkten widerspiegelt. Genauer definieren wir $V(P)$ derart, dass es $\mathbb{E}^d$ in n Zellen unterteilt, eine für jeden Punkt, mit der Eigenschaft, dass x genau dann in der zu p gehörigen Zelle liegt, wenn

$$d(x,p) < d(x,q), \quad \forall q \in P \backslash \{p\}.$$

$d(x,p)$ bezeichnet hierbei die euklidische Distanz der Punkte x und p.

Da allerdings die Berechnung des Delaunay-Triangulations-Netzes nicht immer möglich ist, ermitteln sie die, wie sie sie nennen „lokalen Voronoi-Netz Nachbarn“, die eine Teilmenge der (globalen) Voronoi-Nachbarn sind. Hierzu wandeln sie den Ball-Pivoting Algorithm aus [2] etwas ab. Die Grundidee in [2] ist, dass man, in einer hinreichend dichten Punktwolke, ausgehend von einem Startdreieck, eine Kugel mit Radius ρ, die zunächst alle Punkte des Startdreiecks berührt, um eine der Dreiecksseiten rotieren lässt, bis sie einen weiteren Punkt der Datenmenge berührt, der somit das nächste Dreieck definiert. Dieser Algorithmus wird in [8] dahingehend abgewandelt, dass die rotierende Kugel einen variablen Radius hat und weiters werden nur jene Punkte als Nachbarn in Betracht gezogen, die globale Voronoi-Nachbarn der beiden Punkte der Achse sind. Konkret passiert bei diesem Algorithmus folgendes, um zu einem gegebenen Punkt P die lokalen Voronoi-Nachbarn zu finden:

1. Suche unter den globalen Voronoi-Nachbarn jenen Punkt N, der den geringsten Abstand zu P aufweist.

2. Verwende $\overline{PN}$ als Durchmesser und rotiere eine Kugel mit zunehmenden Radius um $\overline{PN}$, bis diese einen Punkt M, der ein gemeinsamer globaler Voronoi-Nachbar von P und N ist, trifft.

3. Ausgehend vom Dreieck NPM, finde den besten Punkt A, in den gemeinsamen globalen Voronoi-Nachbarn von P und N, um das Netz wachsen zu lassen. Der beste Punkt ist jener, der folgende Kriterien erfüllt:

   - A und M müssen auf verschiedenen Seiten der Ebene PNO liegen, wobei O der am weitesten entfernte Eckpunkt der Voronoi-Zelle von P ist.
   - $\angle PAN$ muss der größtmögliche Winkel sein, damit die Kugel um P, N und A die kleinste ist.

[1] Georgy Fedoseevich Voronoy, geboren am 28.04.1868 in Zhuravka, Russland (heutige Ukraine), gestorben am 20.11.1908 in Warschau

Analog findet man B zu den Punkten P und M.

4. Wiederhole das Verfahren für die Kanten PA und PB, und iteriere diese Vorgehensweise, bis das lokale Netz geschlossen ist oder es überlappt.

Die Autoren weisen auf mögliche, bereits von anderen Wissenschaftern untersuchte Wege hin, die man von den gegebenen Nachbarn aus einschlagen kann, um den Normalvektor in P zu schätzen. Darunter fallen die folgenden Möglichkeiten:

1. Man bestimmt mit Hilfe der Nachbarpunkte eine quadratische Fläche und berechnet die Flächennormale im Punkt P.

2. Man berechnet einen gewichteten Durchschnitt der Normalen der Dreiecke des Netzes um P. Dabei kann die Gewichtung entweder auf den Dreiecksflächen oder auf dem Winkel $\angle P_i P P_{i+1}$ basieren, wobei die P_j die Nachbarpunkte von P sind.

3. Es wird eine Ausgleichsebene ermittelt, die die Nachbarpunkte möglichst gut annähert und die Ebenennormale wird als Schätzer für die Normale in P herangezogen.

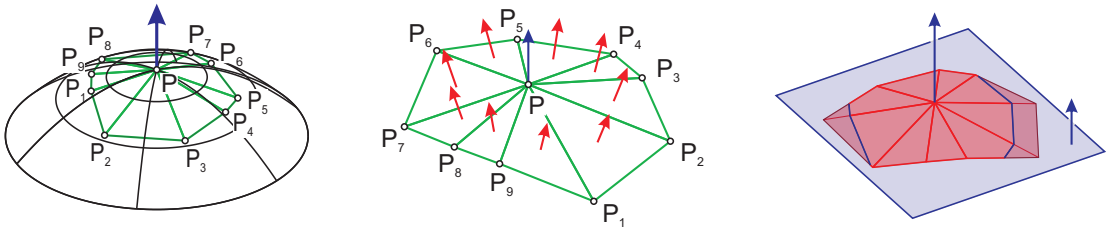

Abb. 2.2: Verschiedene Methoden der Normalvektorenschätzung

Der neue Zugang in der Arbeit [8] führt über quadratische Kurven, die durch die K Nachbarpunkte eines Punktes P_0 gelegt werden. Konkret läuft der Algorithmus wie folgt ab:

1. Finde zu jedem Nachbarpunkt P_j von P_0 den zugehörigen Punkt P_i, sodass der Winkel bei P_0 maximal ist, also

$$\angle P_i P_0 P_j > \angle P_i P_0 P_m, \quad 1 \leq m \leq K, \quad m \neq j. \tag{2.1}$$

2. Bestimme eine quadratische Kurve $\mathbf{P}(u)$ durch P_i, P_0, P_j:

$$\begin{aligned} &\mathbf{P}(u) = \mathbf{a}_0 + \mathbf{a}_1 u + \mathbf{a}_2 u^2, \quad u \in [0,1] \\ &\mathbf{P}(0) = P_i; \quad \mathbf{P}(1) = P_j; \quad \mathbf{P}(u_0) = P_0, \\ &u_0 = \frac{\overline{P_i P_0}}{\overline{P_i P_0} + \overline{P_0 P_j}}. \end{aligned} \tag{2.2}$$

3. Eruiere den normierten Tangentialvektor im Punkt P_0 an die soeben ermittelte Kurve $\mathbf{P}(u)$:

$$\mathbf{v_i} = \left(\frac{\mathbf{a}_1 + 2\mathbf{a}_2 u_0}{|\mathbf{a}_1 + 2\mathbf{a}_2 u_0|} \right). \tag{2.3}$$

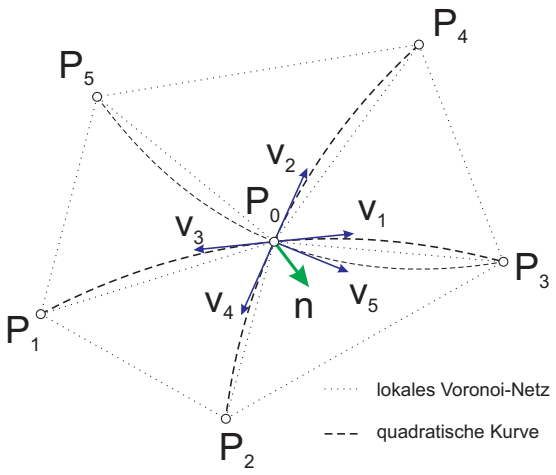

Abb. 2.3: Ermittlung des Normalvektors n in P_0

4. Bestimme den Schätzer für den Normalvektor $\mathbf{n}$ in P_0, durch Minimierung der Varianz s^2 der Skalarprodukte von $\mathbf{n}$ mit den K Tangentialvektoren.

$$s^2 = \frac{\sum_{i=1}^{K} (D_i - \overline{D})^2}{K-1} \tag{2.4}$$

mit

$$D_i = \mathbf{n} \cdot \mathbf{v}_i \quad \text{und} \quad \overline{D} = \frac{\sum_{i=1}^{K} D_i}{K}. \tag{2.5}$$

Alternativ kann man (2.4) auch schreiben als

$$s^2 = \frac{\sum_{i=1}^{K} [\mathbf{n}(\mathbf{v}_i - \overline{\mathbf{v}})]^2}{K-1} \tag{2.6}$$

mit

$$\overline{\mathbf{v}} = \frac{\sum_{i=1}^{K} \mathbf{v}_i}{K}. \tag{2.7}$$

Die Autoren unterstreichen die Vorteile dieser Methode, die unter anderem darin liegen, dass hier lediglich drei Nachbarpunkte benötigt werden um einen Normalvektor zu ermitteln, wohingegen die Methode mit der quadratischen Fläche mindestens neun Punkte benötigt, wodurch womöglich auch Punkte verwendet werden, die nicht in der unmittelbaren Umgebung von P_0 liegen. Dies könnte jedoch negative Auswirkungen auf den Schätzer des Normalvektors haben, da der Normalvektor eine lokale Eigenschaft ist.

2.2 Gewichte für die Berechnung von Normalvektoren

Nelson Max beschreibt in seiner Arbeit [14], wie man innerhalb eines Netzes die Normalen der Nachbardreiecke eines Datenpunktes gewichten kann, um die Normale in besagtem Datenpunkt zu schätzen. Ausgangspunkt dabei ist die folgende Überlegung: Angenommen wir approximieren eine Kugel durch ein ihr eingeschriebenes Polyeder und wollen nun die Normale in einem Eckpunkt Q des Polyeders schätzen. Die korrekte Normale verläuft durch den Mittelpunkt und diesen erhält man, indem man drei Symmetrieebenen, zwischen Q und drei Nachbarpunkten, schneidet. Wir wählen nun den Ursprung unseres Koordinatensystems im Punkt Q. Dann hat die Symmetrieebene zwischen Q und dem Nachbarpunkt $V_i = (x_i, y_i, z_i)$ $(i = 0 \dots n-1)$ die Gleichung

$$x_i x + y_i y + z_i z = \frac{\|V_i\|^2}{2}. \tag{2.8}$$

Für den Fall $n = 3$ erhält man daher drei lineare Gleichungen, deren Lösung die Koordinaten des Mittelpunktes liefern und diese lauten

$$M = \left(\|V_2\|^2 V_0 \times V_1 + \|V_0\|^2 V_1 \times V_2 + \|V_1\|^2 V_2 \times V_0 \right) / (2D), \tag{2.9}$$

wobei D die Determinante der Koeffizientenmatrix ist. Dividieren wir dieses Ergebnis durch $\|V_0\|^2 \|V_1\|^2 \|V_2\|^2 / (2D)$, so erhalten wir ein positives Vielfaches des Normalvektors:

$$\frac{V_0 \times V_1}{\|V_0\|^2 \|V_1\|^2} + \frac{V_1 \times V_2}{\|V_1\|^2 \|V_2\|^2} + \frac{V_2 \times V_0}{\|V_2\|^2 \|V_0\|^2} = cN. \tag{2.10}$$

Um dies nun über die Normalen der angrenzenden Flächen auszudrücken, definieren wir N_i als Normalvektor der i-ten Fläche, die von $Q, V_i, V_{i+1} \pmod n$ aufgespannt wird und

weiters sei α_i der Winkel zwischen V_i und V_{i+1}. Dann gilt

$$V_i \times V_{i+1} = N_i \left\|V_i\right\| \left\|V_{i+1}\right\| \sin \alpha_i$$

und somit

$$\sum_{i=0}^{2} \frac{V_i \times V_{i+1}}{\left\|V_i\right\|^2 \left\|V_{i+1}\right\|^2} = \sum_{i=0}^{2} \frac{N_i \sin \alpha_i}{\left\|V_i\right\| \left\|V_{i+1}\right\|} = cN. \tag{2.11}$$

Damit erhält man also die Gewichte $\sin \alpha_i / (\left\|V_i\right\| \left\|V_{i+1}\right\|)$ für die Normale N_i. Weiters kann man induktiv zeigen, dass

$$\sum_{i=0}^{n-1} \frac{V_i \times V_{i+1}}{\left\|V_i\right\|^2 \left\|V_{i+1}\right\|^2} = cN.$$

Diese Methode geht von einem, einer Kugel eingeschriebenen, Polyeder aus und ist deshalb, wie der Autor erwähnt, bereits für Ellipsoide ungenau. In einem Vergleichstest mit anderen üblichen Gewichtungen zeigt er jedoch, dass diese Methode auch in anderen Fällen gute — und anderen Methoden überlegene — Ergebnisse liefert.

3 D-charts: Quasi-Developable Mesh Segmentation

3.1 Der Lloyd-Algorithmus

Der Algorithmus von Stuart P. Lloyd, der im Jahre 1982 veröffentlicht wurde, kommt aus dem Bereich der Pulscodemodulation, hat aber auch in anderen Bereichen, wie dem nachfolgend Beschriebenen, mit kleinen Abwandlungen, seine Anwendung gefunden. Hier sollen nur die wichtigsten Ideen zusammengefasst werden; bei weiterem Interesse sei der Leser auf [10] verwiesen.

Ausgangspunkt ist eine Funktion, deren Typ bekannt ist und von der eine Stichprobe zur Verfügung steht. Weiters sei $Q_1, Q_2, \ldots, Q_\nu$ eine Zerlegung des Wertebereichs in Teilmengen, und $q_1, q_2, \ldots, q_\nu$ eine Menge von endlichen Werten. Außerdem sei vorausgesetzt, dass wir von den Elementen der Stichprobe nicht den exakten Funktionswert wissen, sondern lediglich welcher der Teilmengen Q_α der Wert angehört. Deswegen verwendet man den Wert q_α als Schätzwert für alle Elemente der Stichprobe, die in dem Bereich Q_α liegen, $\alpha = 1, 2, \ldots, \nu$. Aus den Parameterwerten der Stichprobenelemente und den Schätzwerten q_α wird nun eine Funktion des bekannten Typs berechnet, die als Schätzung für die Ausgangsfunktion dient. Die auf diese Weise ermittelte Funktion wird im Allgemeinen nicht mit der Ausgangsfunktion übereinstimmen und somit stellt sich die Frage, bei welcher Wahl der Q_α und zugehöriger Werte q_α der resultierende Fehler minimal ist.

Dazu stellt der Autor die Frage nach den optimalen Werten $q_1, q_2, \ldots, q_\nu$ bei festen Teilmengen $Q_1, Q_2, \ldots, Q_\nu$. Wie sich herausstellt, ist die Wahl der Massenmittelpunkte der Q_α die ideale Belegung. Ebenso wird der Frage nachgegangen, welche die bestmögliche Wahl der Q_α bei festen Werten q_α ist. Es ist naheliegend, dass unter allen Schätzwerten jener der beste ist, der dem gegebenen Wert am nächsten liegt und auch unter den Voraussetzungen in [10] zeigt sich, dass die Wahl der Halbierungspunkte aufeinanderfol-

gender q_α, als Intervallgrenzen der Q_α, optimal ist.

Der Autor schlägt nun vor, mit einer beliebigen Wahl von Mengen Q_α zu starten und die q_α als die Massezentren der Q_α zu wählen. Diese Werte werden aber im Allgemeinen nicht so liegen, dass die Grenzen der Q_α mit den Mittelpunkten aufeinanderfolgender q_α übereinstimmen. Darum lässt man im nächsten Schritt die q_α fest und ändert die Q_α so ab, dass sie die soeben genannte Eigenschaft aufweisen. Nun werden die q_α nicht mehr Massenmittelpunkte sein, was im darauffolgenden Schritt wieder angepasst wird, usw. Es zeigt sich, dass die einzelnen Schritte den Fehler der Schätzfunktion nicht erhöhen und man somit eine Folge von Fehlern erhält, die monoton fallend ist und aufgrund der Tatsache, dass der Fehler nicht negativ werden kann, muss diese Folge konvergieren. Der Grenzwert muss allerdings nicht mit dem globalen Minimum übereinstimmen.

3.2 Der D-charts-Algorithmus

Die Idee der D-charts, die von Dan Julius, Vladislav Kraevoy und Alla Sheffer im Jahre 2005 in ihrer Arbeit [1] beschrieben wurde, geht von einer Triangulierung einer Fläche aus. Diverese andere Algorithmen, etwa [9], unterteilen ein derartiges Netz in Dreiecksstreifen, welche abwickelbar sind, allerdings sind damit zumeist lange Ränder verbunden, die oftmals unerwünscht sind. Dieser Ansatz baut auf der Vorstellung der Fläche als Vereinigung von Kegelflächen auf, wobei die Kegelfläche als eine Fläche betrachtet wird, für die der Winkel zwischen den Flächennormalen und einer gegebenen Achse konstant ist. Damit sind auch die Spezialfälle des Zylinders und der Ebene inkludiert. Die Grundidee des Algorithmus ist, eine vom Anwender vorgegebene Anzahl von Bereichen zu initialisieren, welche anschließend, unter Bedachtnahme auf die Bedingung für eine Kegelfläche und der Einhaltung gewisser Schwellenwerte, mit angrenzenden Dreiecken erweitert werden. Man kann bei diesem Algorithmus natürlich nicht erwarten, dass man als Produkt eine abwickelbare Fläche erhält, aber die festgelegten Grenzwerte garantieren, dass die resultierenden Bereiche annähernd abwickelbar sind.

Nachdem die Anzahl der Bereiche festgelegt wurde, wird im Anschluss für jeden Bereich ein Startdreieck gewählt, wobei dies laut den Autoren beliebig gewählt werden kann, weil der weitere Algorithmus die Bereiche verrückt, doch kann durch Maximierung der Distanz zwischen den einzelnen Startdreicken die Zeit reduziert werden und ein besseres Ergebnis erzielt werden. Daher sollte ein „farthest point algorithm“ verwendet werden, um die Initialisierung zu optimieren. Anschließend wird für jeden Bereich C ein Paar

$\langle N_C, \theta_C \rangle$ gesucht, wobei N_C ein Einheitsvektor in Richtung der Kegelachse und θ_C der konstante Winkel zwischen Flächennormale und der Kegelachse ist. Dazu wird für jeden

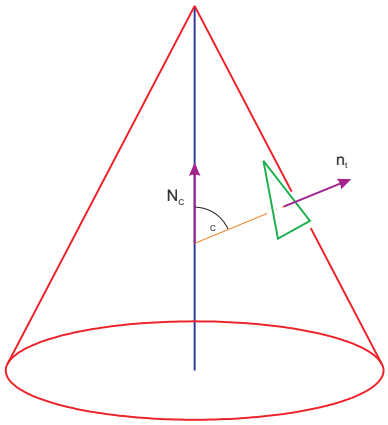

Abb. 3.1: Repräsentant

Eckpunkt des Startdreiecks die Menge der angrenzenden Dreiecke betrachtet und hierzu ein Repräsentantenpaar $\langle N_C, \theta_C \rangle$, in Abhängigkeit der Gesamtfläche A_C des Bereichs C, der Fläche A_t des t-ten Dreiecks und dem sogenannten „fitting error“ $F(C,t)$, welcher in (3.2) definiert ist, durch Lösen des folgenden Optimierungsproblems mit Nebenbedingung berechnet:

$$\min_{N_C, \theta_C} \frac{1}{A_C} \sum_{t \in C} A_t F(C,t), \quad \text{NB: } \|N_C\|^2 = 1. \tag{3.1}$$

Der hierin verwendete Ausdruck

$$F(C,t) = (N_C \cdot n_t - \cos\theta_C)^2 \tag{3.2}$$

für den „fitting error“ ist ein Maß für die Abweichung des untersuchten Dreiecks, mit der Normalen n_t, von der approximierenden Kegelfläche des Bereichs. Schließlich wird jenes Paar $\langle N_C, \theta_C \rangle$ verwendet, welches den niedrigsten fitting error für das Startdreieck hervorbringt.

Danach werden die nun initialisierten Bereiche durch angrenzende Dreiecke vergrößert, sofern diese Nachbarn dafür geeignet sind. Diese Eignung ist durch eine Kostenfunktion (siehe (3.5)) bestimmt, die folgende Gesichtspunkte beinhaltet:

- fitting error $F(C,t)$ des betrachteten Dreiecks,

- bleibt die neue Menge relativ kompakt bzw. rund,
- wie verändert sich der Rand des Bereichs.

Um dafür zu sorgen, dass die derart erzeugten Bereiche nicht zu lange Streifen bilden, berechnet man als Teil der Kostenfunktion den Wert der Funktion

$$C(C,t) = \pi \frac{D(S_C,t)^2}{A_C}, \tag{3.3}$$

wobei S_C das Startdreieck bezeichnet und $D(S_C,t)$ die kürzeste Distanz innerhalb des Bereiches zwischen dem Startdreieck und dem betrachteten Dreieck t. Hierdurch werden Dreiecke, die nahe am Startdreieck liegen mit einem niedrigen Wert bedacht und Dreiecke mit großer Entfernung durch einen dementsprechend höheren Wert benachteiligt. Der letzte Teil der Kostenfunktion besteht aus der Funktion

$$P(C,t) = \frac{l_{out}(C,t)}{l_{in}(C,t)} \tag{3.4}$$

und berechnet das Verhältnis der Länge der Dreieckskanten von t, die außerhalb des Bereichs liegen zur Länge jener innerhalb, was somit die Schaffung von kurzen Rändern fördert.

Die Zusammensetzung dieser drei Faktoren ergibt die Kostenfunktion

$$Cost(C,t) = F(C,t)^{\alpha} C(C,t)^{\beta} P(C,t)^{\gamma} \tag{3.5}$$

deren Schwerpunktsetzung mit den Parametern α, β und γ gesteuert werden kann.

Nun wird jenes Dreieck mit dem niedrigsten Wert der Kostenfunktion zum entsprechenden Bereich hinzugefügt, sofern der Wert des fitting errors unterhalb einer Schranke F_{max} liegt und es nicht bereits zu einem anderen Bereich hinzugefügt wurde. Sollte es zumindest eine dieser Bedingungen nicht erfüllen, so wird es ausgelassen und das nächste Dreieck in der Warteschlange, falls es besagten Bedingungen genügt, dem zugehörigen Bereich hinzugefügt. Mit der Zuweisung eines Dreiecks werden die jeweiligen Nachbardreiecke dieses Dreiecks in die Warteschlange aufgenommen und wie zuvor verfahren. Die Zuordnung eines Dreiecks verändert jedoch die Parameter in der Kostenfunktion eines Bereiches und die Werte der jeweiligen Dreiecke in der Warteschlange müssen daher aktualisiert werden.

Nachdem dieser Prozess abgeschlossen ist, kommt die Idee des Lloyd-Algorithmus in einer abgewandelten Form zu tragen: Die Repräsentanten werden neu ermittelt und ein neues Startdreieck wird gewählt, welches kleinen fitting error und Lage in der Nähe des Zentrums des Bereichs haben sollte. Diese Prozedur wird solange wiederholt, bis nur noch ein geringer Prozentsatz der Dreiecke neu zugeordnet werden oder eine vorgegebene Schranke an Durchläufen erreicht ist.

Einige Dreiecke werden nach Beendigung des Algorithmus noch keinem Bereich zugeordnet sein. Diese werden nun gruppiert und je nach Größe entweder durch einen weiteren Wachstumsprozess, ohne Schranke für den fitting error, einem bestehenden Bereich zugeordnet, oder im Falle größerer Löcher, werden diese durch die Generierung neuer Bereiche gefüllt.

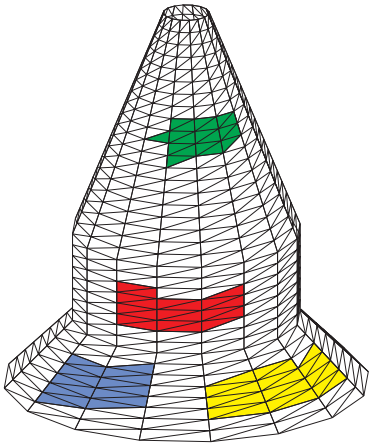

Abb. 3.2: Regionen im Wachstumsprozess

Anschließend werden die Ränder der Regionen nachbearbeitet, wobei das Ziel ist, dass keine ausgefransten Grenzen sondern möglichst glatte Begrenzungen entstehen. Dies wird mittels Definition eines Grenzbereichs initiiert, in welchem die kürzesten Verbindungen zwischen Randteilen bestimmt werden. In einem weiteren Schritt der Nachbearbeitung wird untersucht, ob benachbarte Regionen zu einem einzigen annähernd abwickelbaren Bereich zusammengefasst werden können. Abschließend werden einige Einschnitte in die Bereiche definiert, um Spannungen, die aufgrund der Hinzunahme von Dreiecken mit hohem fitting error entstanden sind, auszugleichen.

4 Mesh Segmentation Driven by Gaussian Curvature

Hitoshi Yamauchi et al. präsentieren in ihrer Arbeit [7] einen Segmentierungsalgorithmus, der auf der Idee der Gaußschen Krümmung und der Tatsache, dass eine Fläche für die $K \equiv 0$ gilt, abwickelbar ist, basiert. Da die Autoren davon ausgehen, dass es nicht möglich sein wird, Bereiche (charts) mit Gauß Krümmung null zu erzeugen, versuchen sie den Fehler gleichmäßig auf die einzelnen Regionen aufzuteilen. Dazu führen sie den Fehler ε für den i-ten Bereich c_i

$$\varepsilon(c_i) = \int_{c_i} |K| \cdot dA \tag{4.1}$$

ein. Sei A der Flächeninhalt eines kleinen Bereiches auf einer Fläche. Wenden wir auf diesen die Gauß-Abbildung an, so erhalten wir eine Region auf der Einheitskugel, deren Flächeninhalt mit A_G bezeichnet werden soll. Wie in Satz 1.2 gezeigt wurde, konvergiert der Quotient A_G/A für $A \to 0$ gegen den Betrag der Gaußschen Krümmung, also $|K| = dA_G/dA$ und zusammen mit (4.1) liefert das

$$\varepsilon(c_i) = \int_{c_i} |dA_G| \,. \tag{4.2}$$

Das zeigt, wie die Autoren bemerken, dass also der Fehler der Region, das Integral des Absolutwertes der Fläche auf der Einheitskugel — in [7] Gaußfläche genannt — ist. Somit erhält man die gewünschte Segmentierung durch gleichmäßiges Aufteilen der Gaußfläche auf die einzelnen Bereiche, was durch Minimierung der Standardabweichung des Fehlers, also

$$\min \sqrt{\sum_i (\overline{\varepsilon} - \varepsilon(c_i))^2} \tag{4.3}$$

erreicht werden kann, wobei $\overline{\varepsilon}$ der durchschnittliche Fehler pro Bereich ist ($\overline{\varepsilon} = \frac{1}{n} \sum_i \varepsilon(c_i)$).

Die Autoren weisen auf den Vorteil, den die Gaußfläche gegenüber der Gaußschen Krümmung

hat hin, denn zweiteres Merkmal ist numerisch instabil und somit bietet sich der einfachere Ansatz über die Gaußfläche an. Während die Werte der Gaußschen Krümmung in manchen Fällen einen sehr großen Bereich umfassen, ist die Gaußfläche durch das Intervall $[0, 4\pi]$ beschränkt.

Zunächst gilt es nun spitze Kanten und Ecken zu diagnostizieren, was durch Einführung einer Schranke erfolgt. Jede Kante, bei der die Normalen der angrenzenden Polygone einen Winkel miteinander einschließen, der größer als die vorgegebene Schranke ist, wird als spitze Kante markiert. Anschließend werden die Normalen in den Datenpunkten geschätzt, wobei besonderes Augenmerk auf den spitzen Elementen liegt. Die Umgebung eines Datenpunktes wird durch die spitzen Kanten in glatte Regionen unterteilt und zu jeder solchen Region wird eine Normale geschätzt und mit einem Gewicht versehen, welches dem Ansatz in [14] folgt und welcher in Kapitel 2.2 beschrieben wurde. Anschließend kann für jedes Element — Punkte, Kanten, Dreiecke — eines Netzes, die Gaußfläche, welche mit ΔA_G bezeichnet wird, berechnet werden (siehe Abb. 4.1).

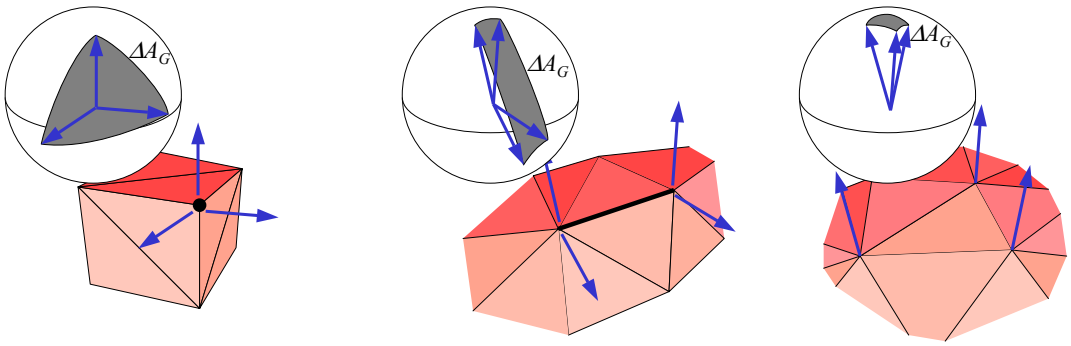

Abb. 4.1: Ermittlung der Gaußfläche (ΔA_G) der verschiedenen Netzelemente. (l) Punkt ΔA_G: Die Normalen (blaue Pfeile) abgebildet auf die Einheitssphäre liefern $\Delta A_G > 0$. (m) Kante ΔA_G: Gaußfläche einer scharfen Kante. (r) Dreieck ΔA_G. [7]

Für Punkte und Kanten, die keine Eckpunkte bzw. scharfen Kanten sind, gibt es pro Datenpunkt nur eine Normale und daher ist ihre Gaußfläche null. Weiters werden die Gaußflächen von Kanten und Punkten erst zum Bereich hinzugerechnet, wenn das ganze Element in dieser Region liegt.

Um die Gaußfläche explizit zu berechnen, benötigt man Formeln aus der sphärischen Geometrie, die wir hier kurz erläutern wollen. Da es für unsere Anwendungen ausreicht

Flächen auf der Einheitskugel zu betrachten, wollen wir uns auf diese beschränken.

Es gilt, dass die Gesamtfläche der Einheitskugel 4π beträgt. Daher ist es naheliegend die Fläche des sphärischen Zweiecks mit Innenwinkel α wie folgt zu berechnen

$$A_Z = \frac{4\pi}{2\pi}\alpha = 2\alpha. \tag{4.4}$$

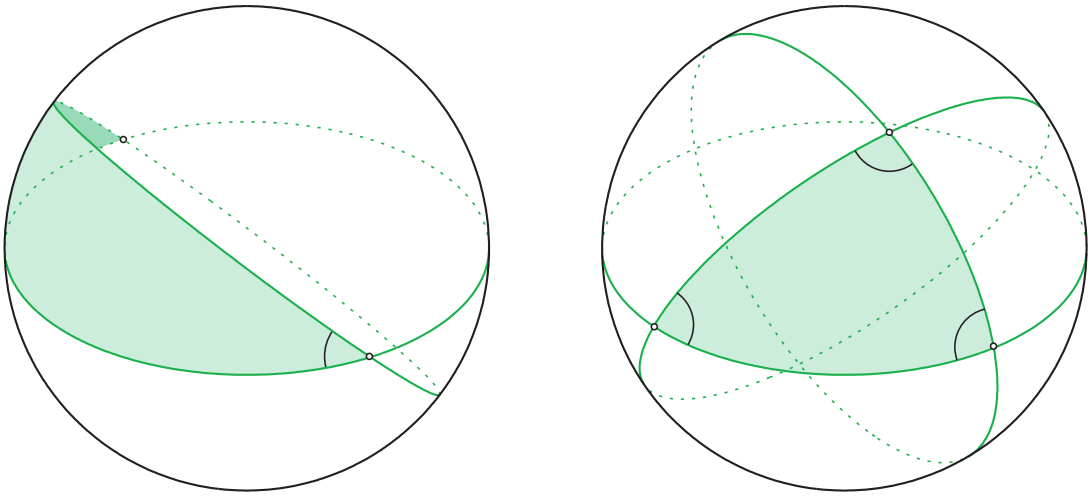

Abb. 4.2: l: sphärisches Zweieck; r: sphärisches Dreieck

Sei A_α der Flächeninhalt des Zweiecks mit Innenwinkel α und analog für andere Winkel, dann gilt für den Flächeninhalt Δ eines sphärischen Dreiecks mit Innenwinkel α, β, γ

$$2\pi = (A_\alpha - \Delta) + (A_\beta - \Delta) + (A_\gamma - \Delta) + \Delta = 2\alpha + 2\beta + 2\gamma - 2\Delta$$

und somit

$$\Delta = \alpha + \beta + \gamma - \pi. \tag{4.5}$$

Die entscheidende, neue Idee, die die Autoren nun anwenden, ist das sogenannte *chart flooding* (*t-flooding*). Hierbei geht es darum, erstens vernünftige Grenzen für die Regionen zu erhalten, und zweitens die Gaußfläche gleichmäßig auf die Bereiche aufzuteilen. Da die Gaußfläche nicht gleichmäßig über die Fläche verteilt ist, führen sie eine Zeitkoordinate t ein (*t-flooding*), um den Flutungsprozess zu parametrisieren. Weiters definieren sie eine beliebige totale Flutzeit t_{total} und bennenen die Gesamtheit der Gaußfläche des Netzes mit $A_{G,total}$ und die Anzahl der Regionen mit k. Ziel ist es nun, dass jeder Bereich mit

konstantem Zufluss α an Gaußfläche, während der Wachstumsphase, vergrößert wird, also dass gilt

$$\alpha := \frac{dA_G}{dt} = const = \frac{A_{G,total}}{k \cdot t_{total}}. \tag{4.6}$$

Bei den typischen Methoden, bei denen die Regionen mit konstanter Geschwindigkeit ihrer Begrenzungen wachsen, berechnet sich, wie die Autoren herausstreichen, der Flächenzufluss bei Zuflussgeschwindigkeit v, durch die Formel $v \cdot p$, wobei p der Umfang des Bereichs ist. Somit hängt der Flächenzufluss insbesondere vom Umfang p ab. Yamauchi et al. legen die Wachstumsgeschwindigkeit v_i der i-ten Region zunächst durch $v_i = \alpha/p_i$ fest, wobei p_i der Umfang des i-ten Bereichs unter der Gaußabbildung ist. Nachdem sie allerdings die Geschwindigkeit vom Umfang lösen wollen, definieren sie für jede Kante E_j am Rand der Region einen lokalen Zufluss α_j.

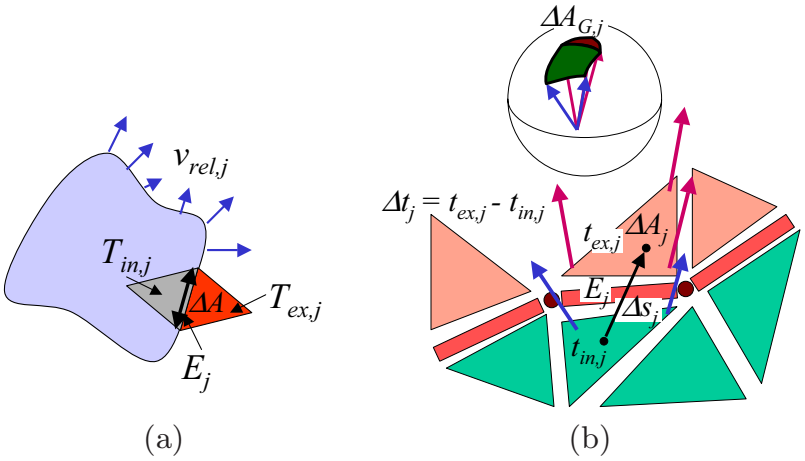

Abb. 4.3: Wachstum eines Bereiches. (a) Zwei Dreiecke $T_{in,j}, T_{ex,j}$, die eine Kante des Randes gemein haben; unterschiedliche relative Geschwindigkeiten $v_{rel,j}$ entlang des Randes. (b) Das Dreieck $T_{ex,j}$, die Kante E_j und dessen Knoten tragen zu $\Delta A_{G,j}$ bei. Die Kugel oben zeigt die Situation unter der Gaußabbildung. [7]

Nehmen wir nun an, dass das bereits in der Region liegende Dreieck $T_{in,j}$, welches die Kante E_j enthält, zum Zeitpunkt $t_{in,j}$ hinzugekommen ist und dass das außerhalb liegende Dreieck $T_{ex,j}$ zum Zeitpunkt $t_{ex,j}$ hinzukommen wird. Die Zeit Δt_j zwischen den beiden Zuwächsen, die E_j betreffen, ist $t_{ex,j} - t_{in,j}$. Sobald das Dreieck $T_{ex,j}$ aufgenommen wird, kommt auch die zugehörige Gaußfläche und ebenso jene der Kante E_j hinzu. Sind mit dem Dreieck $T_{ex,j}$ nun alle Dreiecke um einen gewissen Punkt, in der Region,

so wird auch die Gaußfläche des Punktes aufgenommen. Summiert man nun diese neu hinzugekommenen Gaußflächen, so erhält man den Wert $\Delta A_{G,j}$ und daraus den lokalen Zufluss $\Delta A_{G,j}/\Delta t_j$.

Fassen wir alle lokalen Zuflüsse zusammen, so erhalten wir mit $B_i = \{E_j | E_j \in \partial c_i\}$, also der Menge aller Kanten des Randes der Region c_i, den Ausdruck

$$\alpha = \sum_{j \in B_i} \alpha_j = \sum_j \frac{\Delta A_{G,j}}{\Delta t_j}, \tag{4.7}$$

wobei nur die Δt_j bzw. die Zeiten der Zuwächse $t_{ex,j}$ unbekannt sind. Die Autoren weisen darauf hin, dass es für einen guten Flutungsprozess essenziell ist, dass der Zeitpunkt, zu dem ein Dreieck einem Bereich hinzugefügt wird, berechnet wird, sobald der Rand der Region das Dreieck erreicht. Außerdem ist bei ihrem Ansatz dieser Zeitpunkt abhängig von der Entfernung des Dreiecks zum Startdreieck der Region und vom Umfang des Bereichs. In Gleichung (4.7) sind bis auf ein Δt_j alle übrigen frei wählbar und dies gibt den Wissenschaftern die Möglichkeit, relative Wachstumsgeschwindigkeiten $v_{rel,j}$ entlang des Randes einzuführen. Diese Geschwindigkeiten definieren sie als

$$v_{rel,j} = \frac{1}{f(|K_j|)}, \tag{4.8}$$

wobei der Absolutbetrag der lokalen Gaußschen Krümmung $|K_j|$ durch den Ausdruck $|K_j| = \Delta A_{G,j}/\Delta A_j$ geschätzt wird und ΔA_j die Fläche des Dreiecks $T_{ex,j}$ ist. Die Funktion f kann frei gewählt werden. Der Vorschlag der Autoren lautet

$$f(x) = (x + \varepsilon)^p \text{ mit } p \in [1, 6], \tag{4.9}$$

wobei der Wert ε verhindern soll, dass die Geschwindigkeit in abwickelbaren Regionen unendlich wird. Dieses ε wählten sie im Bereich von 10^{-3} und 10^{-6} der maximalen Gaußschen Krümmung $\max_j |K_j|$ unter allen Dreiecken, Kanten und Punkten, wobei für die letzteren Fälle die kleinste Fläche der Nachbardreiecke verwendet wurde.

Um aus den relativen Geschwindigkeiten $v_{rel,j}$ die echten, lokalen Geschwindigkeiten v_j zu berechnen, führen die Autoren Faktoren λ_i für jeden Bereich ein und definieren $\Delta s_j = \left\| c_{T_{in,j}} - c_{T_{ex,j}} \right\|$ als den Abstand zwischen Dreiecken mit Schwerpunkten c_T. Nun legen sie v_j durch den Ausdruck $v_j = \frac{1}{\lambda_i} v_{rel,j}$ fest und darauf aufbauend, definieren sie die frei wählbaren Zeitverzögerungen als $\Delta t_j = \Delta s_j / v_j$.

Bleibt also noch die Festsetzung der λ_i. Dazu setzt man die Ausdrücke für Δt_j in die Gleichung (4.7) ein und löst nach λ_i. Dieses λ_i ist jedoch, über den zeitlichen Verlauf des Algorithmus gesehen, variabel, da sich der Rand der Regionen ändert und somit ändern sich auch die Ausdrücke $\Delta t_j(\lambda_i)$. Um dieses Problem zu umgehen, legen sie fest, dass die Änderung von λ_i zwar die Ordnung der Aufnahme zwischen den einzelnen Regionen ändert, aber nicht die Ordnung der Aufnahmezeitpunkte innerhalb der i-ten Region.

Die Autoren verwenden nun eine Liste der verschiedenen Bereiche, die nach bereits angehäufter Gaußfläche sortiert ist und für jeden Bereich eine weitere Liste, die die angrenzenden Dreiecke, in der Reihenfolge ihrer Aufnahmezeit enthalten. Als Vorausschau addieren sie den Wert ΔA_G des Dreiecks, das als nächstes aufgenommen werden soll, zu dem Wert in der ersten Liste. Innerhalb der sekundären Listen mit den Dreiecken verwenden sie lokale Zeitskalen mit $t_i = t/\lambda_i$, wodurch sie die Δt_j in der Form

$$\Delta t_j = \Delta s_j \cdot f\left(|K_j|\right) \tag{4.10}$$

ausdrücken können und somit die Unabhängigkeit von λ_i erreichen. Durch diese Methode erreichen sie eine gleichmäßige Flutung der Bereiche.

Nachdem der Ablauf des Algorithmus nun erläutert wurde, bleibt noch die Frage nach der Initialisierungsphase. Die Autoren schreiben, dass sie den Startpunkt für den ersten Bereich beliebig wählen, dann die Region wachsen lassen und anschließend einen weiteren Bereich hinzufügen, der weitestmöglich von den bisherigen Bereichen entfernt ist. Nach Abschluss jeder t-flooding-Phase ermitteln sie jenen Punkt, der die größte Entfernung vom Rand der Region hat. Dazu verwenden sie die Dreiecke und lassen neben den euklidischen Distanzen auch noch die Δt_j als Gewichte einfließen, was dafür sorgt, dass die einzelnen Bereiche nicht zu nahe aneinander liegen. Nachdem die vorgegebene Anzahl von Regionen erreicht wurde, wird deren Positionierung optimiert, wozu unter anderem die Idee des Lloyd-Algorithmus (vgl. 3.1) verwendet wird. Dieser Prozess endet, sobald der berechnete Punkt im Zentrum innerhalb einer Schleife wechselt, eine Höchstanzahl an Iterationsschritten erreicht wurde oder die Verteilung der Gaußfläche in den letzten Schritten nicht verbessert werden konnte.

Wie die Wissenschafter unterstreichen, tritt häufig der Umstand ein, dass die einzelnen Bereiche schon in einer frühen Phase aneinander grenzen und somit ein optimaler Ablauf des Algorithmus behindert wird. Die Verlagerung des Startpunktes reicht nicht immer

aus, um dieses Problem zu beheben. Deshalb führen sie eine zusätzliche, ausgleichende Gewichtung ein, genannt *Gauss area offset* $A_{G,off,i}$, die das schnelle Wachstum von Regionen bremsen soll. Die Gewichtung für die Reihenfolge der Bereiche in der ersten Liste setzt sich nun aus der Gaußfläche und dem neu eingeführten Gewicht zusammen. Zu Beginn hat dieses Zusatzgewicht, welches konstant während jeder t-flooding-Phase ist, den Wert null und danach wird der Wert, nach der folgenden Regel, neu berechnet:

$$A_{G,off,i} := A_{G,off,i} + \beta(A_{G,off,i} - A_{G,off,min}), \quad \beta \in (0,1]. \tag{4.11}$$

$A_{G,off,min}$ bezeichnet hierbei die Gaußfläche der kleinsten Region.

5 Flächenapproximation mittels abwickelbarer Flächen

Die Gruppe Pottmann et al. versuchen in ihrer Arbeit [6], angeregt durch den Schiffsbau, Flächen durch Kegel- und Zylinderflächen zu approximieren. Ihr Ausgangspunkt ist die Tatsache, dass es zu jeder Erzeugenden einer abwickelbaren Fläche F einen Drehkegel gibt, der F von zweiter Ordnung berührt. Zylinder und Ebenen werden dabei als Spezialfälle inkludiert. Drehkegel werden hierbei als Familie unter der Gruppe der Schraubflächen betrachtet, weil diese Sichtweise zu Vereinfachungen im späteren Verlauf des Algorithmus führt.

Ausgangspunkt des beschriebenen Algorithmus ist eine Punktmenge, zu der eine Schraubung gefunden werden soll. Um zu erklären, wie man diese Schraubung charakterisieren kann, betrachten wir zunächst eine allgemeine Bewegung β im dreidimensionalen Raum. Wir denken uns dabei zwei Systeme im $\mathbb{R}^3$, nämlich ein festes System Σ_0, das sogenannte Rastsystem, und ein bewegliches System Σ, das Gangsystem. Seien $\mathbf{u}_0(t)$ die Rastkoordinaten des Ursprungs des Gangsystems zum Zeitpunkt t. Dann hat ein fester Punkt $\mathbf{x}$

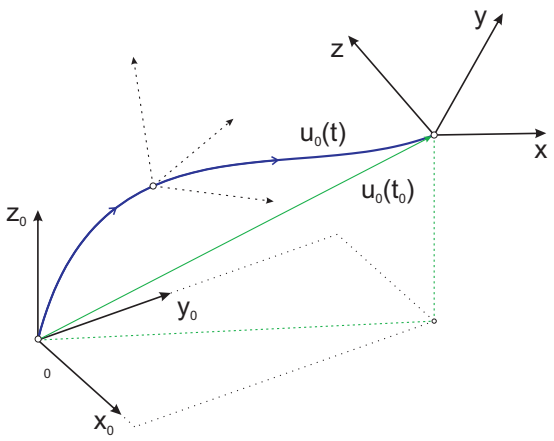

Abb. 5.1: Rast- und Gangsystem

des Gangsystems unter der Bewegung β die Rastkoordinaten

$$\mathbf{x}_0(t) = A(t) \cdot \mathbf{x} + \mathbf{u}_0(t), \tag{5.1}$$

wobei $A(t)$ eine orthogonale Matrix ist, die die Bewegung β beschreibt. Durch Differenziation dieser Gleichung, erhalten wir einen Ausdruck für die Momentangeschwindigkeit

$$\mathbf{v}(\mathbf{x}) = \dot{\mathbf{x}}_0(t) = \dot{A}(t) \cdot \mathbf{x} + \dot{\mathbf{u}}_0(t), \tag{5.2}$$

welche durch Ersetzen von $\mathbf{x}$ durch den Ausdruck $A^T(\mathbf{x}_0 - \mathbf{u}_0)$, welcher aus (5.1) folgt, zu der Darstellung

$$\mathbf{v}(\mathbf{x}) = \dot{A}A^T\mathbf{x}_0 - \dot{A}A^T\mathbf{u}_0 + \dot{\mathbf{u}}_0 \tag{5.3}$$

führt. Da die Matrix $A(t)$ orthogonal ist, d.h. $AA^T = E_3$, folgt daraus durch Differentiation $\dot{A}A^T + A\dot{A}^T = C + C^T = 0$, dass die Matrix $C = \dot{A}A^T$ schiefsymmetrisch ist, also von der Gestalt

$$C = \begin{pmatrix} 0 & -c_3 & c_2 \\ c_3 & 0 & -c_1 \\ -c_2 & c_1 & 0 \end{pmatrix}.$$

Wie man leicht einsieht, gilt mit $\mathbf{c} = (c_1, c_2, c_3)$, dass $C \cdot \mathbf{x} = \mathbf{c} \times \mathbf{x}$ für alle $\mathbf{x}$. Fassen wir nun noch $\dot{\mathbf{u}}_0 - \dot{A}A^T\mathbf{u}_0$ zu $\overline{\mathbf{c}}$ zusammen, so können wir die Gleichung (5.3) in der Form

$$\mathbf{v}(\mathbf{x}) = \overline{\mathbf{c}} + \mathbf{c} \times \mathbf{x} \tag{5.4}$$

schreiben. Somit können wir das Geschwindigkeitsvektorfeld einer Bewegung durch das Paar $(\mathbf{c}, \overline{\mathbf{c}}) \in \mathbb{R}^6$ beschreiben. Die gesuchte Bewegung ist im allgmeinen eine Schraubung, aber als Grenzfälle sind auch eine Drehung oder eine Translation möglich. Die Art der Bewegung kann man aus dem Paar $(\mathbf{c}, \overline{\mathbf{c}})$ ablesen und zwar ergibt sich,

- wenn $\mathbf{c} = 0$ eine Translation parallel zu $\overline{\mathbf{c}}$
- wenn $\mathbf{c} \neq 0$ und $\mathbf{c} \cdot \overline{\mathbf{c}} = 0$ eine Drehung
- sonst eine Schraubung.

Ausgangspunkt für die Suche einer Schraubfläche sind nun die zuvor erwähnten Datenpunkte $\mathbf{d}_i$ und die zugehörigen Flächennormalen $\mathbf{n}_i$, die bereits geschätzt wurden oder aufgrund anderer Umstände zur Verfügung stehen. Nun suchen die Autoren nach einer Bewegung, sodass die Normalen $\mathbf{n}_i$ einen Winkel γ_i nahe $\pi/2$ mit den Geschwindigkeits-

vektoren $\mathbf{v}(\mathbf{x})$ einschließen. Das bedeutet allerdings, dass die Funktion

$$H := \sum_{i=1}^{k} \cos^2 \gamma_i \tag{5.5}$$

zu minimieren ist, wobei

$$\cos \gamma_i = \frac{\mathbf{n}_i \cdot \mathbf{v}(\mathbf{d}_i)}{\|\mathbf{v}(\mathbf{d}_i)\|} = \frac{\overline{\mathbf{c}} \cdot \mathbf{n}_i + \mathbf{c} \cdot \overline{\mathbf{n}_i}}{\|\overline{\mathbf{c}} + \mathbf{c} \times \mathbf{d}_i\|} \tag{5.6}$$

mit $\overline{\mathbf{n}_i} := \mathbf{d}_i \times \mathbf{n}_i$ und somit diese Optimierungsaufgabe nicht linear ist. Deshalb verwenden die Autoren einen Ansatz der in [5] beschrieben wird: Die Abweichung einer Geraden G, welche durch ihre Plücker-Koordinaten $(\mathbf{g}, \overline{\mathbf{g}})$ gegeben ist, von einem linearen Komplex $C = (\mathbf{c}, \overline{\mathbf{c}})$, kann durch

$$m(G, C) := \frac{|\overline{\mathbf{c}} \cdot \mathbf{g} + \mathbf{c} \cdot \overline{\mathbf{g}}|}{\|\mathbf{c}\|} \qquad \text{für } \mathbf{c} \neq 0 \tag{5.7a}$$

$$m(G, C) := \frac{|\overline{\mathbf{c}} \cdot \mathbf{g}|}{\|\overline{\mathbf{c}}\|} \qquad \text{für } \mathbf{c} = 0 \tag{5.7b}$$

ausgedrückt werden. Im Falle $\mathbf{c} \neq 0$, kann man das wie folgt geometrisch interpretieren: Sei $\mathbf{x}$ ein beliebiger Punkt auf G und sei r der Abstand von $\mathbf{x}$ zur Achse der Schraubbewegung, dessen Schraubparameter p sei. Weiters sei $\pi/2 - \alpha$ der Winkel zwischen G und dem Geschwindigkeitsvektor $\mathbf{v}(\mathbf{x})$. Dann gilt

$$m(G, C) = \sqrt{r^2 + p^2} \sin \alpha. \tag{5.8}$$

Eine vernünftige Vereinfachung von (5.5) ist nun

$$F := \sum_{i=1}^{k} (r_i^2 + p^2) \cos^2 \gamma_i, \tag{5.9}$$

was mit (5.7a) und der Bedingung $\|\mathbf{c}\| = 1$ auf das Problem der Minimierung von

$$F(C) = \sum_{i=1}^{k} (\overline{\mathbf{c}} \cdot \mathbf{n}_i + \mathbf{c} \cdot \overline{\mathbf{n}_i})^2 =: C^T \cdot N \cdot C \tag{5.10}$$

unter der Nebenbedingung

$$1 = \|\mathbf{c}\|^2 =: C^T \cdot D \cdot C \tag{5.11}$$

mit $D = \operatorname{diag}(1, 1, 1, 0, 0, 0)$ führt. Unter der Einführung eines Lagrangeschen Multipli-

kators λ gilt es also das verallgemeinerte Eigenwertproblem

$$(N - \lambda D) \cdot C = 0,\ C^T \cdot D \cdot C = 1 \tag{5.12}$$

zu lösen. Seien nun λ_i die Eigenwerte und $\mathbf{v}_i$ die zugehörigen Eigenvektoren, also gelte

$$N\mathbf{v}_i = \lambda_i D\mathbf{v}_i.$$

Nach Normierung der $\mathbf{v}_i$ im Sinne von D zu $\overline{\mathbf{v}_i} = \frac{1}{\sqrt{\mathbf{v}_i^T D\mathbf{v}_i}}\mathbf{v}_i$ gilt

$$\overline{\mathbf{v}_i}^T N\overline{\mathbf{v}_i} = \overline{\mathbf{v}_i}^T(\lambda_i D\overline{\mathbf{v}_i}) = \lambda_i(\overline{\mathbf{v}_i}^T D\overline{\mathbf{v}_i}) = \lambda_i. \tag{5.13}$$

Somit wird die Funktion (5.9) vom Eigenvektor zum kleinsten Eigenwert minimiert.

Die Achse A und der Schraubparameter p lassen sich nun nach (1.18) bestimmen. Die Autoren raten jedoch, falls der Schraubparameter p sehr klein ist im Vergleich zur Größe des Objekts, dieses durch eine Drehfläche zu approximieren. Somit kommt zur Nebenbedingung (5.11) noch die Bedingung

$$0 = \mathbf{c} \cdot \overline{\mathbf{c}} =: C^T \cdot K \cdot C \tag{5.14}$$

und das Problem kann nun, mit Hilfe zweier Lagrangescher Multiplikatoren λ und μ folgendermaßen angeschrieben werden:

$$(N - \lambda D - \mu K) \cdot C = 0,\ C^T \cdot D \cdot C = 1,\ C^T \cdot K \cdot C = 0. \tag{5.15}$$

Für die konkrete Berechnung der Lösung, welche man aus der Gleichung

$$\det(N - \lambda D - \mu K) = 0 \tag{5.16}$$

erhält, schlagen die Autoren einen numerischen Ansatz im Sinne einer Newton-Iteration vor.

Weiters besteht die Möglichkeit, dass die Datenpunkte durch einen Drehkegel approximiert werden können. Sollte dies der Fall sein, so würden die geschätzten Normalen in den Datenpunkten annähernd einen konstanten Winkel, mit der noch unbekannten Achse haben. Deshalb betrachtet die Arbeitsgruppe Pottmann et al. in diesem Fall die

quadratische Form

$$F_q(C,g) := \sum_{i=1}^{k} (\overline{\mathbf{c}} \cdot \mathbf{n}_i + \mathbf{c} \cdot \overline{\mathbf{n}_i})^2 + q \sum_{i=1}^{k} (\mathbf{c} \cdot \mathbf{n}_i - g)^2 =: X^T \cdot N_q \cdot X, \tag{5.17}$$

wobei g der Cosinus des unbekannten Winkels und $X = (C, g)$ ist. Der Parameter $q > 0$ dient als Gewichtung des Einflusses der zweiten Summe und muss vor der Minimierung gewählt werden; die Autoren empfehlen einen Wert von 0.5. Dies führt somit wieder auf die Minimierung einer Funktion mit zwei Nebenbedingungen, was wir, erneut mit zwei Lagrangeschen Multiplikatoren λ und μ, anschreiben können in der Form

$$(N_q - \lambda D' - \mu K') \cdot X = 0,\ X^T \cdot D' \cdot X = 1,\ X^T \cdot K' \cdot X = 0, \tag{5.18}$$

wobei $D' = \mathrm{diag}(1, 1, 1, 0, 0, 0, 0)$ und

$$K' = \left(\begin{array}{cccccc|c} & & & \frac{1}{2} & & & \\ & & & & \frac{1}{2} & & \\ & & & & & \frac{1}{2} & \\ \frac{1}{2} & & & & & & \\ & \frac{1}{2} & & & & & \\ & & \frac{1}{2} & & & & \\ \hline & & & & & & 0 \end{array}\right).$$

Die konkrete Berechnung erfolgt wieder numerisch.

Nachdem die Achse berechnet wurde, wählt man eine beliebige Halbebene durch die Achse und dreht alle Datenpunkte in diese Ebene. Daraufhin berechnet man eine Ausgleichsgerade e durch diese Punktwolke. Die Gerade e — eine Erzeugende des Kegels — ermöglicht uns auch eine Charakterisierung der Approximationsgenauigkeit. Dazu berechnet man die Summe der quadratischen Abstände der gedrehten Punkte von e.

Im Allgemeinen wird jedoch das Problem auftreten, dass die gegebenen Daten nicht durch eine einzige Fläche optimal approximiert werden können. Daher gilt es die gegebenen Daten so zu unterteilen, dass jeder Teil gut durch einen Drehkegel angenähert werden kann. Angenommen wir haben eine Region R mit einer passenden Approximation gefunden, dann werden Punkte anderer Regionen als Ausreißer klassifiziert. Da jedoch zu Beginn keine adäquate Segmentierung zur Verfügung steht, modifizieren die Autoren ihren zuvor beschriebenen Algorithmus, um in einem späteren Schritt die Suche nach

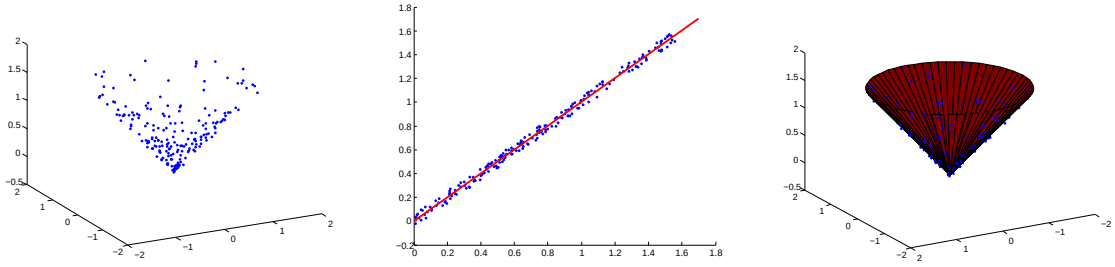

Abb. 5.2: Approximation einer Punktwolke

einer geeigneten Unterteilung zu vereinfachen. Dazu führen sie Gewichte w_i ein und minimieren, im Rahmen eines Iterationsprozesses, nun die Funktion

$$F_{q,w} = \sum_{i=1}^{k} w_i \left((\overline{\mathbf{c}} \cdot \mathbf{n}_i + \mathbf{c} \cdot \overline{\mathbf{n}_i})^2 + q \cdot (\mathbf{c} \cdot \mathbf{n}_i - g)^2 \right) =: X^T \cdot N_{q,w} \cdot X. \tag{5.19}$$

Dabei werden zu Beginn alle Gewichte w_i auf 1 gesetzt und später nach der folgenden Regel adaptiert:

$$w_i = \begin{cases} 1 & \text{wenn } |r_i/\hat{\sigma}| \leq 2.5 \\ 0 & \text{wenn } |r_i/\hat{\sigma}| > 2.5, \end{cases} \tag{5.20}$$

wobei r_i der Abstand des Datenpunktes d_i von der Erzeugenden ist und

$$\hat{\sigma} = \sqrt{\text{med } r_i^2}, \tag{5.21}$$

der Median der Werte r_i, ein Schätzer für die durchschnittliche Fehlergröße. Mit diesen neuen Gewichten wird die Lösung wiederum ermittelt. Durch Iteration erhält man schlussendlich eine gute Achse und jene Punkte, die das Gewicht 0 haben, werden als Ausreißer klassifiziert.

Bei der Approximation mit Drehkegeln kann es passieren, dass die Kegelspitze in der Datenmenge liegt, etwa wenn der betrachtete Teil von einer Tangentenfläche nahe ihrer Gratkurve kommt. Da die Existenz einer Singularität des approximierenden Kegels unerwünscht ist, muss in diesem Fall die Lösung neu berechnet werden. Dazu ermitteln Pottmann et al. zunächst die Kegelspitze, indem sie eine beliebige Halbebene durch die Achse wählen und die Punkte in dieselbige drehen. Anschließend verknüpfen sie ein kartesisches Koordinatensystem, dessen y-Achse die Kegelachse ist, mit der Halbebene. Somit sind die x-Koordinaten der Punkte zugleich die Abstände dieser Punkte zur Ach-

se. Mittels einer Intervallschachtelung mit Startintervall $[y_{min}, y_{max}]$ eruieren sie nun, ob die Kegelspitze außerhalb — die Intervallschachtelung konvergiert gegen einen der Ränder — oder innerhalb — ab einem gewissen Index enthalten die Intervalle keine der ursprünglichen Grenzen — der Datenmenge liegt. Abschließend wird die Achse mit der Methode der kleinsten Quadrate neu berechnet, unter der Nebenbedingung, dass die Kegelspitze in jener der beiden Ebenen $y = y_{min}$ bzw. $y = y_{max}$ liegen muss, welche näher an der ursprünglichen Kegelspitze liegt.

Abgesehen von Kegelflächen können auch ebene Bereiche oder solche Regionen auftreten, die gut durch eine Drehzylinderfläche angenähert werden können. Den Fall eines annähernd ebenen Bereichs kann man über die Betrachtung des Gaußschen Bildes erkennen, denn dort ergibt sich eine sehr kleine Region und im Spezialfall einer Ebene sogar ein einpunktiges Bild. Um zu erkennen, ob sich die Approximation mit einer Zylinderfläche anbietet, folgt die Gruppe dem Ansatz aus [16], der zur Erkennung einer Zylinderfläche zunächst das Gaußsche Bild der Fläche M betrachtet. Weiters legt man einen Raster über die Einheitskugel und betrachtet die Urbilder der Rasterzellen Ω_i. Den Anteil den das Urbild der Zelle Ω_i an der Gesamtfläche von M hat, weist man dem Mittelpunkt P_i von Ω_i als Gewicht zu. Anschließend ermittelt man den bestapproximierenden Großkreis zu diesem gewichteten Gaußbild, der in einer Ebene α liegt. Schließlich projiziert man die Fläche M in die Ebene α, weist erneut Gewichte zu und ermittelt eine approximierende Kurve, die zusammen mit der Normalenrichtung von α den approximierenden Zylinder festlegt.

Der eigentliche Algorithmus in [6] beginnt mit einer groben Segmentierung der Datenpunkte mit Hilfe des Gaußschen Bildes, das, sofern die der Datenmenge zugrundeliegende Fläche, annähernd abwickelbar ist, eine kurvenähnliche Region auf der Einheitskugel liefert. Zunächst wird ein zufälliger Datenpunkt $\mathbf{d}$ gewählt, der allerdings nicht zu nahe am Rand der Datenmenge liegen sollte. Zu diesem Punkt $\mathbf{d}$ wird eine umgebende Region R_0 definiert, die man als Menge aller Punkte, die einen kleineren Abstand als eine vorher definierte Schranke aufweisen, oder unter Zuhilfenahme des Gaußschen Bildes definieren kann. Nun wird ein approximierender Drehkegel Σ_0, wie zuvor beschrieben ermittelt, der auftretende Fehler jedes Punktes berechnet und jene Punkte, deren Abstände eine gewisse Fehlertoleranz übersteigen, als Ausreißer klassifiziert und von der Region R_0 entfernt. Danach werden zwei orientierte Halbebenen $\alpha_0, \overline{\alpha_o}$ durch die Kegelachse gesucht, sodass R_0 im Schnitt der beiden positiven Halbräume liegt und der Winkel zwischen den beiden Halbebenen minimal ist. Dazu wird mit der Achse A_0 des Drehkegels ein zylindrisches

Koordinatensystem (r, ϕ, z) verbunden, sodass es einen Datenpunkt mit dem Wert $\phi = 0$ gibt. Dann wird das Intervall $[a, b] \subset (-\pi, \pi)$, in dem die Polarwinkel der Datenpunkte liegen ermittelt und weiters werden die Polarwinkel der Nachbarpunkte der Region R_0 berechnet. In $[a, b]$ wird nun das größte Intervall $[e, f]$ gesucht, das keinen Polarwinkel eines Nachbarpunktes enthält. Alle Punkte im Intervall $[e, f]$ werden nun der Region R_0 zugerechnet, sofern sie zu Σ_0 passen. Mit der so erhaltenen Region R_0, geht man wieder zum Beginn der Prozedur und iteriert das Verfahren so lange, bis sich R_0 nicht mehr oder nur noch sehr wenig ändert. Nachdem R_0, Σ_0, α_0 und $\overline{\alpha_0}$ fixiert wurden, werden noch jene Punkte zu R_0 hinzugefügt, die gut zu Σ_0 passen und nahe am Rand der Region R_0 liegen.

Angenommen es gibt nun auf der negativen Seite einer Halbebene α_i Datenpunkte, die noch in keiner Region liegen, so nimmt man jene Punkte, die eine kleinere Distanz als d_1 von α_i auf der positiven Seite haben und jene, die auf der negativen Seite einen Abstand kleiner d_2 haben — wobei $d_1 < d_2$ gelten sollte, da die Punkte auf der positiven Seite nur zur Herstellung einer Verbindung dienen sollen — und ermittelt zu diesen Punkten ebenfalls einen Drehkegel und zwei begrenzende Halbebenen. Wendet man diesen Schritt auf beiden Seiten der anfangs initiierten Region an, bis alle Datenpunkte erfasst wurden, so erhält man eine Folge Σ_i von Drehflächen. Abschließend gilt es die Lücken zwischen den einzelnen Drehkegeln zu schließen und zwar derart, dass das Ergebnis abwickelbar bleibt. Die Autoren verweisen hier auf die Arbeit von Leopoldseder und Pottmann [15], die zeigen, wie man Erzeugende $\mathbf{e}_1$ von Σ_1 und $\mathbf{e}_2$ von Σ_2 finden kann, sodass ein Kegel mit G^1-Stetigkeit die Lücke füllt.

6 Approximation von Punktwolken durch abwickelbare Flächen

Ein weiterer Ansatz wird von M. Peternell beschrieben [13]. Er geht von einer Punktwolke im dreidimensionalen Raum aus und versucht zu ermitteln, ob diese Punkte oder zumindest Teilmengen davon, durch abwickelbare Flächen beschrieben werden können. Für diese Anwendung wird eine abwickelbare Fläche als Spezialfall einer Regelfläche

$$\mathbf{f}(u, v) = \mathbf{l}(u) + v\mathbf{r}(u) \tag{6.1}$$

angesehen, wobei $\mathbf{l}(u)$ die sogenannte Leitkurve ist und $\mathbf{r}(u)$ Richtkurve genannt wird. Eine derartige Fläche ist genau dann abwickelbar, wenn $\det(\dot{\mathbf{l}}, \mathbf{r}, \dot{\mathbf{r}}) = 0$ gilt und es ergeben sich die folgenden drei Elementarfälle:

- Zylinder
- Kegel
- Tangentenfläche, d. h. eine Fläche, die aus den Tangenten einer regulären Raumkurve besteht.

Die untersuchte Fläche kann eine Zusammensetzung von Teilflächen sein, wobei die Teilflächen einen der drei Typen von Elementarflächen zugeordnet werden können. Anstatt obiger Definition ist es auch möglich, besagte Elementarflächen als Hüllfläche ihrer Tangentialebenen zu betrachten. So entsteht ein Zylinder, wenn man eine Ebene um eine dazu parallele Achse dreht und ein Kegel wird durch Drehung einer Ebene um eine allgemeine Achse — aber nicht orthogonal zur Ebene — erzeugt.

6.1 Grundidee

M. Peternells Ziel ist es, aufgrund von Schätzungen der Tangentialebenen in den gegebenen Punkten, eine einparametrige Familie von Ebenen zu finden, die die geschätzten

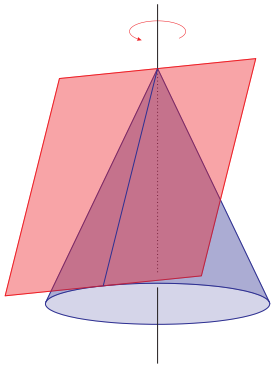

Abb. 6.1: Erzeugung eines Kegels als Hüllfläche

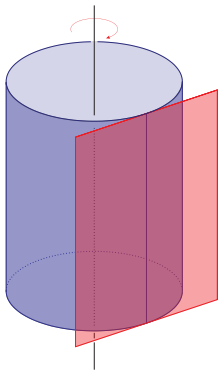

Abb. 6.2: Erzeugung eines Zylinders als Hüllfläche

Tangentialebenen approximieren. Die Tatsache, dass man eine einparametrige Familie von Ebenen erhält, garantiert, dass das Ergebnis abwickelbar ist. Die geschätzte Tangentialebene $T_{\mathbf{p}}$ in einem Punkt $\mathbf{p}$ soll jene Ebene sein, die die Abstände der benachbarten Punkte $\mathbf{q}_i$ zu $T_{\mathbf{p}}$ minimiert, also

$$T_{\mathbf{p}} = \min_{T \ni \mathbf{p}} \sum_i d(T, \mathbf{q}_i),$$

wobei d üblicherweise die l^1 oder l^2-Norm ist.

Nachdem alle Tangentialebenen geschätzt wurden, werden diesen Ebenen Punkte im $\mathbb{R}^4$ zugeordnet. Sei $\varepsilon : n_1x + n_2y + n_3z + d = 0$ mit $n_1^2 + n_2^2 + n_3^2 = 1$ eine Ebene im $\mathbb{R}^3$, so wird dieser, unter der sogenannten Blaschke-Abbildung, der Punkt (n_1, n_2, n_3, d) im $\mathbb{R}^4$ zugeordnet. All diese Punkte, die somit die Gleichung $x_1^2 + x_2^2 + x_3^2 = 1$ erfüllen, bilden den sogenannten Blaschke[1]-Zylinder.

Falls die, den gegebenen Punkten zugrundeliegende Fläche abwickelbar ist, so bilden die Bilder der Tangentialebenen eine Kurve auf dem Blaschke-Zylinder B. Durch die Schätzung der Tangentialebenen kann man keine exakte Kurve erwarten, aber die Punkte sollten in einem engen Bereich um die exakte Kurve liegen. Aus diesem Grund ist es notwendig einen näheren Blick auf die Bilder von Kegel- und Zylinderflächen — repräsentiert durch die einparametrige Familie ihrer Tangentialebenen — zu werfen, um sie später wiedererkennen zu können.

[1]Wilhelm Johann Eugen Blaschke, geboren am 13.9.1885 in Graz, gestorben am 17.3.1962 in Hamburg

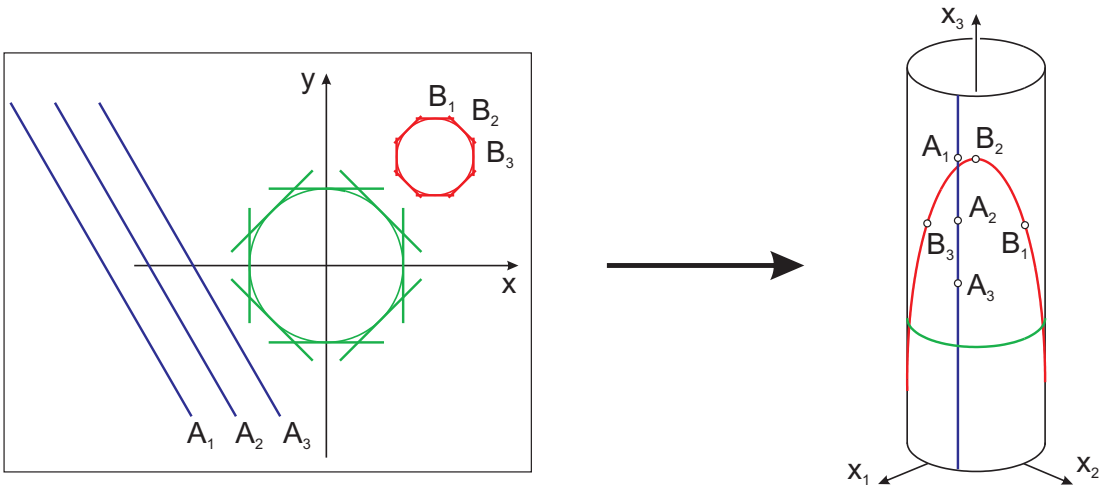

Abb. 6.3: Blaschke-Abbildung des $\mathbb{R}^2$

6.2 Blaschke-Bilder spezieller Flächen

Betrachten wir zunächst einen festen Punkt $\mathbf{p} = (p_1, p_2, p_3)$ im $\mathbb{R}^3$. Alle Ebenen durch diesen Punkt erfüllen die Bedingung

$$p_1 n_1 + p_2 n_2 + p_3 n_3 + d = 0, \tag{6.2}$$

somit liegen die Bilder dieser Ebenen auf dem Blaschke-Zylinder alle in der Hyperebene

$$H : p_1 x_1 + p_2 x_2 + p_3 x_3 + x_4 = 0. \tag{6.3}$$

Als nächstes wenden wir uns den Tangentialebenen einer orientierten Kugel mit Mittelpunkt $\mathbf{m}$ und Radius $r \in \mathbb{R}$ zu. Die zugehörigen Tangentialebenen müssen der Bedingung

$$n_1 m_1 + n_2 m_2 + n_3 m_3 + d = r \tag{6.4}$$

genügen, da sie den orientierten Abstand r vom Kugelmittelpunkt $\mathbf{m}$ haben müssen und ihre Bilder unter der Blaschke-Abbildung gehören daher der Hyperebene

$$H : m_1 x_1 + m_2 x_2 + m_3 x_3 + x_4 - r = 0 \tag{6.5}$$

an. Wenn wir nun diese Tatsache von der anderen Seite betrachten, nämlich derart, dass wir von Punkten auf dem Blaschke-Zylinder ausgehen, die in einer Hyperebene

$$H : a_0 + x_1 a_1 + x_2 a_2 + x_3 a_3 + x_4 a_4 = 0 \tag{6.6}$$

liegen, so tangieren die Urbilder dieser Punkte, falls $a_4 \neq 0$ eine Kugel mit Mittelpunkt $\mathbf{m}$ und Radius r für die gilt

$$\mathbf{m} = \frac{1}{a_4}(a_1, a_2, a_3), \qquad r = -\frac{a_0}{a_4}.$$

Für den Fall $a_0 = 0$ erhalten wir eine Bedingung analog zu (6.3) und die Ebenen haben einen gemeinsamen Punkt. Sollte $a_4 = 0$ sein, so gilt mit $\mathbf{a} = (a_1, a_2, a_3)$ und $\mathbf{n} = (n_1, n_2, n_3)$, dass $\mathbf{a} \cdot \mathbf{n} = -a_0$ also konstant ist und somit die Ebenen einen konstanten Winkel mit der Richtung $\mathbf{a}$ einschließen.

Nun können wir unsere eigentliche Intention wieder aufgreifen, nämlich die Charakterisierung von Kegel- und Zylinderflächen anhand ihrer Bilder am Blaschke-Zylinder. Im Fall einer Kegelfläche verlaufen alle Tangentialebenen durch einen gemeinsamen Punkt, was wir bereits betrachtet haben und was zu der Bedingung (6.3) führt. Die Tangentialebenen einer Zylinderfläche sind alle parallel zur einer gegebenen Richtung, die wir mit Hilfe eines Vektors $\mathbf{a}$ angeben können. Somit sind die Normalvektoren der Tangentialebenen orthogonal zur Richtung $\mathbf{a}$, also gilt $\mathbf{a} \cdot \mathbf{n} = 0$. Das Bild dieser Ebenen liegt daher in der Hyperebene

$$H : a_1 x_1 + a_2 x_2 + a_3 x_3 = 0. \tag{6.7}$$

Abschließend werfen wir noch einen Blick auf die bekannten Spezialfälle Drehzylinder und Drehkegel. Für ersteren Fall betrachten wir zwei, dem Drehzylinder eingeschriebene Kugeln, die daher denselben Radius haben. Da alle Tangentialebenen des Drehzylinders auch diese Kugeln berühren, müssen die zugehörigen Bilder in Hyperebenen der Art (6.5) und somit in deren Schnitt liegen. Haben wir andererseits eine ebene Kurve am Blaschke-Zylinder gegeben, welche sich in zwei Hyperebenen der Art (6.5), mit demselben inhomogenen Anteil r — der den Zylinderradius liefert — befindet, so liefert die Subtraktion der beiden Hyperebenengleichungen die Darstellung $\mathbf{x}(\mathbf{m_1} - \mathbf{m_2}) = 0$, wobei $\mathbf{m_1}$ und $\mathbf{m_2}$ die beiden Kugelmittelpunkte sind und deren Differenz daher die Achsenrichtung wiedergibt.

Für einen Drehkegel gilt ähnliches, allerdings haben die zwei eingeschriebenen Kugeln

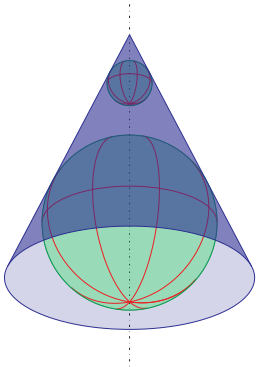

Abb. 6.4: Kegel mit eingeschriebenen Kugeln

in diesem Fall unterschiedliche Radien. Die Bilder der Tangentialebenen befinden sich daher auch wieder im Schnitt zweier Hyperebenen H_1 und H_2 der Gestalt (6.5), nur sind hier, im Gegensatz zum zuvor betrachteten Drehzylinder, die Radien und somit die inhomogenen Komponenten, unterschiedlich. Es gibt daher im Hyperebenenbüschel, das durch H_1 und H_2 bestimmt ist, eine Hyperebene, die den Ursprung enthält, deren Gleichung also homogen ist und somit die Kegelspitze festlegt. Diese Gleichung hat die Form

$$H : (x_1\ x_2\ x_3)(r_2\mathbf{m_1} - r_1\mathbf{m_2})^T + x_4(r_2 - r_1) = 0. \tag{6.8}$$

Analog zum Drehzylinder erhält man die Achsenrichtung $\mathbf{a} = \mathbf{m_1} - \mathbf{m_2}$ und zusätzlich kann man die Kegelspitze $\mathbf{s}$ und den Winkel ϕ zwischen der Achse und den Tangentialebenen folgendermaßen berechnen

$$\mathbf{s} = \frac{1}{r_2 - r_1}(r_2\mathbf{m_1} - r_1\mathbf{m_2}) \qquad \sin\phi = \frac{|r_2 - r_1|}{\|\mathbf{m_2} - \mathbf{m_1}\|}.$$

6.3 Der Algorithmus

Nun liegen, wie bereits erwähnt, für den Fall einer abwickelbaren Fläche, die Bilder der geschätzten Tangentialebenen in einem Bereich um die exakte Kurve, welche durch die exakten Tangentialebenen bestimmt wird. Peternell spricht von einer kurvenähnlichen Region auf dem Blaschke-Zylinder. Für die Überprüfung, ob eine kurvenähnliche Region vorliegt, gilt es jedoch zunächst eine Metrik auf dem Blaschke-Zylinder zu definieren, die möglichst einfach ist und gleichzeitig gute Ergebnisse liefert. Der Autor argumentiert im weiteren für die Wahl der euklidischen Metrik, die unter gewissen Einschränkungen gute

Resultate liefert. Was heißt nun, dass eine beliebige Ebene von einer festen Ebene ε den Abstand r aufweist? Für den Betrachter des vierdimensionalen Raumes bedeutet das, dass der zugehörige Punkt zugleich auf dem Blaschke-Zylinder und auf einer dreidimensionalen Kugel, deren Mittelpunkt der Bildpunkt von ε ist, liegt. Im dreidimensionalen Raum spiegelt sich das durch die Tatsache wider, dass die Ebenen Tangentialebenen einer Fläche der Klasse vier sind. Konkret bedeutet das, dass die Ebenen parallel zu ε im Abstand r sind oder dass der Winkel zwischen den Ebenen beschränkt ist (vgl. dazu auch Abb. 6.5 und Abb. 6.6). Wie bereits angedeutet wurde, liefert diese Herangehensweise nur unter gewissen Einschränkungen gute Ergebnisse und zwar im Falle, dass der Abstand der gegebene Ebene ε vom Ursprung nicht zu groß ist. Daher sollte in der Praxis die Datenmenge so skaliert werden, dass sie innerhalb des Einheitswürfels liegt.

Abb. 6.5: Ebene mit Grenzflächen

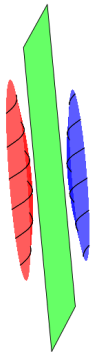

Abb. 6.6: Ebene mit Grenzflächen

Für konkrete Berechnungen auf dem Blaschke-Zylinder unterteilt der Autor dieses Objekt in Zellen. Da Schnitte mit Hyperebenen $x_4 = \text{const.}$ Kugeln mit Radius 1, wie wir sie aus dem dreidimensionalen kennen, liefern, wird zunächst eine Kugel durch Dreiecke angenähert. Die Grundidee dabei ist, dass man mit einem, der Kugel eingeschriebenen, Ikosaeder startet. Anschließend werden die Kanten der Dreiecksflächen halbiert. Die Projektion dieser Halbierungspunkte auf die Kugel liefert die Eckpunkte der verfeinerten Dreiecke. Die dabei entstehenden Dreiecke sind zwar nicht gleich groß, jedoch ist der Unterschied vernachlässigbar. Durch Iteration dieses Verfahrens erhält man nun eine zusehends bessere Approximation der Sphäre durch Dreiecke. Man vergleiche dazu Abb. 6.7, die dieses iterative Verfahren für eine Seite des Ikosaeders darstellt. Die Zerlegung des Blaschke-Zylinders erhält man mittels dreiseitigen Prismen, die aus den eben konstruierten Dreiecken und zugehöriger Höhe in x_4-Richtung entstehen.

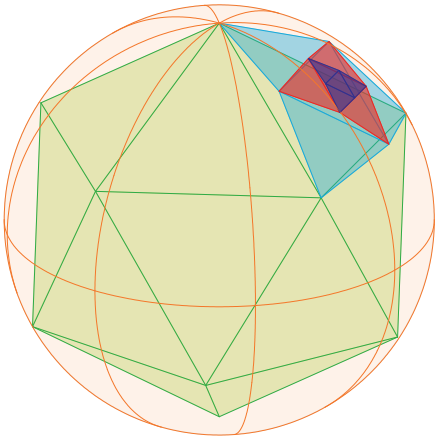

Abb. 6.7: Approximation einer Kugel durch Dreiecke

Im Folgenden geht der Autor von einer Punktmenge aus, die von einer abwickelbaren Fläche D stammt. Es ist anzunehmen, dass speziell für Punkte, die vom Randbereich der Fläche stammen, die geschätzten Tangentialebenen ungenau sein könnten und darum ist zu erwarten, dass es einige Ausreißer gibt, die es zu identifizieren gilt. Aus diesem Grund werden jene Zellen, die nur wenige Punkte beinhalten ignoriert. Nun muss die bereinigte Datenmenge $\mathbf{b}_i$ klassifiziert werden und dies geschieht, indem man zunächst eine Hauptkomponentenanalyse durchführt, wodurch man feststellen kann, ob die gegebenen Punkte annähernd in einer Hyperebene liegen. Es gilt, dass die bestapproximierende Hyperebene den Schwerpunkt der Punktwolke enthalten muss, wie man folgendermaßen einsieht:
Sei $H := \mathbf{x}^T\mathbf{n} + n_0 = 0$ die gesuchte Hyperebene. Gesucht ist also die Lösung des Minimierungsproblems mit Nebenbedingung

$$f(n_0, \mathbf{n}) = \sum_{i=1}^{N} (\mathbf{b}_i^T\mathbf{n} + n_0)^2 \to \min \quad \text{NB: } g(n_0, \mathbf{n}) = \mathbf{n}^T\mathbf{n} - 1 = 0.$$

Mit Hilfe der Lagrange-Funktion kann man dies wie folgt lösen:

$$\begin{aligned} F(n_0, \mathbf{n}, \lambda) =& f(n_0, \mathbf{n}) + \lambda g(n_0, \mathbf{n}) = \sum_{i=1}^{N} (\mathbf{b}_i^T \mathbf{n} + n_0)^2 + \lambda(\mathbf{n}^T \mathbf{n} - 1) \\ \frac{\partial F}{\partial n_0} =& 2 \sum_{i=1}^{N} (\mathbf{b}_i^T \mathbf{n} + n_0) = 0 \\ & \mathbf{n}^T \sum_{i=1}^{N} \mathbf{b}_i + N n_0 = 0 \\ & \mathbf{n}^T \left(\frac{1}{N} \sum_{i=1}^{N} \mathbf{b}_i \right) + n_0 = 0. \end{aligned}$$

Dies zeigt also die Inzidenz des Schwerpunktes der Punktwolke mit der bestapproximierenden Hyperebene. Gesucht ist also, im Koordinatensystem mit Ursprung $\mathbf{s}$, eine Hyperebene der Form

$$h_1 x_1 + h_2 x_2 + h_3 x_3 + h_4 x_4 = 0 \quad \text{mit} \quad h_1^2 + h_2^2 + h_3^2 + h_4^2 = 1,$$

sodass mit $\mathbf{q}_i = \mathbf{b}_i - \mathbf{s}$ und $\mathbf{h} = (h_1, h_2, h_3, h_4)$

$$F(\mathbf{h}) = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{q}_i^T \cdot \mathbf{h})^2 \tag{6.9}$$

minimal ist.

Seien $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \lambda_4$ die Eigenwerte der Matrix C und $\mathbf{v}_i$ die zugehörigen Eigenvektoren. Dann kann man aufgrund der Verteilung der λ_i folgende Aussagen treffen:

1. Zwei kleine Eigenwerte λ_1, λ_2 und verschiedene Koeffizienten h_{10}, h_{20}, ($|h_{10} - h_{20}| > \varepsilon$): D kann durch einen Drehkegel approximiert werden und die Kegelspitze und der Winkel zwischen Achse und Tangentialebenen kann, wie oben beschrieben, bestimmt werden.

2. Zwei kleine Eigenwerte λ_1, λ_2 aber annähernd gleiche Koeffizienten h_{10}, h_{20}, ($|h_{10} - h_{20}| \leq \varepsilon$): Die Fläche D kann durch einen Drehzylinder angenähert werden, dessen Achse und Radius wie zuvor berechnet werden können.

3. Ein kleiner Eigenwert λ_1 und kleiner Koeffizient h_{10}: Dies führt zu einem allgemeinen Kegel mit Spitze $\mathbf{v} = \frac{1}{h_{14}}(h_{11}, h_{12}, h_{13})$.

4. Ein kleiner Eigenwert λ_1 und kleine Koeffizienten h_{10} und h_{14}: Man erhält einen allgemeinen Zylinder mit Erzeugendenrichtung (h_{11}, h_{12}, h_{13}).

5. Ein kleiner Eigenwert λ_1 und kleiner Koeffizient h_{14}: D ist eine Fläche, deren Tangentialebenen einen konstanten Winkel mit einer Achse einschließen. Im Idealfall liefert dies eine Hyperebene der Form $-\gamma + a_1x_1 + a_2x_2 + a_3x_3 = 0$ mit $\|\mathbf{a}\| = 1$, wobei γ den Winkel liefert und $\mathbf{a}$ die gesuchte Achsenrichtung.

6. Ein kleiner Eigenwert λ_1 liefert eine Fläche, dessen Tangentialebenen eine Kugel berühren. Mittelpunkt und Radius der Kugel erhält man folgendermaßen:

$$\mathbf{m} = \frac{1}{h_{14}}(h_{11}, h_{12}, h_{13}), \qquad r = -\frac{h_{10}}{h_{14}}.$$

Diese Herangehensweise kann auch für die Approximation von Flächen verwendet werden, sofern diese beinahe abwickelbar sind, das bedeutet dass zumindest eine der Hauptkrümmungen betragsmäßig klein ist. Der Autor betrachtet Punktmengen von glatten, singularitätenfreien Flächen, die gleichmäßig verteilt sind und deren zugehörige Tangentialebenen eine kurvenähnliche Region auf dem Blaschke-Zylinder liefern, die injektiv parametrisierbar ist. Als erstes muss nun eine Kurve $\mathbf{c}(t)$ an die Punkte am Blaschke-Zylinder angepasst werden, d. h. dass die Kurve der Bedingung

$$c_1(t)^2 + c_2(t)^2 + c_3(t)^2 = 1 \tag{6.10}$$

genügen muss. Dazu werden für die Punkte Parameterwerte geschätzt und eine Approximation mit B-splines durchgeführt. Zum Abschluss wird diese Kurve auf den Blaschke-Zylinder projiziert um der Bedingung (6.10) zu genügen.
Ist nun eine approximierende Kurve $\mathbf{c}(t)$ auf dem Blaschke-Zylinder gefunden, so ist die approximierende Fläche D^* durch die Ebenen

$$E(t) : c_4(t) + c_1(t)x + c_2(t)y + c_3(t)z = 0 \tag{6.11}$$

festgelegt. Die Erzeugenden $L(t)$ von D^* erhält man als Schnittgeraden von $E(t) \cup \dot{E}(t)$. Der Autor nimmt nun an, dass der relevante Teil der Fläche D^* durch zwei Ebenen H_1 und H_2 begrenzt werden kann und dass die Schnittpunkte der Erzeugenden $L(t)$ mit diesen Ebenen eigentliche Punkte sind. Die Randkurven $\mathbf{f}_i(t)$ des Flächenstücks, welches von Interesse ist, berechnet man somit durch

$$\mathbf{f}_i(t) = E(t) \cup \dot{E}(t) \cup H_i, \quad i = 1, 2. \tag{6.12}$$

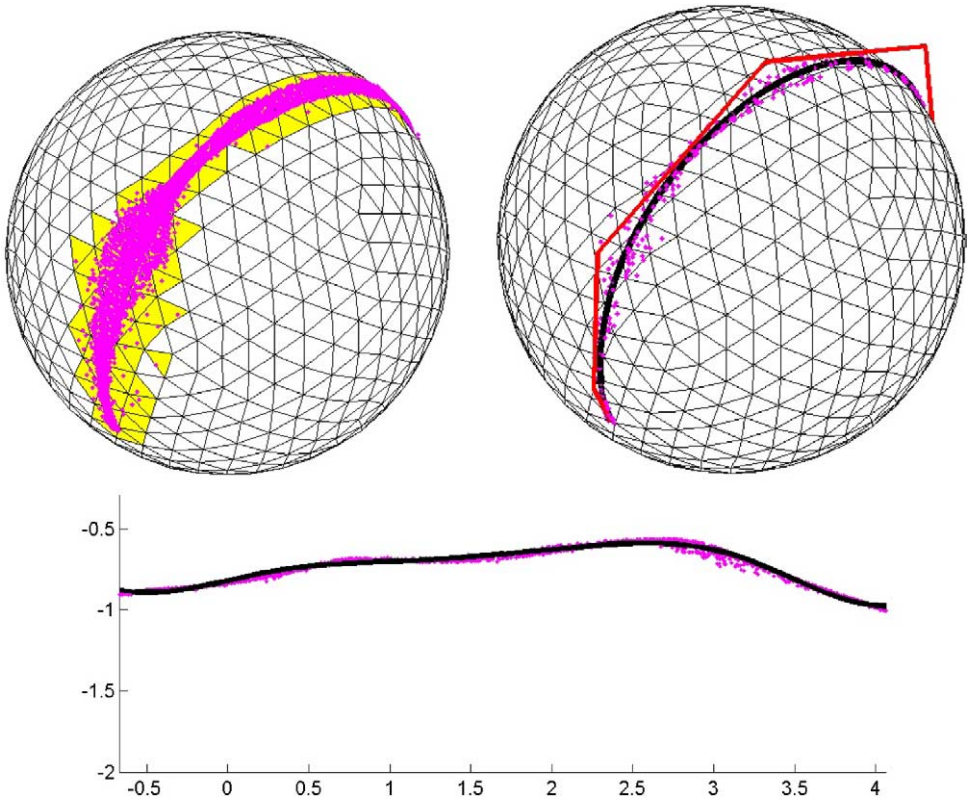

Abb. 6.8: links: Blaschke-Bild (projiziert auf S^2), rechts: bereinigte Punktmenge und approximierende Kurve, unten: 4. Koordinate

Somit kann die Fläche $D*$ in der Form

$$\mathbf{x}(t,u) = (1-u)\mathbf{f}_1(t) + u\mathbf{f}_2(t) \tag{6.13}$$

parametrisiert werden.

Zur Quantifizierung der Qualität der Approximation kann man entweder die Distanzen der geschätzten Ebenen T_i zu den approximierenden Ebenen $E(t_i)$ mit geeignetem t_i wählen, also

$$\mathrm{d}^2(D, D^*) = \frac{1}{N}\sum_{i=1}^{N} \mathrm{dist}^2(T_i, E(t_i)) \tag{6.14}$$

oder, falls man die Punkte $\mathbf{p}_i$ betrachtet

$$\mathrm{d}^2(D, D^*) = \frac{1}{N} \sum_{i=1}^{N} \mathrm{dist}^2(\mathbf{p}_i, E(t_i)). \tag{6.15}$$

7 Curved Folding

Der entscheidende Punkt in der Arbeit [12] von Martin Kilian et al. sind die sogenannten *curved folds*, also Falten, die nicht gerade, sondern gekrümmt sind. Ein bedeutender Unterschied zu den zuvor beschriebenen Werken ist, dass die Wissenschafter hier ein Netz aus Vierecken verwenden, im Gegensatz zu den Triangulierungen, wie wir sie bisher hatten. Dies bietet diverse Vorteile, wie etwa die Garantie der Abwickelbarkeit, falls die Vierecke planar sind, oder die Tatsache, dass die Krümmung leicht geschätzt werden kann, um nur zwei Vorteile zu nennen. Sie verwenden also Streifen von planaren Vierecken, kurz PQ strip (strip of planar quadrilaterals), als diskrete Modelle einer torsalen Regelfläche. Die gemeinsame Kante zweier Vierecke liefert eine diskrete Erzeugende und diese definieren wiederum, zusammen mit den Schnittpunkten aufeinanderfolgender Erzeugenden, im Allgemeinen einen Streckenzug, der als diskrete Variante der Gratkurve aufgefasst werden kann. In speziellen Fällen können diese Erzeugenden allerdings auch parallel oder kopunktal sein. Verfeinert man nun diese PQ strips ad infinitum, unter Beibehaltung der Planarität der Vierecke, so erhält man als Endprodukt eine torsale Regelfläche Σ. Die Erzeugenden von Σ, die aus den diskreten Erzeugenden folgen, hüllen entweder die Gratkurve ein oder sind parallel oder kopunktal.

Wie schon erwähnt, bietet die Verwendung von Viereckснetzen die Möglichkeit die Krümmung zu schätzen. Auf einer abwickelbaren Fläche bilden die Erzeugenden eine Schar von Krümmungslinien zur Hauptkrümmung null. Die andere Schar ist durch die Orthogonaltrajektorien der Erzeugenden gegeben. Zur diskreten Darstellung C der Hauptkrümmungslinien verwenden sie Punkte c_i auf den Erzeugenden R_i, deren Verbindungen auf die Symmetrale S_i aufeinanderfolgender Erzeugenden R_i, R_{i+1} orthogonal steht. In den Mittelpunkten $m_i = (c_i + c_{i+1})/2$ der Strecken $c_i c_{i+1}$ hängen sie den normierten Normalvektor n_i der Ebene, die durch R_i und R_{i+1} bestimmt wird, an. Diese Normalvektoren bestimmen das Gaußsche Bild von C und aus diesem Grund definieren sie die Hauptkrümmung κ_2 im Punkt c_i durch

$$\kappa_2(c_i) := \frac{\|n_i - n_{i-1}\|}{\|c_i - m_{i-1}\| + \|c_i - m_i\|} = \frac{N_i}{L_i}. \tag{7.1}$$

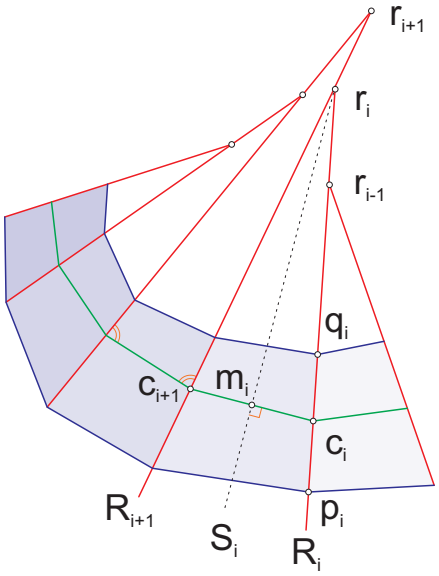

Abb. 7.1: Diskretes Modell einer abwickelbaren Fläche

Die Autoren weisen darauf hin, dass der Nenner L_i dieses Ausdrucks aus der planaren Abwicklung berechnet werden kann und nur der Zähler die räumliche Darstellung benötigt. Weiters streichen sie heraus, dass κ_2 nichtnegativ ist, im Gegensatz zu herkömmlichen Definitionen von Krümmungen, dieser Umstand aber auf Grund der Tatsache, dass dieser Ausdruck lediglich für die Berechnung der Biegeenergie $E_{bend} = \int \kappa_2^2 dA$ benötigt wird, nicht von Bedeutung ist.

Seien nun zwei diskrete Hauptkrümmungslinien $C, \overline{C}$ mit Abstand $h := \|\overline{c}_i - c_i\|$ gegeben, die zusammen mit zwei Symmetralen S_{i-1}, S_i einen Bereich begrenzen, dann ist die Biegungsenergie gegeben durch den Ausdruck

$$E_i = w_i \left\| n_i - n_{i-1} \right\|^2 . \tag{7.2}$$

Die verwendeten Gewichte w_i genügen der Bedingung

$$w_i = h \frac{\ln \overline{L}_i - \ln L_i}{\overline{L}_i - L_i} \tag{7.3}$$

und somit gilt für $\overline{L}_i \to L_i$, dass $w_i = h/L_i$. Für allgemeine PQ strips bedeutet dies, dass die zugehörige Biegeenergie als Summe von Ausdrücken der Form (7.2) angegeben werden kann.

Als nächstes gehen die Autoren von einem Netz M aus, das aus planaren Flächen, hauptsächlich Vierecken, besteht und das die Eigenschaft hat, dass die Summe der Winkel bei jedem Knoten dem Wert 2π entspricht. Damit gibt es eine Bijektion zwischen M und einem planaren Netz, dessen Flächen zu jenen von M isometrisch ist. Nun stellen die Autoren folgende Frage: Angenommen wir kennen die diskrete Darstellung einer gekrümmten Falte zwischen zwei PQ strips D_1 und D_2 und zusätzlich ist D_1 bekannt. Welche Möglichkeiten haben wir nun D_2 zu wählen?

Wir wissen, dass die diskreten Erzeugenden von D_2 so gewählt werden müssen, dass bei jedem Knoten p_i die Winkelsumme einen vollen Winkel ergeben muss. Sei $e_- := (p_{i-1}-p_i)/\left\|p_{i-1}-p_i\right\|$ und $e_+ := (p_{i+1}-p_i)/\left\|p_{i+1}-p_i\right\|$. Weiters nennen wir $\gamma_i = \alpha_i+\beta_i$ die Winkelsumme bei p_i in D_1. Für den Richtungsvektor x der diskreten Erzeugenden in D_2 muss also gelten, dass $\angle(e_-, x) + \angle(x, e_+) = 2\pi - \gamma_i$ ist. Mit $c := \cos\gamma_i$ kann man dies auch in der Form

$$(c^2-1)x^2 + (x\cdot e_-)^2 + (x\cdot e_+)^2 - 2c(x\cdot e_-)(x\cdot e_+) = 0 \tag{7.4}$$

schreiben. Diese Gleichung definiert einen quadratischen Kegel, der die möglichen Richtungen der diskreten Erzeugenden enthält. Die optimale Wahl unter diesen Möglichkeiten liefert der folgende Optimierungsprozess.

Der Optimierungsprozess wirkt gleichzeitig auf eine diskrete abwickelbare Fläche M und auf die zugehörige Abwicklung P. Dabei sei zu Beginn M eine Menge von planaren Vierecken im Raum, die nicht zwingend zusammenhängend sein muss. Die Polygone M^i von M sind isometrisch zu den Polygonen P^i des planaren Netzes P. Im Laufe des Optimierungsalgorithmus wird aus M, durch einen Registrierungsprozess, ein zusammenhängendes Netz. Nun verknüpfen Kilian et al. ein kartesisches Koordinatensystem (o, e_1, e_2) mit der Ebene von P, das unter der Isometrie, welche P^i in M^i überführt, in (o^i, e_1^i, e_2^i) in der Ebene von M^i übergeht. Seien nun (p_x, p_y) die Koordinaten des Knotens p von P_i, dann gilt für den entsprechenden Punkt m von M^i, dass $m = o^i + p_x e_1^i + p_y e_2^i$. Im Verlauf der Optimierung unterliegt das Zweibein (o^i, e_1^i, e_2^i) einer räumlichen Bewegung und ebenso können sich die Koordinaten (p_x, p_y) ändern, da die Polygone P^i nicht starr sind. Unter der Zuhilfenahme eines Geschwindigkeitsvektorfeldes mit $v(x) := \bar{c}^i + c^i \times x$ kann der Punkt m_+ des bewegten Vierecks in der Form

$$m_+ = m + \bar{c}^i + c^i \times o^i + p_x(c^i \times e_1^i) + p_y(c^i \times e_2^i) \tag{7.5}$$

dargestellt werden. Nun wird abwechselnd für die Parameter $\vec{c}^i, c^i$ der Bewegung und den Koordinaten p_x, p_y optimiert.

Die im Folgenden eingeführte Zielfunktion soll einerseits dazu führen, dass M ein zusammenhängendes Netz wird und andererseits, dass M die gegebene Datenmenge gut approximiert und außerdem hat sie noch für ästhetische Eigenschaften zu sorgen. Der erste Teil der Zielfunktion beschäftigt sich mit der Erschaffung des Netzes. Ist p ein Knoten im planaren Netz P der zu k Flächen gehört, so entsprechen p die k verschiedenen Punkte $m^1, \dots m^k$ der zugeordneten Flächen in M, die schlussendlich alle übereinstimmen sollen. Dazu wird ein *vertex agreement term* eingeführt, der die folgende Gestalt hat:

$$F_{vert} := \sum (m^i_+ - m^j_+)^2, \tag{7.6}$$

wobei die Summe über alle $\binom{k}{2}$ Kombinationen für den Knoten $p \in P$ und für alle Knoten in P verläuft.

Um für eine gute Approximation zu sorgen inkludieren sie in die Zielfunktion einen Term F_{fit}. Sei m_c jener Punkt der anzunähernden Fläche, der dem Punkt m am nächsten liegt, und sei weiters n_c die Flächennormale in m_c. F_{fit} besteht nun aus einer Linearkombination des quadratischen Abstandes $(m - m_c)^2$ und der quadratischen Distanz von m zur Tangentialebene in m_c, also $[(m - m_c) \cdot n_c]^2$.

Der letzte Teil, der *fairness term* F_{fair}, besteht für benachbarte Vierecke M^i und M^j aus der Biegeenergie $w_{ij}(n^i_+ - n^j_+)^2$, wobei $n^i_+ = n^i + c^i \times n^i$ die Normale von M^i unter der Bewegung ist. Für Streckenzüge $(p_1, \dots, p_n)$, die eine Falte oder eine Randkurve repräsentiert, besteht der fairness term aus einer Summe quadrierter zweiter Differenzen $\sum (p_{i-1} - 2p_i + p_{i+1})^2$.

Ein Term F_{conv} soll schließlich noch garantieren, dass die Vierecke konvex bleiben und dass keine Vierecke im planaren Netz P umgedreht werden, was gerade bei kleinen Vierecken, die nach mehreren Unterteilungsschritten entstehen, passieren könnte. Dabei ist eine Ecke (p_{i-1}, p_i, p_{i+1}) eines planaren Polygons genau dann konvex, wenn die Determinante $\Delta(p_{i-1}, p_i, p_{i+1})$ positiv ist.

Somit kann die Aufgabe folgendermaßen gestellt werden.

$$\begin{aligned} \text{Minimiere} \quad & F = F_{vert} + \lambda F_{fit} + \mu F_{fair} \\ \text{unter der Nebenbedingung} \quad & F_{conv} \geq 0. \end{aligned} \tag{7.7}$$

Wie bereits erwähnt, wird nun abwechselnd für die Knoten in P und den Geschwindigkeitsvektoren für die Ebenen der Flächen im Raum minimiert.

Geht man nun von einem Netz P und einer Menge M von Polygonen, die eine abwickelbare Fläche annähern aus, so wird abwechselnd für P und M optimiert, bis der *vertex agreement term* unter eine vorgegebenen Schranke fällt. Dann wird das Netz P verfeinert, indem jedes Viereck in zwei neue Vierecke geteilt wird. Die Unterteilung findet dabei nicht auf den Geraden statt, die den Erzeugenden entsprechen. Die neuen Vierecke werden anschließend wieder in den Raum abgebildet und dieser Algorithmus wird solange wiederholt, bis die gewünschte Genauigkeit erreicht wurde.

Dieser Algorithmus bietet somit die Möglichkeit, eine beinahe abwickelbare Fläche, durch eine exakt abwickelbare Fläche zu approximieren. Dazu benötigt man eine planare Abwicklung der Ausgangsdaten, Schätzungen für die Erzeugendenrichtungen in der Datenmenge, eine Zerlegung von P durch Polygone (hauptsächlich Vierecke) und eine Menge von zugehörigen Polygonen, die die Datenmenge annähert. Für den ersten Teil dieser Voraussetzungen verweisen die Autoren auf die Arbeit [11].

Für die Schätzung der Erzeugendenrichtungen ermitteln die Autoren zunächst einen geodätischen Kreis mit Radius r_p um jeden Datenpunkt p. Jeden Punkt q im Inneren dieses Kreises bewerten sie mit der Formel

$$\sigma(q) := n_p \cdot n_q + \nu \left\| p - q \right\| / r_p, \tag{7.8}$$

wobei n_p, n_q die normierten Normalvektoren in p bzw. q sind und für ν wählten sie den Wert $\nu = 0.1$. Im Regelfall treten zwei Maxima in diametral gelegenen Punkten am geodätischen Kreis auf, die die Schätzung der Erzeugenden festlegen. In annähernd planaren Regionen kristallisiert sich jedoch keine eindeutige Richtung heraus und diese werden als planar markiert. Danach werden die ermittelten Erzeugenden so lange verlängert, bis sie durch einen Punkt verlaufen, dessen Normale um mehr als eine vorgegebene Schranke von der Normale n_p abweicht oder sie z.B. in die Nähe des Randes gelangen. Außerdem wird die Standardabweichung der Flächennormalen entlang der Er-

zeugenden ermittelt, um in einem abschließenden Prozess die Menge der Erzeugenden zu verkleinern und in jedem Bereich nur jene zu behalten, die die höchste Qualität aufweisen und eine vernünftige Verteilung auf der Fläche liefern. Diese Erzeugenden bilden sie mit der Abbildung κ, die die Abwicklung festlegt, in die Ebene ab und entwickeln daraus die Zerlegung von P.

Um die Menge der planaren Polygone zu erhalten, die die Fläche annähern, wenden sie zunächst auf die Knoten in P die Abbildung κ^{-1} an. Die dadurch definierten Vierecke werden jedoch im Allgemeinen nicht planar und isometrisch zu jenen in P sein. Daher passen sie eine Kopie des Polygons P^i an die abgebildeten Eckpunkte an.

Nachdem dieser Initialisierungsprozess abgeschlossen ist, kann der zuvor beschriebene Algorithmus angewendet werden.

Literaturverzeichnis

[1] Julius D., Kraevoy V., and Sheffer A. “D-Charts: Quasi-Developable Mesh Segmentation”. In: *Computer Graphics Forum* 24.3 (2005), pp. 581–590.

[2] Bernardini F. et al. “The Ball-Pivoting Algorithm for Surface Reconstruction”. In: *IEEE Transactions on Visualization and Computer Graphics* 5.4 (1999), pp. 349–359.

[3] Brauner H. *Differentialgeometrie*. Vieweg, 1981.

[4] Pottmann H. and Wallner J. *Computational Line Geometry*. Springer-Verlag Berlin Heidelberg, 2001.

[5] Pottmann H. and Randrup T. “Rotational and Helical Surface Approximation for Reverse Engineering”. In: *Computing* 60 (1998), pp. 307–322.

[6] Pottmann H. et al. “On surface approximation using developable surfaces”. In: *Graphical Models and Image Processing* 61.2 (1999), pp. 110–124.

[7] Yamauchi H. et al. “Mesh Segmentation Driven by Gaussian Curvature”. In: *Visual Computer* 21 (2005), pp. 659–668.

[8] Daoshan OuYang und Hsi-Yung Feng. “On the normal vector estimation for point cloud data from smooth surfaces”. In: *Computer-Aided Design* 37 (2005), 1071â1079.

[9] Mitani Jun and Suzuki Hiromasa. “Making papercraft toys from meshes using strip-based approximate unfolding”. In: *ACM Trans. Graph.* 23 (3 2004), pp. 259–263.

[10] Stuart P. Lloyd. “Least squares quantization in pcm”. In: *IEEE Transactions on Information Theory* 28.2 (1982), pp. 129–137.

[11] Kilian M., Mitra N., and Pottmann H. “Geometric Modeling in Shape Space”. In: *ACM Transactions on Graphics (SIGGRAPH)* 26.3 (2007), #64, 1–8.

[12] Kilian M. et al. “Curved Folding”. In: *ACM Transactions on Graphics* 27.3 (2008), 75:1–75:9.

[13] Peternell M. “Developable surface fitting to point clouds”. In: *Computer Aided Geometric Design* 21 (2004), pp. 785–803.

[14] Max N. “Weights for Computing Vertex Normals from Facet Normals”. In: *Journal of Graphics Tools* 4.2 (1999), pp. 1–6.

[15] Leopoldseder S. and Pottmann H. “Approximation of developable surfaces with cone spline surfaces”. In: *Computer-Aided Design* 30 (1998), pp. 571–582.

[16] Randrup T. “Approximation of surfaces by cylinders”. In: *Computer-Aided Design* 30.10 (1998), pp. 807–812.