

Diplomarbeit

3D-Rekonstruktion von DNA-Strukturen

angefertigt am

Institut für Angewandte Mathematik

vorgelegt der

Mathematisch-Naturwissenschaftlichen Fakultät der

RHEINISCHEN FRIEDRICH-WILHELMS-UNIVERSITÄT BONN

Oktober 2001

von

Frank Lenzen aus Brühl

Danksagung:

Ich danke Professor Rumpf für die Betreuung während meiner Arbeit, allen Mitarbeitern und Hilfskräften des IAM, die mir bei Fragen zu UNIX, GRAPE oder Latex weitergeholfen haben sowie meinen Kommilitonen für ihre Anregungen und Hilfe.

Ich möchte Patrick Furrer von der EPFL in Lausanne für seine Hilfe bei Fragen zur Entstehung der DNA-Bilder danken. Jürgen Becker danke ich für das Probelesen meiner Arbeit. Außerdem danke ich meinen Eltern, meinem Bruder und meiner Freundin Bettina, die mich bei meinem Studium allzeit unterstützt haben.

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen	5
2.1	Stereobilder von DNA-Ringen	5
2.1.1	Die DNA	5
2.1.2	Cryo-Elektronenmikroskopie	6
2.1.3	Qualität der Bilder	7
2.2	Diffusionsmodelle	8
2.2.1	Allgemeines Modell	8
2.2.2	Beispiele	9
2.2.3	Lineare Diffusion	9
2.2.4	Modifiziertes Perona-Malik-Modell	10
2.2.5	Allgemeines Modell der Anisotropen Diffusion	12
2.2.6	Anisotrope Diffusion nach Weickert	12
2.3	Stereorekombination	16
2.3.1	Kameramodelle	16
2.3.2	Das Stereoproblem	19
2.3.3	Bilder von Elektronenmikroskopen	25
3	Modell	27
3.1	Einführung	27
3.1.1	Problemstellung	27
3.1.2	Lösungsstrategie	29
3.2	Erkennung schlauchartiger Strukturen	30
3.3	Diffusionsmodell	36
3.3.1	Differentialgleichung	36
3.3.2	Rechte Seite	38
3.3.3	Eine weitere Modifikation der rechten Seite	40
3.4	3D-Rekonstruktion	42
3.4.1	Rekonstruktionsmodell	42
3.4.2	Rekonstruktion verrauschter Bilder	42
3.5	Tracing	46
3.5.1	Gewöhnliche Differentialgleichung	46

3.5.2	Wahl der Metrik	48
3.5.3	Verzweigungen	50
3.5.4	Existenz und Eindeutigkeit	52
4	Wohlgestelltheit des Diffusionsproblems	55
4.1	Lipschitzstetigkeit von $D_\varepsilon(u)$ und $f_\varepsilon(u)$	55
4.2	Existenz und Eindeutigkeit der Lösung	62
5	Implementierung	69
5.1	Gitter	69
5.2	Num. Verfahren zur Lösung der Diffusionsgleichung	71
5.2.1	Zeitdiskretisierung	72
5.2.2	Ortsdiskretisierung mittels FE-Methode	72
5.2.3	Approximation der Diffusionsmatrix	74
5.2.4	Lokale Matrizen	76
5.2.5	Berechnung des Matrixproduktes	77
5.2.6	Löser	78
5.2.7	Adaptivität	78
5.3	3D-Rekonstruktion	82
5.4	Tracing schlauchartiger Strukturen	82
5.4.1	Integration in Gittermetrik	82
5.4.2	Lösung der gewöhnlichen Differentialgleichung	83
5.4.3	Stoppkriterien	85
5.4.4	Überblick über den Algorithmus	86
6	Anwendung	89
6.1	Diffusion	89
6.1.1	Vergleich verschiedener Diffusionsmodelle	89
6.1.2	Anwendung auf DNA-Bilder	95
6.2	Rekonstruktion des 3D-Bildes	100
6.2.1	Rekonstruktion von Ausläufern	100
6.2.2	Rekonstruktion der DNA-Struktur	101
6.3	Tracing eines geschlossenen Schlauchs	102
6.4	Probleme mit Kreuzungspunkten	105
7	Resumé	113
A	Weitere DNA-Ringe	115
A.1	DNA-Ring 1	116
A.2	DNA-Ring 2	118
A.3	DNA-Ring 3	120
A.4	DNA-Ring 4	122

Kapitel 1

Einleitung

Mit der Verbreitung von Computern in der Arbeitswelt wie auch im Privatbereich hat die digitale Bildverarbeitung an Bedeutung gewonnen. Heutzutage sind in den meisten Programmen zur Bildbetrachtung und Formatkonvertierung bereits Werkzeuge zur Bearbeitung der Bilder integriert. In vielen Arbeitsbereichen, in denen Bildinformationen verarbeitet werden, wie z.B. der Medizinischen Diagnostik, finden spezielle Bildverarbeitungsprogramme Anwendung.

Zur Entwicklung neuer Bildverarbeitungsalgorithmen ist ein theoretischer Zugang zur Bildverarbeitung hilfreich. Als brauchbar hat sich hier die Theorie der partiellen Differentialgleichungen erwiesen. Man interpretiert ein Bild als Funktion u_0 von einem zwei- oder dreidimensionalen Gebiet in einen Farbenraum, in Fall von Grauwertbildern ist dies $\mathbb{R}$, und wählt u_0 als Anfangswerte einer partiellen Differentialgleichung. Die Lösung $u(t)$ ausgewertet zu einem bestimmten Zeitpunkt $t = T > 0$ liefert nun ein entsprechend verändertes Bild.

Bekannte Modelle sind hier vor allem das Modell der linearen Diffusion (das inzwischen in allen gängigen Bildbetrachtungsprogrammen implementiert ist) , das Perona-Malik-Modell [9] sowie das Modell der anisotropen Diffusion, welches Weickert in seinem Buch [23] vorgestellt. Sie werden dazu verwendet, Störungen, die bei der Aufnahme der Bilder entstanden sind - wir nennen diese Störungen im folgenden Rauschen - zu entfernen, ohne die wesentlichen Strukturen der Bilder zu verändern.

Das Hauptziel der Bildbearbeitung ist, die Bildqualität zu verbessern oder bestimmte Bildmerkmale zu extrahieren. In beiden Fällen soll die Aussagekraft der Bilder erhöht werden.

Die vorliegende Arbeit beschäftigt sich mit einem Spezialgebiet der Bildverarbeitung, der Stereorekonstruktion, bei der aus zweidimensionalen Projektionen dreidimensionaler Objekte deren Lage und Gestalt in $\mathbb{R}^3$ rekonstruiert wird. In

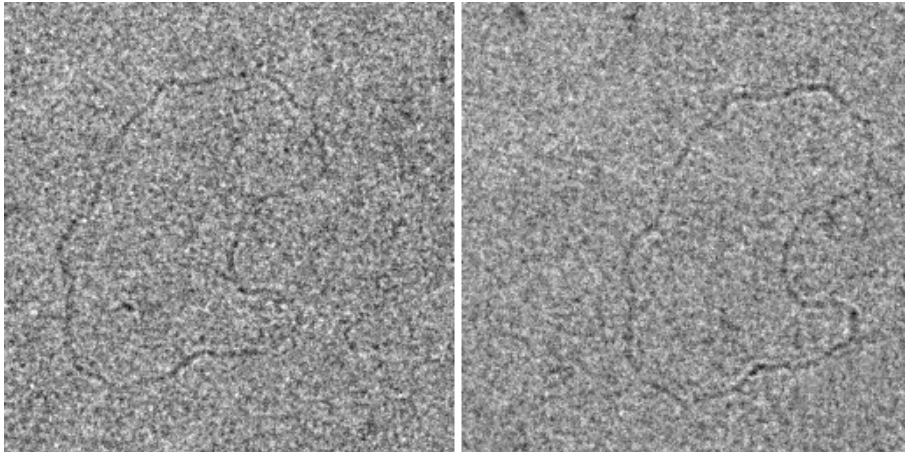


Abbildung 1.1: Beispiel für ein Stereobildpaar

unserem Fall handelt es sich bei den betrachteten Objekten um DNA-Strukturen, die mit Hilfe eines Elektronenmikroskops aus verschiedenen Perspektiven aufgenommen wurden. Abb. 1.1 zeigt ein solches Paar von Stereobildern. Ziel der Arbeit ist es, ein Modell zur Rekonstruktion der dreidimensionalen DNA-Strukturen zu entwickeln.

Ein Teil dieses Modells wird eine partielle Differentialgleichung beinhalten, die eine Anisotrope Diffusion ähnlich der von Weickert vorgestellten Diffusion beschreibt.

Sie wird dazu verwendet werden, das starke Rauschen in den Elektronenmikroskop-Aufnahmen zu reduzieren und gleichzeitig die gesuchte DNA-Struktur hervorzuheben. Nach der anschließenden dreidimensionalen Rekonstruktion, bei der aufgrund des Problems der Nichteindeutigkeit zusätzliche Strukturen entstehen, wird es nötig sein, die gesuchte DNA-Struktur zu isolieren. Wir verwenden dazu ein Verfolgungs-(Tracing-)Verfahren basierend auf einer gewöhnliche Differentialgleichung.

Im folgenden Kapitel werden wir zunächst einige Grundlagen aufgreifen. Wir werden auf einige Details der Elektronenmikroskopie eingehen und die Gründe für die schlechte Qualität der Aufnahmen untersuchen. Desweiteren stellen wir die oben kurz erwähnten Diffusionsmodelle und ihre Eigenschaften vor. Im letzten Teil des Kapitels wenden wir uns dem Stereoproblem zu.

Im dritten Kapitel beschreiben wir dann ausführlich die Problemstellung und das Modell zur Rekonstruktion der DNA-Strukturen, das sowohl die Vorbearbeitung der Bilder mit Hilfe der Diffusionsgleichung als auch das Tracingverfahren nach der Rekonstruktion umfaßt.

In Kapitel vier stellen wir den Beweis der Existenz- und Eindeutigkeit der Lösung der Diffusionsgleichung vor.

Kapitel fünf beschreibt die Implementierung des Modells. Zunächst wird ein adaptiver Löser für die partielle Differentialgleichung vorgestellt, der auf der Methode der Finiten Elemente aufbaut. Wir orientieren uns hier im wesentlichen an der Diplomarbeit von Tobias Preußner [19]. Anschließend beschreiben wir kurz das Rekonstruktionsverfahren. Der letzte Teil des Kapitels beschäftigt sich mit der Implementierung des Tracingverfahrens.

Im sechsten Kapitel werden wir zunächst die verschiedenen Diffusionsmodelle miteinander vergleichen und die Vorzüge des von uns gewählten Modells bei der Erkennung eindimensionaler Strukturen darlegen. Anschließend testen wir das Rekonstruktionsmodell an Hand der vorliegenden DNA-Aufnahmen.

In Kapitel sieben folgt schließlich ein kurzes Resumé, das die Brauchbarkeit des Modells bewertet und einen Ausblick auf weitere Optimierungen gibt.

Im Anhang führen wir noch weitere DNA-Aufnahmen auf, anhand derer eine Rekonstruktion durchgeführt wurde.

Kapitel 2

Grundlagen

2.1 Stereobilder von DNA-Ringen

Die vorliegenden Stereobilder, die uns Eric Larquet¹ freundlicherweise zu Verfügung gestellt hat, sollen dazu dienen, unser Modell in der Anwendung zu testen.

Die Bilder zeigen Elektronenmikroskopaufnahmen von geschlossenen DNA-Ringen. Von jedem DNA-Ring wurde jeweils ein Paar von Stereobildern aufgenommen. Der Winkel zwischen beiden Projektionsrichtungen sowie die durch diese Richtungen bestimmte Rotationsebene sind bekannt. Die Daten liegen als Grauwertbitmaps vor. Die Stereobilder werden im Département de Mathématiques der École Polytechnique Fédérale de Lausanne (EPFL) dazu verwendet, die dreidimensionale Struktur echter DNA mit den Ergebnissen der Simulation von DNA-Strängen zu vergleichen [4].

2.1.1 Die DNA

Die Desoxyribonukleinsäure, kurz DNA, ist ein langes, fadenförmiges Makromolekül. Es besteht aus zwei miteinander verdrehten Strängen, die durch Basenpaare verbunden sind (Abb.2.1). Man bezeichnet diese Struktur als Doppelhelix. In der Abfolge der Basenpaare ist die genetischen Information kodiert. Da DNA-Moleküle eine Vielzahl von Informationen speichern müssen, sind sie entsprechend lang. Um trotzdem auf kleinem Raum Platz zu finden, liegt die DNA oft nicht linear, sondern stark gekrümmt mit vielen Windungen vor.

Geschlossene DNA-Ringe, wie sie oft bei Bakterien vorkommen, können zusätzlich eine sogenannte Superhelixform (engl. Supercoil) besitzen. Hier sind Teile der Doppelhelix nochmal um die Helixachse verdreht [21].

¹Eric Larquet arbeiten momentan am Pasteur-Institut in Paris.

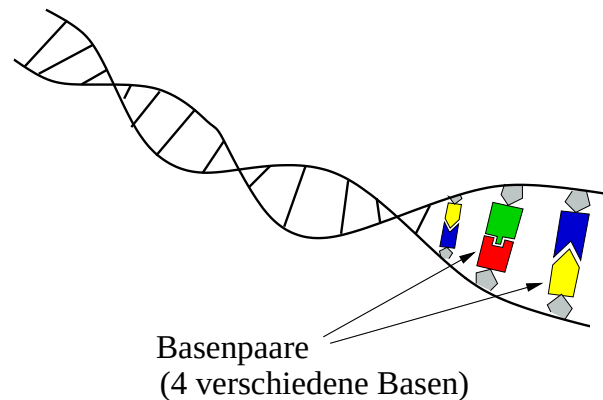


Abbildung 2.1: Modell der DNA-Helix

2.1.2 Cryo-Elektronenmikroskopie

Die Bilder der DNA-Ringe wurden mit Hilfe der Cryo-Elektronenmikroskopie aufgenommen.

Die Elektronenmikroskopie (EM) basiert auf ähnlichen physikalischen Gesetzen wie die Lichtmikroskopie; an Stelle der Lichtstrahlen werden Elektronen verwendet. Die Elektronen besitzen aufgrund ihrer Geschwindigkeit Wellencharakter. Optische Linsen werden hier durch Elektronenlinsen, die mit Hilfe von elektrischen oder magnetischen Feldern arbeiten, ersetzt. Aus der im Vergleich zu Licht wesentlich kleineren Wellenlänge der Elektronenstrahlen resultiert das deutlich höhere Auflösungsvermögen gegenüber dem Lichtmikroskop. Man unterscheidet grundsätzlich zwischen Transmissions-(TEM) und Rasterelektronenmikroskopie (SEM). Bei der TEM wird die Intensität der die Probe durchdringenden Elektronenstrahlen gemessen; beim Rasterelektronenmikroskop wird die Oberfläche eines Objekts mit einem Elektronenstrahl abgetastet und die gestreuten Elektronen werden erfaßt [3].

Da Elektronenstrahlen durch Kollision mit Atomen oder Molekülen abgebremst werden, haben sie in Gasen nur eine geringe Reichweite. In Luft liegt die Reichweite bei etwa 1 cm. Daher ist es nötig, Elektronenstrahl und Probe im Vakuum zu halten. Dies hat jedoch den Nachteil, daß biologisches Material wie z.B. Zellen während der Aufnahme stark austrocknet. Das kann zu unerwünschten strukturellen Veränderungen in der Probe führen. Um dies zu verhindern, wird bei der sogenannten Cryo-Elektronenmikroskopie die Probe in flüssigem Stickstoff auf etwa $-196\text{ }^{\circ}\text{C}$ abgekühlt. Damit keine Kristallbildung einsetzt, was die Qualität der Probe ebenfalls vermindern würde, werden spezielle Verfahren zum schnellen Abkühlen eingesetzt. Eventuell werden die gefrorenen Proben in dünne Scheiben geschnitten (Cryo-sectioning), um weitere Störeffekte zu beseitigen [3].

2.1.3 Qualität der Bilder

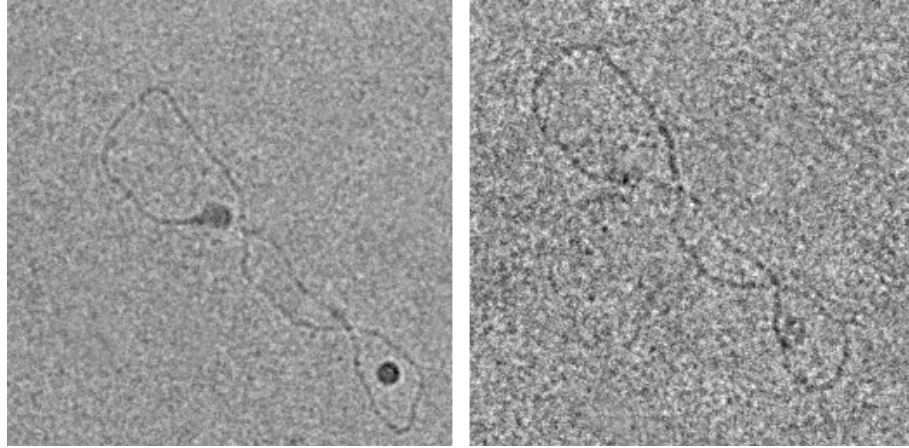


Abbildung 2.2: Beispiel für ein Stereobildpaar

Der große Nachteil der Elektronenmikroskopie liegt in der vergleichsweise schlechten Qualität der Aufnahmen. Ein großes Problem stellen die Beschädigung der Probe durch die Strahlen dar. Während die meisten Elektronen die Probe ungehindert passieren oder durch elastische Stöße abgelenkt werden, geben andere durch unelastische Stöße Energie an die Probe ab, wodurch eventuell Molekülbindungen aufgebrochen werden und chemische Reaktion einsetzen.

In den Bildern macht sich dies durch ein starkes Rauschen bemerkbar. Um solche Schäden zu minimieren, kann die Strahlendosis, d.h. die Anzahl der Elektronen pro Fläche reduziert werden, was jedoch zu einem geringeren Kontrast führt [3].

Aufgrund der Strahlenschäden ist die jeweils zweite Aufnahme eines DNA-Rings deutlich schlechter als die erste. Dies wird sich bei der Bearbeitung der Bilder bemerkbar machen. Bei Aufnahmen der DNA aus mehr als zwei Perspektiven muß ein noch höherer Qualitätsverlustes in Kauf genommen werden (vgl. Abb. 2.2).

Durch die schlechte Qualität der Bilder gestaltet sich eine exakte Erkennung des DNA-Rings schwierig. Betrachtet man kleinere Ausschnitte der Bilder, so ist es z.T. für das menschliche Auge unmöglich, den Verlauf des DNA-Rings zu erkennen. Erst beim Betrachten des gesamten Bildes ist das Auge in der Lage, den Verlauf des DNA-Rings zu erfassen. Wir werden daher auch bei der Anwendung unseres Rekonstruktionsmodells auf einige Schwierigkeiten stoßen.

2.2 Diffusionsmodelle

Ein Teil des Modells, das wir in Kapitel 3 ausführlich beschreiben werden, wird eine partielle Differentialgleichung beinhalten, mit deren Hilfe wir das in den EM-Bildern enthaltene Rauschen reduzieren wollen.

Wir können ein (zweidimensionales) Grauwertbild als Funktionen $u_0 \in (\Omega, \mathbb{R}_0^+)$ auffassen, wobei Ω ein Gebiet in $\mathbb{R}^2$ sei. Wir wählen u_0 als Anfangswert einer parabolischen Differentialgleichung, deren eindeutige Lösung $u : [0, T] \times \Omega \rightarrow \mathbb{R}^+$ bis zur Zeit $T > 0$ unter geeigneten Voraussetzungen eine $C([0, T] \times \Omega)$ – bzw. eine $C^\infty([0, T] \times \Omega)$ -Funktion darstellt. Die Lösung zu einer festen Zeit punktweise ausgewertet interpretieren wir wiederum als Grauwertbild. Die partielle Differentialgleichung wird so zum Bildverarbeitungsmodell.

Wir beschränken uns im folgenden auf Diffusionsgleichungen. In diesem Abschnitt wollen wir die allgemeine Form einer Diffusionsgleichung vorstellen, wobei wir den allgemeineren Fall $\mathbb{R}^d, d = 2, 3$ betrachten. Wir werden einige bekannte Beispiele aus der Literatur aufgreifen, mit denen wir später unser Modell vergleichen wollen.

2.2.1 Allgemeines Modell

Wir betrachten nun folgende allgemeine Form einer Diffusionsgleichung:

Sei $\Omega \subset \mathbb{R}^d, d = 2, 3$.

Seien $u_0 \in L^\infty(\Omega, \mathbb{R}_0^+), f(v) \in L^\infty([0, T] \times \Omega, \mathbb{R})$ für beliebiges $v \in L^\infty(\Omega, \mathbb{R})$.

Sei $\partial\Omega$ der Rand von Ω und $n : \partial\Omega \rightarrow \mathbb{R}^2$ die zugehörige äußere Normale.

Die folgende Differentialgleichung beschreibt einen Diffusionsprozeß:

$$\begin{aligned} \partial_t u - \operatorname{div}(D \nabla u) &= f(u) && \text{in } [0, T] \times \Omega \\ u(0) &= u_0 && \text{in } \Omega \\ \langle D \nabla u, n \rangle &= 0 && \text{auf } [0, \infty] \times \partial\Omega \end{aligned} \quad (2.1)$$

wobei die Diffusionsmatrix $D = D(u)(x)$ (uniform) positiv definit und symmetrisch sei und $D(u)(x) \in L^\infty(\Omega, \mathbb{R}^{d \times d})$ gelte.

Gesucht ist nun eine schwache Lösung $u \in L^2([0, T], H^{1,2}(\Omega))$ der Differentialgleichung im folgenden Sinne:

Für alle $v \in H^{1,2}(\Omega)$ gelte:

$$\langle v, \partial_t u \rangle + \langle \nabla v, D \nabla u \rangle = \langle v, f(u) \rangle \quad (2.2)$$

und $u(0) = u_0$ fast überall in Ω .

Aus der Literatur [23] sind folgende Eigenschaften der Lösung bekannt:

- $u(t)(x) \in C^\infty((0, T] \times \Omega)$, falls $D(u) \in C^\infty(\Omega, \mathbb{R}^{d \times d})$ gilt.
- $u(t)(x) \rightarrow \frac{1}{|\Omega|} \int_{\Omega} u_0(x) dx$ für $t \rightarrow \infty$, falls $f \equiv 0$

2.2.2 Beispiele

Betrachten wir nun einige spezielle Diffusionsgleichungen aus der Literatur. Im folgenden sei immer $f = 0$ gewählt.

Hierbei interessiert es uns, ob die Differentialgleichungen dazu verwendet werden können, die Erkennung von Objekten zu erleichtern.

Wir wollen dies für den Fall $d = 2$ diskutieren. Für den Fall $d = 3$ gilt analoges.

Offene zusammenhängende Teilmengen des $\mathbb{R}^2$, auf denen eine Funktion u in etwa konstant ist, $u \approx c$, bezeichnen wir im folgenden als (konstante) Flächen mit Intensität c . Grenzen sich zwei Flächen unterschiedlicher Intensität durch eine Kurve voneinander ab, wie z.B. bei Treppenfunktionen, so bezeichnen wir diese als Kante. Glättet man eine Treppenfunktion, so grenzen sich die Flächen durch schmale Teilgebieten entlang der ursprünglichen Kante ab, die sich durch einen hohen Betrag des Gradienten auszeichnen. Solche Gebiete bezeichnen wir im folgenden ebenfalls als Kanten.

Unter Objekterkennung verstehen wir das Bestimmen eines Gebiets mit Lipschitz-Rand, das durch Kanten entlang dieses Randes begrenzt ist.

Bei den Beispielen richten wir unser Augenmerk darauf, inwieweit Kanten und Flächen (in 2D) beibehalten werden und wie sich die Diffusion auf Rauschen an Kanten und auf Flächen auswirkt.

Das Analogon einer Kante in 2D ist in 3D die durch ein Levelset von u beschriebene Oberfläche.

2.2.3 Lineare Diffusion

Die einfachste Wahl einer Diffusionsmatrix ist $D = \sigma \cdot \mathbf{1}$. Die Differentialgleichung (2.1) entspricht dann der aus der Physik bekannten Wärmeleitungsgleichung.

Die Diffusion erfolgt in Richtung des negativen Gradienten, daher bezeichnen wir sie als *isotrop*. Zu gegebenen $u_0 \in L^\infty(\Omega)$ ergibt sich die Lösung $u(t)$ durch

$$u(t, x) = \frac{1}{4\sigma^2\pi t} \int_{\mathbb{R}^2} u_0(\tilde{x}) \exp\left(-\frac{(x - \tilde{x})^2}{4\sigma^2 t}\right) d\tilde{x} \quad (2.3)$$

Betrachtet man das Frequenzspektrum von $u(t)$, so stellt man fest, daß $u(t)$ mit steigendem t zunächst die hohen Frequenzen verliert. Daher eignen sich numerische Verfahren zur Lösung dieser Gleichung, die wegen der Eigenschaft $u(t) \in C^\infty(\Omega)$ oft auch als Glätter bezeichnet werden, besonders dazu, auf Flächen mit konstantem Grauwert Rauschen, das ja in der Regel hochfrequent ist, zu entfernen. Der große Nachteil der linearen Diffusion ist, daß auch Kanten mit der Zeit weggeglättet werden. Somit verwischen die Grenzen der zu erkennenden Objekte (vgl. Abb. 2.3).

Im nächsten Abschnitt werden wir ein ähnliches Modell vorstellen, das zusätzlich die Eigenschaft besitzt, Kanten eine Zeit lang zu erhalten oder sogar zu schärfen.

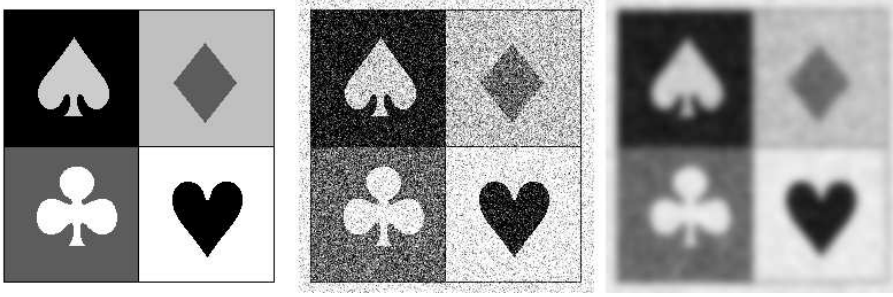


Abbildung 2.3: Testbild, links Original, Mitte verrauscht, rechts nach linearer Diffusion

2.2.4 Modifiziertes Perona-Malik-Modell

Setze

$$D = g(|\nabla u_\sigma|^2) \mathbf{1}$$

Hier sei $g(s) : \mathbb{R} \rightarrow \mathbb{R}^+$, $g(s) = \frac{1}{1 + \frac{s}{\kappa^2}}$.

u_σ entspricht der Funktion gefaltet mit der Gaußfunktion

$G_\sigma(x) := \frac{1}{4\sigma^2\pi} \exp\left(-\frac{x^2}{4\sigma^2}\right)$. Dieses Glätten der Funktion u kann als gewichtete Mittelung von u über $\mathbb{R}^2$ aufgefaßt werden.

Die Funktion g steuert hier punktweise die Stärke der Diffusion in Abhängigkeit vom Gradienten. In Gebieten mit hohem Betrag des Gradienten, also auch an Kanten, ist der Koeffizient klein, in Gebieten mit nahezu konstantem u liegt er nahe eins. Hier ist die Diffusion in etwa linear. Der große Vorteil des Modells gegenüber der linearen Diffusion ist, daß durch diese Steuerung der Diffusion bestimmte Kanten erhalten bleiben bzw. geschärft werden (vgl. [23]).

Wir wollen dies genauer untersuchen. Zu $x_0 \in \mathbb{R}^2$ sei $B_r(x_0) := \{x \in \mathbb{R}^2 \mid |x - x_0|_2 \leq r\}$ ein Ball mit Radius $r > 0$. Sei ∂B_r der Rand von B_r , n die äußere Normale an ∂B_r .

Aus der Differentialgleichung erhalten wir durch Integration über B und Anwenden des Satzes von Gauß:

$$\int_{B_r(x_0)} \partial_t u \, dx = \int_{\partial B_r(x_0)} g(|\nabla u_\sigma|^2) \langle \nabla u, n \rangle \, dx \quad (2.4)$$

Der Term $\int_{\partial B_r(x_0)} g(|\nabla u_\sigma|^2) \langle \nabla u, n \rangle \, dx$ beschreibt also die Diffusion durch den Rand ∂B_r .

Wir betrachten folgendes Beispiel: Die Lösung $u(t, x)$ der Differentialgleichung sei radialsymmetrisch im Punkt $x_0 = 0$. Sei $\tilde{u}(r, \phi) : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$ die Darstellung in Polarkoordinaten. $\tilde{u}(\phi)$ sei streng monoton fallend, d.h. u habe in x_0 eine echtes Maximum.

Sei $f(r, \phi)$ die Darstellung von $|\nabla u_\sigma(t_0, x)|$ in Polarkoordinaten und $f(\cdot, \phi)$ nehme im Punkt $R > 0$ ein Maximum an.

Nun gilt zu beliebigem $r > 0$ auf $\partial B_r(x_0)$ $\langle \nabla u(t_0, x), n \rangle < 0$, d.h. die Diffusion erfolgt in radiale Richtung von x_0 aus nach außen.

Der Term $g(|\nabla u_\sigma|^2)$ steuert nun die Stärke der Diffusion und nimmt auf $\partial B_R(x_0)$ ein Minimum $g_{min} > 0$ an. Die Diffusion durch $\partial B_R(x_0)$ erfolgt daher in einem deutlich geringeren Maße als in einer Umgebung um $\partial B_R(x_0)$. Dies führt zu einer Schärfung der Kante. Da g_{min} echt größer null ist, die Diffusion also nicht völlig gestoppt wird, konvergiert die Lösung $u(t, x)$ für $t \rightarrow \infty$ gegen eine konstante Funktion. Die Kantenschärfung ist also ein zeitlich begrenzter Effekt.

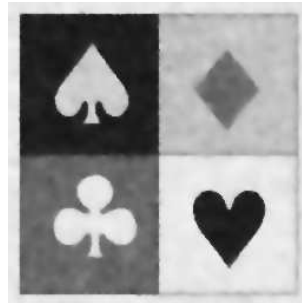


Abbildung 2.4: Testbild nach Perona-Malik-Diffusion

Im ursprünglichen Modell von Perona-Malik ist keine Vorglättung von u vorgesehen ($\sigma = 0$). In dem von Catte, Lions, Morel und Coll [9] modifizierten Modell

($\sigma > 0$) hängt der Diffusionskoeffizient von der vorgeglätteten Funktion u_σ , $\sigma > 0$ ab. Hochfrequente Anteile der Funktion u , also auch Rauschen, gehen nicht in die Berechnung von g ein. Die Vorglättung von u ist jedoch vor allem deswegen von Bedeutung, weil dadurch die Wohlgestelltheit des Problems sichergestellt wird. (Für das vorherige Modell konnte die Existenz einer eindeutigen schwachen Lösung bis zu einem Zeitpunkt $T > 0$ gezeigt werden. [23, Seite 18])

Von Nachteil ist, daß durch die Funktion g nur die Stärke der Diffusion, nicht aber ihre Richtung bestimmt wird. Rauschen an Kanten bleibt somit erhalten, solange dort Kantenschärfung stattfindet. Es wäre an solchen Stellen sinnvoll, zusätzlich zur Schärfung in Gradientenrichtung eine Diffusion in orthogonale Richtung stattfinden zu lassen, die das Rauschen wegglättet. Im nächsten Abschnitt werden wir ein solches Modell vorstellen.

2.2.5 Allgemeines Modell der Anisotropen Diffusion

Wir betrachten folgendes allgemeines Modell einer *anisotropen* Diffusion:

Seien $V_1 \dots V_d : \Omega \rightarrow \mathbb{R}^d$ Vektorfelder, so daß zu jedem $x \in \Omega$ $V_1(x) \dots V_d(x)$ eine Orthonormalbasis des $\mathbb{R}^d$ bilden. Setze

$$B := \begin{pmatrix} V_1(x) & \dots & V_d(x) \end{pmatrix} \quad (2.5)$$

und

$$D = B \begin{pmatrix} \eta_1(x) & 0 \\ 0 & \eta_2(x) \mathbf{1}_{(d-1) \times (d-1)} \end{pmatrix} B^T \quad (2.6)$$

mit $\eta_1(x), \eta_2(x) \in C^\infty(\Omega)$ beliebig.

Die Diffusionmatrix ist so gewählt, daß wir einen Diffusionskoeffizienten in Richtung $V_1(x)$ sowie einen orthogonal dazu wählen können. Wir werden im folgenden Abschnitt Beispiele dazu betrachten. Das in Kapitel 3 vorgestellte Modell hat ebenfalls diese Form.

Bemerkung:

Offensichtlich ist die Differentialgleichung unabhängig von der Wahl der Vorzeichen von $V_1(x) \dots V_d(x)$ und im dreidimensionalen Falls von der Wahl der Basis $V_2(x), V_3(x)$ des Unterraums $\mathbb{R}^3 \setminus \{t V_1(x) | t \in \mathbb{R}\}$. Aufgrund der geforderten Orthogonalität ist die Lösung der Differentialgleichung durch Wahl von $V_1(x)$, $\|V_1(x)\| = 1$ eindeutig bestimmt. Im folgenden beschränken wir uns daher auf die Vorgabe des Vektorfeldes $V_1(x)$.

2.2.6 Anisotrope Diffusion nach Weickert

Wir wollen V_1 nun so wählen, daß $V_1(x)$ im Fall $d = 2$ in der Nähe von Kanten orthogonal, im Fall $d = 3$ nahe Oberflächen, normal zu ihnen liegt. Dazu

betrachten wir den Strukturtensor (vgl. [23])

$$J_0(\nabla u_\sigma(x)) := \nabla u_\sigma(x) \otimes \nabla u_\sigma(x)$$

Offensichtlich ist J_0 symmetrisch.

$E_1(x) := \{c \cdot \nabla u_\sigma(x), c \in \mathbb{R}\}$ bildet den eindimensionalen Eigenraum zum Eigenwert $|\nabla u_\sigma(x)|^2$. $E_2(x) := \mathbb{R}^2 \setminus E_1(x)$ ist der Eigenraum zum Eigenwert 0. $J_0(\nabla u_\sigma(x))$ ist also positiv semidefinit.

Wir sind nun an einer Glättung bzw. Mittelung des Strukturtensors interessiert und falten daher J_0 komponentenweise mit einem Gaußkern $G_\rho, \rho > 0$:

$$J_\rho(\nabla u_\sigma(x)) := G_\rho * J_0(\nabla u_\sigma(x))$$

wobei ρ die Gewichtung dieser Mittelung bestimmt.

J_ρ ist ebenfalls symmetrisch und positiv semidefinit [23]. Es existiert also eine Orthonormalbasis aus Eigenvektoren $v_1(x) \dots v_d(x)$ zu den Eigenwerten $\lambda_1(x) \dots \lambda_d(x)$. Wir wählen die Numerierung so, daß $0 \leq \lambda_1(x) \leq \dots \leq \lambda_d(x)$ gilt. Im allgemeinen gilt nun nicht mehr $v_1 \parallel \nabla u_\sigma$.

Betrachten wir nun folgendes Beispiel:

$$u : \Omega \rightarrow \mathbb{R}$$

$$u(x, y) := \begin{cases} 1 & \text{falls } x, y \geq 0 \\ 0 & \text{sonst} \end{cases}$$

Wir untersuchen nun den Strukturtensor

$$J_0 = \begin{pmatrix} j_{11} & j_{12} \\ j_{12} & j_{22} \end{pmatrix} \text{ bzw. } J_\rho = \begin{pmatrix} j_{\rho,11} & j_{\rho,12} \\ j_{\rho,12} & j_{\rho,22} \end{pmatrix}$$

sowie die Eigenwerte λ_1, λ_2 von J_ρ in folgenden Teilgebieten von Ω :

- Die Umgebung des Eckpunktes $(0, 0)$:
 $E := \{(x, y) \in \Omega | x^2 + y^2 \leq \varepsilon^2\}, \varepsilon > 0$
- Die Umgebung der Kanten entfernt vom Eckpunkt:
 $K_1 := \{(x, y) \in \Omega | x \gg 1, |y| \approx 0\}$
 $K_2 := \{(x, y) \in \Omega | |x| \approx 0, y \gg 1\}$
- Die Flächen: $F_1 := \{(x, y) \in \Omega | x \ll -1 \text{ oder } y \ll -1\}$
und $F_2 := \{(x, y) \in \Omega | x \gg 1 \text{ und } y \gg 1\}$

Auf den Flächen ist u konstant bzw. u_σ annähernd konstant, daher sind die Einträge von J_0, J_ρ klein, also $\lambda_1, \lambda_2 \approx 0$.

Betrachten wir einen Punkt in K_1 , hier ist $a(x, y) := \partial_x u|_{(x,y)} u$ annähernd null, $b(x, y) := \partial_y u|_{(x,y)} \approx \frac{1}{\pi\sigma^2} \exp(-\frac{y^2}{\sigma^2}) \approx \frac{1}{\pi\sigma^2} \gg 0$ und $j_{11} = a^2(x, y) \approx 0$. Der Strukturtensor J_0 zeichnet sich daher durch $j_{11}, j_{12} \approx 0$ und $j_{22} \gg 0$ in einer Umgebung der Kante aus. Nach der Glättung gilt nun ebenfalls $j_{\rho,22} \gg 0$ und $j_{\rho,11}, j_{\rho,12} \approx 0$. Daher hat J_ρ (nach Vereinbarung über die Numerierung) den Eigenvektor v_1 annähernd parallel zu $(0, 1)$ zu einem Eigenwert $\lambda_1 \gg 0$, während λ_2 annähernd null ist.

Am Eckpunkt $(0, 0)$ hat ∇u_σ aufgrund der Symmetrie die Gestalt $\begin{pmatrix} a \\ a \end{pmatrix}$ mit $a = a(0, 0) \in \mathbb{R}$. Daher hat $J_0 = \begin{pmatrix} a^2 & a^2 \\ a^2 & a^2 \end{pmatrix}$ den Eigenvektor $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ zum Eigenwert $2a^2$, der andere Eigenwert ist null. Glättet man nun J_0 komponentenweise, so gilt aufgrund der Symmetrie

$$J_\rho|_{(0,0)} = \int_{\{(x,y) \in \Omega | x > y\}} \begin{pmatrix} a^2(x, y) + b^2(x, y) & 2a(x, y)b(x, y) \\ 2a(x, y)b(x, y) & a^2(x, y) + b^2(x, y) \end{pmatrix} G_\rho(x, y) dx dy$$

also $j_{\rho,11} = j_{\rho,22} > j_{\rho,12} > 0$. Die Eigenvektoren sind $v_{1,2} = \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$ zu den Eigenwerten $\lambda_{1,2} = j_{\rho,11} \pm j_{\rho,12}$

Zusammenfassend gilt also:

$$\begin{array}{ll} \lambda_1 \approx \lambda_2 \approx 0 & \text{auf den Flächen} \\ \lambda_1 \gg \lambda_2 \approx 0 & \text{an den Kanten} \\ \lambda_1 > \lambda_2 > 0 & \text{am Eckpunkt} \end{array}$$

Das motiviert folgende Definition (vgl. [23]):

Seien die Kohärenz sowie die normierte Kohärenz von u definiert durch

$$\begin{aligned} coh(u)(x) &:= \lambda_1(x) - \lambda_2(x) \\ \overline{coh}(u)(x) &:= \begin{cases} \frac{\lambda_1(x) - \lambda_2(x)}{\lambda_1(x)} & \text{für } \lambda_1 > 0 \\ 0 & \text{für } \lambda_1 = 0 \end{cases} \end{aligned} \quad (2.7)$$

Also gilt $coh(u)(x) \in [0, \infty)$ und $\overline{coh}(u)(x) \in [0, 1]$. Die Kohärenz ist nach dem oben gesagten ein Kriterium zur Erkennung von Kanten. (Zu Beachten ist, daß in dieser Definition implizit eine Vorglättung von u enthalten ist.)

Im Fall $d = 3$ gibt es verschiedene Alternativen, eine Kohärenz zu definieren. Wir behalten die Definition (2.7) für coh bzw. $\overline{coh}$ bei. In [14] werden aus den Eigenwerten drei verschiedene ‐Kohärenz‐-Werte ermittelt, die die Anisotropie der Struktur beschreiben.

Im Diffusionsmodell nach Weickert [23] setzt man nun in (2.5) $V_1 := v_1$.

Für die Wahl der Diffusionskoeffizienten η_1 und η_2 werden zwei Möglichkeiten vorgestellt.

a) Diffusion mit Kantenschärfung

Sei $g_1 : \mathbb{R} \rightarrow \mathbb{R}_0^+$ definiert durch:

$$g_1(s) = \begin{cases} 1 & (s \leq 0) \\ 1 - \exp\left(\frac{-C_m}{(s/\kappa)^m}\right) & (s > 0) \end{cases} \quad \text{mit } m \in \mathbb{N}, C_m > 0, \lambda > 0$$

Die Konstante C_m wird in Abhängigkeit zu m so gewählt, daß eine Schärfung von Kanten mit $|\nabla u_\sigma| > \kappa$ stattfindet, also:

$$\begin{aligned} \partial_s(s * g(s)) &> 0 \text{ für } s \in [0, \kappa) \\ \partial_s(s * g(s)) &< 0 \text{ für } s \in (\kappa, \infty) \end{aligned}$$

Die Wahl

$$\eta_1 = g_1(\lambda_1), \quad \eta_2 = 1,$$

impliziert eine Diffusionsgleichung mit konstantem Diffusionskoeffizienten in Richtung des Vektorfeldes v_1 , wohingegen die Diffusion in orthogonale Richtung durch g_1 ähnlich wie bei Perona-Malik gesteuert wird. Wie bei Perona-Malik werden im Fall $d = 2$ in diesem Modell Flächen geglättet, während Kanten geschärft werden (vgl. Abschnitt 2.2.4). Der Vorteil liegt jedoch in der konstanten kantenparallelen Glättung: Rauschen an Kanten bleibt bei einer Kantenschärfung nicht erhalten, wie beim Perona-Malik-Modell, sondern wird durch die Diffusion in Kantenrichtung geglättet. Analoges gilt für $d = 3$ an Levelset-Oberflächen von u .

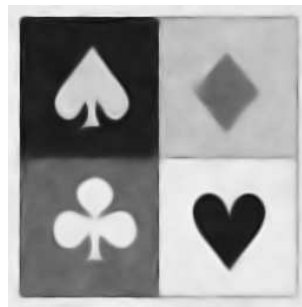


Abbildung 2.5: Testbild nach Anisotroper Diffusion mit Kantenschärfung

b) Diffusion mit Kohärenzverstärkung

Sei $g_2 : \mathbb{R} \rightarrow \mathbb{R}_0^+$ definiert durch:

$$g_2(s) = \begin{cases} \alpha & (s = 0) \\ \alpha + (1 - \alpha) \exp\left(\frac{-C}{s^{2m}}\right) & (s > 0) \end{cases} \quad \text{mit } m \in \mathbb{N}, C > 0$$

Man setzt nun η_1, η_2 wie folgt:

$$\eta_1 := \alpha, \quad \eta_2 := g_2(\text{coh}(u)),$$

wobei α dasselbe wie in der Definition von g_2 ist. Der Diffusionskoeffizient in Richtung v_2 nimmt mit steigender Kohärenz zu. Der Diffusionskoeffizient normal zu Kanten bzw. Oberflächen ist konstant. Kanten bzw. Oberflächen werden also nicht geschärft.

2.3 Stereorekombination

2.3.1 Kameramodelle

In diesem Abschnitt werden wir zwei Kameramodelle vorstellen, das Modell der orthogonalen Projektion und das der Lochkamera. Wir beschränken uns dabei zunächst auf optische Kameras, die die Intensität des einfallenden Lichts aufnehmen. In Abschnitt 2.3.3 übertragen wir diese Modelle auf die Elektronenmikroskopie.

Sei eine Anordnung von dreidimensionalen Objekten gegeben, die von einer Lichtquelle beleuchtet werden. Jeder Punkt auf der Oberfläche gibt nun einen Teil der einfallenden Lichtstrahlen in jede Richtung ab. Diese Lichtstrahlen werden mit einer Kamera erfaßt. Mit den folgenden Kameramodellen beschreiben wir den Weg, den die Lichtstrahlen ausgehend von der Oberfläche eines Objekts bis zum Auftreffen auf der Fotoplatte der Kamera zurücklegen.

Wir gehen davon aus, daß die Objekte undurchsichtig sind und sich daher die Lichtstrahlen nicht überlagern. Demnach fällt auf der Fotoplatte nur Licht eines bestimmten Oberflächenpunktes ein. Einen solchen Punkt bezeichnen wir als sichtbaren Punkt, Oberflächenpunkte, deren ausgesendetes Licht nicht von der Kamera erfaßt werden, als verdeckte Punkte. Punkte, die zu einem Objekt gehören, aber nicht auf dessen Oberfläche liegen, bezeichnen wir als innere Punkte.

Wir stellen uns nun einen Lichtstrahl ausgesendet im Punkt $M \in \mathbb{R}^3$ vor, der auf die Fotoplatte trifft. Die Fotoplatte beschreiben wir im Modell durch eine zweidimensionale Ebene, der Projektionsebene, versehen mit einem Koordinatensystem.

Den Auftreffpunkt des Lichtstrahls notieren wir mit den zweidimensionalen Koordinaten m . Zu einem Kameramodell erhalten wir somit eine Projektionsvorschrift P , die jedem von der Kamera erfaßten Punkt M des $\mathbb{R}^3$ den entsprechenden Punkt m in $\mathbb{R}^2$ zuordnet.

Sei nun Ω ein Gebiet in $\mathbb{R}^2$. In jedem Punkt $m \in \Omega$ wird nun die Intensität des einfallenden Lichtstrahls erfaßt. Sei M der Ausgangspunkt des einfallenden Lichtstrahls. Dann bezeichnen wir M als den in m *abgebildeten Punkt* und m als die *Abbildung* bzw. das *Abbild* von M . Falls in einem Punkt $m \in \Omega$ keine Lichtstrahlen auftreffen, setzen wir die Intensität gleich null. Der maximalen Intensität weisen wir den Wert 1 zu.

Ein aufgenommenes Bild interpretieren wir so als Funktion $I : \Omega \rightarrow [0, 1]$. Nach Fortsetzen von I durch null auf $\mathbb{R}^2 \setminus \Omega$ können wir $\Omega = \mathbb{R}^2$ annehmen.

Wir wählen eine Kameraposition $C \in \mathbb{R}^3$ und eine Kamerarichtung $v \in \mathbb{R}^3$. Position und Richtung werden in beiden Modellen die Lage der Projektionsebene eindeutig festlegen. Kameraposition, Kamerarichtung und das für die Projektionsebene gewählte Koordinatensystem bezeichnen wir als Kameraparameter.

Zur Vereinfachung der Notation betrachten wir die sogenannten Kamerakoordinaten in $\mathbb{R}^3$:

Nach Anwenden einer affinen Transformation, bestehend aus einer Translation und einer Drehung können wir $C = (0, 0, 0)^T$ und $v = (0, 1, 0)^T$ annehmen. Im Fall von mehr als einer Kamera müssen die verschiedenen Koordinatenwechsel berücksichtigt werden.

Wir gehen davon aus, daß nur Punkte in Kamerarichtung, also aus $D := \{x \in \mathbb{R}^3 \mid x \cdot v > 0\}$ durch die Kamera abgebildet werden.

Orthogonale Projektion

In diesem Modell verlaufen alle in die Kamera einfallenden Lichtstrahlen in Kamerarichtung, also parallel zur y -Achse, bis sie auf die (x, z) -Ebene (mit kanonischem Koordinatensystem) auftreffen (Abb.2.6). Der Auftreffpunkt eines vom Punkt M ausgehenden Lichtstrahls ergibt sich durch die orthogonale Projektion

$$P_{OP} : \mathbb{R}^3 \rightarrow \mathbb{R}^2, P(x, y, z) := (x, z)$$

Die Projektionsvorschrift ist offensichtlich von der Kameraposition unabhängig. [10]

Lochkameramodell

Alle in die Kamera einfallenden Lichtstrahlen laufen auf den Punkt C zu, der in

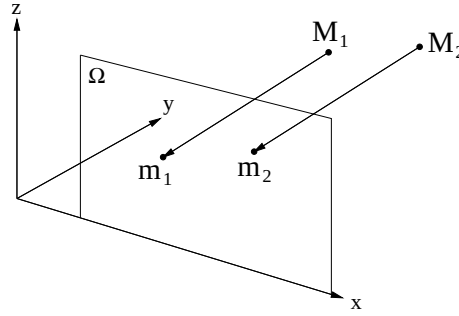


Abbildung 2.6: Kameramodell mit orthogonaler Projektion

diesem Fall auch als Fokus bezeichnen wird. Als Projektionsebene wählen wir die zur (x, z) -Achse parallele Ebene durch den Punkt $(0, s, 0)$, $s > 0$ mit dem 2D-Koordinatensystem $(x, z) \hat{=} (x, s, z)$ (Abb.2.7). Der Auftreffpunkt des Lichtstrahls von Punkt M ergibt sich aus dem Schnittpunkt S der Geraden $\overline{MC}$ mit der Projektionsebene: $S = (\frac{sx}{y}, s, \frac{sz}{y})$.

Wir definieren daher

$$P_{LK} : \mathbb{R}^3 \rightarrow \mathbb{R}^2, P_{LK}(x, y, z) := \left(\frac{sx}{y}, \frac{sz}{y} \right)$$

Nach Skalierung des zweidimensionalen Koordinatensystems mit $\frac{1}{s}$ können wir $s = 1$ annehmen.

Dieses Projektionsmodell ist invariant unter Skalierungen $A := \alpha \mathbf{1} \in \mathbb{R}^{3 \times 3}$, $\alpha \in \mathbb{R}$. [10, 12]

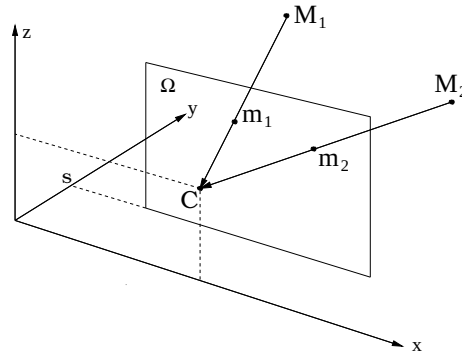


Abbildung 2.7: Lochkamera-Modell

Projektive Koordinaten

Nachteil der Definition von P_{LK} beim Lochkameramodell ist die Nichtlinearität von P . Um diese zu umgehen, definiert man die sogenannten *projektiven Koordinaten* für $\mathbb{R}^d$, $d = 2, 3$. Man verwendet eine zusätzliche Koordinate und identifiziert einen Punkt $p = (x_1, \dots, x_d)$ mit den projektiven Koordinaten $(l \cdot x_1, \dots, l \cdot$

$x_d, l)$ mit beliebigem $l \in \mathbb{R} \setminus \{0\}$. Punkte mit projektiven Koordinaten $(x, y, z, 0)$ interpretiert man als Punkte im Unendlichen.

Wir suchen nun die entsprechende Darstellung der Projektion P_{LK} in projektiven Koordinaten. Ein Punkt mit den projektiven Koordinaten

$(lx, ly, lz, l), l > 0, z > 0$ wird unabhängig von l auf einen Punkt $(x/z, y/z)$ mit projektiven Koordinaten (x, y, z) abgebildet. Die Unabhängigkeit von l liegt in der Skalierungsinvarianz von P begründet.

In projektiven Koordinaten hat unsere Projektion also die Gestalt

$$P_{LK} \begin{pmatrix} x \\ y \\ z \\ l \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ l \end{pmatrix}$$

Somit erhalten wir ein Modell, welches mit Hilfsmitteln aus der linearen Algebra beschrieben werden kann.[12]

2.3.2 Das Stereoproblem

Wir betrachten nun ein System von zwei Kameras mit unterschiedlichen Kamerarichtungen und wählen eines der oben definierten Kameramodelle. Die Kameraparameter seien bekannt, dadurch ergeben sich die entsprechenden Projektionen $P_1, P_2 : \mathbb{R}^3 \rightarrow \mathbb{R}^2$. Wir beschränken uns im folgenden auf ein Gebiet $\Omega^{3D} \subset \mathbb{R}^3$. Seien $\Omega_i := P_i \Omega^{3D}$, $i = 1, 2$ und $I_i : \Omega_i \rightarrow \mathbb{R}^2$, $i = 1, 2$ die Intensitätsfunktionen. Die Aufgabe besteht nun darin, zu jedem Punkt $m_1 \in \Omega_1$ bzw. $m_2 \in \Omega_2$ den Punkt $M \in \mathbb{R}^3$ zu bestimmen, der auf m_1 bzw. m_2 abgebildet wurde.

Das Korrespondierenden-Problem

Kennt man zu einem Punkt $m \in \Omega_1$ den Punkt $m' \in \Omega_2$, so daß beide Punkte Abbildungen ein und desselben Punktes $M \in \mathbb{R}^3$ sind, so lassen sich die dreidimensionalen Koordinaten von M berechnen. Dazu bestimmt man den Schnittpunkt der Geraden $P_1^{-1}(m)$ und $P_2^{-1}(m')$. Das Paar (m, m') bezeichnen wir als *korrespondierende Punkte*.

Zur Entscheidung darüber, ob zwei Punkte m, m' zueinander korrespondieren, können nur die Werte der Intensitätsfunktionen I_i an diesen Punkten bzw. in einer geeigneten Umgebung dieser Punkte verwendet werden. Wir setzen voraus, daß für korrespondierende Punkte (m, m') $I_1(m) = I_2(m')$ gilt.

Betrachten wir nun einen Punkt $M \in \mathbb{R}^3$, der auf der (C^∞ -) Oberfläche eines Objektes (d.h. eine Untermannigfaltigkeit) liegt. Wir nehmen ferner an, daß die Oberflächennormale in M nicht orthogonal zu einer der Kamerarichtungen ist.

Seien $m_i = P_i(M) \in \Omega_i$, $i = 1, 2$. Nun existiert eine Umgebung $U(m_1)$ von m_1 , so daß zu jedem $m \in U(m_1)$ der Schnittpunkt der Geraden $P_1^{-1}(m)$ mit der Oberfläche eindeutig bestimmt ist. Wir bezeichnen ihn mit $\phi_1(m)$. Dann ist durch $\phi := P_2 \circ \phi_1$ eine Bijektion von $U(m_1)$ auf eine Umgebung $U'(m_2)$ definiert. Nach unserer Annahme gilt nun $I_1(m) = I_2(\phi(m))$ für alle $m \in U(m_1)$.

Diese Betrachtung legt nahe, bei der Suche nach korrespondierenden Punkten die Intensitäten zweier möglicher korrespondierender Punkte m, m' auf geeigneten Umgebungen $U(m), U'(m')$ zu vergleichen.

Die Aufgabe, anhand der vorliegenden Funktionen I_i korrespondierende Punkte zu bestimmen, nennen wir im folgenden das *Korrespondierenden-Problem* (engl. 'correspondence problem').

Kern des Stereoproblems ist also das Korrespondierenden-Problem.

Unabhängig vom gewählten Kameramodell gilt:

Der Ausgangspunkt eines in $m \in \Omega_1$ einfallenden Lichtstrahls liegt auf der Geraden $L := P_1^{-1}(m)$ in $\mathbb{R}^3$. Diese Gerade, die wir im folgenden auch als *Projektionsgerade* bezeichnen, wird, da die Kamerarichtungen unterschiedlich sind, von der zweiten Kamera auf eine Gerade $l'(m)$ in Ω_2 abgebildet, die wir als die korrespondierende Gerade zu m bezeichnen. Zwei Geraden l, l' heißen zueinander korrespondierende Geraden, falls für jeden Punkt $m \in l$ $l'(m) = l'$ gilt, bzw. für jeden Punkt $m' \in l'$ $l(m') = l$. Das Suchen eines korrespondierenden Punktes zu m beschränkt sich also auf die Gerade $l'(m)$.

Zum Finden korrespondierender Punkte wird in der Literatur [12],[8] meist ein Fehlerfunktional aufgestellt, in dem zu einem möglichen korrespondierenden Paar m, m' geeignete Umgebungen $U(m), U'(m')$ miteinander verglichen werden.

Kamerakalibrierung

Zur Lösung des Korrespondierenden-Problems ist wie oben erwähnt die Kenntnis der Kameraparameter erforderlich. Diese Parameter lassen sich, falls nicht bekannt, berechnen, sofern genügend korrespondierende Paare bekannt sind. Man spricht von der Kamerakalibrierung. Longuet-Higgins [17] stellt für das Lochkameramodell einen Algorithmus zur Kamerakalibrierung vor, der die Kenntnis von acht geeigneten korrespondierenden Punkten voraussetzt.

Im Fall des Lochkameramodells besitzen die Geraden $\{l'(m)\}_{m \in \Omega_1} \subset \Omega_2$ einen gemeinsamen Schnittpunkt e' und werden als epipolare Geraden bzw. Linien bezeichnet. Sei e der entsprechende Epipol in Ω_1 . Bei Verwendung projektiver Koordinaten mit Ursprung e bzw. e' ist eine Abbildung, die zu $m \in \Omega_1$ einen Vektor v' parallel zu $l'(m)$ liefert, also $e' + tv' \in l'(m)$ für alle t gilt, linear. Die zugehörige Matrix wird in [12] als Fundamentalmatrix bezeichnet. Mit Kenntnis der Fundamentalmatrix ist eine Kamerakalibrierung nicht mehr notwendig.

Soll die Fundamentalmatrix bis auf Skalierung bestimmt werden, ergibt sich ein Gleichungssystem mit acht Unbekannten, die mit Kenntnis von acht korrespondierenden Paaren berechnet werden können [18].

Verdeckung und Nichteindeutigkeit:

Das Korrespondierenden Problem kann nur für Punkte gelöst werden, die in beiden Bildern sichtbar sind. Aussagen über verdeckte Punkte lassen sich nicht treffen. (Verdeckte Objekte lassen sich nicht rekonstruieren.) Wir bezeichnen dies als das Problem der Verdeckung.

Das Finden korrespondierender Punkte ist i.a. ein nicht eindeutiges Problem, da als alleinige Information die Intensitätsfunktionen I_1, I_2 vorliegen.

Suchen wir den korrespondierenden Punkt zu m , so kann es mehrere Punkte auf der korrespondierenden Geraden $l'(m)$ geben, die aufgrund gleicher Intensität als korrespondierende Punkte zu m in Frage kommen. Sei $K(m) \subset l'(m)$ die Menge dieser Punkte. (Sie muß keineswegs diskret sein.) Es läßt sich nicht eindeutig feststellen, welcher Punkt $m' \in K(m)$ zu m korrespondiert. Wir bezeichnen dies im folgenden als das *Nichteindeutigkeitsproblem*.

Trifft man eine falsche Wahl eines korrespondierenden Punktes zu m , so konstruiert man einerseits einen falschen Punkt in $\mathbb{R}^3$, andererseits wird der gesuchte Punkt M , der auf m abgebildet wurde, nicht rekonstruiert. Letzterer liegt auf der Geraden $l'(m)$ nicht zwangsläufig hinter dem falsch rekonstruierten Punkt (von der Position der ersten Kamera aus betrachtet). Wir können also nicht entscheiden, welche der dreidimensionalen Punkte, die aus den möglichen korrespondierenden Punkten rekonstruiert werden können, verdeckt sind und welche nicht.

Eine Alternative besteht darin, zu *jedem* Paar möglicher korrespondierender Punkte (m, m') , $m' \in K(m)$ einen dreidimensionalen Punkt M zu rekonstruieren. Man erhält dadurch mehr Strukturen als tatsächlich von den Kameras erfaßt wurden. Dieses Vorgehen hat den Vorteil, daß in jedem Fall die tatsächlich abgebildeten Objekte rekonstruiert werden. Sie müssen jedoch nach der Rekonstruktion aus der Menge aller rekonstruierten Objekte isoliert werden.

Betrachten wir zum Nichteindeutigkeitsproblem folgende Beispiele:

Beispiel 2.1: (Diskret verteilte punktförmige Objekte)

Wir wählen das Modell der orthogonalen Projektion. Das Beispiel läßt sich ohne weiteres auf das Lochkameramodell übertragen. Den Winkel zwischen beiden Kamerarichtungen, der im offenen Intervall zwischen 0 und π liege, bezeichnen wir mit α . Nach Wahl des 3D-Koordinatensystems können wir annehmen, daß

die P_i folgende Gestalt haben (Abb 2.8):

$$\begin{aligned} P_1(x, y, z) &:= (x, z) \\ P_2(x, y, z) &:= (\cos(\alpha)x + \sin(\alpha)y, z) \end{aligned}$$

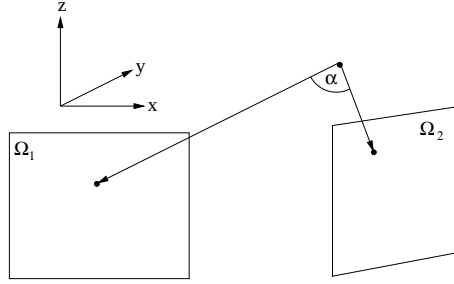


Abbildung 2.8: Anordnung der Kameras

Die korrespondierende Gerade zu einem Punkt $m = (x_0, y_0) \in \Omega_1$ liegt horizontal in Ω_2 ; sie ist gegeben durch $\{(x, y_0)\}_{x \in \mathbb{R}} \in \Omega_2$.

Wir nehmen an, daß in $\mathbb{R}^3$ n Punkte liegen, die Licht gleicher Intensität c ausstrahlen. Die Punkte seien gegeben durch die Koordinaten $M_i := (i, i, 0)$, $i = 1 \dots n$, d.h. die Punkte liegen auf der Diagonalen der $(x, y, 0)$ -Ebene in Ω^{3D} .

Beide Kameras bilden diese Punkte injektiv auf die x -Achse des jeweiligen zweidimensionalen Koordinatensystems ab:

$$\begin{aligned} m_i &:= P_1(M_i) = i \\ m'_i &:= P_2(M_i) = (\sin(\alpha) + \cos(\alpha))i \end{aligned}$$

mit $i = 1 \dots n$. Als Intensitätsfunktionen erhalten wir

$$I_1 = \left\{ \begin{array}{ll} c & \text{auf } \{m_i\}_{i=1 \dots n} \\ 0 & \text{sonst} \end{array} \right\}$$

definiert auf $\Omega_1 = \mathbb{R}^2$ und I_2 analog.

Sei $m := m_{i_0}$ für ein festes $i_0 \in \{1 \dots n\}$. Als zu m korrespondierender Punkt, der auf der korrespondierenden Geraden $l(m) = \{(x, 0)_{x \in \mathbb{R}}\} \subset \Omega_2$ liegen muß, kommen nun alle Punkte m'_i in Frage, da die Intensitätsfunktion I_2 an m'_i denselben Intensitätswert wie I_1 im Punkt m annimmt. Es kann also keine eindeutige Wahl eines korrespondierenden Paares (m, m'_i) getroffen werden.

Aus jedem Paar (m, m'_j) ließe sich ein eindeutiger Punkt M_j in $\mathbb{R}^3$ rekonstruieren. Von den M_j kann jedoch nur einer in der zweiten Kamera sichtbar sein. Abb.2.9

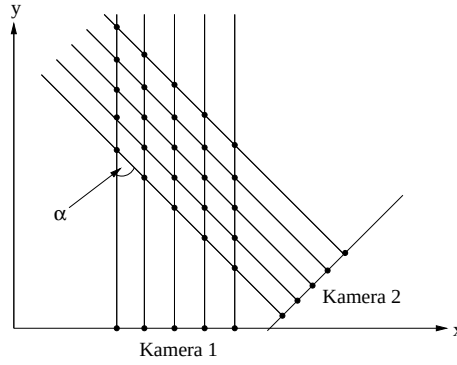


Abbildung 2.9: Rekonstruierbare Punkte im Beispiel 1

zeigt die Punkte, die sich aus den Paaren $(m_i, m'_j)_{ij=1\dots n}$ rekonstruieren lassen.

Beispiel 2.2:

Betrachten wir die gleiche Kameraanordnung wie im ersten Beispiel. Als Objekt im R^3 wählen wir einen (unendlich dünnen) Ring R , der in der Ebene $\{(x, x, z) | (x, z) \in \mathbb{R}^2\}$ liegt:

$$R := \{(x, y, z) \in \mathbb{R}^3 | x = y, x^2 + y^2 + z^2 = 1\}$$

Die Punkte $(0, \pm 1, 0) \in R$ bezeichnen wir im folgenden als Pole.

Jeder Punkte von R sende Licht der gleichen Intensität c aus.

Der Ring wird von beiden Kamera als Ellipsen

$$\begin{aligned} E_1 &= \{(x, y) \in \Omega_1 | 2x^2 + y^2 = 1\} \\ E_2 &= \{(x, y) \in \Omega_2 | 2\left(\frac{x}{\cos(\alpha) + \sin \alpha}\right)^2 + y^2 = 1\} \end{aligned}$$

abgebildet, d.h. die Intensitätsfunktionen I_i nehmen auf E_i den Wert c , sonst den Wert 0 an.

Zur Rekonstruktion des Ringes betrachten wir nun die korrespondierenden Geraden $\{(x, y)\}_{y=y_0}$ in Ω_1 und Ω_2 . Für $y_0 = \pm 1$ schneidet eine solche Gerade die Ellipse an nur einem Punkt $m = (0, y_0) \in \Omega_1$ bzw. $m' = (0, y_0) \in \Omega_2$, d.h. die Intensitätsfunktionen nehmen auf diesen korrespondierenden Geraden nur an m bzw m' die Intensität c an, (m, m') sind somit eindeutig korrespondierende Punkte, aus denen sich die Pole des Rings rekonstruieren lassen.

Die korrespondierenden Geraden zu $y_0 \in (-1, 1)$ schneiden die Ellipsen in jeweils zwei Punkten $m, -m \in \Omega_1$ bzw. $m', -m' \in \Omega_2$ als Schnittpunkte der Ellipsen E_1 bzw. E_2 mit den Geraden $\{(x, y_0)\}_{x \in \mathbb{R}}$ im jeweiligen Koordinatensystem. Es ergeben sich also vier Paare möglicher korrespondierender Punkte $M_1 \dots M_4$, die in der Ebene $\{(x, y, y_0)\}_{(x, y) \in \mathbb{R}^2}$ eine Raute bilden (Abb.2.10).

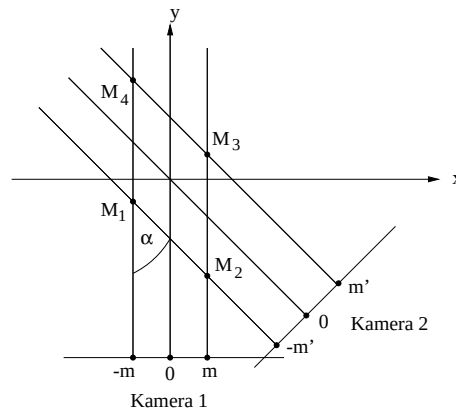


Abbildung 2.10: Ebenenweise Rekonstruktion eines Torus

Fügen wir die ebenenweise rekonstruierten Punkte zusammen, so erhalten wir ein Objekt, das aus ursprünglichen Ring R und einer Ellipse E besteht, der sich durch 90° -Drehung von R um die z -Achse mit anschließender Skalierung ergibt (Abb.2.11).

Fazit:

Verzichtet man bei der Rekonstruktion auf das eindeutige Lösen des Korrespondierendenproblems, so rekonstruiert man mehr Strukturen, als von den Kameras erfaßt wurden. Dies hat den Vorteil, daß die gesuchten Strukturen auf jeden Fall rekonstruiert werden. Sie müssen jedoch anschließend von den falsch rekonstruierten getrennt werden.

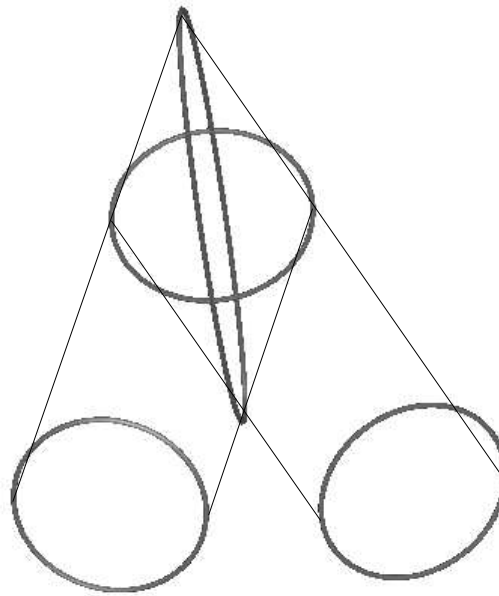


Abbildung 2.11: Rekonstruktion aus den 2D-Projektion eines 3D-Rings

2.3.3 Bilder von Elektronenmikroskopen

Anders als im Fall einer normalen Fotoaufnahme emittieren beim Elektronenmikroskop die Proben kein indirektes Licht, sondern werden von den Elektronenstrahlen durchleuchtet. Daher gilt unsere Annahme, dass die aufgenommene Intensitätswerte wie bei Lichtstrahlen einem bestimmten Ausgangspunkt zugeordnet werden und sich nicht überlagern, i.a. nicht. Wir gehen jedoch davon aus, dass die aufgenommenen DNA-Strukturen für Elektronenstrahlen fast undurchlässig sind, während die Elektronenstrahlen mit restliche Gewebe nicht interagieren. Wir erwarten daher eine niedrige Intensität in einem Punkt des Bildes, falls der zugehörige Elektronenstrahl die DNA-Struktur durchdringt und dadurch eine hohe Streuung erfolgt. Falls der Elektronenstrahl die Probe ohne Streuung an einem DNA-Moldekül durchdringt, ist die Intensität hoch.

Den Weg, den die Elektronenstrahlen zurücklegen, beschreiben wir wie bei der Fotokamera durch eines der beiden beschriebenen Kameramodelle.

Im nächsten Kapitel werden wir ein Kriterium vorstellen, mit dem dünne Strukturen in den Bildern erkannt werden können. Wir werden die DNA-Strukturen mit einer hohen Dichte, den Hintergrund mit einer niedrigen Dichte gleichsetzen. Dazu werden wir die Intensitätsfunktionen invertieren, d.h. wir gehen von Funktionen $u(x) \in [0, 1]$ zu Funktionen $1 - u(x)$ über.

Kapitel 3

Modell

3.1 Einführung

Im folgenden werden wir ein mathematisches Modell vorstellen, daß die Rekonstruktion dreidimensionaler schlauchartiger Strukturen aus ihren zweidimensionalen Stereoprojektionen beschreibt. Dabei wollen wir davon ausgehen, daß nur zwei Projektionen vorliegen. Das Modell läßt sich ohne weiteres auf den Fall mit mehr als zwei Projektionen übertragen, wir werden dies jedoch nicht weiter ausführen.

3.1.1 Problemstellung

Wir betrachten zunächst eine dreidimensionale schlauchartige Struktur, die entlang einer geschlossenen Kurve verläuft.

Sei Ω^{3D} ein Gebiet in $\mathbb{R}^3$. Sei $\gamma : [a, b] \rightarrow \Omega^{3D}$ eine geschlossene, sich nicht selbst überschneidende C^∞ -Kurve. Um diese Kurve liege nun ein Schlauch mit festem Radius $\varepsilon > 0$. Sei $\Gamma := \{x \in \Omega^{3D} | \exists t \in [a, b] \text{ } dist(x, \gamma(t)) < \varepsilon\}$ die Menge aller Schlauchpunkte. Wir wollen Γ als Urbildmenge einer L^∞ -Funktion S auf Ω^{3D} darstellen. Dazu wählen wir Werte $c_{Objekt}, c_{Hintergrund} \in [0, 1]$, wobei wir $c_{Objekt} > c_{Hintergrund}$ annehmen. S sei dann definiert durch:

$$\begin{aligned} S(x) &\geq c_{Objekt} && \text{falls } x \in \Gamma \\ S(x) &= c_{Hintergrund} && \text{falls sonst} \end{aligned}$$

Also gilt $\Gamma = S^{-1}([c_{Objekt}, \infty))$.

Wir betrachten nun zweidimensionale Abbilder des Schlauchs Γ . Hierbei wählen wir das Kameramodell der orthogonalen Projektion.

Sei $P : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ eine orthogonale Projektion. Wir müssen nun die entsprechende Intensitätsfunktion $u \in L^\infty(P\Omega^{3D})$ ermitteln. Wir nehmen an, daß wir in einem Punkt $x \in P\Omega^{3D}$ eine hohe Intensität ($u(x) \geq c_{\text{Objekt}}$) erhalten, falls die Gerade $P^{-1}(x)$ den Schlauch schneidet, und andernfalls die Intensität $c_{\text{Hintergrund}}$.

Wir definieren daher zu P folgenden Operator $\Pi : L^\infty(\Omega^{3D}) \rightarrow L^\infty(P\Omega^{3D})$:

$$\Pi v(x) := \max_{z \in P^{-1}(x)} v(z)$$

(Dieses Modell beschreibt die tatsächliche Entstehung der EM-Bilder natürlich nur ungenau. Dies hat jedoch keinen Einfluß auf das vorgestellte Rekonstruktionsmodell.)

Betrachten wir nun ein System von zwei Kameras. Seien P_1, P_2 zwei orthogonale Projektionen und Π_i die zugehörigen Operatoren. Der Winkel zwischen beiden Kamerarichtungen liege im Intervall $(0, \frac{\pi}{2})$. Seien $\Omega_i := P_i(\Omega^{3D})$, $i = 1, 2$ die Projektionen des dreidimensionalen Gebiets. O.b.d.A. wählen wir das Koordinatensystem des $\mathbb{R}^3$ so, daß die z-Achse orthogonal zu den Projektionsrichtungen liegt.

Sei $\alpha \in (0, \pi)$ der Winkel zwischen beiden Kamerarichtungen. Dann gilt:

$$\begin{aligned} P_1(x, y, z) &:= (x, z) \\ P_2(x, y, z) &:= (\cos(\alpha)x + \sin(\alpha)y, z) \end{aligned}$$

Wir erhalten zwei Funktionen $u_i : \Omega_i \rightarrow [0, 1]$ durch $u_i := \Pi_i S$.

Seien $K_i := \{x \in \Omega_i | u_i(x) \geq c_{\text{Objekt}}\}$. Wir bezeichnen K_i im folgenden als Abbilder des Schlauchs Γ . Sie sind schlauchförmige Mengen entlang den Kurven $\gamma_i(t) := P_i \gamma(t)$ in $\mathbb{R}^2$, die sich anders als γ selbst überschneiden können. Schnittpunkte von $\gamma_i(t)$ nennen wir im folgenden Kreuzungspunkte.

Die Werte $c_{\text{Objekt}}, c_{\text{Hintergrund}}$ bezeichnen wir im folgenden als Objekt- und Hintergrundintensität.

Anstelle der Funktionen u_i betrachten wir nun durch Störungen modifizierte $\tilde{u}_i \in L^\infty(\Omega_i)$. Wir nehmen an, daß die Werte der $\tilde{u}_i$ trotz der Störung im Intervall $[0, 1]$ liegen. Die Störung soll im wesentlichen aus Rauschen bestehen, d.h. einen hochfrequenten Anteil aufweisen. Der Schlauch Γ ist in diesen Bildern aufgrund der Störung nicht mehr exakt durch den Grauwert zu ermitteln, d.h. die Mengen $\tilde{K}_i := \{x \in \Omega_i | \tilde{u}_i(x) \geq c_{\text{Objekt}}\}$ stellen keine guten Approximationen der K_i dar.

Die Aufgabe besteht nun darin, aus den vorliegenden $\tilde{u}_i$ die Menge Γ zu rekonstruieren.

3.1.2 Lösungsstrategie

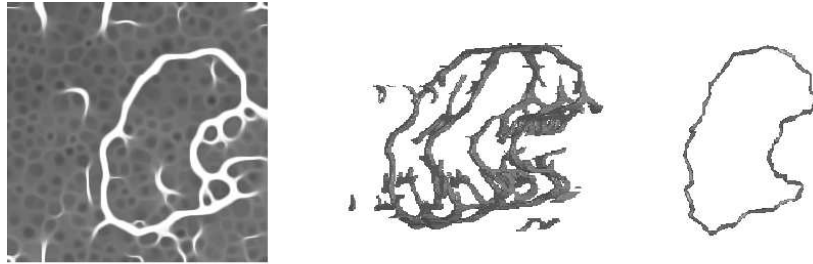


Abbildung 3.1: Strategie: Glätten in 2D, 3D-Rekonstruktion und Tracing

Unser Modell besteht aus drei Teilen.

Zunächst werden wir die Funktionen $\tilde{u}_i$ mit Hilfe einer Diffusionsgleichung glätten, um den Einfluß der Störung zu reduzieren. Wir erhalten zu den Anfangswerten $\tilde{u}_i$ Lösungen $u_i(t)$ der Differentialgleichung, die wir zu einem Zeitpunkt $t = T > 0$ auswerten. Durch Wahl einer speziellen rechten Seite werden wir eine Kontrastschärfung erreichen, sodaß auf den erkannten Strukturen $u(T)(x) \approx 1$ gelten wird (vgl. Abb. 3.1 links). Wir werden einen Schwellenwert $c_{approx} \in (0, 1]$ wählen und durch $K_i^* := u_i(T)^{-1}([c_{approx}, 1])$ die zweidimensionalen Abbilder des Objekts approximieren.

Im zweiten Schritt erfolgt die 3D-Rekonstruktion. Aus den geglätteten Funktionen $u_i(T)$ werden wir eine Funktion $S^* : \mathbb{R}^3 \rightarrow [0, 1]$ ähnlich S generieren. Dazu werden wir einen Operator $R : L^\infty(\Omega_1, [0, 1]) \times L^\infty(\Omega_2, [0, 1]) \rightarrow L^\infty(\Omega^{3D}, [0, 1])$ definieren und $S^* = R(u_1(T), u_2(T))$ wählen.

Wir werden bei der Rekonstruktion auf das Suchen korrespondierender Punkte verzichten. Wie beim Beispiel des Torus in 2.3.2 werden dadurch mehrere schlauchartige Strukturen entstehen. Zudem gehen in den Rekonstruktionsprozeß zusätzliche Strukturen ein, die durch den vorangegangenen Diffusionsprozeß im Hintergrund erkannt und hervorgehoben wurden (Abb. 3.1 Mitte). Um aus diesen Strukturen den gesuchten Schlauch Γ zu isolieren, ist ein weiterer Bearbeitungsschritt nötig. Wir werden die gesuchte Struktur stückweise durch ε -Schläuche entlang Lösungen einer gewöhnlichen Differentialgleichung approximieren (Abb. 3.1 rechts).

Diese Strategie sieht also das Glätten in 2D und eine anschließende 3D-Rekonstruktion vor. Eine Alternative dazu wäre, zunächst die Rekonstruktion und anschließend die Diffusion in $\mathbb{R}^3$ durchzuführen. Um unser Vorgehen zu begründen, werden wir, nachdem wir das Diffusions- sowie das Rekonstruktionsmodell vorgestellt haben, in Abschnitt 3.4.2 auf die Vor- und Nachteile beider Alternativen eingehen.

Im letzten Abschnitt wird die gewöhnliche Differentialgleichung erläutert, mit deren Hilfe wir den gesuchten Schlauch isolieren.

3.2 Erkennung schlauchartiger Strukturen

Wir werden nun einige Hilfsgrößen einführen, die in den folgenden Abschnitten dazu verwendet werden, um in Funktionen $u \in L^\infty(\Omega, [0, 1])$ schlauchartige Strukturen zu erkennen. Hierbei sei Ω ein Gebiet in $\mathbb{R}^d$, $d = 2, 3$.

Wir können eine solche Funktion u als Massenverteilungen bzw. Dichte auf Ω auffassen. Somit lassen sich physikalische Modelle, z.B. aus der Theorie starren Körper, anwenden. Wir sind dabei nur an lokalen Eigenschaften der Dichteverteilung interessiert und betrachten daher zu jedem Punkt x_0 des Gebiets lediglich die Dichte in einer kleinen Umgebung $B_\varepsilon(x_0) := \{x \in \Omega \mid |x - x_0|_2 < \varepsilon\}$. Dabei sei ε unabhängig von x_0 gewählt. Wir bestimmen nun den Schwerpunkt der in $B_\varepsilon(x_0)$ liegenden Masse, wobei wir x_0 als Bezugspunkt wählen. Desweiteren bestimmen wir bzgl. dieser Masse ein Massenmoment $M(x_0) \in \mathbb{R}^{d \times d}$, dessen Eigenwerte und -vektoren Aufschluß über die Gestalt dieser Masse geben.

Ziel ist es, Strukturen mit hoher Intensität zu finden. Um Punkte hoher Intensität, also Punkte x mit $u(x) \approx 1$, gegenüber Punkten niedriger Intensität stärker zu gewichten, werden wir zur Berechnung von Schwerpunkt und Massenmoment von der Funktion $u(x)$ zur Funktion $u(x)^n$ mit $n \in \mathbb{N}$ übergehen (wobei der Fall $n = 1$ weiterhin zugelassen sei).

Um auch nahe des Randes von Ω die Funktion u auf den ε -Umgebungen auswerten zu können, setzen wir sie außerhalb von Ω geeignet fort. Sei $\hat{\Omega} := \{\bigcup_{x \in \Omega} B_\varepsilon(x)\}$. u sei auf $\hat{\Omega} \setminus \Omega$ durch Spiegelung am Lipschitz-Rand fortgesetzt.

Wir definieren zu festem $\varepsilon > 0$ die *lokale Masse*, den *lokalen Schwerpunkt* sowie ein *lokales Massenmoment*, das mit dem Trägheitsmoment für starre Körper zusammenhängt, wie folgt:

a) die lokale Masse um $x_0 \in \Omega$:

$$m(u)(x_0) := \int_{B_\varepsilon(x_0)} u^n(x) \, dx \quad (3.1)$$

b) den lokalen Schwerpunkt (falls $m(u)(x_0) \neq 0$):

$$sp(u)(x_0) := \frac{1}{m(u)(x_0)} \int_{B_\varepsilon(x_0)} u^n(x)(x - x_0) \, dx \quad (3.2)$$

c) das lokale Erste Massenmoment (falls $m(u)(x_0) \neq 0$):

$$M(u)(x_0) := \frac{1}{m(u)(x_0)} \int_{B_\varepsilon(x_0)} u^n(x) (x - x_0 - sp(u)(x_0)) \otimes (x - x_0 - sp(u)(x_0)) \, dx \quad (3.3)$$

Hierbei verwenden wir als Ursprung des (lokalen) Koordinatensystems den Mittelpunkt der betrachteten Umgebung.

Der Faktor $\frac{1}{m(u)(x_0)}$ in den Formeln für Schwerpunkt und Massenmoment führt zu Singularitäten von $sp(u)(x)$ und $M(u)(x)$, falls für $x \in \Omega$ $m(u)(x) = 0$ gilt. Wir ersetzen $\frac{1}{m(u)(x_0)}$ in (3.2) und (3.3) daher durch $\frac{1}{m(u)(x_0) + \delta}$ mit festem $\delta > 0$. Wir motivieren dies wie folgt:

Wir gehen von der Funktion $u(x)^n$ zu der Funktion $u(x)^n + c$ mit $c := \frac{\delta}{\text{area}(B_\varepsilon(0))}$ über. Dadurch ändert sich die lokale Masse:

$$m_\delta(u)(x_0) := \int_{B_\varepsilon(x_0)} (u(x)^n + c) \, dx = m(u)(x_0) + \delta \quad (3.4)$$

Wir setzen

$$\begin{aligned} sp_\delta(u)(x_0) &:= \frac{1}{m_\delta(u)(x_0)} \int_{B_\varepsilon(x_0)} (u(x)^n + c)(x - x_0) \, dx \\ &= \frac{1}{m(u)(x_0) + \delta} \int_{B_\varepsilon(x_0)} u(x)^n(x - x_0) \, dx \\ &= \frac{m(u)(x_0)}{m(u)(x_0) + \delta} sp(u)(x_0) \end{aligned} \quad (3.5)$$

$sp_\delta(u)$ unterscheidet sich von $sp(u)$ also nur durch einen Faktor. Es gilt

$$\|sp(u)(x) - sp_\delta(u)(x)\| \leq \varepsilon \left(1 - \frac{m(u)(x_0)}{m(u)(x_0) + \delta}\right) \leq \varepsilon \frac{\delta}{m(u)(x_0) + \delta}$$

Wir definieren $M_\delta(u)$ durch

$$\begin{aligned} M_\delta(u)(x_0) &:= \frac{1}{m_\delta(u)(x_0)} \int_{B_\varepsilon(x_0)} u(x)^n(x - x_0 - sp_\delta(x_0)) \otimes (x - x_0 - sp_\delta(x_0)) \, dx \\ &= \frac{1}{m(u)(x_0) + \delta} \int_{B_\varepsilon(x_0)} u(x)^n(x - x_0 - sp_\delta(x_0)) \otimes (x - x_0 - sp_\delta(x_0)) \, dx \end{aligned} \quad (3.6)$$

Zu beachten ist, daß im Integral $u(x)^n$ statt $u(x)^n + c$ auftritt. Andernfalls würden wir einen zusätzlichen Summanden

$$X = c \int_{B_\varepsilon(x_0)} \chi_\varepsilon (x - x_0)(x - x_0 - sp_\delta(u)(x_0)) \otimes (x - x_0 - sp_\delta(u)(x_0)) dx$$

erhalten. Wir vernachlässigen diesen Term jedoch. (Dieser Term läßt abschätzen durch $\|X\|_2 \leq C \frac{\delta \varepsilon^2}{m(x_0) + \delta}$.)

Mit $|x - x_0 - sp(u)(x_0)|, |x - x_0 - sp_\delta(u)(x_0)| \leq 2\varepsilon$ gilt:

$$\begin{aligned} & \|M_\delta(u)(x_0) - M(u)(x_0)\|_2 \\ & \leq \left(\frac{1}{m(u)(x_0)} - \frac{1}{m(u)(x_0) + \delta} \right) \\ & \quad \left\| \int_{B_\varepsilon(x_0)} u(x)^n (x - x_0 - sp(u)(x_0)) \otimes (x - x_0 - sp(u)(x_0)) dx \right\|_2 \\ & \quad + \frac{1}{m(u)(x_0) + \delta} \left\| \int_{B_\varepsilon(x_0)} u(x)^n (sp(u)(x_0) - sp_\delta(u)(x_0)) \otimes (x - x_0 - sp(u)(x_0)) dx \right\|_2 \\ & \quad + \frac{1}{m(u)(x_0) + \delta} \left\| \int_{B_\varepsilon(x_0)} u(x)^n (sp(u)(x_0) - sp_\delta(u)(x_0)) \otimes (x - x_0 - sp_\delta(u)(x_0)) dx \right\|_2 \\ & \leq C \left(\frac{1}{m(u)(x_0)} - \frac{1}{m(u)(x_0) + \delta} \right) + C \frac{m}{m + \delta} \|sp(u)(x_0) - sp_\delta(u)(x_0)\|_2 \\ & \leq C \frac{\delta}{m(u)(x_0) + \delta} \end{aligned}$$

Im folgenden verwenden wir $sp_\delta(u)$ und $M_\delta(u)$ statt $sp(u)$ und $M(u)$, verzichten aber bei der Notation auf den Index δ .

In Abschnitt 3.1.1 hatten wir angenommen, daß die ungestörten Funktionen $u_i, i = 1, 2$ auf den Abbildern des Schlauches Intensitäten größer gleich c_{Objekt} , sonst die Intensität $c_{Hintergrund}$ annehmen. Obige Betrachtungen zeigen (für $n = 1$), daß die von null verschiedene Intensität $c_{Hintergrund}$ einen Einfluß auf den Schwerpunkt und das Massenmoment hat. Dies gilt ebenfalls für die gestörten Funktionen $\tilde{u}_i$, da wir erwarten, daß die Störung die Hintergrundintensität im Mittel unverändert läßt. Durch das Potenzieren der Intensitätsfunktion u mit Exponent n gehen nun hohe Intensitäten stärker in Berechnung ein als niedrige, der Einfluß der Hintergrundintensität wird also reduziert.

Wir werden in der Notation der Größen $m(u)(x)$, $sp(u)(x)$ und $M(u)(x)$ im folgenden auf den Parameter u verzichten, wenn aus dem Kontext klar ist, welche Funktion zu ihrer Berechnung ausgewertet wird.

$M(u)$ ist symmetrisch und positiv semidefinit (Die positive Semidefinitheit folgt durch Approximation mit Treppenfunktionen, deren Integral in diesem Fall einer Linearkombination positiv semidefiniter Matrizen mit positiven Koeffizienten entspricht). Daher existiert eine Orthonormalbasis $v_i(u)(x_0), i = 1 \dots d$ aus Eigenvektoren von $M(u)(x_0)$ zu den Eigenwerten $\lambda_i(u)(x_0) \in \mathbb{R}_0^+$. Wir nehmen

o.B.d.A. $\lambda_1(u)(x_0) \geq \dots \geq \lambda_d(u)(x_0)$ an.

Um die Bedeutung der Matrix M und ihrer Eigenvektoren zu verstehen, betrachten wir zunächst den Trägheitstensor aus der Physik. (vgl.[16]). Seien im folgenden $d = 3$ und $\rho \in L^\infty(\Omega, [0, 1])$ eine Dichteverteilung auf Ω mit Gesamtmasse $m := \int_\Omega \rho(x) dx > 0$, die einen 'starren Körper' beschreibe, d.h. die Masse besitze eine gemeinsame Translations- bzw. Rotationsbewegung. Sei $sp := \frac{1}{m} \int_\Omega \rho(x) dx$ der Schwerpunkt der in Ω liegenden Masse. Definiere den Trägheitstensor T wie folgt:

$$T := \int_\Omega \rho(x) (|x - sp|^2 Id - (x - sp) \otimes (x - sp)) dx$$

Falls sich der starre Körper mit einer Winkelgeschwindigkeit ω um eine Achse $v \in \mathbb{R}^3$, $|v| = \omega$ dreht, ist die Rotationsenergie des rotierenden Körpers durch $W = \frac{1}{2} v^T T v$ gegeben. Da der Trägheitstensor symmetrisch ist, existiert eine Basis aus Eigenvektoren, die auch als Hauptträgheitsachsen des starren Körpers bezeichnet werden. Die zugehörigen Eigenwerte sind ein Maß für die Energie, die für eine Rotation um die betreffende Achse nötig wäre. Die Masse trägt dabei umso mehr zur Rotationsenergie bei, je weiter sie von der Rotationsachse entfernt ist. Bei einer Masse, die entlang einer Geraden $v_0 \in \mathbb{R}^3$ angeordnet ist, erwartet man also, daß v_0 ein Eigenvektor zum Eigenwert 0 ist. Kennen wir nur den Trägheitstensor T eines starren Körpers und besitzt T zwei Eigenvektoren mit relativ großen Eigenwerten und einen Eigenvektor mit relativ kleinem Eigenwert, dann kann davon ausgegangen werden, daß der starre Körper in etwa die Gestalt eines dünnen Stabes besitzt.

Das Massenmoment

$$\tilde{M} := \int_\Omega \rho(x) (x - sp) \otimes (x - sp) dx$$

besitzt dieselben Eigenvektoren wie T . Die Differenz der jeweiligen Eigenwerte ist ebenfalls gleich.

Betrachten wir nun im dreidimensionalen Fall das lokale Massenmoment $M(\rho)(x_0)$, so entsprechen dessen Eigenvektoren also den Hauptträgheitsachsen der in $B_\varepsilon(x_0)$ liegenden Masse.

Die Resultate lassen sich auf den zweidimensionalen Fall übertragen.

Die Differenz der Eigenwerte von $M(\rho)(x_0)$ ist dann ein Maß dafür, inwieweit die Masse in der betrachteten Umgebung linear angeordnet ist. Wir definieren daher die Kohärenz im Punkt $x \in \Omega$:

$$\text{coh}(u)(x) := \lambda_1(u)(x) - \lambda_2(u)(x) \tag{3.7}$$

Die *normierte* Kohärenz sei definiert durch

$$\overline{\text{coh}}(u)(x) := \begin{cases} 0 & \text{falls } \lambda_1(u)(x) = 0 \\ \frac{\text{coh}(u)(x)}{\lambda_1(u)(x)} & \text{sonst} \end{cases} \quad (3.8)$$

$\overline{\text{coh}}(u)(x)$ beschreibt die Anordnung der Masse in der ε -Umgebung von x wie folgt:

Bei $\overline{\text{coh}}(x) \approx 0$ liegt keine eindimensionale Struktur vor, bei $\overline{\text{coh}}(x) \approx 1$ ist die Masse stabförmig angeordnet. Die Anordnungsrichtung wird durch den ersten Eigenvektor v_1 beschrieben.

Im Fall $d = 3$ erhalten wir drei Eigenwerte von $M(u)(x)$. Im Fall einer stabförmig angeordneten Masse erwarten wir $\lambda_1 \gg 0$ und $\lambda_2, \lambda_3 \approx 0$. Ist die Masse dagegen nicht entlang einer Geraden angeordnet, so erhält man mindestens zwei Eigenwerte $\lambda_1, \lambda_2 > 0$, deren Differenz $\lambda_1 - \lambda_2$ klein ist. Aus diesem Grund verwenden wir oben vorgestellte Definition der Kohärenz auch im Fall $d = 3$ und erhalten dieselbe Aussage wie im zweidimensionalen Fall.

Eine Alternative besteht darin, die Größen

$$c_l := \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} \quad c_p := \frac{2\lambda_2 - \lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \quad c_i := \frac{3\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$$

mit $c_l, c_p, c_s \in [0, 1]$, $c_l + c_p + c_s = 1$ zu definieren, die beschreiben inwieweit die Masse linear, planar oder isotrop im dreidimensionalen Raum angeordnet ist (vgl.[14]).

Desweiteren definieren wir die Distanz $\text{dist}(x_0) := |sp(x_0)|$ des Punktes x_0 zum Schwerpunkt der Masse in $B_\varepsilon(x_0)$. Allein durch die Distanz läßt sich keine Aussage über die Anordnung der Masse in $B_\varepsilon(x_0)$ treffen. Nehmen wir jedoch an, daß die Masse in $B_\varepsilon(x_0)$ die Form eines dünnen Schlauches besitzt, so approximiert $\text{dist}(x_0)$ die Distanz von x_0 zum Schlauch.

Betrachten wir zwei Beispiele (vgl. Abb. 3.2):

Beispiel 3.1: Dünner Stab

Im ersten Beispiel betrachten wir einen dünnen Stab als Spezialfall eines dünnen Schlauches. Der Stab sei durch eine Lücke unterbrochen, die deutlich größer als das gewählte ε sei. Wir betrachten daher nur ein Stabende.

Wir wollen das Massenmoment M und dessen Eigenwerte und -vektoren in der Nähe des Stabendes bestimmen. Den Stab beschreiben wir durch eine Dichtefunktion ρ . Seien dazu $r > 0$ und $S := \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq r^2\} \times [0, \infty)$. Die Dichtefunktion sei gegeben durch $\rho := \chi_S$.

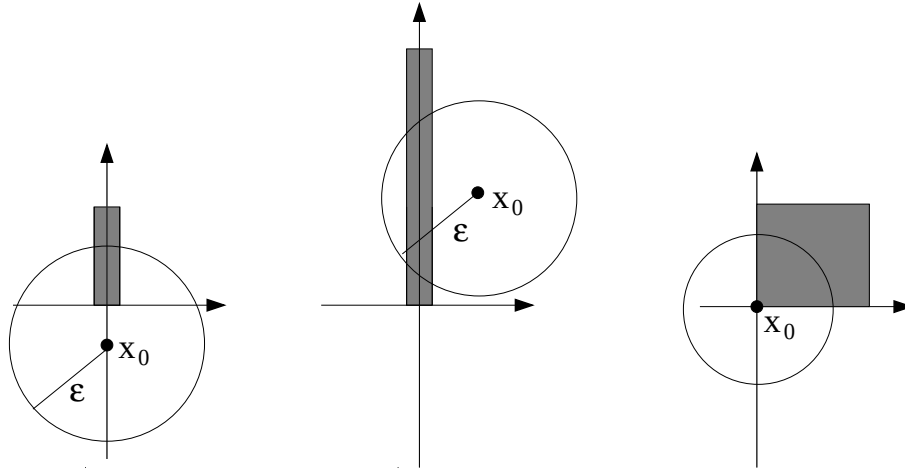


Abbildung 3.2: Beispiele für Dichteverteilungen: Ein dünner Stab, Ecke (Querschnitte)

Sei im folgenden $\varepsilon \gg r$ fest gewählt. Wir bestimmen zu ρ den Schwerpunkt $sp(x)$ und das Massenmoment $M(x)$ bezüglich der ε -Umgebung eines Punktes $x_0 = (0, 0, -a)$ $a \geq 0$, wobei a so gewählt sei, daß $a < \varepsilon$ sei. a ist also die Distanz zwischen Stab und Auswertungspunkt x (vgl. Abb.3.2).

Zur Vereinfachung approximieren wir $B_\varepsilon(x_0) \cap S$ durch $\tilde{S} := \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq r^2\} \times [0, \varepsilon - a]$, d.h. wir vergrößern das Integrationsgebiet geringfügig. Diese Änderung können wir, gemessen im Lebesgue-Maß, durch $\pi r^2 h(r)$ mit $h(r) := \varepsilon - \sqrt{\varepsilon^2 - r^2}$ abschätzen. Mit der Taylorformel erhält man die Abschätzung $|h(r)| \leq C(a, \varepsilon) r^2$ für $r \leq a < \varepsilon$. Der Integrationsfehler hängt also von r in vierter Ordnung ab.

Der Schwerpunkt ist dann gegeben durch $sp(x_0) = (0, 0, \frac{\varepsilon+a}{2})^T$, das Massenmoment durch $M(x_0) = \text{diag}(\frac{r^2}{4}, \frac{r^2}{4}, \frac{(\varepsilon-a)^2}{12})$.

Ist nun $a < \varepsilon - \sqrt{3}r^2$, d.h. das Massenmoment wird in der Nähe des Stabes ausgewertet, so gilt $\lambda_1(x_0) \gg \lambda_2(x_0) = \lambda_3(x_0)$ und der Eigenvektor $v_1(x_0) = (0, 0, 1)^T$ beschreibt die Ausrichtung des Stabes. Ist der Auswertungspunkt dagegen weiter entfernt von dem Stab, so sind wir nicht mehr in der Lage, diese Ausrichtung anhand von M zu erkennen, da nur noch ein kleiner Teil des Stabes in der ε -Umgebung liegt. Der Eigenvektor $v_1(x_0)$ approximiert in diesem Fall nicht die Ausrichtung des Stabes.

Die Distanz $\text{dist}(x_0) = \frac{\varepsilon+a}{2}$ approximiert den Abstand a von x_0 zum Stab i.A. nur ungenau.

Betrachtet wir nun einen Stab ohne Lücke. Der Einfachheit halber wählen wir dieselbe Funktion ρ , werten Schwerpunkt und Masse jedoch weit entfernt von der Lücke aus. Sei also $x_1 = (R, 0, h)$ einen Punkt mit $\varepsilon \gg R > r$, $h > \varepsilon$. Mit Approximation von $B_\varepsilon(X_0) \cap S$ durch $\{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq r^2\} \times [h - \varepsilon, h + \varepsilon]$ erhalten wir $sp(x_1) = (-R, 0, 0)$ und $v_1(x_1) || (0, 0, 1)$.

$v_1(x_1)$ beschreibt also die Ausrichtung des Stabes, $dist(x_1) = R$ den Abstand von x_1 zum Stab.

Beispiel 3.2

Im zweiten Beispiel betrachten wir nun eine dünne Platte mit einer Ecke. Wir werden auf dieses Beispiel in Abschnitt 3.5.2 zurückkommen. Sei $h > 0$ die Plattendicke und $D := \{(x, y, z) \in \mathbb{R}^3 | x, y \geq 0, |z| \leq h\}$. Die Dichtefunktion u sei gegeben durch $u = \chi_D$. Wir bestimmen $m(x)$, $sp(x)$ und $M(x)$ bezüglich einer ε -Umgebung des Punktes $x_0 = 0$ mit $\varepsilon > 0$.

$$\text{Es gilt } m(0) = \frac{h\pi\varepsilon^2}{2}, \quad sp(0) = \frac{4\varepsilon}{3\pi} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad M(0) = r^2 \begin{pmatrix} \varepsilon^2 a & \varepsilon^2 b & 0 \\ \varepsilon^2 b & \varepsilon^2 a & 0 \\ 0 & 0 & \frac{h^2}{3} \end{pmatrix}$$

mit $a = \frac{1}{4} + \frac{16}{3\pi^2} > b = \frac{1}{2\pi} + \frac{16}{9\pi^2} > 0$. Falls $h \ll \varepsilon$ gilt, so gilt die Eigenwerte $\lambda_1 = \varepsilon^2(a + b) > 0$, $\lambda_2 = \varepsilon^2(a - b) > 0$ und $\lambda_3 = \frac{h^2}{3}$ zu den Eigenvektoren $v_1 = \frac{1}{\sqrt{2}}(1, 1, 0)^T$, $v_2 = \frac{1}{\sqrt{2}}(-1, 1, 0)^T$ und $v_3 = (0, 0, 1)^T$ von $M(0)$: $\lambda_1 > \lambda_2 \gg \lambda_3$.

3.3 Diffusionsmodell

Das Diffusionsmodell beschreiben wir für den allgemeinen Fall $\mathbb{R}^d, d = 2, 3$. Sei Ω wie oben ein Gebiet in $\mathbb{R}^d$. Sei $u_0 \in L^\infty(\Omega, [0, 1])$ eine Intensitätsfunktion, die Rauschen beinhalte. Das Diffusionsmodell soll so konzipiert sein, daß in u_0 lineare Strukturen hoher Intensität durch den Diffusionsprozess erkannt und hervorgehoben werden. Dazu müssen die Ausgangsbilder eventuell invertiert werden. Gleichzeitig soll durch die Diffusion das Rauschen weggeglättet werden.

3.3.1 Differentialgleichung

Wir greifen nun das Diffusionsmodell (2.1) aus Abschnitt 2.2.1 auf. Sei $\partial\Omega$ der Rand von Ω und $n : \partial\Omega \rightarrow \mathbb{R}^2$ die zugehörige äußere Normale. Sei $f(u) \in L^\infty([0, T] \times \Omega, \mathbb{R})$ für beliebiges $u \in L^2([0, T] \times \Omega, \mathbb{R})$.

Gesucht ist eine schwache Lösung $u \in L^2((0, T], H^{1,2}(\Omega))$ folgender Differentialgleichung:

$$\begin{aligned}
\partial_t u - \operatorname{div}(D_\varepsilon(u) \nabla u) &= f(u) && \text{in } [0, T] \times \Omega \\
u(0) &= u_0 && \text{in } \Omega \\
\langle D_\varepsilon(u) \nabla u, n \rangle &= 0 && \text{auf } [0, T] \times \partial\Omega
\end{aligned} \tag{3.9}$$

Die Matrix $D_\varepsilon(u)$ sei wie folgt definiert:

Seien $V_1 \dots V_d : \Omega \rightarrow \mathbb{R}^d$ Vektorfelder, so daß zu jedem $x \in \Omega$ $V_1(x) \dots V_d(x)$ eine Orthonormalbasis aus den Eigenvektoren der Massenmoments $M(x)$ darstellen. Setze

$$B := \begin{pmatrix} V_1(x) & \dots & V_d(x) \end{pmatrix}$$

und

$$D = B \begin{pmatrix} 1 & 0 \\ 0 & g(\operatorname{coh}(u) |\nabla u_\sigma|^2) \mathbf{1}_{(d-1) \times (d-1)} \end{pmatrix} B^T$$

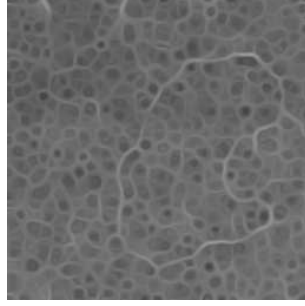
Hier ist $g(s) := \frac{1}{1 + \frac{s}{\kappa^2}}$ die aus dem Perona-Malik-Modell bekannte Funktion. $\operatorname{coh}(u)$ sei wie im vorigen Abschnitt definiert.

Weickert setzt in seinem Modell als Argument von g den größeren Eigenwert des Strukturtenors ein (vgl. Abschnitt 2.2.6). Dieser entspricht im ungeglätteten Fall ($\rho = 0$) $|\nabla u_\sigma|^2$. Durch die gradientengesteuerte Diffusion wird eine Kantenschärfung möglich. Um Kantenschärfung auch in unserem Fall zu erreichen, verfahren wir analog zum Perona-Malik-Modell und verwenden $|\nabla u_\sigma|^2$ als Argument von g . Es soll aber nur dort Kantenschärfung auftreten, wo eine schlauchartige Struktur erkannt wird. Wir fügen daher zu $|\nabla u_\sigma|^2$ den Faktor $\operatorname{coh}(u)$ hinzu. Im Fall $\operatorname{coh}(x) = 0$, d.h. $\lambda_1 = \lambda_2$, gilt dann $D_\varepsilon(u)(x) = \mathbf{1}$. Die Gestalt von $D_\varepsilon(u)(x)$ ist dann unabhängig von der Numerierung der Eigenwerte $\lambda_i, i = 1, 2$.

Modelle, in denen die Diffusionskoeffizienten in Abhängigkeit vom berechneten Eigenwert λ_1 oder der Kohärenz gewählt wurden, haben in der Praxis keine so gute Kantenschärfung wie die Kopplung von g mit $|\nabla u_\sigma|$ geliefert.

Die Diffusionsmatrix $D_\varepsilon(u)$ ist nun so gewählt, daß in Richtung des Vektorfeldes $V_1(x)$, also entlang der erkannten schlauchartigen Strukturen eine Diffusion mit konstantem Koeffizienten, in orthogonale Richtung in Abhängigkeit von Kohärenz und Gradienten eine glättende oder eine kantenschärfende Diffusion stattfindet. Wir bezeichnen $V_1(x)$ im folgenden als *Hauptdiffusionsrichtung*.

Die Diffusion wird nun zu einer Verstärkung von Strukturen führen, die "heller" als ihre Umgebung ist, d.h. daß $u(t)$ an diesen Stellen $x \in \Omega$ eine höhere Intensität als auf einer Umgebung besitzt. Insbesondere wird die Diffusion infolge der Störung entstandene Lücken in den Strukturen schließen.

Abbildung 3.3: Diffusion mit rechter Seite $f = 0$

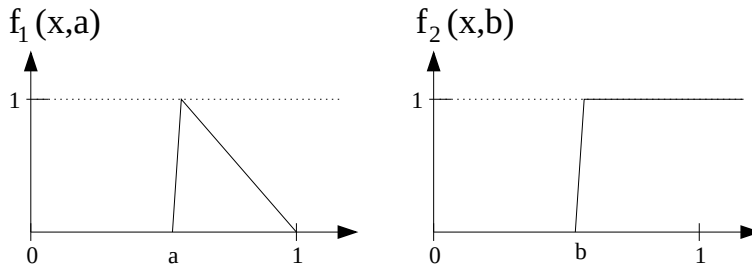
3.3.2 Rechte Seite

Die hier vorgestellte Differentialgleichung hat die Eigenschaft, daß bei einer rechten Seite $f = 0$ die Lösungen mit wachsendem t an Kontrast verliert.

Zur Erkennung der schlauchartigen Strukturen ist ein hoher Kontrast jedoch wünschenswert. Wir werden daher eine rechte Seite $f \neq 0$ definieren, die die Lösung an erkannten Strukturen auf den Wert 1 zwingt. Gleichzeitig soll die Lösung auf das Bildintervall $[0, 1]$ beschränkt bleiben.

Sei $u(t)(x) \in L^2([0, T] \times \Omega)$. Zur Erkennung heller schlauchartiger Strukturen wählen wir $c_1 \in [0, 1]$, $c_2 > 0$ und definieren $K^1(u(t)) := \{x \in \Omega | u(t)(x) > c_1 \text{ und } \text{coh}(u(t))(x) > c_2\}$. Die Menge K^1 umfaßt also alle Punkte mit einer genügend hohen Intensität, in denen lineare Strukturen erkannt werden. $f(u(t))$ soll nun an einem Punkt $x \in K^1$ größer null sein, falls $u(t)(x)$ kleiner als 1 ist, sonst gleich null. Auf diese Weise stellen wir sicher, daß $u(t)(x) \leq 1$ fast überall auf $[0, T] \times \Omega$. Außerhalb von $K^1(u(t))$ sei f ebenfalls null.

Um die Existenz und Eindeutigkeit der Lösung der Differentialgleichung zu garantieren, wählen wir nun die rechte Seite der Differentialgleichung so, daß sie in jedem Punkt $x \in \Omega$ Lipschitz-stetig von $u(t)(x)$ und $\text{coh}(u(t))(x)$ abhängt.

Abbildung 3.4: Die Funktionen f_1, f_2

Seien dazu die Hilfsfunktionen f_1, f_2 zu beliebigen $\delta > 0$, $0 < a < 1 - \delta$ und $b > 0$ wie folgt definiert (vgl. Abb. 3.4):

$$f_1(x, a) := \begin{cases} 0 & \text{falls } x \leq a \\ \frac{x-a}{\delta}(1-a-\delta) & \text{falls } x \in (a, a+\delta) \\ 1-x & \text{falls } x \in [a+\delta, 1] \\ 0 & \text{falls } x > 1 \end{cases}$$

$$f_2(x, b) := \begin{cases} 0 & \text{falls } x \leq b \\ \frac{x-b}{\delta} & \text{falls } x \in (b, b+\delta) \\ 1 & \text{falls } x \geq b+\delta \end{cases}$$

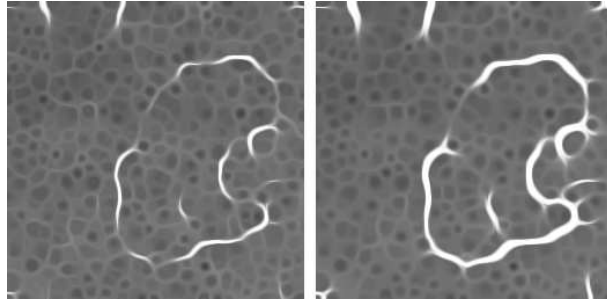


Abbildung 3.5: Diffusion mit rechter Seite $f \neq 0$ zu verschiedenen Zeitpunkten

Zu $0 < c_1 < 1$ sei das oben definierte δ so gewählt, daß $c_1 + \delta < 1$ gelte. Sinnvoll ist hier eine Wahl $c_1 \in (c_{\text{Hintergrund}}, c_{\text{Objekt}}]$, da wir nach Möglichkeit nur die Punkte hervorheben wollen, die zum gesuchten Schlauch gehören. Sei $c_2 > 0$. Wir definieren $f(u) = f_\varepsilon(u) \in L^\infty([0, T] \times \Omega)$ durch

$$f_\varepsilon(u)(t, x) := c f_1(u(t, x), c_1) f_2(\text{coh}(u(t)))(x), c_2) \quad (3.10)$$

mit einem beliebigen Faktor $c \geq 0$. Nach der obigen Definition hängt $f_\varepsilon(u)(t_0, x)$ für festes t_0 nicht nur von $u(t_0, x)$, sondern auch von $\text{coh}(u(t_0))(x)$ und somit von $u(t_0, x)|_{B_\varepsilon(x)}$ ab.

Da $f_\varepsilon(u)(x) = 0$, falls $u(x) \geq 1$, gilt nach dem Maximumsprinzip $u(t, x) \in [0, 1]$ fast überall auf $[0, T] \times \Omega$.

Wir nennen $K^1(u(t))$ im folgenden die zum Zeitpunkt t erkannten Punkte.

Abb. 3.5 zeigt, wie durch diese Kontrastschärfung die gesuchte Struktur hervorgehoben wird. Die anisotrope Diffusion sorgt dafür, daß sich die Lücken mit der Zeit schließen.

Zum Beweis der Existenz und Eindeutigkeit einer Lösung der Differentialgleichung (3.9) mit der hier vorgestellten rechten Seite verweisen wir auf Kapitel 4.

3.3.3 Eine weitere Modifikation der rechten Seite

Sei u_0 im folgenden eine der ungestörten Funktionen $u_i, i = 1, 2$ aus Abschnitt 3.1.1 und $\tilde{u}_0$ die zugehörige gestörte Funktion. Sei $K := \{x \in \Omega | u_0(x) > c_{Objekt}\}$ das ungestörte Abbild des Schlauches. Die Menge K soll nun durch Anwenden der oben beschriebenen Differentialgleichung im Fall $\mathbb{R}^d, d = 2$ hervorgehoben werden.

Im vorherigen Abschnitt hatten wir

$K^1(w) := \{x \in \Omega | w(x) > c_1 \text{ und } \text{coh}(w)(x) > c_2\}$ für beliebiges $w \in L^2(\Omega)$ zur Definition der rechten Seite benutzt. $K^1(\tilde{u}_0)$ ist eine erste Approximation der Menge K zum Zeitpunkt $t = 0$.

In der Praxis hat sich gezeigt, daß diese Approximation noch verbessert werden kann. Entsprechend werden wir $f(u)$ für den Zeitpunkt $t = 0$ mit Hilfe der verbesserten Approximation definieren. Wir stellen nun diese Modifikation vor.

Sei also im folgenden $d = 2$. Wir wollen voraussetzen, daß $u_0, \tilde{u}_0$ in $C^\infty(\Omega)$ liegen. Dies würde nach Faltung mit einem Gaußkern gelten. Um unser Vorgehen besser verdeutlichen zu können, nehmen wir jedoch an, daß u_0 folgende Gestalt hat:

$$u_0(x) = \begin{cases} c_{Objekt} & \text{für } x \in K \\ c_{Hintergrund} & \text{für } \text{dist}(x, K) \geq \varepsilon \end{cases}$$

u_0 ist also in diesen Teilgebieten konstant.

Auf $\Delta K := \{x \in \Omega | 0 < \text{dist}(x, K) < \varepsilon\}$ sei u_0 geeignet fortgesetzt, so daß $u_0 \in C^\infty(\Omega)$. $\tilde{u}_0$ sei wie oben definiert durch Störung der Funktion u_0 und liege ebenfalls in $C^\infty(\Omega)$.

Wir betrachten nun die Ableitungen von u_0 bzw. $\tilde{u}_0$.

Nach Definition der ungestörten Funktion u_0 verschwindet die erste Ableitung sowohl auf K als auch auf dem Hintergrund $\Omega \setminus (K \cup \Delta K)$ und ist daher zur Unterscheidung beider Mengen ungeeignet.

Betrachten wir nun die gestörte Funktion $\tilde{u}_0$, so können wir daher nicht erwarten, daß $|\nabla \tilde{u}_0|$ ein gutes Kriterium zur Unterscheidung zwischen Objekt K und dem Hintergrund $\Omega \setminus (K \cup \Delta K)$ ist.

Mit Hilfe der zweiten Ableitungen können wir die Oberflächenkrümmung von u_0 bzw. $\tilde{u}_0$ berechnen.

Sei für beliebiges $v \in C^2(\Omega)$ die Matrix Q definiert durch

$$Q(v)(x) := -\frac{1}{\sqrt{1 + (\partial_x v)^2 + (\partial_y v)^2}} \begin{pmatrix} \partial_{xx} v & \partial_{xy} v \\ \partial_{xy} v & \partial_{yy} v \end{pmatrix} \Big|_x$$

Seien $\lambda^1(w)(x)$ und $\lambda^2(w)(x)$ die Eigenwerte von $Q(v)(x)$. Diese existieren aufgrund der Symmetrie von Q . O.B.d.A. gelte $\lambda^1(v)(w) \geq \lambda^2(w)(x)$.

Für u_0 gilt offensichtlich:

$$\begin{aligned} \lambda^1(u_0)(x) &>> \lambda^2(u_0)(x) \geq 0 && \text{falls } x \in K \\ \lambda^1(u_0)(x) &= 0 \quad \lambda^2(u_0)(x) = 0 && \text{falls } x \text{ außerhalb einer } \varepsilon\text{-Umgebung von } K \end{aligned}$$

Für die gestörte Funktion $\tilde{u}_0$ erwarten wir ähnliche Eigenschaften von $Q(\tilde{u}_0)$. Um den Einfluß von Rauschen zu minimieren und die Informationen zu mitteln, glätten wir nun $Q(v)$ komponentenweise: $Q_\sigma(v) := G_\sigma * Q(v)$. $Q_\sigma(v)$ ist ebenfalls symmetrisch. Seien $\lambda_\sigma^1(w), \lambda_\sigma^2(w)$ die zugehörigen Eigenwerte mit $\lambda_\sigma^1(w) \geq \lambda_\sigma^2(w)$.

Wir wählen nun $\lambda_\sigma^1(w)$ als Kriterium zur Erkennung von K und definieren $K^2 := \{x \in \Omega \mid w(x) > c_1, \lambda_\sigma^1(w)(x) > c_3\}$, wobei c_1 dasselbe wie in der Definition von $K^1(w)$ ist und $c_3 > 0$. Die rechte Seite der Differentialgleichung zum Zeitpunkt $t = 0$ definieren wir durch:

$$f(u)(0, x) := c f_1(u(0, x), c_1) f_2(\lambda_\sigma^1(u(0))(x), c_3) \quad (3.11)$$

mit beliebigem $c \geq 0$. Abb. 3.6 zeigt eines der Originalbilder sowie die Bildpunkte entsprechend den Mengen K^1 und K^2 . Durch die in diesem Abschnitt vorgestellten Kriterien wird die DNA-Struktur deutlich besser erkannt.

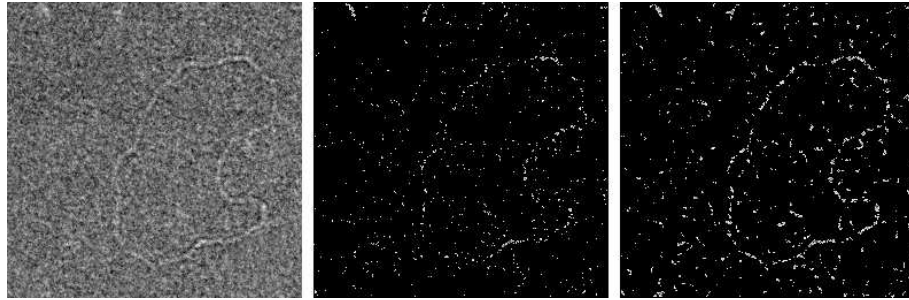


Abbildung 3.6: Links: invertiertes Originalbild, Mitte: die Menge K^1 , Rechts: die Menge K^2

Für Zeitpunkte $t > 0$ sind die erkannten schlauchartigen Strukturen bereits soweit hervorgehoben, daß wir die im vorigen Abschnitt definierte rechte Seite verwenden können.

3.4 3D-Rekonstruktion

3.4.1 Rekonstruktionsmodell

Wir wollen nun das Modell vorstellen, mit dem wir aus zwei gegebenen 2D-Funktionen eine 3D-Funktion rekonstruieren.

Wir verwenden das Kameramodell der orthogonalen Projektion und setzen voraus, daß die Kameraparameter bekannt sind. Seien $P_i : \Omega^{3D} \rightarrow \Omega_1, i = 1, 2$ die zugehörigen orthogonalen Projektionen eingeschränkt auf Ω^{3D} .

Wir definieren nun einen Operator

$$R : L^\infty(\Omega_1, [0, 1]) \times L^\infty(\Omega_2, [0, 1]) \rightarrow L^\infty(\Omega^{3D}, [0, 1]),$$

den wir im weiteren Sinne als Umkehrung der Projektionsoperatoren Π_i interpretieren können.

Zu $f_i \in L^\infty(\Omega_i, [0, 1]), i = 1, 2$ sei $R(f_1, f_2)(x)$ definiert durch

$$R(f_1, f_2)(x) := \min(f_1(P_1(x)), f_2(P_2(x)))$$

Dieses Rekonstruktionsmodell wertet die vorliegenden Bilder punktweise aus. Denkbar wäre ein Modell, daß zur Rekonstruktion eines Punktes $x \in \Omega^{3D}$ geeignete Umgebungen der Projektionen $P_i(x)$ betrachtet.

Das Projektionsmodell ist nicht in der Lage, das Korrespondierenden-Problem eindeutig zu lösen. Wie wir in den Beispielen in Kapitel 2 gesehen haben, entstehen bei der Rekonstruktion aufgrund der Nichteindeutigkeit mehr Strukturen, als tatsächlich vorhanden sind. Im nächsten Abschnitt werden wir sehen, welche Probleme dadurch bei der Rekonstruktion verrauschter Bilder auftreten.

3.4.2 Rekonstruktion verrauschter Bilder

In diesem Abschnitt wollen wir darlegen, warum wir uns in Abschnitt 3.1.2 dafür entschieden haben, zunächst die Originaldaten zu glätten und dann eine Funktion in $\mathbb{R}^3$ zu rekonstruieren.

Seien $u_i, i = 1, 2$, wie in Abschnitt 3.1.2 definiert, zwei (ungestörte) Abbilder eines dreidimensionalen Schlauchs, $\tilde{u}_i$ die entsprechend gestörten Intensitätsfunktionen.

Eine Rekonstruktion der Originaldaten mit anschließender Diffusion in $\mathbb{R}^3$ hat den Vorteil, daß in der richtigen Metrik geglättet wird. Die Diffusion kann entlang einer schlauchartigen Struktur erfolgen. Wendet man dagegen den zweidimensionalen Diffusionsprozeß auf die Funktionen u_i an, so vernachlässigt man die "Tiefe" der Struktur, die durch die Projektion verloren gegangen ist. Besonders bemerkbar macht sich dies an Kreuzungspunkten. Hier überlagern sich die Abbildungen

zweier Teile des 3D-Schlauchs, die in $\mathbb{R}^3$ einen gewissen Abstand zueinander haben, der jedoch durch die 2D-Diffusion nicht berücksichtigt werden kann. Wie wir in Abschnitt 6.4 sehen werden, verfälscht die zweidimensionale Glättung an diesen Kreuzungspunkten die 3D-Rekonstruktion.

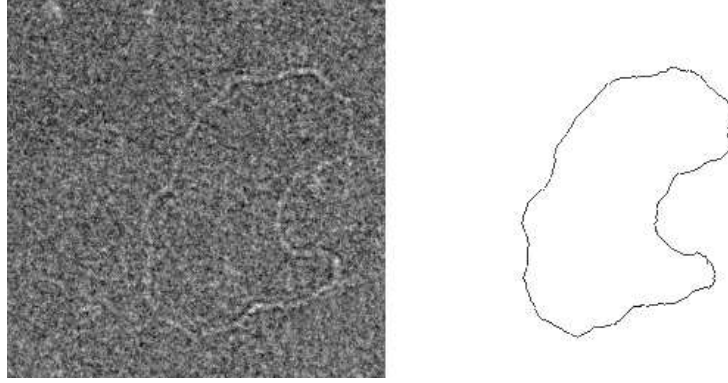


Abbildung 3.7: Invertiertes Originalbild und approximierter Verlauf des DNA-Rings

Betrachten wir die 3D-Rekonstruktion basierend auf den verrauschten Funktionen $\tilde{u}_i$. Das oben beschriebene Diffusionsmodell ist so konzipiert, daß helle lineare Strukturen hervorgehoben werden. Soll die Diffusion nach der 3D-Rekonstruktion erfolgen, so ist es wichtig, daß die in den 2D-Funktionen erkennbaren Strukturen bei der Rekonstruktion erhalten bleiben und keine zusätzlichen Strukturen aus dem Rauschen rekonstruiert werden. Wir werden dies im folgenden genauer untersuchen.

In unserem Modell nehmen die Funktionen u_i Werte größer gleich c_{Objekt} auf den Abbildern des Schlauchs, sonst den Wert $c_{\text{Hintergrund}}$ an. Wäre die Störung moderat, d.h. $|u_i - \tilde{u}_i| \ll c_{\text{Objekt}} - c_{\text{Hintergrund}}$, so ließe sich ein Schwellenwert $c_{\text{thres}} \in (c_{\text{Hintergrund}}, c_{\text{Objekt}})$ finden, so daß die Abbilder des Schlauchs durch $\tilde{K}_i := \tilde{u}_i^{-1}((c_{\text{thres}}, \infty))$ bestimmen ließen. Rekonstruierten wir den dreidimensionalen Datensatz $R(\tilde{u}_1, \tilde{u}_2)$, so würden wir mit $R^{-1}((c_{\text{thres}}, \infty))$ ein Objekt erhalten, daß ähnlich wie in Beispiel 2.2 aus dem gesuchten Schlauch und aufgrund des Nichteindeutigkeitsproblems aus zusätzlichen Strukturen besteht.

Eine Betrachtung der vorliegenden DNA-Bilder zeigt, daß wir diese Moderatheit der Störung nicht vorausgesetzt können. Vielmehr müssen wir annehmen, daß die Störung stark ist, d.h. an manchen Stellen des Hintergrundes die Intensität von $\tilde{u}_i$ höher ist als auf dem Objekt und umgekehrt die $\tilde{u}_i$ an Punkten auf dem Objekt einen niedrigeren Wert als auf dem Hintergrund annehmen.

Betrachten wir als Beispiel das DNA-Bild in Abb 3.7. Wir approximieren den Verlauf des DNA-Rings wie abgebildet. Abb. 3.8 zeigt die normierten Grauwertverteilungen auf dem Ring sowie auf dem Hintergrund. Beide Verteilungen haben die Form von Gaußverteilungen mit im Vergleich zur den Varianzen dicht beieinanderliegenden Mittelwerten. Man erkennt, daß der Grauwert vieler Hintergrundpunkte über dem Mittelwert des Objekt und umgekehrt der Grauwert vieler Objektpunkte unter dem Mittel des Hintergrundes liegt.

Seien K_i (wie oben) die Abbilder des gesuchten Rings. Die Punkte der oben definierten Mengen $\tilde{K}_i$ nennen wir die erkannten Punkte. In unserem Modell müssen wir also davon ausgehen, daß die Mengen $\tilde{K}_i$ die zweidimensionalen Abbilder K_i nur lückenhaft beschreiben und zusätzlich Hintergrundpunkte enthalten. Punkte aus K_i , die nicht in $\tilde{K}_i$ liegen, nennen wir im folgenden *Lücken*, Punkt in $\tilde{K}_i$, die nicht zu K_i gehören, *falsch erkannte Punkte*.

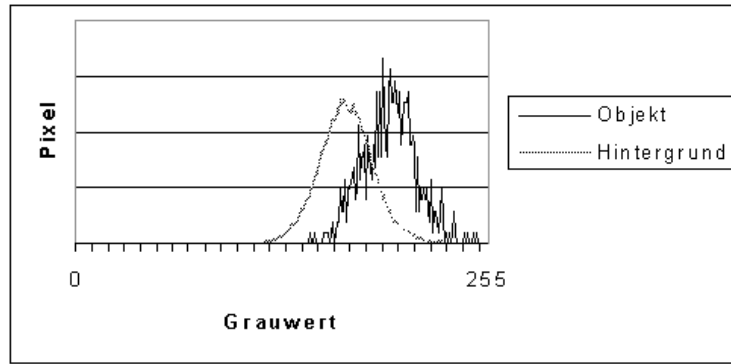


Abbildung 3.8: Grauwertverteilung von Objekt und Hintergrund

Betrachten wir nun nun die Menge $\tilde{\Gamma} := R(\tilde{u}_1, \tilde{u}_2)^{-1}[c_{thres}, \infty)$. Die Punkte aus $\tilde{\Gamma}$ nennen wir die rekonstruierten Punkte. Rekonstruierte Punkte, die nicht zu der gesuchten Struktur gehören, bezeichnen wir als falsch rekonstruierte Punkte. Punkte des gesuchten Rings, die nicht in $\tilde{\Gamma}$ liegen, nennen wir wie im zweidimensionalen Fall Lücken.

Wir rekonstruieren einen Punkt $M \in \Omega^{3D}$ nur dann, wenn die Projektionen $P_i M$ in $\tilde{K}_i$ liegen, d.h. wenn sie erkannte Punkte sind. Offensichtlich ist ein Punkt $M \in \Omega^{3D}$ ein falsch rekonstruierter Punkt, falls nur einer der beiden Punkte $(P_1 M, P_2 M)$ ein falsch erkannter Punkt ist.

Da in den verrauschten Bildern eine Vielzahl von Punkten falsch erkannt werden ist, rekonstruieren wir also in Ω^{3D} viele falsche Punkte. Diese sind oft entlang einer der Projektionsgeraden angeordnet. Solche Punkte werden eventuell bei einer anschließenden Diffusion als Struktur erkannt und diese Struktur wird dann

weiter ausgeprägt.

Wir können einen Punkt M des gesuchten Objekts *nicht* rekonstruieren, falls nur einer der Punkte $P_i M$ eine Lücke darstellt. Die Zahl der Lücken in $\tilde{\Gamma}$ ist also wesentlich größer als die Anzahl der Lücken in den Abbildern des Objekts.

Insgesamt ist also die Qualität der rekonstruierten Funktion $R(\tilde{u}_1, \tilde{u}_2)$ schlechter als die der zweidimensionalen Funktionen $\tilde{u}_i$. Wir werden sehen, daß wir bereits bei der Erkennung der Strukturen in den 2D-Bildern auf Probleme stoßen. Daher erwarten wir nicht, daß eine Erkennung der dreidimensionalen Strukturen nach der Rekonstruktion verrauschter Bilder erfolgreich durchgeführt werden kann. Wir wählen daher die erste Alternative, zuerst die zweidimensionalen Bilder vorzuglätten und anschließend die 3D-Rekonstruktion durchzuführen. Abb. 3.9 zeigt solche vorgeglätteten Bilder. Wir erwarten, daß nach der Vorgeglättung die Abbilder des Objekts deutlich zu erkennen sind. In der aus den vorgeglätteten Funktionen rekonstruierte Funktion $S^* := R(u_1(T), u_2(T))$, ist nun deutlich das gesuchte schlauchförmige Objekt zu erkennen. Gleichzeitig entstehen durch das Vorgeglätten und die Nichteindeutigkeit der Rekonstruktion zusätzliche Strukturen. Abb. 3.10 zeigt die Oberfläche einer Menge $(S^*)^{-1}([c_{thres}, 1])$ zu einem geeigneten $c_{thres} > 0$. Die zu Beginn des Abschnitts dargelegten Nachteile dieser Reihenfolge von Diffusion und Rekonstruktion nehmen wir in Kauf.

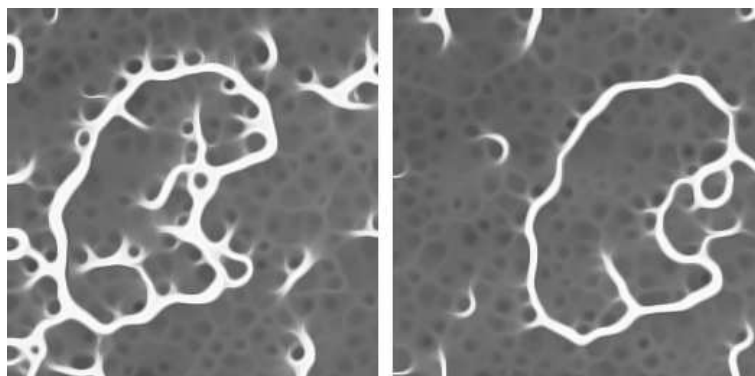


Abbildung 3.9: Geglättete Bilder



Abbildung 3.10: Ansichten der rekonstruierten Menge aus verschiedenen Perspektiven

3.5 Tracing

Wie wir im vorherigen Abschnitt gesehen haben, enthält eine aus den geglätteten Funktionen $u_i(T)$, $i = 1, 2$ rekonstruierte Menge $R(u_1(T), u_2(T))^{-1}([c, 1])$ neben dem gesuchten schlauchförmigen Objekt Γ weitere Strukturen. Wir wollen nun versuchen, den Schlauch Γ zu isolieren. Dazu stellen wir eine gewöhnliche Differentialgleichung vor, deren Lösung dem gesuchten Schlauch folgt. Zu einem geeigneten ε stellt der ε -Schlauch um diese Lösung dann eine Approximation von Γ dar. Abb. 3.11 zeigt den ε -Schlauch um eine solche Lösungskurve.

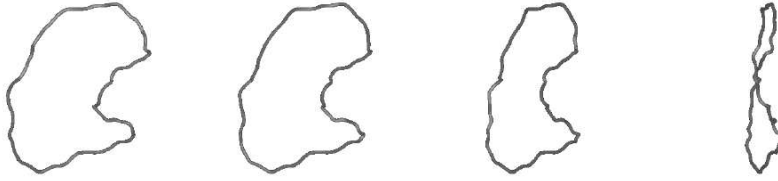


Abbildung 3.11: Verschiedene Ansichten eines dreidimensionalen Polygonzugs

3.5.1 Gewöhnliche Differentialgleichung

Seien $\Omega^{3D} \subset \mathbb{R}^3$ und $S^* \in L^\infty(\Omega^{3D}, [0, 1])$. Wir nehmen an, daß ein $c_{thres} \in [0, 1]$ existiert, so daß der gesuchte Schlauch Γ in $\tilde{\Gamma} := (S^*)^{-1}[c_{thres}, 1]$ enthalten ist. Sei $x_0 \in \Omega$ ein bekannter Punkt des gesuchten Objekts.

Wir wollen eine Differentialgleichung vorstellen, die die Bewegung eines Massenpunktes beschreibt. Die Beschleunigung soll so gesteuert werden, daß der Massenpunkt den gesuchten Schlauch in $\tilde{\Gamma}$ durchläuft, bis er den Ausgangspunkt (bis auf einen $\tilde{\varepsilon}$ -Abstand) wieder erreicht, so daß wir eine geschlossene Kurve erhalten. Mit $y : [t_0, t_1] \rightarrow \Omega^{3D}$ bezeichnen wir die Trajektorie dieses Punktes.

Ob der Massenpunkt den Ausgangspunkt wieder erreicht, bestimmen wir durch folgendes Kriterium: Zu vorgegebenem $\tilde{\varepsilon} > 0$ existieren $\tau_0 < \tau_1 \in (t_0, t_1)$ mit $|y(t_0) - y(\tau_0)| >> \tilde{\varepsilon}$ und $|y(t_0) - y(\tau_1)| < \tilde{\varepsilon}$. Wir verwenden dann $y|_{[t_0, \tau_1]}$ zur Approximation des gesuchten Schlauches. Wir werden dies in Abschnitt 3.5.3 detailliert beschreiben.

An Punkten, an denen schlauchartige Strukturen aneinanderstoßen oder sich durchschneiden, erwarten wir, daß der Massenpunkt aufgrund seiner Trägheit $\tilde{\Gamma}$ möglichst gradlinig durchläuft, also nicht in andere Strukturen abzweigt.

Betrachten wir zunächst den allgemeinen Fall eines in einem Kraftfeld $F(x) : \Omega^{3D} \rightarrow \mathbb{R}^3$ beschleunigten Massenpunktes $y(t)$. Die zugehörige Differentialgleichung ist $\ddot{y} = F(x)$. Als Anfangswerte wählen wir den Startpunkt x_0 und

einen Vektor $r_0 \neq 0 \in \mathbb{R}^3$ für die Startgeschwindigkeit.

Im allgemeinen wird die Kraft F einen nicht verschwindenden Anteil in Bewegungsrichtung beinhalten, so daß die Geschwindigkeit zunimmt. Damit die Geschwindigkeit des Massenpunktes beschränkt bleibt, fügen wir einen weiteren Term hinzu. Die Geschwindigkeit des Massenpunktes soll nahe einer vorgegebenen Geschwindigkeit v_0 liegen. Der Term, den wir zufügen, soll zu einer positiven Beschleunigung in Bewegungsrichtung führt, falls $\dot{y}(t)$ kleiner als v_0 , und zu einer negativen, falls $\dot{y}(t)$ größer als v_0 ist.

Wir erhalten folgende Differentialgleichung:

Seien $0 < t_0 < t_1$, $\alpha > 0$. Gesucht ist $y(t) : [t_0, t_1] \rightarrow \Omega$ mit

$$\begin{aligned} \ddot{y} + \alpha (|\dot{y}| - v_0)\dot{y} &= F(y, \dot{y}) \\ y(t_0) &= x_0 \\ \dot{y}(t_0) &= r_0 \end{aligned} \quad (3.12)$$

F soll nun so gewählt werden, daß die Lösung $y(t)_{t \geq 0}$ eindimensionalen Strukturen in S^* folgt. Sei zu festem $\varepsilon > 0$ $v(x)$ der Eigenvektor von $M(S^*)(x)$ zum größten Eigenwert $\lambda_1(x)$. Sei $\text{coh}(x) = \text{coh}(S^*)(x)$. Aus Abschnitt 3.2 wissen wir, daß v tangential an schlauchartige Strukturen anliegt. Ein erster Ansatz wäre

$$F(x, r) = F(x) := \text{coh}(x)v(x)$$

Hierbei ist zu beachten, daß $v(x)$ nur bis auf das Vorzeichen bestimmt ist. Wir definieren daher die Projektion auf den Eigenraum zu $\lambda_1(x)$ durch $P(x)r := \langle v(x), r \rangle v(x)$ für beliebiges $r \in \mathbb{R}^3$ ($v(x)$ ist nach Definition normiert), und gehen zu $F(x, r) := \text{coh}(x)P(x)r$ über.

Falls λ_1 gleich λ_2 , ist die Wahl von v nicht eindeutig. Durch den Faktor $\text{coh}(x)$ ist jedoch $F(x, r)$ eindeutig bestimmt.

Die Trägheit des Massenpunktes führt dazu, daß auf den Massenpunkt an gekrümmten Strukturen eine Zentrifugalkraft wirkt. Um zu verhindern, daß $y(t)$ die Menge $\tilde{K}$ verläßt, fügen wir zur Funktion F einen weiteren Term zu. Dazu betrachten wir den Schwerpunkt $sp(x) := sp(S^*)(x)$ in der ε -Umgebung von x . Der Massenpunkt $y(t)$ soll nun zusätzlich in Richtung des Schwerpunkts $sp(y(t))$ beschleunigt werden. Wir setzen daher:

$$F(x, r) := \begin{cases} \beta \text{coh}(x)P(x)r + \gamma sp(x)(t) & \text{für } x \in \Omega \\ 0 & \text{für } x \notin \Omega \end{cases} \quad (3.13)$$

mit $\beta, \gamma > 0$.

Die rechte Seite (3.13) der Differentialgleichung hat also die Form $F(x, r) = F_1(x)r + F_2(x)$, wobei $F_1(x) := \beta \text{coh}(x)P(x)$ eine matrixwertige, $F_2(x) :=$

$\gamma sp(x)(t)$ eine vektorwertige Funktion ist. Es existieren Konstanten $C_1, C_2 > 0$, so daß in der Operatornorm zur 2-Norm $\sup_{x \in \Omega^{3D}} \|F_1(x)\|_{Op} < C_1$ und in der 2-Norm auf $\mathbb{R}^3$ $\sup_{x \in \Omega^{3D}} \|F_2(x)\|_2 < C_2$ gilt (vgl. Lemma 4.1 in Kapitel 4). Wendet man auf beide Seiten der Gleichung (3.12) das Funktional $\langle \cdot, \dot{y} \rangle$ an, so erhält man (mit $F(x, r) = F_1(x)r + F_2(x)$):

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\dot{y}\|^2 &= \langle F_1(y)\dot{y}, \dot{y} \rangle + \langle F_2(y), \dot{y} \rangle + \alpha v_0 \|\dot{y}\|^2 - \alpha \|\dot{y}\|^3 \\ &\leq \|F_1(y)\|_{Op} \|\dot{y}\|_2^2 + \|F_2(y)\|_2 \|\dot{y}\|_2 + \alpha v_0 \|\dot{y}\|^2 - \alpha \|\dot{y}\|^3 \end{aligned} \quad (3.14)$$

Mit der Beschränktheit der F_i folgt nun, daß die rechte Seite negativ ist, falls $\|\dot{y}\|$ genügend groß.

Daher ist die Geschwindigkeit des Massenpunktes nach oben beschränkt.

Mit Hilfe der Parameter β und γ können wir sowohl die Stärke der Gesamtbeschleunigung als auch den Anteil der Beschleunigungen in tangentialer Richtung zur Struktur und in Richtung des Schwerpunktes festlegen. Mit den Parametern β und v_0 kann die durchschnittliche Geschwindigkeit des Massenpunktes beeinflußt werden. Wir sind einerseits daran interessiert, daß der Massenpunkt die Struktur möglichst gradlinig durchläuft. Daher ist eine hohe Geschwindigkeit sinnvoll. Andererseits sollte der Massenpunkt innerhalb der Struktur bleiben. Bei einer zu großen Geschwindigkeit gibt es aufgrund der Trägheit jedoch Probleme an den Stellen, an denen die Strukturen zu stark gekrümmt ist.

3.5.2 Wahl der Metrik

Wir wählen bei der Berechnung von $sp(S^*)(x)$ und $M(S^*)(x)$ eine spezielle Metrik, die *Tensormetrik* bzw. ein spezielles Koordinatensystem des $\mathbb{R}^3$. Diese Metrik entspricht der kanonischen Metrik des $\mathbb{R}^3$ unter der Annahme, daß die Kamerarichtungen orthogonal zueinander liegen. Bevor wir diese Metrik genauer beschreiben, präzisieren wir zunächst die Kameraanordnung. Seien $P_1, P_2 : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ zwei orthogonale Projektionen. Der Winkel $\alpha \in (0, \pi)$ zwischen beiden Projektionsrichtung sei bekannt. Wie in Beispiel 2.1 können wir nach Wahl der Koordinatensysteme annehmen, daß die P_i folgende Gestalt haben:

$$\begin{aligned} P_1(x, y, z) &:= (x, z) \\ P_2(x, y, z) &:= (\cos(\alpha)x + \sin(\alpha)y, z) \end{aligned}$$

Sei $A \in \mathbb{R}^{3 \times 3}$ gegeben durch:

$$A := \begin{pmatrix} 1 & 0 & 0 \\ -\frac{\cos(\alpha)}{\sin(\alpha)} & \frac{1}{\sin(\alpha)} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Dann gilt offensichtlich

$$P_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A^{-1}$$

Wir wählen nun folgende Metrik:

$$d(x, y) := |A^{-1}(x - y)|$$

Sei $B_\varepsilon(x_0) := \{x \in \mathbb{R}^3 | d(x, x_0) < \varepsilon\}$.

Lokale Masse, Schwerpunkt und Massenmoment im x_0 berechnen wir nun durch

$$\begin{aligned} m(x_0) &:= \int_{B_\varepsilon(x_0)} (S^*)^n(x) \, dx \\ sp(x_0) &:= \frac{1}{m(x_0) + \delta} \int_{B_\varepsilon(x_0)} (S^*)^n(x) \cdot A^{-1}(x - x_0) \, dx \\ M(x_0) &:= \frac{1}{m(x_0) + \delta} \int_{B_\varepsilon(x_0)} (S^*)^n(x) (A^{-1}(x - x_0) - sp(x_0)) \otimes (A^{-1}(x - x_0) - sp(x_0)) \, dx \end{aligned}$$

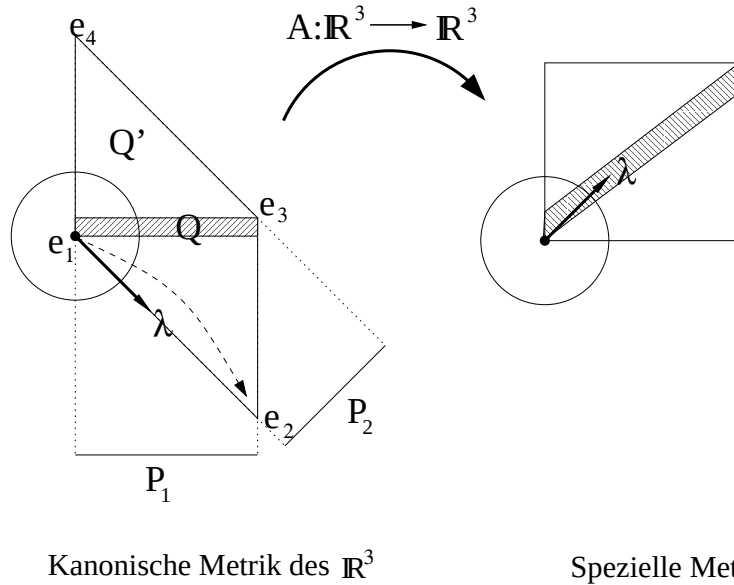


Abbildung 3.12: Beispiel eines dünnen Balkens (Querschnitt)

Um die Wahl dieser Metrik zu rechtfertigen, betrachten wir folgendes Beispiel: Sei Q ein langgestreckter dünner Balken in $\mathbb{R}^3$ mit Höhe h , der durch die Projektionen P_i in $\mathbb{R}^2$ auf Rechtecke R_i abgebildet wird. Die zugehörigen Intensitätsfunktionen nehmen auf R_i einen konstanten Wert $c > 0$, sonst den Wert null an. Abb. 3.12 zeigt einen horizontalen Querschnitt des Balkens, der Kameraanordnung und der beiden Projektionen. Aus den Intensitätsfunktionen wird durch unser Rekonstruktionsmodell nicht der Balken Q , sondern ein Parallelepiped Q' mit rautenförmiger Grundfläche rekonstruiert.

Wir identifizieren wir nun die beiden linken Seiten der Rechtecke R_i miteinander. Wir wählen an diesen Seiten zwei Punkte $m_i = (x_i, y_i)$ auf gleicher Höhe, o.B.d.A. $y_1 = y_2 = 0$. Aus diesen zueinander korrespondierenden Punkten rekonstruieren wir einen Punkt $M \in \mathbb{R}^3$, den wir als Startpunkt unserer gewöhnlichen Differentialgleichung wählen. Betrachten wir nun den horizontalen Querschnitt durch M . Seien $e_1 \dots e_4$ die Eckpunkte der Querschnittsfläche von Q' , wie in der Skizze (Abb. 3.12) dargestellt, $e_1 := M$.

Zur Bestimmung von v betrachten wir nun eine ε -Umgebung von M . In Abb. 3.12 ist der horizontale Querschnitt dieser Umgebung durch einen Kreis angedeutet. Bei der Berechnung von v wird nur der Teil von Q' berücksichtigt, der innerhalb dieser Umgebung liegt. $v(M)$ beschreibt nun die Hauptanordnungsrichtung der Masse im Gebiet $Q' \cap B_\varepsilon(M)$. Diese Anordnungsrichtung verläuft nun in etwa parallel zur zweiten Projektionsrichtung. Die Massenpunkt wird startend vom Punkt M in diese Richtung beschleunigt. Zusätzlich erfolgt eine Beschleunigung in Richtung Schwerpunkt. Den vertikalen Anteil der Beschleunigung vernachlässigen wir. Die Beschleunigung führt dazu, daß der Massenpunkt nicht dem Balken Q folgt und auf den Punkt e_3 zuläuft, sondern in etwa in Richtung der zweiten Projektionsrichtung auf den Punkt e_2 zu. Er folgt also einem falsch rekonstruierten Teil des Objekts. Der Verlauf der Lösung ist in der Skizze durch eine gestrichelte Linie angedeutet.

Wendet man das Modell auf die DNA-Ringe an, so besitzt die rekonstruierte Menge $\tilde{\Gamma}$ an den Stellen, an denen der DNA-Strang in der $(x, y, 0)$ -Ebene verläuft, einen solchen rautenförmigen Querschnitt. Numerische Versuche haben bestätigt, daß die Lösung der gewöhnlichen Differentialgleichung (3.12) in einem solchen Fall trotz der Trägheit des bewegten Massenpunktes in "falsche" Strukturen hineinläuft.

Das Problem besteht darin, daß die Gestalt des rekonstruierten Objekts Q' stark in eine bestimmte Richtung ausgeprägt ist, die bei der Bestimmung der Eigenvektoren v_i , $i = 1 \dots 3$ entscheidend ist. Verwenden wir bei der Berechnung des Schwerpunktes und des Massenmoments die oben beschriebene Metrik, so weisen wir Q' eine andere Gestalt, die eines Quaders zu. Der Eigenvektor v zum größten Eigenwert von M im Eckpunkt e_1 liegt im entsprechenden Koordinatensystem nun parallel zum Vektor $e_3 - e_1$ (vgl. Beispiel 3.2 in Abschnitt 3.2). Entsprechend verläuft die Lösungskurve $y(t)$ nun von Startpunkt e_1 bis zum Punkt e_3 , und folgt somit dem dünnen Balken.

3.5.3 Verzweigungen

Wie wir bereits gesehen haben, können wir i.a. nicht erwarten, daß die rekonstruierten dreidimensionalen Strukturen einen einzelnen geschlossenen Schlauch

bilden. Wir müssen davon ausgehen, daß zum einen durch die Nichteindeutigkeit, zum anderen durch den Diffusionsprozeß weitere schlauchartige Strukturen entstehen können. Dabei entstehen auch Verzweigungs- und Kreuzungspunkte.

Als Verzweigungspunkt bezeichnen wir den Punkt, in dem sich mehrere schlauchartige Strukturen berühren. (Dieser Begriff ist, da sich der Berührungspunkt aufgrund der Dicke der schlauchartigen Struktur nicht eindeutig bestimmen läßt, abstrakt zu sehen.) Durchkreuzen sich zwei (oder mehr) schlauchartige Strukturen, so nennen wir dies einen Kreuzungspunkt (vgl. Abschnitt 3.1.1).

Aufgrund dieser Verzweigungspunkte ist es sinnvoll, auch bei der Lösung der gewöhnlichen Differentialgleichung Verzweigungen zu erlauben, d.h. wir verfolgen die Lösung $y(t) =: y_0(t)$ ausgehend von einem Startpunkt x_0 , den wir im folgenden als den *ersten Startpunkt* bezeichnen bis zu einem möglichen Verzweigungspunkt $y(t_1)$, den wir als neuen Startpunkt zum Zeitpunkt $t'_0 = t_1$ wählen und nun zu verschiedenen Verzweigungsrichtungen $V_0 \dots V_n, n > 0$ Lösungen $y_{00}(t) \dots y_{0n}(t)$ der Differentialgleichung suchen, Hierbei sei y_{0i} die Lösung der Differentialgleichung mit Startpunkt $y_0(t_1)$ und Startvektor V_i . Als Kriterium für mögliche Verzweigungspunkte verwenden wir die in Abschnitt 3.2 definierte Kohärenz. Dazu wählen wir einen Schwellenwert $c_{bifurc} > 0$, und verzweigen in einem Punkt $y_0(t)$, falls $\text{coh}(S^*(y(t))) < c_{bifurc}$.

Indem wir sukzessive weiter verzweigen, erhalten wir Lösungen $y_k(t)$ mit Multiindex $k = (0, k_1 \dots k_l) \in \mathbb{N}^l, l > 0$. Eine Lösung $y_k(t)$ bezeichnen wir im folgenden als Teillösung.

In unserem Verfahren werden wir die Berechnung einer Teillösung abbrechen, falls die Lösung den ersten Startvektor wieder erreicht oder sich abzeichnet, daß diese Teillösung nicht der gesuchten Struktur folgt. Wir werden die Kriterien dafür detailliert in Kapitel 5 beschreiben.

Die Teillösungen interpretieren wir als einen Graphen mit Baumstruktur, wobei wir die Teillösungen als Kanten, ihren jeweiligen Start- und Endpunkt als Knoten auffassen. Die Teillösungen, die nicht weiter verzweigen, bezeichnen wir als Blätter. Sei $y_k(t)$ eine Teillösung mit Startpunkt x , dann nennen wir die Teillösung, die x als Endpunkt besitzt, die vorangegangene Lösung.

Fassen wir ausgehend von einem Blatt dessen Teillösung und rekursiv alle vorangegangenden Teillösungen zusammen, so bezeichnen wir die Menge dieser Teillösungen als Ast. Die Teillösungen eines Astes ergeben hintereinandergesetzt eine stückweise differenzierbare Kurve.

Wir erwarten, daß eine der Teillösungen wieder zum ersten Startpunkt x_0 zurückkehrt. Die durch die Teillösungen dieses Astes gegebene geschlossene Kurve $y^*(t)$

verwenden wir zur Approximation des gesuchten Schlauches Γ :

$$\tilde{\Gamma} := \{x \in \Omega^{3D} | \text{dist}(x, y^*(t)) < \varepsilon\}$$

Eine Alternative besteht darin, alle Äste des Graphen zur Approximation von $\tilde{\Gamma}$ zu verwenden, deren Teillösungen eine geschlossene Kurve bilden. Dies ist dann sinnvoll, falls eine rekonstruierte Menge $(S^*)^{-1}([c_{thres}, \infty))$ Kreuzungspunkte enthält. d.h. Punkte an denen sich schlauchartige Strukturen durchkreuzen (vgl. Abschnitt 6.4).

3.5.4 Existenz und Eindeutigkeit

Wir beweisen nun die Existenz und Eindeutigkeit der Lösung der gewöhnlichen Differentialgleichung 3.12 mit der rechten Seite $F(x, r)$ wie in (3.13) definiert, in dem wir die Lipschitz-Stetigkeit von $F(x, r)$ zeigen und auf die Theorie gewöhnlicher Differentialgleichungen zurückgreifen.

Sei $u \in C^0(\Omega)$ fest gewählt.

Wir setzen $m(x) := m(u)(x)$, $sp_i(x) := sp_i(u)(x)$ und $M_{ij}(x) := M_{ij}(u)(x)$ für $i, j = 1 \dots d$.

Seien $(\lambda_i(x))$ die Eigenwerte von $M(x)$, $P(x)$ die Projektion auf den Eigenraum zu $\lambda_1(x)$ und $\text{coh}(x) := \text{coh}(M(x))$.

Lemma 3.1 (Lipschitz-Stetigkeit von m , sp , M)

Es existieren positive Konstanten L_1, L_2, L_3 , so daß

$$\begin{aligned} |m(x) - m(y)| &\leq L_1 |x - y| \\ \max_{i=1\dots d} |sp_i(x) - sp_i(y)| &\leq L_2 |x - y| \\ \max_{i,j=1\dots d} |M_{ij}(x) - M_{ij}(y)| &\leq L_3 |x - y| \end{aligned} \quad (3.15)$$

mit L_1, L_2, L_3 unabhängig von x, y .

Dann gilt auch

$$\|M(x) - M(y)\| \leq C |x - y| \quad (3.16)$$

für eine Konstante $C > 0$.

(Hierbei ist $\|\cdot\|$ die Operatornorm bzgl. der 2-Norm auf $\mathbb{R}^d$.)

Beweis:

Sei $f(x, x_0)$ Lipschitz-stetig in x_0 und beschränkt für festes x_0 und $x \in B_\varepsilon(x_0)$. (f dient als Platzhalter für die Integranden in den Formeln für m , sp_i und M_{ij} .)

Seien

$$\begin{aligned} B_1 &:= B_\varepsilon(x_0) \setminus B_\varepsilon(y_0) \\ B_2 &:= B_\varepsilon(x_0) \cap B_\varepsilon(y_0) \\ B_3 &:= B_\varepsilon(y_0) \setminus B_\varepsilon(x_0) \end{aligned}$$

mit $|B_1| = |B_3| \leq \varepsilon |x_0 - y_0|$. Es gilt

$$\begin{aligned}
& \left| \int_{\Omega} u(x) f(x, x_0) \chi_{\varepsilon}(x - x_0) - u(x) f(x, y_0) \chi_{\varepsilon}(x - y_0) dx \right| \\
& \leq \left| \int_{B_1} u(x) f(x, x_0) dx \right| \\
& \quad + \left| \int_{B_2} u(x) f(x, x_0) - f(x, y_0) dx \right| \\
& \quad + \left| \int_{B_3} u(x) f(x, y_0) dx \right| \\
& \leq C(|B_1| + |B_2|) + C \int_{B_3} u(x) dx |x_0 - y_0| \\
& \leq \varepsilon C |x_0 - y_0| + C |B_3| \|u\|_{L^2(\Omega)} |x_0 - y_0|
\end{aligned}$$

also ist $\int_{\Omega} u(x) f(x, x_0) \chi(x - x_0)_{\varepsilon} dx$ Lipschitz-stetig in x_0 .

Mit $f \equiv 1$ folgt die Lipschitz-Stetigkeit von $m(x)$ und daraus die Lipschitz-Stetigkeit von $g(x) := \frac{1}{m(x)+\delta}$. Sei L_1 die Lipschitz-Konstante zu $m(x)$.

Mit $f(x, x_0) = x^i - x_0^i$ und der Tatsache, daß $g(x) \leq C$ und $\int_{\Omega} u(x) f(x, x_0) \chi(x - x_0)_{\varepsilon} dx < C$ unabhängig von x_0 , folgt die Lipschitz-Stetigkeit von

$$sp_i(x_0) = g(x_0) \cdot \int_{\Omega} u(x) f(x, x_0) \chi(x - x_0)_{\varepsilon} dx$$

mit $|sp_i(x_0) - sp_i(y_0)| \leq c_i |x_0 - y_0|$. Wähle $L_2 := \max_i \{c_i\}$.

Mit $f(x, x_0) = (x^i - x_0^i - sp_i(x_0))(x^j - x_0^j - sp_j(x_0))$ und dem gleichen Argument folgt die Lipschitz-Stetigkeit der $M_{ij}(x_0)$.

mit Lipschitz-Konstante c_{ij} . Setze $L_3 := \max_{i,j} \{c_{ij}\}$.

Nun gilt

$$\begin{aligned}
\|M(x) - M(y)\| & \leq C \|M(x) - M(y)\|_2 \\
& \leq C \sqrt{\sum_{i,j} |M_{ij}(x) - M_{ij}(y)|^2} \\
& \leq C \sqrt{\sum_{i,j} L_3^2 |x - y|^2} \leq C L_3 |x - y|
\end{aligned} \tag{3.17}$$

□

Lemma 3.2 (Lipschitzstetigkeit von $\text{coh}(x) P(x)$)

Es existiert eine Konstante $C > 0$ unabhängig von x, y mit

$$\|\text{coh}(x) P(x) - \text{coh}(y) P(y)\| \leq C \|M(x) - M(y)\| \tag{3.18}$$

Mit Lemma 3.1 ist $\text{coh}(x)P(x)$ Lipschitz-stetig.

($P(x)$ ist als Projektion auf den Eigenraum zum Eigenwert $\lambda_1(x)$ von $M(x)$ definiert (vgl. 3.5.1).)

Beweis:

Wir verwenden die Resultate (4.1), (4.2) und (4.3) aus [5], die wir zum Beweis der Existenz der Lösung zur Diffusionsgleichung auf Seite 56 vorgestellt haben.

Es gilt:

$$|\text{coh}(x)P(x) - \text{coh}(y)P(y)| \quad (3.19)$$

$$\begin{aligned} &\leq |\text{coh}(x) - \text{coh}(y)| \|P(x)\| + |\lambda_1(y) - \lambda_2(y)| \|P(x) - P(y)\| \\ &\stackrel{(4.1)}{\leq} 2\|M(x) - M(y)\| \quad (3.20) \\ &\quad + |\lambda_1(x) - \lambda_2(y)| \|P(x) - P(y)\| + \underbrace{|\lambda_1(x) - \lambda_1(y)| \|P(x) - P(y)\|}_{\text{vgl. (4.2)}} \end{aligned}$$

$$\leq C\|M(x) - M(y)\| + |\lambda_1(x) - \lambda_2(y)| \|P(x) - P(y)\|$$

Falls $r := |\lambda_1(x) - \lambda_2(y)| = 0$ folgt die Behauptung, andernfalls gilt mit (4.2), (4.3):

$$\begin{aligned} &|\lambda_1(x) - \lambda_2(y)| \|P(x) - P(y)\| \quad (3.21) \\ &\leq r\|P(x)P(y) - P(y)\| + r\|P(x)P^\perp(y)\| \\ &\leq r \frac{2C}{r} \|M(x) - M(y)\| = 2C\|M(x) - M(y)\| \end{aligned}$$

□

Satz 3.3 (Existenz und Eindeutigkeit)

Sei $(x_0, y_0) \in \mathbb{R}^d \times \mathbb{R}^d$ und

$$F(x, y) = \beta \text{coh}(x)P(x)y + \gamma sp(x) \quad (3.22)$$

Dann existiert eine eindeutige Lösung $x \in C^1([0, T], \mathbb{R}^d)$ des Anfangswertproblems

$$\ddot{x} + \alpha(|\dot{x}| - v_0)\dot{x} = F(x, \dot{x}) \quad (3.23)$$

Beweis:

Die Aussage folgt nach [2], falls $F(x, y)$ Lipschitz-stetig ist ($\alpha(|\dot{x}| - v_0)\dot{x}$ ist Lipschitz-stetig für beschränktes $\dot{x}$). Die Lipschitz-Stetigkeit von F bezüglich y folgt aus der Linearität des zweiten Summanden, da $\text{coh}(x)P(x)$ beschränkt ist.

Die Lipschitz-Stetigkeit bezüglich x folgt aus Lemma 3.1 und 3.2.

□

Kapitel 4

Wohlgestelltheit des Diffusionsproblems

In diesem Kapitel werden wir die Existenz und Eindeutigkeit einer Lösung der im letzten Kapitel vorgestellten partiellen Differentialgleichung (3.9) mit der in Abschnitt 3.3.2 beschriebenen rechten Seite f_ε beweisen. Dazu werden wir zunächst die Lipschitz-Stetigkeit von $D_\varepsilon(u)$ und $f_\varepsilon(u)$ in der L^∞ -Norm zeigen.

Wir beweisen die Existenz und Eindeutigkeit zunächst für den Fall, daß der Exponent n in den Formeln für lokale Masse, Schwerpunkt und Massenmoment den Wert 1 annimmt. Für den allgemeinen Fall verweisen wir auf die Bemerkung 1 am Ende dieses Abschnitts.

4.1 Lipschitzstetigkeit von $D_\varepsilon(u)$ und $f_\varepsilon(u)$

Im folgenden sei C eine universelle Konstante.

Betrachten wir die Eigenwerte μ_i , $i = 1 \dots d$ einer symmetrischen Matrix A , so seien diese immer so numeriert, daß sie mit absteigendem Betrag geordnet sind. Zu einer Matrix A sei $\|A\|_2$ die 2-Norm auf $\mathbb{R}^{d \times d}$ und $\|A\|$ die Operatornorm zur 2-Norm auf $\mathbb{R}^d$. Die Kohärenz der Eigenwerte von A bezeichnen wir mit $\text{coh}(A) := \mu_1 - \mu_2$.

Sei $v_1 \dots v_m$ eine Orthonormalbasis eines Unterraums des $\mathbb{R}^d$. Die Projektion P auf diesen Unterraum definieren wir durch $Px = \sum_{i=1 \dots m} \langle v_i, x \rangle v_i$.

Sei $\chi_\varepsilon := \chi_{B_\varepsilon(0)}$ die charakteristische Funktion auf $B_\varepsilon(0)$.

$\hat{\Omega}$ sei die ε -Umgebung von Ω (vgl. Definition auf Seite 30).

Resultate aus der Literatur:

Nach [5] erhalten wir folgende Aussagen:

1. Seien A, B zwei symmetrische Matrizen, $(\mu_i(A))_i$ Eigenwerte von A , $(\mu_i(B))_i$ die Eigenwerte von B . Dann gilt:

$$\max_i |\mu_i(A) - \mu_i(B)| \leq \|A - B\| \quad (4.1)$$

2. Sei $S(A) \subset \{\mu_i(A)\}_i$, $S(B) \subset \{\mu_i(B)\}_i$ mit $\text{dist}(S(A), S(B)) = r > 0$. Seien $P(A), P(B)$ die Projektionen auf die Eigenräume zu $S(A), S(B)$. Dann gilt für die Operatornorm mit $c = \frac{\pi}{2}$:

$$\|P(A)P(B)\| \leq \frac{c}{r} \|A - B\| \quad (4.2)$$

3. Da $(P(A) + P(A)^\perp)P(B) = P(B)$, gilt auch

$$\|P(A)^\perp P(B) - P(B)\| \leq \frac{c}{r} \|A - B\| \quad (4.3)$$

Wir beweisen zunächst einige Abschätzungen, die wir später benötigen.

Lemma 4.1 (Abschätzungen für sp_j , M_{ij} , $\mu_i(u)$, $\text{coh}(M(u))$)

Sei $u \in L^2(\Omega)$. Es gilt:

$$\begin{aligned} \|sp_i(u)\|_{L^\infty} &\leq \varepsilon \quad \text{für } i, j = 1, 2 \\ \|M_{ij}(u)\|_{L^\infty} &\leq 4\varepsilon^2 \quad \text{für } i, j = 1, 2 \\ \|\mu_i\|_{L^\infty} &\leq C \quad \text{für } i = 1, 2 \\ \|\text{coh}(M(u))\|_{L^\infty} &\leq C \end{aligned} \quad (4.4)$$

wobei C unabhängig von u ist.

Beweis:

$$\begin{aligned} |sp_i(u)(x_0)| &= \frac{1}{m(u)(x_0)+\delta} \int_{\hat{\Omega}} u(x)(x^i - x_0^i) \chi_\varepsilon(x - x_0) dx \\ &\leq \|(x^i - x_0^i) \chi_\varepsilon(x - x_0)\|_{L^\infty} \frac{1}{m(u)(x_0)+\delta} \int_{\hat{\Omega}} u(x) \chi_\varepsilon(x - x_0) dx \\ &\leq \varepsilon \frac{m(u)(x_0)}{m(u)(x_0)+\delta} \leq \varepsilon \end{aligned}$$

$$\begin{aligned} |M_{ij}(u)(x_0)| &= \frac{1}{m(u)(x_0)+\delta} \int_{\hat{\Omega}} u(x)(x^i - x_0^i - sp_i(x_0))(x^j - x_0^j - sp_j(x_0)) \chi_\varepsilon(x - x_0) dx \\ &\leq \|(x^i - x_0^i - sp_i(x_0))\|_{L^\infty}^2 \frac{m(u)(x_0)}{m(u)(x_0)+\delta} \leq 4\varepsilon^2 \end{aligned}$$

Punktweise gilt nun:

$$\begin{aligned} |\mu_2(u)| &\leq |\mu_1(u)| = \|M(u)\| \leq C \|M(u)\|_2 \leq C \\ |\text{coh}(u)| &= |\mu_1(u) - \mu_2(u)| \leq |\mu_1| \end{aligned}$$

□

Lemma 4.2 (Abschätzungen für $m(u)$, $sp(u)$, $M(u)$ und $\text{coh}(u)$)

Seien $u, v \in L^2(\Omega, \mathbb{R})$ beliebig, dann gilt für alle $x \in \Omega$:

$$\begin{aligned} |m(u)(x) - m(v)(x)| &\leq C \|u - v\|_{L^2} \\ |sp_i(u)(x) - sp_i(v)(x)| &\leq C \|u - v\|_{L^2} \quad i = 1 \dots d \\ |M_{ij}(u)(x) - M_{ij}(v)(x)| &\leq C \|u - v\|_{L^2} \quad i, j = 1 \dots d \end{aligned} \quad (4.5)$$

(Die Konstanten sind unabhängig von x) und

$$\|M(u) - M(v)\|_{L^\infty} \leq C \|u - v\|_{L^2} \quad (4.6)$$

$$\|\text{coh}(M(u)) - \text{coh}(M(v))\|_{L^\infty} \leq C \|u - v\|_{L^2} \quad (4.7)$$

Beweis:

$$\begin{aligned} |m(u)(x_0) - m(v)(x_0)| &= \left| \int_{\hat{\Omega}} (u(x) - v(x)) \chi_\varepsilon(x - x_0) dx \right| \\ &\leq \int_{\hat{\Omega}} |1 \cdot (u - v)| dx \leq \sqrt{|\hat{\Omega}|} \|u - v\|_{L^2} \end{aligned}$$

$$\begin{aligned} &|sp_i(u)(x_0) - sp_j(v)(x_0)| \\ &= \left| \frac{1}{m(u)(x_0) + \delta} \int_{\hat{\Omega}} u(x) (x^i - x_0^i) \chi_\varepsilon(x - x_0) dx \right. \\ &\quad \left. - \frac{1}{m(v)(x_0) + \delta} \int_{\hat{\Omega}} v(x) (x^i - x_0^i) \chi_\varepsilon(x - x_0) dx \right| \\ &\leq \left| \frac{1}{m(u)(x_0) + \delta} - \frac{1}{m(v)(x_0) + \delta} \right| \left| \int_{\hat{\Omega}} u(x) (x^i - x_0^i) \chi_\varepsilon(x - x_0) dx \right| \\ &\quad + \left| \frac{1}{m(v)(x_0) + \delta} \int_{\hat{\Omega}} (u(x) - v(x)) (x^i - x_0^i) \chi_\varepsilon(x - x_0) dx \right| \\ &\leq \frac{|m(u)(x_0) - m(v)(x_0)|}{(m(u)(x_0) + \delta)(m(v)(x_0) + \delta)} \int_{\hat{\Omega}} u(x^i - x_0^i) \chi_\varepsilon(x - x_0) dx \\ &\quad + \frac{1}{\delta} \|u - v\|_{L^2} \|(x^i - x_0^i) \chi_\varepsilon(x - x_0)\|_{L^2} \\ &\leq \frac{1}{\delta} |m(u)(x_0) - m(v)(x_0)| |sp_i(u)(x_0)| + \frac{\varepsilon \sqrt{|\hat{\Omega}|}}{\delta} \|u - v\|_{L^2} \\ &\leq C \|u - v\|_{L^2} \end{aligned}$$

Mit

$$\begin{aligned} f_{ij}(u, v, w)(x, x_0) &:= \\ &u(x)(x^i - x_0^i - sp_i(v)(x_0))(x^j - x_0^j - sp_j(w)(x_0)) \chi_\varepsilon(x - x_0) \end{aligned} \quad (4.8)$$

gilt:

$$\begin{aligned}
& |M_{ij}(u)(x_0) - M_{ij}(v)(x_0)| \\
&= \left| \frac{1}{m(u)(x_0)+\delta} \int_{\hat{\Omega}} f(u, u, u)(x, x_0) dx - \frac{1}{m(v)(x_0)+\delta} \int_{\hat{\Omega}} f(v, v, v)(x, x_0) dx \right| \\
&\leq \frac{|m(u)(x_0)-m(v)(x_0)|}{(m(u)(x_0)+\delta)(m(v)(x_0)+\delta)} \left| \int_{\hat{\Omega}} f(u, u, u) dx \right| \\
&+ \left| \frac{1}{m(v)(x_0)+\delta} \right| \left| \int_{\hat{\Omega}} f(u, u, u)(x, x_0) - f(v, v, v)(x, x_0) dx \right| \\
&= |m(u)(x_0) - m(v)(x_0)| \frac{1}{\delta} |M_{ij}(u)(x_0)| \\
&+ \frac{1}{m(v)(x_0)+\delta} \left| \int_{\hat{\Omega}} f(u, u, u)(x, x_0) - f(v, v, v)(x, x_0) dx \right|
\end{aligned}$$

Der ersten Term der rechten Seite ist mit Lemma 4.1 und der Abschätzung für $|m(u)(x) - m(v)(x)|$ kleiner gleich $C \|u - v\|_{L^2}$. Den zweiten schätzen wir wie folgt ab:

$$\begin{aligned}
& \left| \frac{1}{m(v)(x_0)+\delta} \int_{\hat{\Omega}} f(u, u, u)(x, x_0) - f(v, v, v)(x, x_0) dx \right| \\
&\leq \left| \frac{1}{m(v)(x_0)+\delta} \int_{\hat{\Omega}} f(u, u, u)(x, x_0) - f(v, u, u)(x, x_0) dx \right| \\
&+ \left| \frac{1}{m(v)(x_0)+\delta} \int_{\hat{\Omega}} f(v, u, u)(x, x_0) - f(v, v, u)(x, x_0) dx \right| \\
&+ \left| \frac{1}{m(v)(x_0)+\delta} \int_{\hat{\Omega}} f(v, v, u)(x, x_0) - f(v, v, v)(x, x_0) dx \right| \\
&\leq C \|u - v\|_{L^2} \underbrace{\|(x^i - x_0^i - sp_i(u)(x_0))\chi_\varepsilon(x - x_0)\|_{L^\infty}}_{\leq 2\varepsilon} \\
&\quad \underbrace{\|(x^j - x_0^j - sp_j(u)(x_0))\chi_\varepsilon(x - x_0)\|_{L^2}}_{\leq 2\varepsilon\sqrt{|\hat{\Omega}|}} \\
&+ \left(\|sp_i(u)(x_0) - sp_i(v)(x_0)\|_{L^\infty} \|(x^j - x_0^j - sp_j(u)(x_0))\chi_\varepsilon(x - x_0)\|_{L^\infty} \right. \\
&\quad \left. + \|(x^i - x_0^i - sp_i(v)(x_0))\chi_\varepsilon(x - x_0)\|_{L^\infty} \|(sp_j(u)(x_0) - sp_j(v)(x_0))\|_{L^\infty} \right) \\
&\quad \frac{1}{m(v)(x_0)+\delta} \int_{\hat{\Omega}} v \chi_\varepsilon(x - x_0) dx \\
&\leq C \|u - v\|_{L^2}
\end{aligned}$$

Es folgt $\|M(u)(x) - M(v)(x)\|_2 \leq C \|u - v\|_{L^2}$ für alle x und wegen der Äquivalenz der Operatornorm zur 2-Norm und der Unabhängigkeit der Konstante von x : $\|M(u) - M(v)\|_{L^\infty} \leq C \|u - v\|_{L^2}$.

(4.7) folgt mit (4.1). Für die Kohärenz gilt:

$$\begin{aligned}
\|\text{coh}(u) - \text{coh}(v)\|_{L^\infty(\Omega)} &\leq \|\mu_1(u) - \mu_1(v)\|_{L^\infty(\Omega)} + \|\mu_2(u) - \mu_2(v)\|_{L^\infty(\Omega)} \\
&\stackrel{(4.1)}{\leq} 2\|M(u) - M(v)\|_{L^\infty(\Omega)} \\
&\stackrel{\text{Lemma 4.2}}{\leq} C\|u - v\|_{L^2}
\end{aligned}$$

□

Die Diffusionsmatrix $D_\varepsilon(u)$ ist wie folgt definiert:

Definition 1

Seien $E_1(u), E_2(u)$ die Eigenräume zu $\mu_1(u)$ bzw. $\{\mu_2(u) \dots \mu_d(u)\}$.

Seien $P_i(u)$ die Projektionen auf $E_i(u)$.

Wir definieren (punktweise)

$$\tilde{\mu}_1(u) := 1, \quad \tilde{\mu}_2(u) := \frac{1}{1 + \coth(M(v)) \frac{|\nabla u_\sigma|^2}{\kappa^2}} \quad (4.9)$$

Dann gilt für die oben definierte Diffusionsmatrix $D_\varepsilon(u)$:

$$D_\varepsilon(u) = \tilde{\mu}_1(u)P_1(u) + \tilde{\mu}_2(u)P_2(u) \quad (4.10)$$

Lemma 4.3 (Lipschitz-Stetigkeit von $D_\varepsilon(u)$)

Seien $u, v \in L^2(\Omega, \mathbb{R}) \cap L^\infty(\Omega, \mathbb{R})$ mit $\|u\|_{L^\infty}, \|v\|_{L^\infty} \leq c_a$.

Dann gilt

1. Punktweise fast überall :

$$|\tilde{\mu}_i(u) - \tilde{\mu}_i(v)| \leq C \|u - v\|_{L^2} \quad (4.11)$$

$$|\tilde{\mu}_i(u) - \tilde{\mu}_j(v)| \leq C |\mu_i(u) - \mu_j(v)| \quad (i \neq j) \quad (4.12)$$

2.

$$\|D_\varepsilon(u) - D_\varepsilon(v)\|_{L^\infty} \leq C \|u - v\|_{L^2} \quad (4.13)$$

Die Konstanten hängen nur von c_a ab.

Beweis:

In den folgenden Abschätzungen werden wir benutzen, daß nach der Faltungsabschätzung gilt:

$$\sup_{x \in \Omega} |\nabla u_\sigma(x)| \leq C(\sigma) \|u\|_{L^\infty} \leq C(\sigma) c_a \quad (*)$$

1.

1. Abschätzung für $|\tilde{\mu}(u)_1 - \tilde{\mu}(v)_1|$:

Folgt sofort mit $\tilde{\mu}_1(u) = \tilde{\mu}_1(v)$

2. Abschätzung für $|\tilde{\mu}(u)_2 - \tilde{\mu}(v)_2|$:

Punktweise gilt:

$$\begin{aligned}
|\tilde{\mu}_2(u) - \tilde{\mu}_2(v)| &= \left| \frac{1}{1+\text{coh}(u)\frac{|\nabla u_\sigma|^2}{\kappa^2}} - \frac{1}{1+\text{coh}(v)\frac{|\nabla v_\sigma|^2}{\kappa^2}} \right| \\
&\leq \frac{1}{\kappa^2} \frac{1}{\left| \left(1+\text{coh}(u)\frac{|\nabla u_\sigma|^2}{\kappa^2}\right) \left(1+\text{coh}(v)\frac{|\nabla v_\sigma|^2}{\kappa^2}\right) \right|} |\text{coh}(u)| |\nabla u_\sigma|^2 - \text{coh}(v) |\nabla v_\sigma|^2| \\
&\leq C (|\text{coh}(u) - \text{coh}(v)| |\nabla u_\sigma|^2 + |\text{coh}(v)| (|\nabla u_\sigma|^2 - |\nabla v_\sigma|^2)) \\
&\stackrel{L.4.1, (*)}{\leq} C (\|u - v\|_{L^2} + |(\nabla u_\sigma)^2 - (\nabla v_\sigma)^2|) \\
&\leq C \|u - v\|_{L^2} + |\nabla u_\sigma + \nabla v_\sigma| |\nabla u_\sigma - \nabla v_\sigma| \\
&\stackrel{(*)}{\leq} C \|u - v\|_{L^2}
\end{aligned}$$

3. Abschätzung für $|\tilde{\mu}(u)_i - \tilde{\mu}(v)_j|$, $i \neq j$:

Punktweise gilt:

$$\begin{aligned}
|\tilde{\mu}_1(u) - \tilde{\mu}_2(v)| &= \left| 1 - \frac{1}{1+\text{coh}(v)\frac{|\nabla v_\sigma|^2}{\kappa^2}} \right| \\
&\leq \frac{1}{\kappa^2} \frac{1}{\left| \left(1+\text{coh}(v)\frac{|\nabla v_\sigma|^2}{\kappa^2}\right) \left(1+\text{coh}(v)\frac{|\nabla v_\sigma|^2}{\kappa^2}\right) \right|} |\text{coh}(v)| |\nabla v_\sigma|^2 \\
&\stackrel{(*)}{\leq} C |\text{coh}(v)| = C |\mu_1(v) - \mu_2(v)| \\
&\leq C \underbrace{|\mu_1(u) - \mu_1(v)|}_{=0} + C |\mu_1(u) - \mu_2(v)|
\end{aligned}$$

Die Abschätzung für $|\tilde{\mu}_2(u) - \tilde{\mu}_1(v)|$ folgt mit Vertauschen von u und v .

Die rechten Seiten der Abschätzungen sind unabhängig von x .

2.

Punktweise gilt:

$$\begin{aligned}
\|D_\varepsilon(u) - D_\varepsilon(v)\| &= \left\| \sum_{i=1,2} \tilde{\mu}_i(u) P_i(u) \underbrace{(P_1(v) + P_2(v))}_{=1} - \tilde{\mu}_i(v) \underbrace{(P_1(u) + P_2(u))}_{=1} P_i(v) \right\| \\
&\leq \sum_{i,j=1,2} |\tilde{\mu}_i(u) - \tilde{\mu}_j(v)| \|P_i(u) P_j(v)\| \\
&\leq C \sum_{i=1,2} \|u - v\|_{L^2} \underbrace{\|P_i(u) P_i(v)\|}_{\leq 1} + C \sum_{i,j=1,2, j \neq i} |\mu_i(u) - \mu_j(v)| \|P_i(u) P_j(v)\|
\end{aligned}$$

Falls $\mu_i(u) - \mu_j(v) = 0$ folgt die Behauptung.

Falls für feste Indizes $i \neq j$ $r := |\mu_i(u) - \mu_j(v)| > 0$, so gilt

$$\begin{aligned} |\mu_i(u) - \mu_j(v)| \|P_i(u)P_j(v)\| &\leq C |\mu_i(u) - \mu_j(v)| \frac{1}{r} \|M(u) - M(v)\| \\ &\leq C \|u - v\|_{L^2} \end{aligned}$$

Da die Konstanten ebenfalls von x unabhängig sind, folgt die Behauptung.

□

Lemma 4.4 (Lipschitzstetigkeit von $f_\varepsilon(u)$)

Es existiert eine Konstante $C > 0$, so daß für alle $u, v \in L^2(\Omega, [0, 1])$ gilt:

$$\|f_\varepsilon(u) - f_\varepsilon(v)\|_{L^2} \leq C \|u - v\|_{L^2} \quad (4.14)$$

Beweis:

Die rechte Seite der Differentialgleichung 3.9 hat die Gestalt

$f_\varepsilon(u)(x) = f_1(u(x)) + f_2(\text{coh}(u)(x))$, wobei f_1, f_2 beschränkt und Lipschitz-stetig bzgl. $u(x)$ bzw. $\text{coh}(u)(x)$ sind. $u(x)$ und $\text{coh}(u)(x)$ sind fast überall beschränkt.

Es gilt daher punktweise fast überall

$$\begin{aligned} |f_\varepsilon(u)(x) - f_\varepsilon(v)(x)| &\leq |f_1(u(x))(f_2(\text{coh}(u)(x)) - f_2(\text{coh}(v)(x)))| \\ &\quad + |(f_1(u(x)) - f_1(v(x)))f_2(\text{coh}(v)(x))| \\ &\leq C|u(x) - v(x)| + C|\text{coh}(u)(x) - \text{coh}(v)(x)| \end{aligned}$$

Nach Lemma 4.2 gilt:

$$\|\text{coh}(u)(x) - \text{coh}(v)(x)\|_{L^\infty} \leq C \|u - v\|_{L^2}$$

und daher

$$|f_\varepsilon(u)(x) - f_\varepsilon(v)(x)| \leq C|u(x) - v(x)| + C\|u - v\|_{L^2}$$

Mit $(a + b)^2 \leq 2a^2 + 2b^2$ folgt

$$\begin{aligned} \int_{\Omega} |f(u) - f(v)|^2 dx &\leq C \int_{\Omega} |u - v|^2 dx + C\|u - v\|_{L^2}^2 \\ &\leq C\|u - v\|_{L^2}^2 \end{aligned}$$

□

4.2 Existenz und Eindeutigkeit der Lösung

Für den Beweis der Existenz benötigen wir noch ein Resultat aus der Funktionalanalysis.

Wir definieren zu einem $u \in L^2(0, T; H^{1,2}(\Omega))$ die verallgemeinerte Zeitableitung $\frac{du}{dt}$ als das Element des Raumes $L^2(0, T; H^{-1,2}(\Omega))$, für das gilt:

$$\int_0^T \left\langle \frac{du}{dt}, v \right\rangle \phi(t) dt = - \int_0^T \langle u(t), v \rangle \phi'(t) dt$$

für alle $\phi \in C_0^\infty(0, T; \mathbb{R})$ und $v \in H^{1,2}(\Omega)$. Dabei sei $\phi'(t) = \partial_t \phi(t)$.

Definition 2 Der Raum $W(0, T)$

Zu $T > 0$ sei $W(0, T)$ definiert durch

$$W(0, T) := \left\{ w \in L^2(0, T; H^{1,2}(\Omega)), \frac{dw}{dt} \in L^2(0, T; H^{-1,2}(\Omega)) \right\},$$

$W(0, T)$ ist mit der Norm

$$\|w\|_W := \|w\|_{L^2(0, T; H^{1,2}(\Omega))} + \left\| \frac{dw}{dt} \right\|_{L^2(0, T; H^{-1,2}(\Omega))}$$

ein Banach-Raum [13].

Lemma 4.5

Es existiert eine kompakte Einbettung von $W(0, T)$ in $L^2(0, T; L^2(\Omega))$.

Beweis:

Nach [13] existieren stetige Einbettungen von $W(0, T)$ nach $C([0, T], H^{-1,2}(\Omega))$ und $C([0, T], L^2(\Omega))$, d.h. es existiert eine Konstante K , so daß

$$\sup_{0 \leq t \leq T} \|u(t)\|_{H^{-1,2}(\Omega)} \leq K \|u\|_W \quad (4.15)$$

$$\sup_{0 \leq t \leq T} \|u(t)\|_{L^2(\Omega)} \leq K \|u\|_W \quad (4.16)$$

Desweiteren gilt für alle $s, t \in [0, T]$ die folgende Gleichung:

$$\frac{1}{2} (\|u(t)\|_{L^2}^2 - \|u(s)\|_{L^2}^2) = \int_s^t \langle u'(\tau), u(\tau) \rangle d\tau \quad (4.17)$$

Diese Gleichung wird in [13] bewiesen, in dem $u \in W(0, T)$ durch eine Folge von Funktionen $u_n \in C^1([0, T], H^{1,2}(\Omega))$ mit $u_n \rightarrow u$ in der $\|\cdot\|_W$ -Norm approximiert

$(C^1([0, T], H^{1,2}(\Omega)))$ liegt dicht in $W(0, T)$ und zum Grenzübergang übergeht.

Aus (4.17) erhält man

$$\begin{aligned}
\|u(t+s) - u(t)\|_{L^2}^2 &\leq \left| \|u(t+s)\|_{L^2}^2 - \|u(t)\|_{L^2}^2 \right| \\
&\leq 2 \int_0^s \left\| \frac{du}{dt} \right\|_{H^{-1,2}} \|u\|_{H^{1,2}} \\
&\leq 2K \|u\|_W \int_0^s 1 \|u\|_{H^{1,2}} \\
&\leq 2K s \|u\|_W \|u\|_{L^2(0,T;H^{1,2}(\Omega))} \\
&\leq K s (\|u\|_W + \|u\|_{L^2(0,T;H^{1,2}(\Omega))}) \leq C s \|u\|_W
\end{aligned} \tag{4.18}$$

und damit für festes $s \geq 0$

$$\begin{aligned}
&\int_0^T \|u(t+s) - u(t)\|_{L^2}^2 dt \\
&\leq C s \|u\|_W + \int_{[0,s] \cup [T-s,T]} \|u(t)\|_{H^{1,2}}^2 dt \leq C s \|u\|_W
\end{aligned} \tag{4.19}$$

wobei $u(t)$ auf $[T, T+s]$ durch null fortgesetzt sei.

Sei X eine beschränkte Teilmenge von $W(0, T)$. Wir beweisen die Präkompaktheit von X in $L^2([0, T]; L^2(\Omega))$, in dem wir zeigen, daß zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so daß X in einer ε -Umgebung um eine präkompakte Menge X_δ von Treppenfunktionen mit Werten in $H^{1,2}(\Omega)$ liegt.

Wähle mit $N \in \mathbb{N}$ und $\delta := \frac{T}{N}$ eine Zerlegung des Intervalls $[0, T]$ in Teilintervalle $I_i := [i\delta, (i+1)\delta]$ mit $i = 0 \dots N-1$. Zu $u \in X$ definieren wir eine Treppenfunktion u_δ , die konstant auf jedem der I_i ist. Sei dazu $v_i := \frac{1}{\delta} \int_{I_i} u(s) ds$ für $i = 0 \dots N-1$, also $v_i \in H^{1,2}(\Omega)$. Wir definieren nun u_δ durch $u_\delta(t) := \sum_i \chi_{I_i}(t) v_i$.

Dann gilt für $t \in I_i \subset [0, T]$ mit geeignetem i

$$\begin{aligned}
\|u_\delta(t) - u(t)\|_{L^2}^2 &= \left\| \frac{1}{\delta} \int_{I_i} u(s) - u(t) ds \right\|_{L^2}^2 \\
&\leq \frac{1}{\delta} \int_{I_i} \|u(s) - u(t)\|_{L^2}^2 ds \\
&\leq \frac{1}{\delta} \int_{-\delta}^{\delta} \|u(t+s) - u(t)\|_{L^2}^2 ds
\end{aligned} \tag{4.20}$$

Mit (4.19) folgt nun

$$\begin{aligned}
&\int_0^T \|u_\delta(t) - u(t)\|_{L^2}^2 dt \\
&\leq \frac{1}{\delta} \int_{-\delta}^{\delta} \int_0^T \|u(t+s) - u(t)\|_{L^2}^2 dt ds \\
&\leq \frac{1}{\delta} C \|u\|_W \int_{-\delta}^{\delta} s ds = C \|u\|_W \delta
\end{aligned} \tag{4.21}$$

Wir wählen nun $\delta < \frac{\varepsilon^2}{C\|u\|_W}$. Dann gilt also $\|u_\delta - u\|_{L^2(L^2)} \leq \varepsilon$. Mit $X_\delta := \{u_\delta | u \in X\}$ folgt also $X \subset B_\varepsilon(X_\delta)$ bzgl. der $L^2(0, T; L^2(\Omega))$ -Norm.

Wir zeigen nun, daß X_δ präkompakt ist, in dem wir eine Überdeckung durch endlich viele $\tilde{\varepsilon}$ -Kugeln zu beliebigem $\tilde{\varepsilon} > 0$ konstruieren. Es gilt

$$\begin{aligned} \|v_i\|_{H^{1,2}}^2 &\leq \frac{1}{\delta} \int_{I_i} \|u(s)\|_{H^{1,2}}^2 ds \\ &\leq \frac{1}{\delta} \int_0^T \|u(s)\|_{H^{1,2}}^2 ds \leq \frac{C}{\delta} =: R \end{aligned} \quad (4.22)$$

Die Funktionen $\|v_i\|$ liegen also in $B_R^{H^{1,2}}(0) \subset H^{1,2}(\Omega)$, wobei wir die durch die $H^{1,2}$ -Norm induzierte Metrik verwenden. Nach dem Satz von Rellich ist die Einbettung von $H^{1,2}(\Omega)$ nach $L^2(\Omega)$ kompakt, also ist $B_R^{H^{1,2}}(0)$ präkompakt in $L^2(\Omega)$. Daher existieren zu beliebigem $r > 0$ Funktionen $w_0 \dots w_m$ in $L^2(\Omega)$, so daß gilt:

$$B_R^{H^{1,2}}(0) \subset \bigcup_{j=0 \dots m} B_r^{L^2}(w_j) \quad (4.23)$$

Zu v_i sei die Funktion w_{ji} so gewählt, daß $v_i \in B_r^{L^2}(w_{ji})$ liegt. Wir setzen $w(t) := \sum_{i=0 \dots N} \chi_{I_i}(t) w_{ji}$. Dann gilt

$$\begin{aligned} \int_{[0,T]} \|u_\delta(t) - w(t)\|_{L^2}^2 dt &= \sum_{i=0 \dots N} |I_i| \|v_i - w_{ji}\|_{L^2}^2 \\ &\leq \sum_{0 \leq i < \frac{T}{\delta}} \delta r^2 \leq T r^2 \end{aligned} \quad (4.24)$$

Wir wählen nun $r < \sqrt{\tilde{\varepsilon}T}$. Dann liegt jedes u_δ in $B_{\tilde{\varepsilon}}(w) \subset L^2(0, T; L^2(\Omega))$. Da die Anzahl der möglichen Funktionen w endlich ist, folgt die Präkompaktheit von X_δ .

□

Wir kommen nun zum Beweis der Existenz und Eindeutigkeit.

Satz 4.6 *Existenz und Eindeutigkeit*

Zu $u_0 \in L^\infty(\Omega)$ existiert eine schwache Lösung $u \in L^2((0, T], H^{1,2}(\Omega))$ folgender Differentialgleichung:

$$\begin{aligned} \partial_t u - \operatorname{div}(D_\varepsilon(u) \nabla u) &= f_\varepsilon(u) && \text{auf } (0, T] \times \Omega \\ u(0) &= u_0 && \text{in } \Omega \\ \langle D_\varepsilon(u) \nabla u, n \rangle &= 0 && \text{auf } [0, T] \times \partial\Omega \end{aligned} \quad (4.25)$$

Beweis ($n = 1$, vgl. Bemerkung 1):

1. Existenz:

Wir gehen analog zu [9] vor. Wir betrachten den Hilbertraum

$$W(0, T) := \left\{ w \in L^2(0, T; H^{1,2}(\Omega)), \frac{dw}{dt} \in L^2(0, T; H^{-1,2}(\Omega)) \right\},$$

versehen mit der Norm $\|w\|_W := \|w\|_{L^2(0,T;H^{1,2}(\Omega))} + \left\| \frac{dw}{dt} \right\|_{L^2(0,T;H^{-1,2}(\Omega))}$. Sei $w \in W(0, T) \cap L^\infty(0, T, L^\infty(\Omega))$, so daß

$$\|w\|_{L^\infty(0,T;L^\infty(\Omega))} \leq c_1 := 1$$

Mit P_w bezeichnen das Problem

$$\begin{aligned} \forall v \in H^{1,2}(\Omega) \quad \text{f.ü. in } [0, T] : \\ \langle v, \partial_t u(t) \rangle + \int_{\Omega} \nabla v(t) D_\varepsilon(w) \nabla u(t) = \int_{\Omega} v f_\varepsilon(w) \quad (4.26) \\ u(0) = u_0 \end{aligned}$$

Da $w \in L^\infty(0, T; L^\infty(\Omega))$, existiert eine Konstante c_2 , die nur von G_σ und c_1 abhängt, so daß $|\nabla G_\sigma * w|^2 \leq c_2$ gilt, und eine Konstante s unabhängig von w , so daß $\|\text{coh}(u)\|_{L^\infty} \leq s$ gilt. Sei $\nu := \frac{1}{1+s \frac{c_2}{\kappa^2}}$. Für die Eigenwerte $(\tilde{\mu}_i)_{i=1\dots d}$ von $D_\varepsilon(u)$ gilt dann

$$\tilde{\mu}_i \geq \nu \quad \text{f.ü. in } (0, T) \times \Omega \quad (4.27)$$

Nach Resultaten aus der Theorie linearer parabolischer Differentialgleichungen existiert eine eindeutige Lösung $U(w)$ in $W(0, T)$ mit

$$\|U(w)\|_{L^2(0,T;H^{1,2}(\Omega))} \leq c_2 \quad (4.28)$$

$$\left\| \frac{d}{dt} U(w) \right\|_{L^2(0,T;H^{-1,2}(\Omega))} \leq c_3 \quad (4.29)$$

c_2, c_3 sind unabhängig von w . (Hierbei nutzen wir aus, daß $\|D_\varepsilon(u)\|_{L^\infty([0,T],L^\infty(\Omega))}$ und $\|f(u)\|_{L^\infty([0,T],L^\infty(\Omega))}$ unabhängig von u abgeschätzt werden können.) Die rechte Seite der Differentialgleichung ist so gewählt, daß für beliebiges $x \in \Omega$ $f(u)(x) = 0$ gilt, falls $u(x) \notin (0, 1)$. Daher gilt nach dem Maximumsprinzip

$$\|U(w)\|_{L^\infty(0,T;L^\infty(\Omega))} \leq 1 =: c_1 \quad (4.30)$$

Wir führen nun eine Untermenge W_0 von $W(0, T)$ ein:

$$\begin{aligned} W_0 := \{ \quad & w \in W(0, T) \mid \|w\|_{L^2(0,T;H^{1,2}(\Omega))} \leq c_2, \\ & \|w\|_{L^\infty(0,T;L^\infty(\Omega))} \leq c_1, \\ & \left\| \frac{dw}{dt} \right\|_{L^2(0,T;H^{-1,2}(\Omega))} \leq c_3, \quad w(0) = u_0 \quad \} \end{aligned} \quad (4.31)$$

U ist Abbildung von W_0 nach W_0 . Wir wollen den Fixpunktsatz von Schauder anwenden, W_0 ist konvex, nicht leer und ein schwach kompakter Unterraum von $W(0, T)$. Wir zeigen, daß U schwach stetig ist. Da $W(0, T)$ nach Lemma 4.5 kompakt in $L^2(0, T; L^2(\Omega))$ eingebettet ist, sind die Voraussetzungen des Fixpunktsatzes erfüllt und es existiert ein $u \in W_0$, so daß $u = U(u)$, d.h. u ist Lösung der Differentialgleichung (4.25).

Beweis der Stetigkeit:

Sei (w_j) eine Folge in W_0 , die schwach gegen ein $w \in W_0$ konvergiert.

Sei $u_j := U(w_j)$. Mit den obigen Schranken existiert eine Teilfolge, ebenfalls mit (w_j) bezeichnet, so daß:

$$u_j \rightharpoonup u \text{ in } L^2(0, T; H^{1,2}(\Omega)) \quad (4.32)$$

$$\frac{du_j}{dt} \rightharpoonup \frac{du}{dt} \text{ in } L^2(0, T; H^{-1,2}(\Omega)) \quad (4.33)$$

$$\frac{\partial u_j}{\partial x_i} \rightharpoonup \frac{\partial u}{\partial x_i} \text{ in } L^2(0, T; L^2(\Omega)), \quad i=1 \dots d \quad (4.34)$$

Wegen der Kompaktheit von $W(0, T)$ in $L^2(0, T; L^2(\Omega))$ (4.5) existiert eine weitere Teilfolge (w_j) , so daß:

$$\|w_j(t) - w(t)\|_{L^2(\Omega)} \rightarrow 0 \text{ fast überall in } (0, T) \quad (4.35)$$

$$\|u_j(t) - u(t)\|_{L^2(\Omega)} \rightarrow 0 \text{ fast überall in } (0, T) \quad (4.36)$$

und damit nach Lemma 4.3 und 4.4

$$\begin{aligned} \|D_\varepsilon(u_j(t)) - D_\varepsilon(u(t))\|_{L^2(\Omega)} &\rightarrow 0 \quad \text{fast überall in } (0, T) \\ \|f_\varepsilon(u_j(t)) - f_\varepsilon(u(t))\|_{L^2(\Omega)} &\rightarrow 0 \quad \text{fast überall in } (0, T) \end{aligned} \quad (4.37)$$

Mit der Beschränktheit der $w_j(t), u_j(t), D(u_j), f(u_j)$ folgt nach dem Lebesgue-Konvergenzsatz die Konvergenz von $D_\varepsilon(u_j)$ und $f_\varepsilon(u_j)$ in $L^2(0, T; L^2(\Omega))$. Wir können also in (4.26) den Grenzübergang durchführen und erhalten $u = U(w)$. Aufgrund der Eindeutigkeit der Lösung zu P_w konvergiert die ganze Folge schwach gegen $u = U(w)$.

Offensichtlich gilt auch $\|u\|_{L^\infty(\mathbb{L}^\infty)} \leq c_1$, also $u \in W_0$.

2. Eindeutigkeit:

Seien $\bar{u}, \hat{u}$ zwei Lösungen der Differentialgleichung (4.25) zum selben Anfangswert u_0 .

Für beliebige Testfunktion $\phi \in H^{1,2}(\Omega)$ und fast alle t in $[0, T]$ gilt dann

$$\begin{aligned} \langle \phi, \partial_t \bar{u}(t) - \operatorname{div}(D_\varepsilon(\bar{u}(t)) \nabla \bar{u}(t)) \rangle &= \langle \phi, f_\varepsilon(\bar{u}(t)) \rangle \quad \langle D_\varepsilon(\bar{u}(t)) \nabla \bar{u}(t), N \rangle = 0 \\ \langle \phi, \partial_t \hat{u}(t) - \operatorname{div}(D_\varepsilon(\hat{u}(t)) \nabla \hat{u}(t)) \rangle &= \langle \phi, f_\varepsilon(\hat{u}(t)) \rangle \quad \langle D_\varepsilon(\hat{u}(t)) \nabla \hat{u}(t), N \rangle = 0 \end{aligned}$$

Daraus erhält man

$$\begin{aligned} & \langle \phi, \partial_t(\bar{u}(t) - \hat{u}(t)) - \operatorname{div}[D_\varepsilon(\bar{u}(t))(\nabla \bar{u}(t) - \nabla \hat{u}(t))] \rangle \\ &= \langle \phi, \operatorname{div}[(D_\varepsilon(\bar{u}(t)) - D_\varepsilon(\hat{u}(t)))\nabla \hat{u}(t)] + f_\varepsilon(\bar{u}(t)) - f_\varepsilon(\hat{u}(t)) \rangle \end{aligned} \quad (4.38)$$

Als Testfunktion würden wir nun gerne $\phi := \bar{u}(t) - \hat{u}(t)$ wählen. Diese Funktion liegt jedoch nicht in $H^{1,2}(0, T; H^{1,2}(\Omega))$. Wir werden den Beweis der Eindeutigkeit formal mit dieser Testfunktion führen. Durch Einsetzen von ϕ in (4.38) erhält man:

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} (\|\bar{u}(t) - \hat{u}(t)\|_{L^2(\Omega)}^2) + \nu \|\nabla \bar{u}(t) - \nabla \hat{u}(t)\|_{L^2(\Omega)}^2 \\ & \leq \|D_\varepsilon(\bar{u}(t)) - D_\varepsilon(\hat{u}(t))\|_{L^\infty(\Omega)} \|\nabla \hat{u}(t)\|_{L^2(\Omega)} \|\nabla \bar{u}(t) - \nabla \hat{u}(t)\|_{L^2(\Omega)} \\ & \quad + \|f_\varepsilon(\bar{u}(t)) - f_\varepsilon(\hat{u}(t))\|_{L^2(\Omega)} \|\bar{u}(t) - \hat{u}(t)\|_{L^2(\Omega)} \end{aligned} \quad (4.39)$$

Mit Lemma 4.3 und 4.4 folgt:

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} (\|\bar{u}(t) - \hat{u}(t)\|_{L^2(\Omega)}^2) + \nu \|\nabla \bar{u}(t) - \nabla \hat{u}(t)\|_{L^2(\Omega)}^2 \\ & \leq C \|\bar{u}(t) - \hat{u}(t)\|_{L^2(\Omega)} \|\nabla \hat{u}(t)\|_{L^2(\Omega)} \|\nabla \bar{u}(t) - \nabla \hat{u}(t)\|_{L^2(\Omega)} \\ & \quad + C \|\bar{u}(t) - \hat{u}(t)\|_{L^2(\Omega)}^2 \\ & \stackrel{\text{Youngsche Ungl.}}{\leq} \frac{2}{\nu} C \|\bar{u}(t) - \hat{u}(t)\|_{L^2(\Omega)}^2 \|\nabla \hat{u}(t)\|_{L^2(\Omega)} \\ & \quad + \frac{\nu}{2} \|\nabla \bar{u}(t) - \nabla \hat{u}(t)\|_{L^2(\Omega)}^2 + C \|\bar{u}(t) - \hat{u}(t)\|_{L^2(\Omega)}^2 \end{aligned} \quad (4.40)$$

Also

$$\begin{aligned} & \frac{d}{dt} (\|\bar{u}(t) - \hat{u}(t)\|_{L^2(\Omega)}^2) \\ & \leq \frac{4}{\nu} C \|\bar{u}(t) - \hat{u}(t)\|_{L^2(\Omega)}^2 \|\nabla \hat{u}(t)\|_{L^2(\Omega)} + C \|\bar{u}(t) - \hat{u}(t)\|_{L^2(\Omega)}^2 \end{aligned} \quad (4.41)$$

Daraus folgt mit $\bar{u}(0) = \hat{u}(0)$ und dem Gronwall-Lemma

$$\|\bar{u}(t) - \hat{u}(t)\|_{L^2(0,T;L^2(\Omega))} = 0$$

Eine zulässige Testfunktion erhält man wie folgt:

Man wählt zu beliebigem $\delta > 0$ eine Zerlegung des Intervalls $[0, T]$ in Teilintervalle $I_i := [\delta, (i+1)\delta]$ mit $i = 0 \dots \frac{T}{\delta} - 1$ und setzen $\phi|_{I_i} := \text{konst.} = \int_{I_i} \bar{u}(t) - \hat{u}(t)$. Dann ist $\phi|_{I_i} \in H^{1,2}(I_i; H^{1,2}(\Omega))$. Durch Einsetzen in (4.38) lassen ähnliche, zeitdiskrete Abschätzungen wie im formalen Beweis erbringen. Ein diskretes Gronwall-Argument liefert dann $\|\bar{u}(t) - \hat{u}(t)\| = 0$.

□

Bemerkung 1 *Beweis im Fall $n > 1$*

Im Existenzbeweis hatten wir den Raum W_0 so definiert, daß für alle $u \in W_0$ $\|u\|_{L^\infty([0,T] \times \Omega)} \leq 1$ gilt. Die in Lemma 4.2 und 4.1 formulierten Abschätzungen lassen sich analog für den Fall $n > 1$ herleiten, da wir für $u, v \in W_0$ Terme der Form $\int_\Omega u^n \dots dx$ durch $\|u^{n-1}\|_{L^\infty([0,T] \times \Omega)} \int_\Omega u \dots dx$ und Terme der Form $\int_\Omega (u^n - v^n) \dots dx$ durch

$n \max\{\|u\|_{L^\infty([0,T] \times \Omega)}, \|v\|_{L^\infty([0,T] \times \Omega)}\} \int_\Omega (u - v) \dots dx$ abschätzen können.

Kapitel 5

Implementierung

5.1 Gitter

Seien $P_1, P_2 : \Omega^{3D} \rightarrow \mathbb{R}^2$ zwei auf das Gebiet Ω^{3D} eingeschränkte orthogonale Projektionen. Der Winkel $\alpha \in (0, \pi)$ zwischen beiden Projektionsrichtungen sei bekannt. Nach Wahl des 3D-Koordinatensystems können wir annehmen, daß die P_i folgende Gestalt haben:

$$\begin{aligned} P_1(x, y, z) &:= (x, z) \\ P_2(x, y, z) &:= (\cos(\alpha)x + \sin(\alpha)y, z) \end{aligned}$$

Abbildung 5.1 zeigt die Anordnung der Projektionsrichtungen.

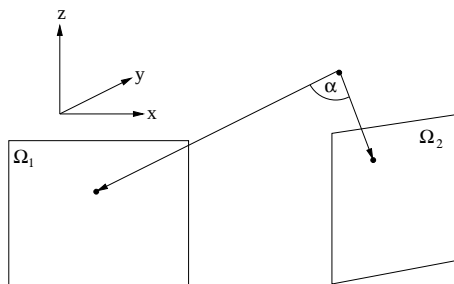
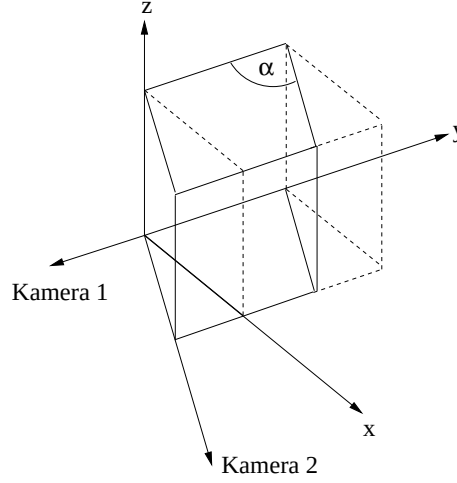


Abbildung 5.1: Anordnung der Kameras

Wir wollen annehmen, daß Ω^{3D} durch die beiden Projektionen auf rechteckige Gebiete Ω_i abgebildet wird. Die Abmessungen beider Rechtecke Ω_i seien gleich: $\Omega_i = [0, a] \times [0, b], i = 1, 2$ mit $a, b > 0$.

Ω^{3D} ist daher ein Parallelepiped mit senkrecht auf der Grundfläche stehenden Kanten (Abb. 5.2). Die Eckpunkte $p_1 \dots p_8$ des Parallelepipeds sind

Abbildung 5.2: Das Gebiet Ω^{3D}

$$\begin{aligned}
 p_1 &= A(0, 0, 0) &= (0, 0, 0) \\
 p_2 &= A(a, 0, 0) &= (a, -a \cot(\alpha), 0) \\
 p_3 &= A(a, a, 0) &= (a, a(\frac{1}{\sin(\alpha)} - \cot(\alpha)), 0) \\
 p_4 &= A(0, a, 0) &= (0, \frac{a}{\sin(\alpha)}, 0) \\
 p_i &= p_{i-4} + (0, 0, b) \text{ für } i = 5 \dots 8
 \end{aligned}$$

wobei A die in 3.5.2 vorgestellte Matrix ist.

Die zweidimensionalen Gebiete versehen wir nun mit gleichmäßigen Gittern mit einer gitterunabhängigen Zahl $N+1$ an Gitterpunkten in jede Richtung. Seien $h_x := a/N, k = 1, 2$ und $h_y := b/N$. Die Gitterpunkte sind also durch $G_k = \{(i h_x, j h_y) \in \mathbb{R}^2 | i, j \in \{0 \dots N\}\}, k = 1, 2$ gegeben.

Sei

$$\begin{aligned}
 V_h(\Omega_k) &:= \\
 &\{f \in C^0(\Omega_k) | f|_E \text{ ist Polynom 1. Grades für jede Gitterzelle } E \text{ von } G\}
 \end{aligned}$$

Die vorliegenden DNA-Bilder interpretieren wir nun als stetige Funktionen $\tilde{u}_k, k = 1, 2$ auf Ω_k , indem wir die auf das Intervall $[0, 1]$ skalierten Grauwerte der Pixel den einzelnen Gitterpunkten zuordnen und die zu diesen Gitterpunkten eindeutig bestimmte Funktion aus $V_h(\Omega_k)$ wählen.

Aus den zweidimensionalen Gittern konstruieren wir nun ein dreidimensionales Gitter auf Ω^{3D} , dessen Gitterpunkte durch die Projektionen P_1, P_2 auf die Gitterpunkte der beiden 2D-Gitter abgebildet werden: Sei G die Menge der Gitterpunkte

$$G := \{x_{ijk} | P_1(x_{ijk}) = x_{ij}^1, P_2(x_{ijk}) = x_{kj}^2\}_{i,j,k \in \{0 \dots N\}}$$

wobei x_{ij}^k die Koordinaten der Gitterpunkte von G_k bezeichnet.
Ein Gitterpunkt $x_{ijk} \in G$ besitzt dementsprechend die Koordinaten

$$A(i h_x, j h_x, i h_y)^T$$

Abb.5.3 zeigt das dreidimensionale Gitter.

Die Koordinaten $A^{-1}x$ eines Punktes $x \in \Omega^{3D}$ bezeichnen wir im folgenden als die *Gitterkoordinaten*. Die in Abschnitt 3.5.2 vorgestellte Tensormetrik entspricht also der kanonischen Metrik der Gitterkoordinaten. Die Gitterzellen bezeichnen wir mit $E_i, i = 1 \dots N_{Zellen}$, ihre Mittelpunkte mit $z_i, i = 1 \dots N_{Zellen}$.

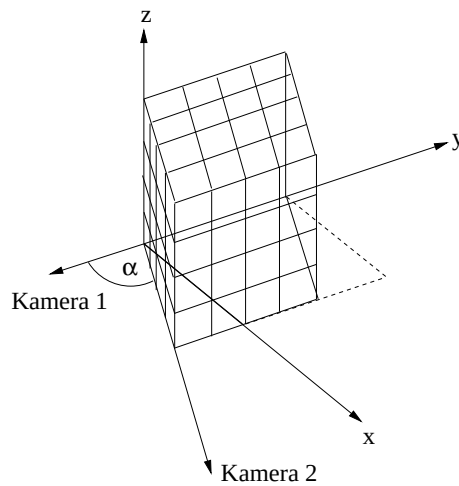


Abbildung 5.3: Das dreidimensionale Gitter

Sei

$$V_h(\Omega^{3D}) := \{f \in C^0(\Omega^{3D}) | f|_E \text{ ist Polynom 1. Grades für jede Gitterzelle } E \text{ von } G\}$$

5.2 Num. Verfahren zur Lösung der Diffusionsgleichung

Mit Hilfe des Diffusionsprozesses sollen, wie in Abschnitt 3.1.2 dargelegt, die Stereobilder vorgeglättet werden. Wir beschränken uns daher bei der Beschreibung des numerischen Verfahrens auf den Fall $d = 2$.

5.2.1 Zeitdiskretisierung

Betrachten wir die schwache Formulierung der Differentialgleichung (3.9):

Problem (P):

Zu $u_0 \in L^\infty(\Omega)$ ist $u \in L^2((0, T], H^{1,2}(\Omega))$ gesucht, so daß gilt:

$$\begin{aligned} \langle w, \partial_t u \rangle - \langle \nabla w, D_\varepsilon(u) \nabla u \rangle &= \langle w, f_\varepsilon(u) \rangle && \text{auf } (0, T] \times \Omega \quad \forall w \in H^{1,2}(\Omega) \\ u(0) &= u_0 && \text{f.ü. in } \Omega \\ \langle D_\varepsilon(u) \nabla u, n \rangle &= 0 && \text{f.ü. auf } [0, T] \times \partial\Omega \end{aligned} \quad (5.1)$$

mit $D_\varepsilon(u), f_\varepsilon(u)$ wie in 3.3 definiert.

Seien $N \in \mathbb{N}$ und $\tau := T/N$.

Wir wollen das Problem (P) näherungsweise für die Zeitpunkte $t_n := n \cdot \tau, n = 1 \dots N$ lösen. Dazu ersetzen wir die Zeitableitung durch einen Rückwärtsdifferenzenquotienten. Die Nichtlinearität wird aufgelöst, indem die Terme $D_\varepsilon(u)$ und $f_\varepsilon(u)$ zum zuletzt berechneten Zeitpunkt ausgewertet werden. Das Verfahren ist somit semi-implizit.

Setze $u^0 := u_0$.

Wir erhalten eine Folge $\{u^n\}_{0 \leq n \leq N} \in H^{1,2}(\Omega)$ durch sukzessives Lösen der folgenden Differentialgleichung:

Problem (P^τ)

Zu gegeben $u^n, n \geq 0$ ist $u^{n+1} \in H^{1,2}(\Omega)$ gesucht, so daß für alle $w \in H^{1,2}(\Omega)$ gilt:

$$\langle w, u^{n+1} \rangle + \tau \langle \nabla w, D_\varepsilon(u^n) \nabla u^{n+1} \rangle = \langle w, u^n + \tau \cdot f_\varepsilon(u^n) \rangle \quad (5.2)$$

Bemerkung:

Die Existenz und Eindeutigkeit einer Lösung zu (P^τ) ist durch den Satz von Lax-Milgram [1] gesichert.

5.2.2 Ortsdiskretisierung mittels FE-Methode

Mit Hilfe des Finite-Elemente-Ansatzes werden wir das Problem (P^τ) auf einem endlich-dimensionalen Unterraum $V_h \subset H^{1,2}(\Omega)$ lösen.

Durch das in 5.1 vorgestellte Gitter erhalten wir eine reguläre Zerlegung $\cup_{i=1 \dots N_E} E_i = \Omega$ des Gebiets Ω in kleine Rechtecke E_i . Sei $h > 0$ die maximale Gitterweite. Setze $\mathcal{T}_h := \{E_i\}_{i=1 \dots N_E}$.

Definiere

$$V_h = V_h(\Omega, \mathcal{T}_h) := \{v \in C^0(\Omega) \mid v|_{E_i} \text{ ist Polynom 1. Grades } \forall E_i \in \mathcal{T}_h\}$$

(Wir werden später neben $\mathcal{T}_h$ andere Triangulierungen verwenden.)

Dann bildet V_h mit der $H^{1,2}$ -Norm einen Unterraum von $H^{1,2}(\Omega)$. Es gilt zudem $V_h \subset C^0(\Omega)$.

Sei $\Pi_h : H^{1,2}(\Omega) \rightarrow V_h$ die Projektion auf den Unterraum.

Sei $\mathcal{N}(\mathcal{T}_h) := \{x_j\}_{j=1 \dots N_K}$ die Menge der Eckpunkte der $\{E_i\}_{i=1 \dots N_E}$, die wir im folgenden auch als Knoten bezeichnen. $v \in V_h$ ist eindeutig bestimmt durch Vorgabe der Werte $v(x_j), j \in \{1 \dots N_K\}$.

Zu $i \in \{1 \dots N_K\}$ sei $\phi_i \in V_h$ definiert durch $\phi_i(x_j) := \delta_{ij}$ für alle $j = 1 \dots N_K$. $\{\phi_i\}_{i=1 \dots N_K}$ ist somit Basis des V_h .

Beim Lösen des Problems (P^τ) beschränken wir uns nun auf V_h als Lösungs- und Testraum. Da die Differentialgleichung linear bezüglich der Testfunktion ist, reicht es aus, mit den Basisfunktionen zu testen.

Zur Vereinfachung der Notation verwenden wir die Einsteinsche Summenkonvention, die besagt, daß über doppelt auftretende Indizes summiert wird.

Wir approximieren $u^n, n \geq 0$ durch $u_h^n := \Pi_h u^n =: U_i^n \phi_i$ mit geeignetem Vektor $(U_i^n)_i \in \mathbb{R}^{N_K}$.

Den Term $D_\varepsilon(u^n)$ ersetzen wir durch eine Treppenfunktion $D_{\varepsilon,h}(U_i^n \phi_i), D_h|_E = \text{konst.}$ für alle $E \in \mathcal{T}_h$. Auf die Wahl dieser Approximation werden im nächsten Abschnitt eingehen. Den Term $f_\varepsilon(u^n)$ ersetzen wir durch $f_{\varepsilon,h}(u_h^n) := \Pi_h f_\varepsilon(u_h^n) =: F_i^n \phi_i$ mit $(F_i^n)_i \in \mathbb{R}^{N_K}$.

Wir erhalten dann folgendes Gleichungssystem:

Problem (P_h^τ) :

Zu gegebenem $(U_i^n)_i \in \mathbb{R}^{N_K}$ suche $(U_i^{n+1})_i \in \mathbb{R}^{N_K}$ so daß für alle $j = 1 \dots N_K$ gilt:

$$\langle \phi_j, \phi_i \rangle U_i^{n+1} + \tau \langle \nabla \phi_j, D_{\varepsilon,h}(U_k^n \phi_k) \nabla \phi_i \rangle U_i^{n+1} = \langle \phi_j, \phi_i \rangle (U_i^n + \tau F_i^n) \quad (5.3)$$

Wir definieren nun die globale Massematrix $M \in \mathbb{R}^{N_K \times N_K}$ durch $M_{ij} := \langle \phi_i, \phi_j \rangle$ und $L(V) \in \mathbb{R}^{N_K \times N_K}$ durch $L_{ij}(V) := \langle \nabla \phi_i, D_{\varepsilon,h}(V_k \phi_k) \nabla \phi_j \rangle$ für beliebiges $V \in \mathbb{R}^{N_K}$. Dadurch erhalten wir folgende kompakte Formulierung:

$$MU^{n+1} + \tau L(U^n)U^{n+1} = M(U^n + \tau F^n) \quad (5.4)$$

wobei U^n und F^n die Vektoren $(U_i^n)_i$ und $(F_i^n)_i$ bezeichnen.

Die Matrizen M und L sind durch Integrale über Ω definiert. Da eine Basisfunktion ϕ den Träger $T_i := \{E \in \mathcal{T}_h | x_i \in E\}$ besitzt, beschränkt sich die Integration bei der Berechnung von M_{ij} und L_{ij} auf das Gebiet $T_i \cap T_j = \{E \in \mathcal{T}_h | x_i, x_j \in E\}$. Da $T_i \cap T_j = \emptyset$ ist, falls x_i, x_j zu keinem gemeinsamen Rechteck gehören, sind die Matrizen dünn besetzt. Zur Berechnung der Matrizen durch numerische Integration sowie das Verwenden gelumpfter Massen verweisen wir auf die Literatur.

5.2.3 Approximation der Diffusionsmatrix

Definition der Treppenfunktion $D_{\varepsilon,h}$

Sei $w \in V_h$ beliebig. Wir wollen eine Treppenfunktion $D_{\varepsilon,h}(w)$ definieren, $D_{\varepsilon,h}(w)|_{E_i} = \text{konst.}$, so daß $D_{\varepsilon,h}(w)(x) \rightarrow D(w)(x)$ für $h \rightarrow 0$ in jedem Punkt $x \in \Omega$. Seien $z_i \in \Omega$ die Mittelpunkte der $E_i \in \mathcal{T}_h$.

Betrachten wir ein festes Rechteck E mit Mittelpunkt z .

O.B.d.A. habe E den Mittelpunkt $z = (0, 0)$, die Seitenlängen h_x, h_y und folgende Eckpunkte $e_1 \dots e_4$:

$$\begin{aligned} e_1 &= \left(-\frac{h_x}{2}, -\frac{h_y}{2} \right) \\ e_2 &= \left(\frac{h_x}{2}, -\frac{h_y}{2} \right) \\ e_3 &= \left(\frac{h_x}{2}, \frac{h_y}{2} \right) \\ e_4 &= \left(-\frac{h_x}{2}, \frac{h_y}{2} \right) \end{aligned}$$

Wir berechnen nun näherungsweise die lokale Masse $m_h(w)$, den Schwerpunkt $sp_h(w)$ und das Massenmoment $M_h(w)$ im Punkt z , indem wir in den Formeln für die lokale Masse, Schwerpunkt und Massenmoment aus Abschnitt 3.2 die Integration über $B_\varepsilon(0)$ durch eine endliche Summe über alle Gitterpunkte $B_h(z) := \{x_j | \|x_j - z\|_2 < \varepsilon\}$ mit geeigneten Gewichten ersetzen.

Die Elemente der Triangulierung sind kongruent mit Flächeninhalt $\text{area}(E) = \text{konst.} = h_x h_y$. Als Gewichte wählen wir $\text{area}(E)$. (Dies entspricht der Schwerpunktsintegration über ein versetztes Gitter, zu dem die bisherigen Knoten die Elementmitten darstellen.) Nach Übergang von δ zu $\tilde{\delta} := \frac{\delta}{\text{area}(E)}$ läßt sich $\text{area}(E)$ kürzen. Wir erhalten:

$$\begin{aligned} m_h(z) &:= \sum_{x \in B_h(z)} w^n(x) \, dx \\ sp_h(z) &:= \frac{1}{m_h(z) + \tilde{\delta}} \sum_{x \in B_h(z)} w^n(x) \cdot (x - z) \, dx \\ M_h(z) &:= \frac{1}{m_h(z) + \tilde{\delta}} \sum_{x \in B_h(z)} w^n(x) (x - z - sp_h(z)) \otimes (x - z - sp_h(z)) \, dx \end{aligned}$$

Zu Definition von B_h verwenden wir die 2-Norm statt der ∞ -Norm, die zwar leichter zu implementieren wäre, aber aufgrund der quadratischen Form des Integrationsgebiets zu Artefakten in den Funktionen führen.

Die Eigenwerte von $M_h =: \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ sind, falls $a_{12} = a_{21} \neq 0$ ist, durch

$$\lambda_{1,2}^h = \frac{1}{2} \left(a_{11} + a_{22} \pm \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2} \right) \quad (5.5)$$

zu den Eigenvektoren $v_1^h \perp v_2^h$ mit

$$v_1^h \parallel \begin{pmatrix} 2a_{12} \\ (a_{22} - a_{11} + \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2}) \end{pmatrix} \neq 0 \quad (5.6)$$

gegeben. Ist M_h Diagonalmatrix, so ist die Bestimmung der Eigenwerte und -vektoren trivial.

Wir definieren $D_{\varepsilon,h}(w)(z)$ im Mittelpunkt z durch Vorgabe der Eigenvektoren v_1^h, v_2^h zu den Eigenwerten $\eta_1(z) := 1$ und $\eta_2(z)$, wobei $\eta_2(z) \approx g(\text{coh}(w)|\nabla w_\sigma(z)|^2)$ noch geeignet bestimmt werden muß.

Sei dazu $\tilde{w}$ eine Approximation von w_σ . Die Wahl von $\tilde{w}$ beschreiben wir am Ende dieses Abschnitts. Zur Approximation der ersten Ableitungen in x - und y -Richtung verwenden wir Differenzenquotienten:

$$\begin{aligned} D_x \tilde{w}(z) &:= \frac{1}{2h_x} ((\tilde{w}(e_1) + \tilde{w}(e_2)) - (\tilde{w}(e_0) + \tilde{w}(e_3))) \\ D_y \tilde{w}(z) &:= \frac{1}{2h_y} ((\tilde{w}(e_3) + \tilde{w}(e_2)) - (\tilde{w}(e_0) + \tilde{w}(e_1))) \end{aligned}$$

Wir approximieren den Gradienten von $\tilde{w}$ durch

$$\nabla_h \tilde{w}(z) := \begin{pmatrix} D_x \tilde{w}(z) \\ D_y \tilde{w}(z) \end{pmatrix}$$

und setzen $\eta_2(z) := g(\text{coh}(w)|\nabla_h \tilde{w}(z)|^2)$

Somit ist die Diffusionsmatrix $D_{\varepsilon,h}$ auf den Elementmitten definiert. Wir setzen sie auf ganz E konstant fort.

Es gilt offensichtlich $D_{\varepsilon,h}(w) \rightarrow D_\varepsilon(w)$ punktweise für $h \rightarrow \infty$.

Es ist nicht nötig, die Matrix $D_{\varepsilon,h}(w)(z)$ explizit aufzustellen. Wir setzen $v = \begin{pmatrix} v^{(1)} \\ v^{(2)} \end{pmatrix} := v_1^h$ und η_1, η_2 elementweise konstant fort und erhalten so Treppenfunktionen auf Ω .

Die Matrix $L_{ij}(w), w \in V_h$ berechnen wir wie folgt:

$$L_{ij}(w) = \sum_E L_{ij}^E(w)$$

$$\begin{aligned}
L_{ij}^E(w) &:= \int_E \nabla \phi_i D_{\varepsilon,h}(w) \nabla \phi_j dx \\
&= \int_E \nabla \phi_i D_{\varepsilon,h}(w) \underbrace{((\nabla \phi_j \cdot v)v + (\nabla \phi_j - (\nabla \phi_j \cdot v)v))}_{\text{Projektion von } \nabla \phi_j \text{ auf die Eigenräume zu } \eta_1 \text{ und } \eta_2} dx \\
&= \int_E \eta_1 (\nabla \phi_i \cdot v)(\nabla \phi_j \cdot v) + \eta_2 (\nabla \phi_i \cdot \nabla \phi_j - (\nabla \phi_i \cdot v)(\nabla \phi_j \cdot v)) \\
&= \int_E (\eta_1 - \eta_2)(\nabla \phi_i \cdot v)(\nabla \phi_j \cdot v) + \eta_2 (\nabla \phi_i \cdot \nabla \phi_j) \\
&= (\eta_1 - \eta_2) (b_{ij}^{11}(v^{(1)})^2 + b_{ij}^{12}v^{(1)}v^{(2)} + b_{ij}^{22}(v^{(2)})^2) + \eta_2 \tilde{L}_{ij}
\end{aligned}$$

mit

$$\begin{aligned}
b_{ij}^{11} &= \int_{\Omega} \nabla \phi_i^{(1)} \nabla \phi_j^{(1)} dx \\
b_{ij}^{12} &= \int_{\Omega} \nabla \phi_i^{(1)} \nabla \phi_j^{(2)} + \nabla \phi_i^{(2)} \nabla \phi_j^{(1)} dx \\
b_{ij}^{22} &= \int_{\Omega} \nabla \phi_i^{(2)} \nabla \phi_j^{(2)} dx \\
\tilde{L}_{ij} &= \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j dx
\end{aligned}$$

Hierbei seien $\nabla \phi_i^{(k)}$, $k = 1, 2$ die Komponenten von $\nabla \phi_i$. $\tilde{L}_{ij}$ ist die aus der Literatur bekannte globale Steifigkeitsmatrix. Die Matrizen $b_{ij}^{11}, b_{ij}^{12}, b_{ij}^{22}$ bezeichnen wir ebenfalls als globale Matrizen (vgl. den folgenden Abschnitt “Lokale Matrizen”).

Numerische Vorglättung

Zur Berechnung der Treppenfunktion $D_{\varepsilon,h}(w)$ benötigen wir noch eine Approximation $\tilde{w}$ von w_{σ} . Wir bestimmen diese Approximation, in dem wir numerisch die zugehörige Differentialgleichung zur Zeit $T = 1$ lösen:

Sei $W^0 \in \mathbb{R}^{N_K}$ so gewählt, daß $w = W_k^0 \phi_k$ gilt. Zu beliebigem $N \in \mathbb{N}$ und $\tau = \frac{1}{N}$ bestimmen wir iterativ die Lösung $\tilde{W}^n \in \mathbb{R}^{N_K}$, $n = 1 \dots N$ des Gleichungssystems

$$(M_{ij} + \tau \tilde{L}_{ij}) \tilde{W}_j^{n+1} = M_{ij} W_j^n$$

und setzen $\tilde{W} := \tilde{W}_k^N \phi_k$.

5.2.4 Lokale Matrizen

Die globalen Matrizen $M_{ij}, b_{ij}^{11}, b_{ij}^{12}, b_{ij}^{22}$ und $\tilde{L}_{ij}$ hängen nur von der Geometrie des Gitters ab, sie brauchen also nur einmal berechnet zu werden. Die Integration erfolgt elementweise und kann nach der Transformationsregel auf Integration auf das Referenzelement $\hat{E} := [0, 1] \times [0, 1]$ zurückgeführt werden.

Wir definieren zu den Basisfunktionen von $V_h(\hat{E}, \{\hat{E}\})$ die *lokalen Matrizen*

$\hat{M}_{ij}, \hat{b}_{ij}^{11}, \hat{b}_{ij}^{12}, \hat{b}_{ij}^{22}, \hat{\tilde{L}}_{ij}$ wie oben.

Die Berechnung einer globalen Matrix läßt sich nun auf eine Summierung der Einträge der entsprechenden lokalen Matrix mit geeigneten Faktoren zurückführen. Zur ausführlichen Erläuterung verweisen wir auf [6],[19].

5.2.5 Berechnung des Matrixproduktes

Bei der Anwendung unseres Verfahrens arbeiten wir in der Regel mit 257×257 Gitterpunkten, also ca. 66000 Unbekannten des Gleichungssystems (5.1). Die Matrix $M + \tau L$ ist, wie oben dargelegt, dünn besetzt; sie besitzt etwa 600000 von null verschiedene Einträge. Da sich diese Einträge leicht aus den Einträgen der lokalen Matrizen berechnen lassen, ist eine Speicherung der einer solchen Matrix nicht sinnvoll. (Bei der Implementierung des Diffusionsmodells in 3D ist aufgrund der großen Zahl an Unbekannten ein sparsamer Gebrauch des Speicherplatzes wichtig. Da wir uns auf die Implementierung in 2D konzentrieren, gehen wir auf diese Schwierigkeiten nicht genauer ein.)

Wir umgehen das Speichern der Matrix, indem wir ein iteratives Verfahren zur Lösung des Gleichungssystems, hier ist es das Konjugierte-Gradienten-Verfahren, verwenden, das keine explizite Darstellung der Matrix, sondern lediglich das Matrix-Vektor-Produkt zu beliebigen Vektoren benötigt.

Die Berechnung des Matrix-Vektor-Produkts erfolgt in unserem Verfahren “on-the-fly”, d.h. durch einen Durchlauf über alle Gitterelemente. Bei jedem Durchlauf werden die Einträge der Matrix neu ermittelt. Daß die Matrizen dünn besetzt sind, ist dabei von großem Vorteil für die Berechnungsdauer.

Wir beschreiben die “On-the-fly”-Berechnung am Beispiel der Massenmatrix. Zu $E \in \mathcal{T}_h$ mit den Kantenlängen h_x, h_y definieren wir die Matrix $M_{ij}^E := \int_E \phi_i \phi_j dx$ mit acht von null verschiedenen Einträgen, die sich durch Skalierung der Einträge der lokalen Matrix mit dem Faktor $h_x h_y$ ergibt.

(Hierbei ist auf die richtige Identifizierung der global numerierten Knoten von E mit den lokal numerierten Knoten von $\hat{E}$ zu achten.)

Es gilt:

$$(Mv)_i = \sum_j \left(\sum_E M_{ij}^E \right) v_j = \sum_E \sum_j (M_{ij}^E v_j)$$

Nun machen wir uns zunutze, daß zu festem E $M_{ij}^E = 0$ gilt, falls $x_j \notin E$. Es gilt also

$$(Mv)_i = \sum_E \left(\sum_{j, x_j \in E} M_{ij}^E v_j \right)$$

In die äußere Summe gehen jedoch nur die Rechtecke E ein, für die $x_i \in E$ gilt. Die Berechnung des Matrixprodukts kann also durch folgenden Algorithmus erfolgen:

```

Initialisiere Vektor  $((Mv)_i)_{i=1\dots N_K}$  mit 0.
Durchlaufe alle  $E \in \mathcal{T}$ 
  Durchlaufe die vier Knoten  $x_i \in E$ 
    Durchlaufe die vier Knoten  $x_j \in E$ 
       $(Mv)_i = (Mv)_i + M_{ij}^E v_j$ 

```

Die Berechnung der übrigen globalen Matrizen erfolgt analog. Die Einträge der Matrizen berechnen sich aus den entsprechend der Transformationsregel skalierten Einträgen der lokalen Matrizen.

5.2.6 Löser

Wir verwenden das Konjugierte-Gradienten-Verfahren. Für Details verweisen wir auf [19][20].

5.2.7 Adaptivität

Triangulierung

Um die Effizienz des Lösungsverfahrens zu steigern, gehen wir von der bisher beschriebenen Triangulierung zu einer adaptiven Triangulierung über. Da die Ausgangsdaten diskret vorliegen, ist eine grundlegende Triangulierung gegeben, die nicht weiter verfeinert werden soll. Vielmehr können wir durch Vergrößerung dieser Triangulierung an geeigneten Stellen die Zahl der Unbekannten reduzieren, ohne die wichtigen Strukturen in Ω zu verlieren.

Die Aufstellung eines adaptiven Gitters erfolgt gemäß [19], wir fassen die wichtigsten Punkte im folgenden zusammen:

Zur Vereinfachung der Notation nehmen wir an, daß Ω quadratisch mit Seitenlängen h_0 ist. Die grundlegende Triangulierung sei durch ein gleichmäßiges Gitter mit $2N + 1$ Knoten in jede Richtung gegeben.

Zunächst beschreiben wir den rekursiven Aufbau der grundlegenden Triangulierung. Ausgehend von einer Makrotriangulierung $\mathcal{T}_{h_0}^0$ verfeinern wir jedes Quadrat E mit Seitenlänge h_i einer Triangulierung $\mathcal{T}_{h_i}^0$ durch vier Teilquadrate der Seitenlänge $h_{i+1} := \frac{h_i}{2}$ und erhalten so eine neue Triangulierung $\mathcal{T}_{h_{i+1}}^{i+1}$ mit dem Verfeinerungslevel $i + 1$. Mit Verfeinerungslevel N erhalten wir die grundlegende Triangulierung $\mathcal{T}_{h_N}^N$. Im folgenden sei $\mathcal{T}^i := \mathcal{T}_{h_i}^i$.

Adaptivität erreichen wir nun, indem die Verfeinerung durch ein Stoppkriterium S steuern: Ein Element $E \in \mathcal{T}^i, i < N$ wird nur dann weiter verfeinert, falls eine

Aussage $S(E)$ nicht wahr ist. Auf diese Weise erhalten wir eine Triangulierung $\mathcal{T}_S$. Abb. 5.4 zeigt ein Beispiel für eine adaptive Triangulierung.

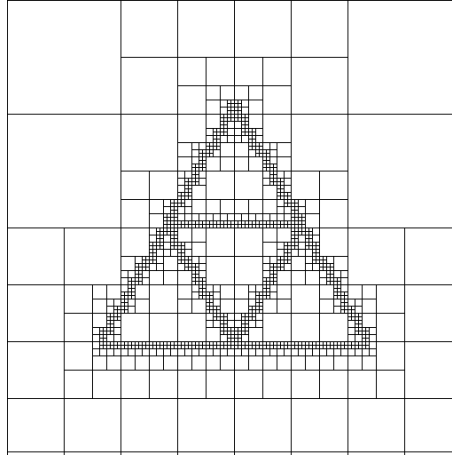


Abbildung 5.4: Beispiel für ein adaptives Gitter

Das Stoppkriterium werden wir in Abhängigkeit von der zuletzt berechneten Lösung $u_h^n = U_k^n \phi_k$ wählen, wir erhalten somit für jeden Zeitschritt ein neues adaptives Gitter.

Hängende Knoten

In einem nichtadaptiven Gitter ist jeder Knoten Eckpunkt der angrenzenden Elemente. Bei einem wie oben konstruierten adaptiven Gitter sind Knoten auf einer Kante zwischen zwei Elementen E_1, E_2 unterschiedlichen Levels keine Eckpunkte des groberen Elements. Solche Knoten nennen wir *hängende Knoten*, alle übrigen *freie Knoten*.

Als Beispiel betrachten wir die in Abb. 5.5 dargestellte Situation. Der Knoten b liegt auf der Mitte der Kante zwischen a und c und ist kein Eckpunkt von E_1 , also ein hängender Knoten. Sei $f \in V_h(\Omega, \mathcal{T}_S)$. Durch die Stetigkeit von f ergibt sich folgende Abhängigkeit:

$$f(b) = \frac{1}{2} (f(a) + f(c)) \quad (5.7)$$

Sind a und c freie Knoten, so nennen wir b einfach hängenden Knoten. (5.7) ist somit hinreichende Bedingung für Stetigkeit im Knoten b . Ist dagegen mindestens einer der beiden Knoten a, c hängender Knoten, so nennen wir b einen mehrfach hängenden Knoten. Neben (5.7) ergeben sich durch die Stetigkeit von f weitere Abhängigkeiten von $f(a), f(c)$ zu Funktionswerten an den entsprechend benachbarten Knoten.

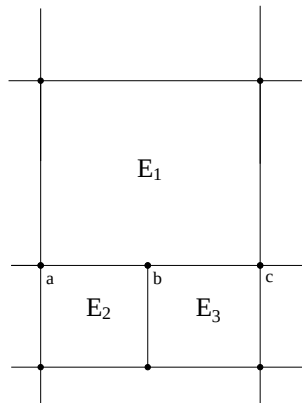


Abbildung 5.5: Beispiel für einen hängenden Knoten

Um solche mehrfachen Abhängigkeiten zu verhindern, fordern wir, daß in der adaptiven Triangulierung nur einfach hängende Knoten existieren. Abb. 5.6 zeigt ein adaptives Gitter, das keine mehrfach hängenden Knoten besitzt.

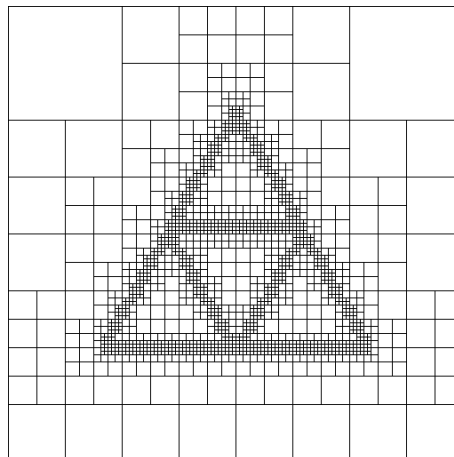


Abbildung 5.6: Adaptives Gitter mit einfach hängenden Knoten

Sei N_F die Menge der Indizes der freien Knoten von $\mathcal{N}(\mathcal{T}_S)$. Geben wir an den freien Knoten $\{x_i\}_{i \in N_F}$ Funktionswerte F_i vor, so existiert eine eindeutige Funktion $f \in V_h(\Omega, \mathcal{T}_S)$ mit $f(x_i) = F_i, i \in N_F$, wobei sich die Funktionswerte an den hängenden Knoten gemäß (5.7) aus den vorgegebenen Werten ergeben. Das Gleichungssystem muß an diese Situation angepasst werden, indem alle Unbekannten, also die Funktionswerte der gesuchten Lösung, zu hängenden Knoten gemäß der Zusatzbedingung (5.7) eliminiert werden.

Die Forderung, daß nur einfach hängende Knoten existieren, werden wir durch eine Eigenschaft des Stoppkriteriums erfüllen.

Dazu gehen wir zunächst zu einem nodalen Stoppkriterium über:

Sei im folgenden $\mathcal{N}(E)$ die Menge der Eckknoten von E . Zu einem Knoten $x \in \mathcal{N}(\mathcal{T}_S)$ sei eine Aussage $s(x)$ gegeben. Wir definieren die *Verfeinerungsknoten* von E durch

$$\mathcal{N}_C(E) := \{\cup_{\tilde{E} \in \mathcal{T}_h, \tilde{E} \subset E} \mathcal{N}(\tilde{E})\} \setminus \mathcal{N}(E)$$

$\mathcal{N}_C(E)$ ist also die Menge aller Knoten der Elemente, die bei der rekursiven Verfeinerung von E bis zum maximalen Level entstehen, mit Ausnahme der Ecken von E .

Wir setzen

$$S(E) := \bigwedge_{x_i \in \mathcal{N}_C(E)} s(x_i) \quad (5.8)$$

Gilt nun die **Saturierungsbedingung**

$$\begin{aligned} s(x_i) \text{ ist wahr für alle } x_i \in \mathcal{N}(E), \text{ dann} \\ \text{ist es wahr für alle } x_i \in \mathcal{N}_C(E) \end{aligned}$$

so treten in der durch das Stoppkriterium S gegebenen Triangulierung nur einfach hängende Knoten auf (vgl.[19]). Die Saturierungsbedingung erleichtert zudem die Verifizierung des Stoppkriteriums, da in (5.8) nur die Verfeinerungsknoten des nächsten Levels, also die Knoten auf den Kantenmitten und der Elementmitte berücksichtigt werden müssen.

Wahl des Stoppkriteriums

Wir wählen nun das nodale Stoppkriterium $s(x)$ für alle Knoten x des grundlegenden Gitters, $x \in \mathcal{N}(\mathcal{T}_h)$, in Abhängigkeit von einer beliebigen Funktion $v \in V_h(\Omega, \mathcal{T}_h)$. Im Algorithmus wählen wir dann zur Bestimmung des adaptiven Gitters für den n -ten Zeitschritt als v die Lösung zum zuletzt berechneten Zeitschritt: $v = u_h^{n-1}$.

Der Approximationsfehler, der beim Vergrößern einer Triangulierung auftritt, indem v auf den entsprechenden neuen Raum $V_h(\Omega, \mathcal{T}_S)$ projiziert wird, ist abhängig vom Gradienten von v . Eine Vergrößerung ist nur auf Elementen, auf denen v annähernd konstant ist, sinnvoll. Wir verwenden daher folgendes Stoppkriterium:

$$s(x) := (|\nabla v(x)| < c_{adapt}) \quad (5.9)$$

mit festem Schwellenwert $c_{adapt} > 0$, wobei wir die Komponenten des Gradienten durch zentrale Differenzen auf dem grundlegenden Gitter approximieren.

Numerische Integration im adaptiven Fall

Nachdem die Lösung $u_h^n \in V(\Omega, \mathcal{T}_S)$ zum Zeitpunkt t_n berechnet wurde, interpolieren wir diese auf dem grundlegenden Gitter. (Die interpolierte Funktion wird eventuell bereits beim Speichern der Daten benötigt). Wir können nun wie im nicht-adaptiven Fall vorgehen und die Integration durch eine Summation über alle Knoten in den B_ε -Umgebungen approximieren.

Dies stellt keine Einschränkung an die Adaptivität des Verfahrens dar, da die der Zahl der Knoten $x \in B_h(x_0) := \{x \in \mathcal{N}(\mathcal{T}_h) \mid |x - x_0|_2 \leq \varepsilon\}$ klein im Vergleich zur Zahl der Unbekannten ist und wir die Integrationen nur auf den Elementenmitten der adaptiven Triangulierung durchführen. Zudem kann die Menge B_h numerisch vergleichsweise leicht bestimmt werden (z.B. indem man die Offsetkoordinaten der $x \in B_h(x_0)$ zu x_0 speichert).

5.3 3D-Rekonstruktion

Die Funktionen $\tilde{u}_k, k = 1, 2$ verwenden wir als Anfangswerte der Differentialgleichung (5.1) und bestimmen nach dem oben beschriebenen Verfahren zwei Lösungen $u_k := u_k(T) \in V_h(\Omega_k)$ zum gemeinsamen Zeitpunkt $T > 0$.

Seien U_{ij}^k die entsprechenden Knotenwerte von u_k auf $x_{ij}^k \in G_k$. Aus diesen Werten rekonstruieren wir nun eine Funktion $S \in V_h(\Omega^{3D})$. Definiere folgende Gitterwerte

$$S_{ijk} := \min(u_1(x_{ij}^1), u_2(x_{kj}^2)) = \min(U_{ij}^1, U_{kj}^2)$$

S ist nun durch $S(x_{ijk}) = S_{ijk}$ eindeutig bestimmt. Den so definierten Operator $V_h(\Omega_1) \times V_h(\Omega_2) \rightarrow V_h(\Omega^{3D})$ bezeichnen wir mit R_h .

Da der Rekombinationsprozeß numerisch nicht aufwendig ist, empfiehlt es sich, die Werte S_{ijk} bei jedem Zugriff neu aus den $U_{ij}^k, k = 1, 2$ zu berechnen.

5.4 Tracing schlauchartiger Strukturen

5.4.1 Integration in Gittermetrik

Gegeben sei eine Funktion $S \in V_h(\Omega^{3D})$ auf dem dreidimensionalen Gitter G . Zu S berechnen wir Schwerpunkt, Massenmoment und dessen Eigenvektoren und -werte in der Tensormetrik. Sie werden zunächst in den Mittelpunkten der Gitterzellen berechnet und zellenweise konstant fortgesetzt. Sei E eine Gitterzelle mit Mittelpunkt z und $A \in \mathbb{R}^{3 \times 3}$ wie in 3.5.2 definiert.

Sei $B_h(z) := \{x_{ijk} \in G \mid d(x_{ijk}, z) < \varepsilon\}$, $n \in \mathbb{N}$ beliebig.

$$\begin{aligned} m_h(z) &:= \sum_{x \in B_h(z)} A^{-1} S^n(x) \, dx \\ sp_h(z) &:= \frac{1}{m_h(z) + \delta} \sum_{x \in B_h(z)} A^{-1} S^n(x) A^{-1}(x - z) \, dx \\ M_h(z) &:= \frac{1}{m_h(z) + \delta} \sum_{x \in B_h(z)} A^{-1} S^n(x) \\ &\quad (A^{-1}(x - z) - sp_h(z)) \otimes (A^{-1}(x - z) - sp_h(z)) \, dx \end{aligned}$$

Die Eigenwerte und -vektoren der Matrix $M_h(z)$ bestimmen wir numerisch durch eine einfache Vektoriteration [20]. Das Verfahren könnte an dieser Stelle noch optimiert werden.

Ein Vergleich der errechneten Eigenwerte mit dem aus dem charakteristischen Polynom exakt bestimmten hat gezeigt, daß diese Iteration gute Resultate liefert.

5.4.2 Lösung der gewöhnlichen Differentialgleichung

Die Differentialgleichung (3.12) wird zu einem Differentialgleichungssystem erster Ordnung der Form

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= F(x, v) \end{aligned}$$

umformuliert.

Zur Lösung verwenden wir ein einfaches explizites Euler-Verfahren.

Die Lösung erfolgt nun ebenfalls in den Gitterkoordinaten mit der zugehörigen Metrik. Die Punkte des gefundenen Streckenzugs werden anschließend in die ursprünglichen Koordinaten transformiert.

Zunächst wählen wir einen Startpunkt, indem wir in beiden zweidimensionalen Bildern ein Paar korrespondierender Punkte festlegen, wodurch ein eindeutiger Punkt in $X_0 = X_0^* \in \Omega^{3D}$ bestimmt ist, sowie einen Startvektor $V_0 \in \mathbb{R}^3$ für die Geschwindigkeit.

Wir bestimmen nun rekursiv die einzelnen Teillösungen des in 3.5.3 vorgestellten Graphen. Seien dazu ein Startpunkt X_0 und eine Startrichtung $V_0 \in \mathbb{R}^3$ gegeben. Seien $N \in \mathbb{N}, T > 0, \tau := \frac{T}{N}$. Wir approximieren die zugehörige Teillösung $(X_i, V_i) \in \mathbb{R}^6$ zum Zeitpunkt $t = i\tau, i = 1 \dots N$ durch folgende Iterationsvorschrift:

$$\begin{aligned} X_i &:= X_{i-1} + \tau V_{i-1} \\ V_i &:= V_{i-1} + \tau F(X_{i-1}, V_{i-1}) \end{aligned}$$

Bei der Wahl der Zeitschrittweite ist zu beachten, daß zur Berechnung der Beschleunigung im Punkt X_i die Funktion S nur auf der (durch Gitterzellen approximierten) Umgebung $B_\varepsilon(X_i)$ ausgewertet wird, also nur in $B_\varepsilon(X_i)$ liegende Strukturen erkannt werden. Daher ist es sinnvoll, daß $\bigcup_i B_\varepsilon(X_i)$ die berechnete Lösungskurve überdeckt, also der Abstand zweier Punkte X_i, X_{i-1} kleiner ε ist,

z.B. $|X_i, X_{i-1}| \approx \frac{\varepsilon}{2}$. Die Zeitschrittweite τ sollte daher so gewählt werden, daß $\tau V_i \leq \frac{\varepsilon}{2}$ für alle $i = 0 \dots N$ gilt. (Eine Alternative besteht darin, eine variable Zeitschrittweite τ_i mit $\tau_i V_i \leq \frac{\varepsilon}{2}, i = 0 \dots N$ einzuführen.)

Wir werden im folgenden entscheiden müssen, ob die Lösung $\{X_i\}_{i=0 \dots N}$ einen vorgegebenen Punkt Z , zum Beispiel den ersten Startpunkt, erreicht, d.h. ob ein $i > 0$ und ein $s \in [0, 1]$ existieren, so daß $\|X_{i-1} + s(X_i - X_{i-1}) - Z\|_2 < \tilde{\varepsilon}$ für ein vorgegebenes $\tilde{\varepsilon} > 0$ gilt. Hier ist es sinnvoll, einen Wert in der Größenordnung der Gitterweite zu wählen.

Zur Vereinfachung der Notation nehmen wir an, daß die Abstände je zweier Punkte X_{i-1}, X_i deutlich klein als $\tilde{\varepsilon}$ sind, so daß wir lediglich zu prüfen brauchen, ob $\|X_i - Z\|_2 < \tilde{\varepsilon}$ gilt.

Zur Berechnung von $F(X_{i-1}, V_{i-1})$ werden der Schwerpunkt und die Eigenvektoren und -werte von M_h im Punkt X_{i-1} benötigt. Da wir beim Lösen der Differentialgleichung nur in den im Vergleich zur Anzahl der Gitterpunkte wenigen Punkten $\{X_i\}_{i=0 \dots N}$ arbeiten und diese Punkte in der Regel paarweise disjunkt sind, berechnen wir coh , sp und v_k erst zu dem Zeitpunkt, zu dem wir sie benötigen. Da diese Werte eingeschränkt auf eine Gitterzelle konstant sind, ist es sinnvoll, lediglich die zuletzt berechneten Werte $coh(X_{i-1}), sp(X_{i-1}), v_k(X_{i-1})$ zu speichern und darauf im folgenden Schritt zurückzugreifen, falls X_{i-1} und X_i zu derselben Gitterzelle gehören.

Nach Berechnung eines neuen Lösungspunktes $X_i, i > 0$ schätzen wir ab, ob die durch die Funktion S beschriebene Struktur eine Verzweigung in diesem Punkt besitzt. Dazu verwenden wir wie in 3.2 beschrieben die Kohärenz.

Erreicht die Lösung einen solchen Verzweigungspunkt $X_{i_0}, i_0 > 0$, so wird das Lösungsverfahren unterbrochen und die Punkte $\{X_i\}_{i=0 \dots i_0}$ des gefundenen Polygonzugs werden gespeichert. Wir bestimmen n Verzweigungsrichtungen $V_0^1 \dots V_0^n$ und setzen das Lösungsverfahren n mal rekursiv mit neuem Startpunkt $X_0 = X_{i_0}$ und mit Startvektoren $V_0^i, i = 1 \dots n$ fort. Als neue Richtungen verwenden wir:

$$\begin{aligned} V_0^1 &:= v_0 v_1 \\ V_0^2 &:= v_0 v_2 \\ V_0^3 &:= -v_0 v_2 \end{aligned}$$

(Hierbei ist v_0 ein vorgegebener skalarer Geschwindigkeitswert; v_1, v_2 sind die Eigenvektoren von M_h .) Wir geben für die Baumstruktur, die wir durch diese Verzweigen erhalten, eine maximale Tiefe vor.

Das Lösungsverfahren zu einem Startpunkt X_0 wird abgebrochen, falls die Teillösung den ersten Startpunkt wieder erreicht oder sich abzeichnet, daß die Lösung

nicht die gesuchte Struktur approximiert. Die genauen Kriterien dazu führen wir im nächsten Abschnitt auf.

Im ersten Fall fügen wir die einzelnen Polygonzüge des betreffenden Astes zu einem geschlossenen Polygonzug zusammen. Seien $\{X_0 \dots X_l = X_0\}, l > 0$ die zugehörigen Punkte. Sei

$$\mathcal{X} := \{x \in \Omega^{3D} | \exists i \in \{0 \dots l-1\}, t \in [0, 1] \ x = X_i + t(X_{i+1} - X_i)\}$$

Eine Approximation des gesuchten Schlauches erhalten wir nun, in dem wir einen ε -Schlauch $\tilde{\Gamma}$ um $\mathcal{X}$ mit $\varepsilon > 0$ betrachten:

$$\tilde{\Gamma} := \{x \in \Omega^{3D} | \text{dist}(x, \mathcal{X}) < \varepsilon\}$$

Eine Alternative besteht, wie in 3.5.3 erwähnt, darin, Teillösungen zu suchen, deren Endpunkte an andere Teillösungen anliegen.

Durch Zusammenfassen aller dieser Lösungen erhalten wir ein Gebilde, daß aus der Vereinigung geschlossener Polygonzüge besteht.

5.4.3 Stoppkriterien

Sei zum Startpunkt X_0 die Lösung bis zu einem Zeitpunkt $i_0 \in \{1 \dots N\}$ bereits berechnet. Es ist sinnvoll, das Lösungsverfahren abbrechen zu lassen, falls die Lösung zum ersten Startpunkt X_0^* zurückkehrt oder sich abzeichnet, daß $\{X_i\}_{i \in \{0 \dots i_0\}}$ nicht den gesuchten Schlauch approximiert.

Wir verwenden dazu folgende Kriterien:

Das Lösungsverfahren wird abgebrochen, falls

- N Zeitschritte berechnet wurden: $i_0 = N$.
- die Lösung das Gebiet Ω^{3D} verläßt: $X_{i_0} \notin \Omega^{3D}$
- die Lösung die Struktur verläßt.
Als Kriterium verwenden wir $|sp(X_i)| > c_{dist} > 0$.
- die Lösung zum eigenen Startpunkt zurückführt:
 $|X_{i_0} - X_0|_2 < \tilde{\varepsilon}$ für $i_0 \gg 0$.
- die Lösung zum ersten Startpunkt zurückführt:
 $|X_{i_0} - X_0^*|_2 < \tilde{\varepsilon}$ für $X_0 \neq X_0^*$.
- die Lösung einer bereits berechneten Lösung $\{\tilde{X}_i\}_{i \in I}$ (mit $I \subset \mathbb{N}$) folgt,
d.h. $\text{dist}(X_i, \tilde{X}_j) < \tilde{\varepsilon}$ für genügende viele $(i, j) \in [0, i_0] \times I$

5.4.4 Überblick über den Algorithmus

Wir beschreiben in Grundzügen den Tracing-Algorithmus:

Zu einem Knoten K des Baumes speichern wir folgende Werte:

$K.X_0$	Startpunkt der Teillösung zum Knoten K
$K.V_0$	Startrichtung, Startgeschwindigkeit
$K.Liste$	Liste der berechneten 3D-Punkte der Lösung
$K.s$	Resultat des Lösungsprozesse (s.u.)

Sei GL sei eine (zunächst leere) Liste von dreidimensionalen Punkten, wir werden in dieser Liste alle berechneten Punkte der Teillösungen zusammenfassen.

Wir wählen nun einen Startpunkt X_0 und eine oder mehrere Startrichtungen V_i $i = 0 \dots n-1$. Im Fall $n = 1$ legen wir einen Knoten K als Wurzel des Baumes an mit

$$K.X_0 = X_0, \quad K.V_0 = V_0, \quad K.Liste = \langle \text{leer} \rangle, \quad K.s = 0,$$

andernfalls setzen wir

$$K.X_0 = X_0, \quad K.V_0 = V_0, \quad K.Liste = \langle \text{leer} \rangle, \quad K.s = 1$$

und fügen an K n Kinderknoten K_i , $i = 0 \dots n-1$ an mit:

$$K_i.X_0 = X_0, \quad K_i.V_0 = V_i, \quad K_i.Liste = \langle \text{leer} \rangle, \quad K_i.s = 0$$

Die Teillösung zu einem Knoten K berechnen wir wie folgt:

Funktion Solve(K):

$x_0 = K.X_0, v_0 = K.V_0$

für $n = 1 \dots N_{max}$:

berechne $(x_n, v_n) = (x_{n-1}, v_{n-1}) + \tau(v_{n-1}, F(x_{n-1}, v_{n-1}))$

 zu festem $\tau > 0$

speichere x_n in $K.Liste$

gebe 0 zurück, falls:

- $x_n \notin \Omega$
- $|sp(x_n)| > c_{dist} > 0$ ¹
- $\#\{x \in K.Liste \mid \text{dist}(x, GL) < \tilde{\varepsilon}\} > N_1$ ²
- $n > n_{min} \wedge |x_n - K.X_0|_2 < \tilde{\varepsilon}$ ³

falls $|x_n - X_0|_2 < \tilde{\varepsilon}$ (dabei $X_0 \neq K.X_0!$)⁴ **gebe** 2 zurück

falls $n > n_{min} \wedge \text{coh}(x_n) < c_{bifurc}$

wähle neue Startrichtungen V_i , $i = 0 \dots n-1$

 falls max. Baumtiefe noch nicht erreicht:

füge n neue Kinderknoten K_i von K in den Baum ein mit:

$$K_i.X_0 = x_n, \quad K_i.V_0 = V_i, \quad K_i.Liste = \langle \text{leer} \rangle, \quad K_i.s = 0$$

gebe 1 zurück

gebe 1 zurück

Alle im folgenden vorgestellten traverse-Routinen sollen ausgehend von der Wurzel den Baum rekursiv durchlaufen. Der Aufbau des Baumes erfolgt mit Hilfe folgender Routine

Funktion Traverse1(K):

```
falls  $K.s = 0$  :
     $K.s = \text{solve}(K)$ 
    falls  $K.s = 0$  :
        lösche Knoten  $K$ 
    sonst
        füge Kopie von  $K$ .Liste an  $GL$  an
falls  $K.s = 1$  und max. Baumtiefe noch nicht erreicht:
    traverse1( $K_i$ ) für alle Kinder  $K_i$  von  $K$ 
```

Falls genau eine Teillösung zum ersten Startpunkt zurückkehrt, findet man den entsprechenden geschlossenen Polygonzug durch folgende Routine:

Funktion traverse2(K):

```
falls  $K$  Blatt des Baumes ist:
    falls  $K.s = 2$  :
        gebe 1 zurück
    sonst gebe 0 zurück
sonst:
    für alle Kinderknoten  $K_i$ :
         $v_i := \text{traverse2}(K_i)$ 
        falls  $v_i = 0$  entferne  $K_i$ 
    gebe  $\max_i v_i$  zurück
```

Diese Routine läßt sich ebenfalls verwenden, falls die Teillösungen zu mehreren Blättern des Baumes zum ersten Startpunkt zurückkehren.

Sollen alle Teillösungen isoliert werden, die Teil eines geschlossenen Polygonzugs sind (unabhängig davon, ob der erste Startpunkt Teil dieses Polygonzugs ist) , so lassen sich die übrigen Teillösungen mit Hilfe folgender Routine entfernen:

¹Teillösung außerhalb der Struktur

²Teillösung folgt bereits berechneter Lösung

³Teillösung kehrt zu eigenem Startpunkt zurück

⁴Teillösung erreicht ersten Startpunkt

Funktion $\text{traverse3}(K)$:

$\text{traverse3}(K_i)$ für alle Kinderknoten von K

falls K Blatt:

 sei x letzter Punkt der Liste $K.\text{Liste}$.

 bestimme $d = \min_{\tilde{K} \in \mathcal{K}, \tilde{K} \neq K} \text{dist}(x, \tilde{K}.\text{Liste})$

 (wobei $\mathcal{K}$ die Menge aller Knoten des Baumes ist)

 falls $d > \tilde{\varepsilon}$ entferne K

Mit anderen Worten, wir entfernen rekursiv alle Blätter, deren Teillösungen nicht an andere Teillösungen angrenzen.

Kapitel 6

Anwendung

6.1 Diffusion

6.1.1 Vergleich verschiedener Diffusionsmodelle

In diesem Abschnitt wollen wir anhand von Beispielen die Vorteile des neuen Verfahrens zur Erkennung und Hervorhebung schlauchartiger bzw. linearer Strukturen aufzeigen. Zunächst werden wir die verschiedenen Diffusionsmodelle miteinander vergleichen.

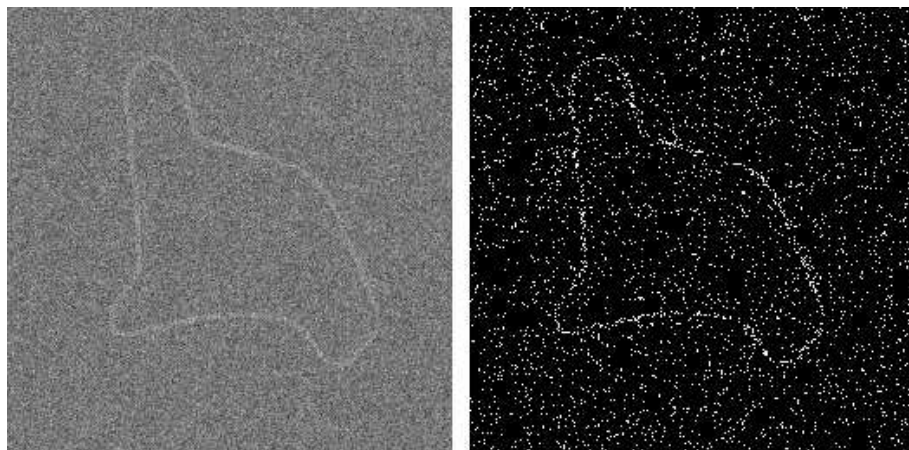


Abbildung 6.1: Originalbild und Grauwertsegmentierung

Dazu betrachten wir als Beispiel ein Bild mit schlauchförmiger Struktur, - die wir im folgenden auch als Ring bezeichnen -, und künstlichem Rauschen (Abb.6.1), und wenden darauf lineare Diffusion, Perona-Malik-Diffusion und anisotrope Diffusion nach dem Modell von Weickert so wie anisotrope Diffusion nach dem auf dem Massenmoment basierenden Modell an. Als rechte Seite wählen wir in allen

vier Fällen $f = 0$. Anschließend führen wir eine Grauwertsegmentierung mit geeignetem Schwellenwert $c \in [0, 1]$ durch.

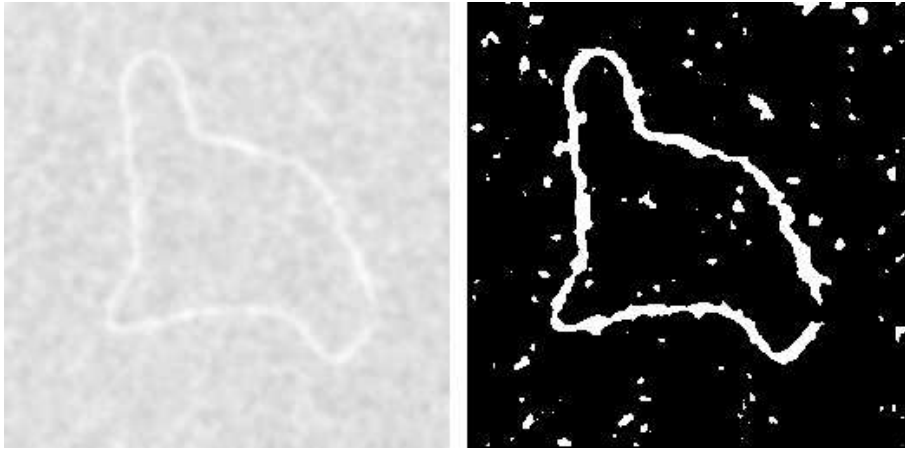


Abbildung 6.2: Resultat der linearen Diffusion

Abb.6.2 zeigt das Ergebnis der linearen Diffusion. Wie erwartet sorgt die Diffusion für eine gleichmäßige Glättung des Bildes. Dadurch wird zum einen das Rauschen reduziert, zum anderen verschwimmt der Bereich zwischen Struktur und Hintergrund, die Struktur wird unscharf. Bei einer Grauwertsegmentierung ($c = 0,91$) wird ein deutlich breiteres Objekt erkannt als im Original vorhanden (Abb.6.2). Das erkannte Objekt enthält etliche Lücken.

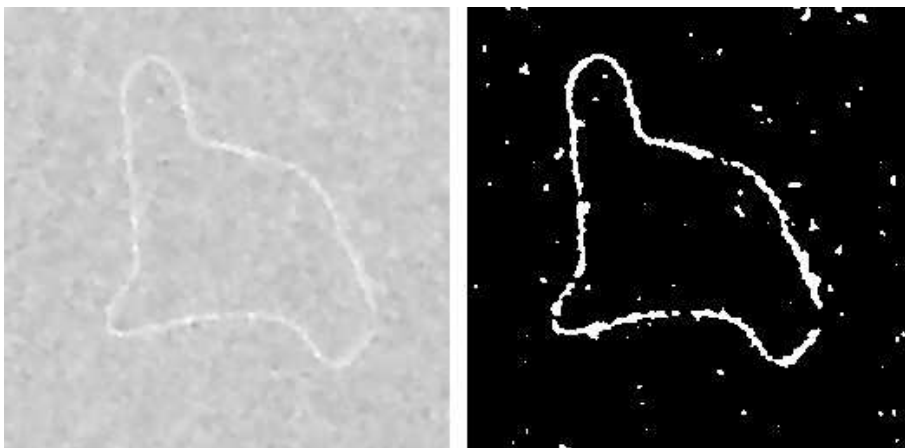


Abbildung 6.3: Resultat der Perona-Malik-Diffusion

Im Vergleich dazu ist in Abb. 6.3 deutlich die punktweise Steuerung der Diffusion zu erkennen. An einigen Stellen tritt eine Kontrastschärfung ein. Insgesamt ist

die Struktur besser zu erkennen als bei der linearen Diffusion. Die Grauwertsegmentierung ($c = 0,85$) liefert ein Objekt, daß ebenfalls etliche Lücken enthält. Man erkennt, daß Teile des Hintergrundes durch die Schärfung ebenfalls hervorgehoben werden.

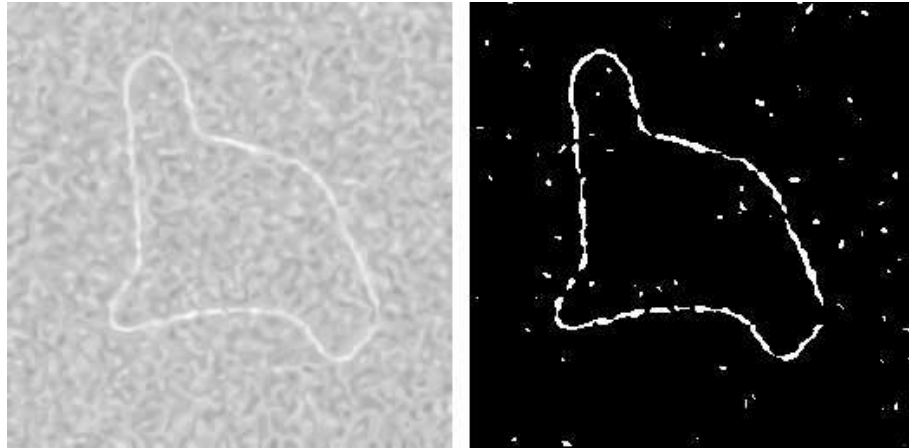


Abbildung 6.4: Resultat der Anisotropen Diffusion nach Weickert

Beim Ergebnis der anisotropen Diffusion (Abb. 6.4) ist deutlich die Kantenparallele Glättung zu erkennen. Die Kanten des Schlauchs werden dadurch hervorgehoben, und die Abgrenzung zum Hintergrund wird verstärkt. Im Hintergrundrauschen bilden sich größere Strukturen heraus. Diese werden jedoch mit fortlaufender Diffusion weggeglättet. Die Grauwertsegmentierung ($c = 0,86$) liefert ein im Vergleich zu den vorherigen Resultaten glatter berandetes Objekt, daß infolge der Kantenparallelen Glättung weniger Lücken aufweist.

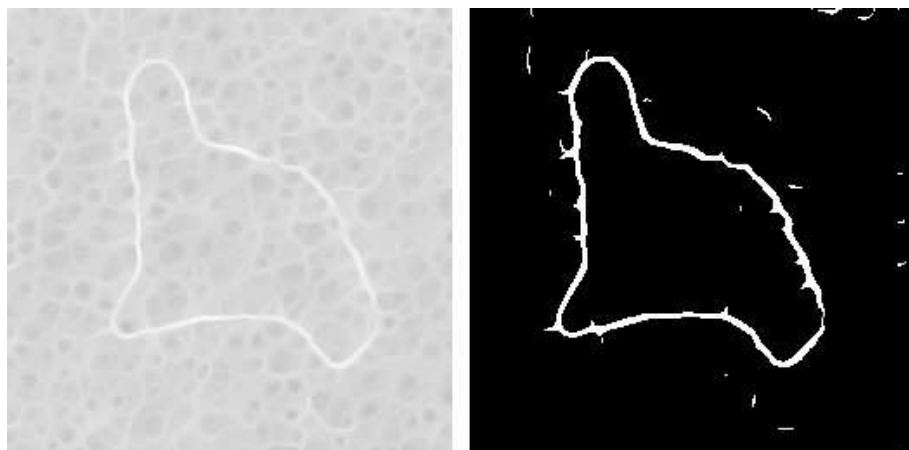


Abbildung 6.5: Resultat der Anisotropen Diffusion nach dem neuen Modell

Das Ergebnis des Diffusionsprozesses nach dem neuen Modell (Abb.6.5) zeigt noch deutlicher eine Diffusion in tangentielle Richtung an eindimensionalen Strukturen, so daß die Struktur gleichmäßiger geglättet wird als beim Modell nach Weickert. Auch hier werden Strukturen im Hintergrund erkannt und hervorgehoben. Entlang des Rings bilden sich kleine Ausläufer, die in die Hintergrundstrukturen übergehen. Diese Ausläufer lassen sich auch nach der Grauwertsegmentierung ($c = 0,87$) erkennen. Das segmentierte Objekt besitzt einen glatteren Rand als das der Segmentierung des Bildes, das nach dem von Weickert [23] vorgestellten Modell geglättet wurde, und weist zudem weniger Lücken auf.

Ziel der Diffusion ist es, eine *lückenlose* Erkennung des geschlossenen Rings durch die Grauwertsegmentierung zu ermöglichen. Würden Lücken bestehen bleiben, so müßten sie in einem Nachbearbeitungsschritt nach der Rekonstruktion geschlossen werden, da sonst das Tracingverfahren an diesen Lücken stoppt. Um eine solche Nachbearbeitung zu vermeiden, sollen durch die Diffusion alle Lücken geschlossen werden. Die lineare Diffusion sowie das Perona-Malik-Modell sind dazu ungeeignet. Im folgenden werden wir uns daher auf einen Vergleich der Modelle mit anisotroper Diffusion konzentrieren. Wir werden sehen, daß das neue Modell wesentlich besser in der Lage ist, Lücken zu schließen, als das Modell von Weickert.

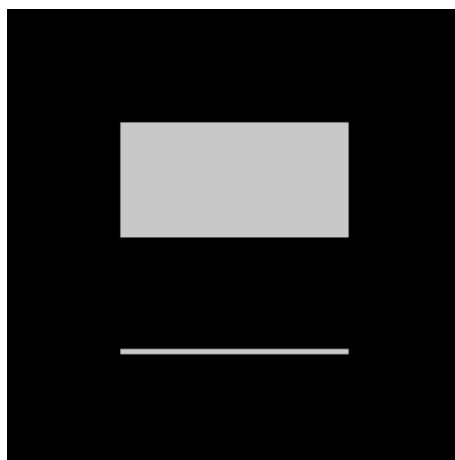


Abbildung 6.6: Testbild

Wir betrachten dazu ein Beispiel. Als Ausgangsbild wählen wir ein Bild, das eine schmale Linie L sowie ein Rechteck R vor schwarzem Hintergrund zeigt (Abb.6.6). Die Linie hat eine von null verschiedene Breite und ist daher mathematisch als schmales Rechteck zu betrachten. Dennoch werden wir weiterhin von Linie sprechen. Ihre beiden schmalen Seiten bezeichnen wir als Linienenden. Die langen Seiten von L und die Seiten von R bezeichnen wir im folgenden auch als Kanten

(vgl. 2.2.2). Wir werden untersuchen, inwiefern diese Objekte durch die beiden Modelle der anisotropen Diffusion verändert werden. Auch hier sei wieder $f = 0$ gewählt.

Abb. 6.7 zeigt die Ergebnisse beider Diffusionen. Beide Verfahren erhalten wie erwartet die Kanten des großen Rechtecks, dessen Form dadurch weitgehend unverändert bleibt. Auch die Kanten der Linie werden in beiden Fällen durch Kantenschärfung erhalten. Die beiden Verfahren unterscheiden sich deutlich in der Weise, wie die Enden der Linie durch die Diffusion verändert werden.

Die Diffusion nach dem von Weickert vorgestellten Modell verläuft an den Enden nicht parallel zur Linienrichtung, sondern auf die horizontale Symmetrieachse zu. Dadurch bilden die Enden der Linie im Laufe der Diffusion die Form von spitzen Dreiecken aus. Orthogonal an den langen Seiten dieser Dreiecke findet fortan Kantenschärfung statt. Nur an den äußeren Spitzen der Dreiecke sorgt die Diffusion in Linienrichtung für ein “Wachsen” der Linie. Das Wachstum wird also durch das Zuspitzen der Ecken und die Kantenschärfung an den Seiten beschränkt.

Um die oben beschriebene Ausprägung von Spitzen zu verstehen, betrachten wir die Eigenvektoren v_{J_ρ} des Strukturtenors J_ρ , $\rho > 0$, der die Hauptdiffusionsrichtung bestimmt, und v_{J_0} des ungeglätteten Strukturtenors J_0 zum kleineren Eigenwert. v_{J_0} liegt an Kanten orthogonal zum Gradienten, also orthogonal zu den Kanten. An den Enden von L liegt v_{J_0} also orthogonal zur Linienrichtung (Abb. 6.8). Ein Diffusionsprozeß mit $\rho = 0$ würde, da in Richtung v_0 Kantenschärfung stattfinden würde, alle vier Kanten von L erhalten. Beide Objekte L, R würden gleich behandelt. Beim Glätten des Strukturtenors J_0 mit Parameter $\rho > 0$ wird nun v_{J_0} gemittelt. Wählt man ρ genügend groß, so überwiegt an den Enden von L der Einfluß der Daten der angrenzenden langen Kanten, so daß v_{J_ρ} dort nicht mehr kantenparallel ist (vgl. Abb 6.8), sondern schräg zur horizontalen Symmetrieachse von L liegt. Durch die Diffusion in Richtung v_{J_ρ} prägen sich dann die oben beschriebenen spitz zulaufenden Kanten aus.

Im Vergleich dazu betrachten wir nun den Eigenvektor v_M des Massenmoments M zum größeren Eigenwert, der im neuen Modell die Hauptdiffusionsrichtung bestimmt. v_M liegt, wie wir oben gesehen haben, parallel zu linearen Strukturen. Vorteil ist, daß diese Strukturen nicht im betrachteten Punkt x , sondern nur in der ε -Umgebung B_ε liegen müssen, die in Berechnung von $M(x)$ eingeht. Die Daten brauchen daher nicht wie im Modell nach Weickert durch eine Glättung gemittelt werden. Abb. 6.9 skizziert den Eigenvektor v_M in einer Umgebung der Linie. Die Kantenschärfende Diffusion orthogonal zu v_M ist in Abb. 6.7 zu erkennen. Die Linie wächst durch die Diffusion in Linienrichtung. Dieses Wachsen erfolgt dadurch, daß v_M an den Linienenden parallel ist, schneller als bei der

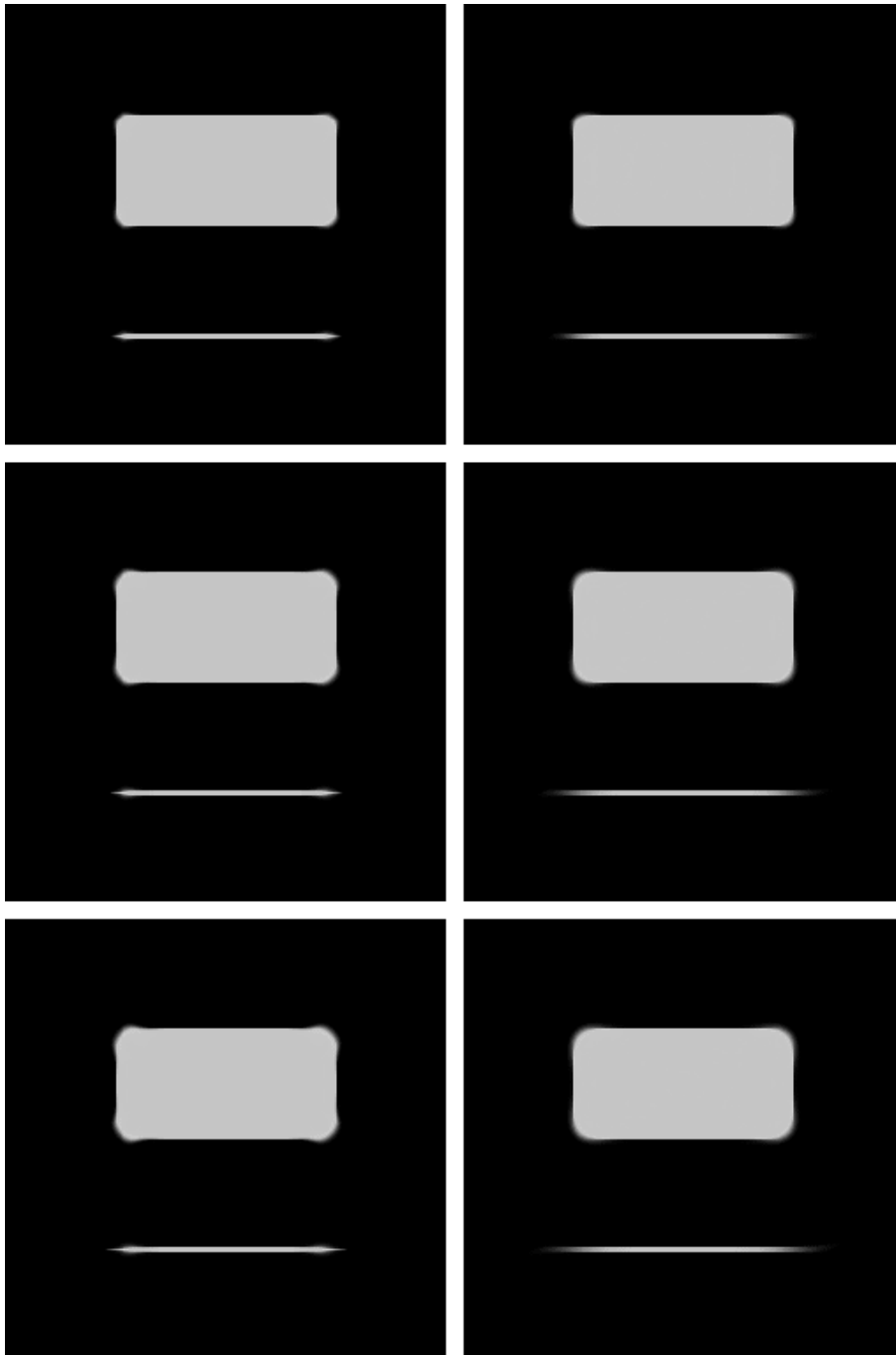


Abbildung 6.7: Links: Modell nach Weickert, rechts: neues Modell zu drei verschiedenen Zeitschritten

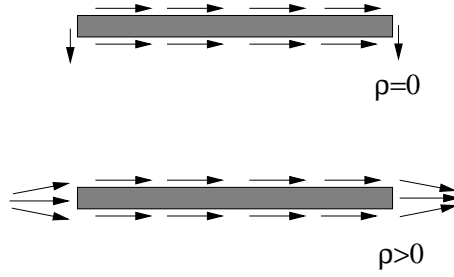


Abbildung 6.8: Eigenvektoren von J_0 (links) bzw. J_ρ (rechts) zum kleineren Eigenwert

Diffusion nach dem von Weickert vorgestellten Modell. Bei der Auswertung des Massenmoments im nächsten Zeitschritt wird dann die verlängerte Linie erkannt, sodaß sich die Diffusion in Linienrichtung weiter fortsetzt.

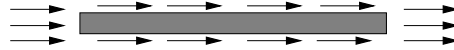


Abbildung 6.9: Eigenvektor v des Massenmoments

Im oben gezeigten Beispiel sehen wir, daß durch die anisotrope Diffusion mit der Zeit der Kontrast vermindert wird, da wir bisher $f = 0$ vorausgesetzt haben. Um diesen Verlust des Kontrastes zu verhindern, verwenden wir in unserem Modell eine rechte Seite $f \neq 0$.

Eine solche rechte Seite ließe sich natürlich auch bei Weickert [23] einsetzen, um auch hier den Kontrast zu erhöhen.

6.1.2 Anwendung auf DNA-Bilder

Wir betrachten nun das in Abb. 6.10 gezeigte Stereopaar von DNA-Bildern. Wir glätten beide Bilder mit derselben Zeitschrittweite bis zu einem Zeitpunkt $T > 0$, zu dem alle Lücken beider DNA-Ringe geschlossen sind.

Abb. 6.11 zeigt die Ergebnisse der ersten vier Zeitschritte zum rechten Bild sowie die Entwicklung der rechten Seite $f(u)$.

Anhand der rechten Seite $f_\varepsilon(t, x)$ läßt sich nachvollziehen, wie die Erkennung der Struktur voranschreitet. An Bildpunkten, in denen eindimensionale Strukturen erkannt werden, findet eine Kontrastschärfung statt, so daß diese Strukturen hervorgehoben werden. Mit der Zeit wachsen diese nun entlang der Hauptdiffusionsrichtung. An diesen Strukturen findet im nächsten Zeitschritt u.U. eine erneute Kontrastschärfung statt, es werden weitere Punkte am Rand der Struk-

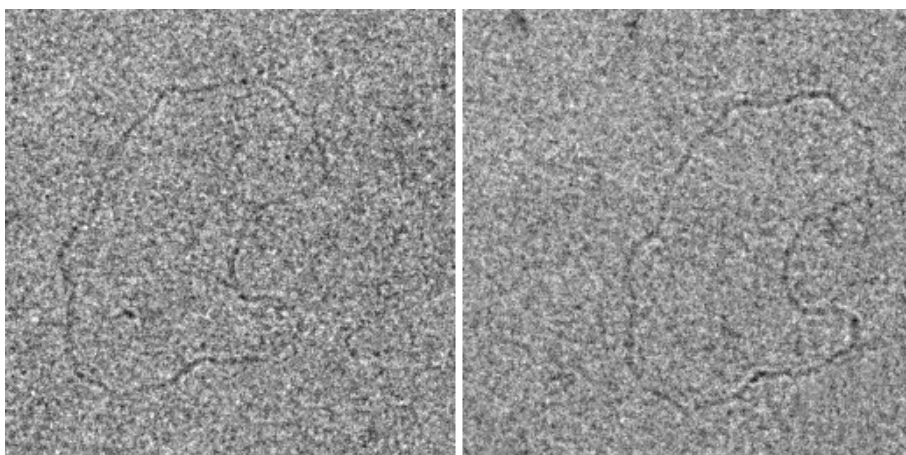


Abbildung 6.10: EM-Stereobilder eines DNA-Rings

turen erkannt und hervorgehoben.

Ausgangspunkt für das Wachsen dieser Strukturen sind die Punkte, die zum Zeitpunkt $t = 0$ erkannt wurden. Die rechte Seite zum Zeitpunkt $t = 0$ ist daher von besonderer Bedeutung für das Ergebnis der Diffusion. Im vorliegenden Bild erfolgt die Kontrastschärfung zum Zeitpunkt $t = 0$ sowohl an Punkten, die zur gesuchten Struktur gehören, als auch an Punkten im Hintergrund. Daher werden auch Strukturen im Hintergrund hervorgehoben. Wir nennen diese Strukturen im folgenden auch *zusätzliche* Strukturen.

Diese Hintergrundstrukturen sind dann problematisch, wenn sie mit der DNA-Struktur verschmelzen und sich auf diese Art Verzweigungspunkte bilden. Bei dem anschließenden dreidimensionalen Tracing des Rings werden wir an solchen Verzweigungspunkten auf Probleme stoßen.

An Stellen, an denen im ersten Schritt bereits genügend Punkte des DNA-Rings hervorgehoben wurden, werden bereits nach wenigen Zeitschritten große Teile der Struktur erkannt. Es gibt jedoch relativ große Teile des DNA-Rings an denen keine Kontrastschärfung stattfindet, so daß hier große Lücken entstehen, die sich erst nach einigen Zeitschritten schließen. So ist nach dem vierten Zeitschritt die DNA-Struktur zum Großteil hervorgehoben, lediglich zwei Lücken schließen sich nach dem sechsten Zeitschritt.

Die rechte Seite zum Zeitpunkt $t = 0$ sollte also so gewählt werden, daß die Kontrastschärfung nach Möglichkeit nur an Punkten erfolgt, die zur gesuchten Struktur gehören, und daß diese Punkte möglichst dicht in der Struktur liegen. Wir verwenden daher im ersten Zeitschritt statt der rechten Seite $f_1(u)$ die in Abschnitt 3.3.3 vorgestellt rechte Seite $f_2(u)$.

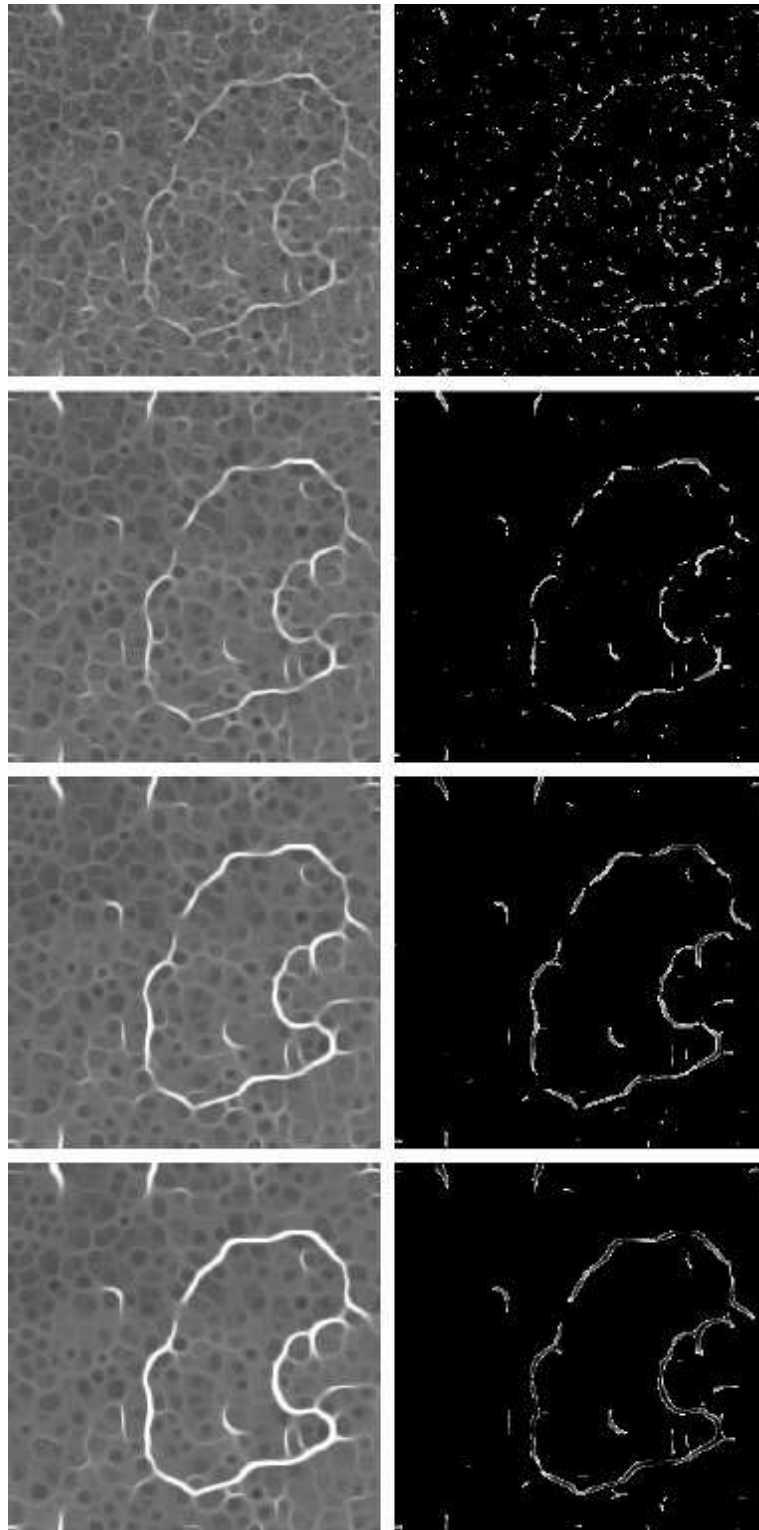


Abbildung 6.11: Ergebnis und rechte Seite beim Glätten des rechten Bildes in den ersten vier Schritten

Ein weiterer Nachteil der Kontrastschärfung neben der Hervorhebung von Hintergrundstrukturen ist, daß die Dicke der Struktur mit der Zeit zunimmt. Grund dafür ist, daß die Kontrastschärfung auch an den Längsseiten der Struktur auftritt (vgl Abb. 6.11 , rechte Spalte).

Wir führen im Anschluß an die Diffusion eine Grauwertsegmentierung durch, indem wir einen Schwellenwert bestimmen und alle Bildpunkte mit einem höheren Grauwert als Teil der (approximierten) Struktur interpretieren (Abb. 6.12).

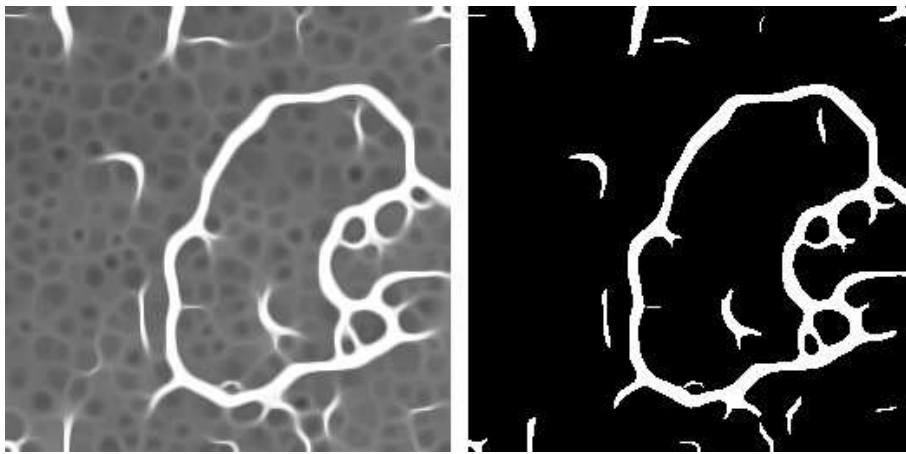


Abbildung 6.12: links: Ergebnis des siebten Glättungsschrittes, rechts: zugehörige Grauwertsegmentierung

Betrachten wir nun die Ergebnisse bei der Bearbeitung des zweiten Bildes. Wir lassen das Verfahren mit denselben Parametern ablaufen. Lediglich die Parameter zur ersten Erkennung der Struktur, d.h. die Parameter der rechten Seite zum Zeitpunkt $t = 0$, passen wir optimal an das neue Bild an.

Abb.6.13 zeigt die Ergebnisse nach den ersten vier und nach dem sechsten und achten Zeitschritt. Man erkennt deutlich, daß die Qualität des zweiten Bildes schlechter als die des ersten ist. Es bleiben in den ersten Zeitschritten mehr Lücken der Struktur offen, gleichzeitig entstehen mehr zusätzliche Strukturen. Es dauert dementsprechend länger, bis alle Lücken geschlossen sind. Dies ist erst nach dem achten Zeitschritt der Fall.

Rechts neben dem DNA-Ring wird eine Struktur hervorgehoben, die aus zwei Schleifen verbunden mit einer langen horizontalen Struktur besteht. Es handelt sich hierbei möglicherweise ebenfalls um einen DNA-Ring, der stark um die horizontale Achse verdrillt ist. Im Originalbild ist jedoch nicht eindeutig zu erkennen,

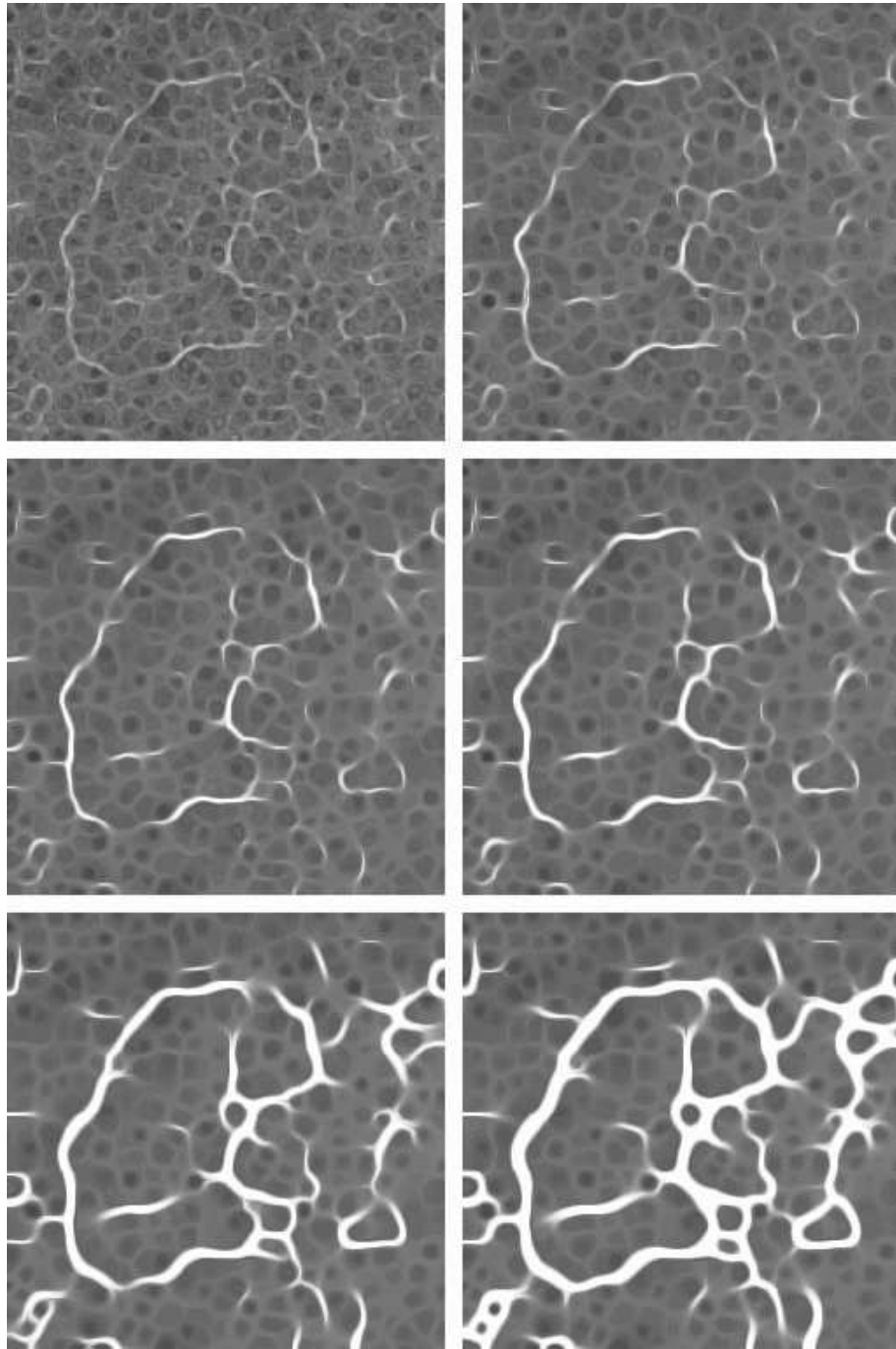


Abbildung 6.13: Ergebnis der Glättung des linken Bildes, Schritte 1-4,7,8

ob hier ein weiterer Ring vorhanden ist.

Trotz der Tatsache, daß die DNA-Ringe zu unterschiedlichen Zeitpunkten vollständig erkannt sind, verwenden wir zur Rekonstruktion des 3D-Bildes die Ergebnisse *desselben* Zeitschritts, damit sich die Strukturen in ihrer Dicke nicht zu stark unterscheiden. In diesem Fall wählen wir also die Ergebnisse nach dem achten Zeitschritt.

In den folgenden Abschnitten werden wir sehen, wie sich aus diesem Paar zunächst einen 3D-Datensatz gewinnen und anschließend der dreidimensionalen DNA-Ring von den zusätzlich entstandenen Strukturen trennen läßt.

Weitere DNA-Bilder sind im Anhang angefügt.

6.2 Rekonstruktion des 3D-Bildes

6.2.1 Rekonstruktion von Ausläufern

Wir haben gesehen, daß bei der Vorglättung der zweidimensionalen Bilder neben dem DNA-Ring zusätzliche Strukturen hervorgehoben werden. Diese Strukturen sind, da sie aus dem Hintergrundrauschen entstehen, in der Regel nur in einem der beiden Bilder enthalten, d.h. es gibt kein Pendant dazu im entsprechend anderen Bild. Wir wollen untersuchen, wie sich dies auf die Rekonstruktion auswirkt.

Rekonstruiert wird eine solche Struktur S eines Bildes nur, wenn sich im anderen Bild “korrespondierende” Strukturen befinden, d.h. im zweiten Bild liegen auf den korrespondierenden Geraden zu den Punkten von S Bildpunkte mit mindestens gleichhoher Intensität. Mit anderen Worten, zu den Punkten von S existieren mögliche korrespondierende Punkte. Als korrespondierende Struktur zu den zusätzlichen kann unter anderem der DNA-Ring des anderen Bildes in Frage kommen.

Uns interessieren nun gerade die dreidimensionalen Strukturen, die aus den zusätzlichen Strukturen mit dem DNA-Ring als korrespondierende Struktur rekonstruiert werden, da sie unmittelbar an den DNA-Ring angrenzen.

Betrachten wir dazu ein modellhaftes Beispiel. Abb. 6.14 zeigt zwei Projektionen eines Torus. Die Lage der Projektionsrichtungen sei die gleiche wie in den Beispielen 2.1. und 2.2. In der rechten Projektion wurden zwei zusätzliche Strukturen A und B zugefügt.

Zur Struktur A existieren im anderen Bild keine korrespondierenden Strukturen,



Abbildung 6.14: Projektionen eines Torus mit zusätzlichen Strukturen

sie geht also nicht in die Rekonstruktion ein. Als korrespondierende Strukturen von B kommen im zweiten die beiden Teile des Torus in Frage, die auf derselben Höhe liegen. Da diese im Vergleich zur Struktur A dünn sind, entsteht bei der Rekonstruktion von A mit einem dieser Teile eine Struktur, die in Richtung der ersten Projektion verläuft (Abb. 6.15) und in stumpfen Winkel an den Torus anliegt.



Abbildung 6.15: Rekonstruktion des Torus

Allgemein entstehen also aus den zusätzlichen Strukturen in den zweidimensionalen Bildern bei der Rekonstruktion unter anderem Strukturen, die im stumpfen Winkel am dreidimensionalen DNA-Ring anliegen. Wir werden uns dies im Tracingverfahren zu Nutze machen. Die Lösungskurve wird aufgrund der Trägheit des bewegten Punktes nicht in diese Strukturen hineinlaufen.

6.2.2 Rekonstruktion der DNA-Struktur

Aus den vorgeglätteten Bildern läßt sich nun wie in Kapitel 5 beschrieben eine Funktion $S \in V_h(\Omega^{3D})$ rekonstruieren.

Wir wählen nun ein Grauwert $c \in [0, 1]$ und stellen die Oberfläche $S^{-1}(c)$ mit dem Graphikprogramm GRAPE dar. Abb. 6.16 zeigt diese Oberfläche. Man erkennt

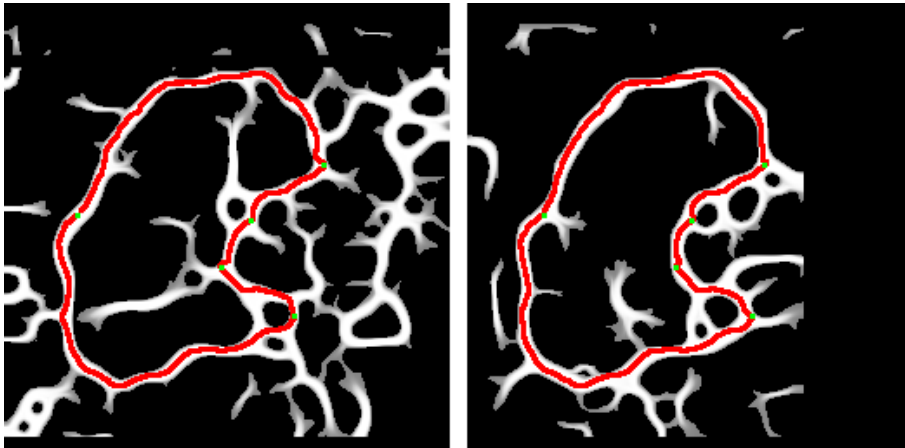


Abbildung 6.16: Verschiedene Ansichten der rekonstruierten Menge

u.a. zwei schlauchartige geschlossene Strukturen, die eine approximiert den gesuchten DNA-Ring, die andere ergibt sich aus der Nichteindeutigkeit der Rekombination (vgl. Beispiel 2.2). Die Strukturen, die während des Diffusionsprozesses in den 2D-Bildern entstanden sind, gehen, wie wir im vorigen Abschnitt gesehen haben, ebenfalls in den Rekonstruktionsprozeß mit ein; es entstehen zusätzliche Strukturen in 3D.

Aufgrund dieser zusätzlichen Strukturen und des durch die Nichteindeutigkeit entstandenen zweiten Rings ist das Resultat der dreidimensionalen Rekonstruktion unbefriedigend. Wir werden die Strukturen, die den gesuchten DNA-Ring approximieren, mit Hilfe der in Abschnitt 3.5 vorgestellten gewöhnliche Differentialgleichung isolieren.

6.3 Tracing eines geschlossenen Schlauchs

Abbildung 6.17: Projektionen der Mengen Γ_h und des Polygonzugs

Sei $S \in V_h(\Omega^{3D})$ die im letzten Abschnitt betrachtete Funktion. Wir wollen nun den gesuchten DNA-Ring, der in der Menge $\tilde{\Gamma} := S^{-1}([c, \infty))$ mit einem $c > 0$ wie im vorhergehenden Abschnitt enthalten ist, isolieren. Wir wählen per Hand ein

Paar korrespondierender Punkte, die zu den DNA-Ringen in beiden 2D-Bildern gehören und rekonstruieren daraus einen Punkt $M \in \Omega^{3D}$, den wir als ersten Startpunkt des Tracing-Verfahrens verwenden.

Um den Verlauf der Teillösungen, die wir mit Hilfe unseres Verfahrens ermitteln, innerhalb der Menge $\tilde{\Gamma}$ zu verfolgen, projizieren wir $\tilde{\Gamma}$ sowie den Polygonzug mittels der Projektionen P_1, P_2 zurück nach $\mathbb{R}^2$.

Abb. 6.17 zeigt die Projektionen der Menge $\tilde{\Gamma}$ als weiße Punkte, den projizierten Polygonzug in roter Farbe. Die grünen Punkte markieren jeweils den Startpunkt einer Teillösung. Die Lösung startet im grünen Punkt in der linken Hälfte des DNA-Rings und verläuft im Uhrzeigersinn. Nachdem das Lösungsverfahren viermal unterbrochen und mit neuen Startwerten fortgesetzt wurde, erreicht eine der Teillösungen wieder den ersten Startpunkt. Nur dieser geschlossene Polygonzug wurde abgebildet, die übrigen Teillösungen wurden verworfen.

Eine Unterbrechung erfolgte an den Stellen mit $\text{coh}(x) < c_{\text{bifurc}}$. Wir sehen, daß dies in der Regel an Stellen geschieht, an denen die Struktur Y-förmige Verzweigungen besitzt. Das Verfahren ist nicht in der Lage, die beiden Verzweigungsrichtungen zu erkennen, da der Eigenvektor v_1 von M nur die Hauptträgheitsachse wiedergibt (vgl. Abb. 6.18). Aus diesem Grund und aufgrund der Trägheit des bewegten Massenpunktes folgt die Lösung dieser Achse bis zum Rand der Struktur, wo das Verfahren aufgrund der hohen Kohärenz abbricht.

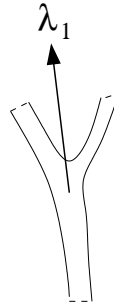


Abbildung 6.18: Modell eines Verzweigungspunktes

Die Verzweigungsrichtungen sind in der Regel nicht parallel zu einem der Eigenvektoren von M . Das Verfahren könnte optimiert werden, indem man ein Kriterium verwendet, daß den Verzweigungspunkt früher erkennt und die Verzweigungsrichtungen exakt ermittelt, so daß dort die Verzweigung erfolgen kann. Wir werden auf diese Problem im nächsten Abschnitt zurückkommen.

In Abschnitt 6.2.1 haben wir beschrieben, wie die Ausläufer, die bei der Vorglättung entstanden sind, rekonstruiert werden. Diese liegen in der Regel in einem stumpfen Winkel am DNA-Ring an. In der Praxis hat sich wie erwartet gezeigt, daß

die Lösungskurve aufgrund der Trägheit des Massenpunktes nur selten in diese falschen Strukturen hineinläuft. Werden die Parameter der gewöhnlichen Differentialgleichung optimal angepaßt, so läßt sich ein geschlossener Polygonzug finden, ohne daß eine Verzweigung notwendig ist. (Dies betrifft den hier dargestellten DNA-Ring sowie die Ringe 1 und 2 im Anhang.)



Abbildung 6.19: Verschiedene Ansichten des dreidimensionalen Polygonzugs

Abb. 6.19 zeigt verschiedene Ansichten eines dreidimensionalen ε -Schlauches um den errechneten Polygonzug.

Wir wollen untersuchen, wie weit der errechnete Polygonzug von der echten DNA-Struktur abweicht. Da wir keine Kenntnis über deren exakten Verlauf besitzen, können wir den Approximationsfehler nur schätzen. Im folgenden beschränken wir uns jedoch darauf, den ermittelten Polygonzug auf die Originalbilder zurückzuprojizieren.

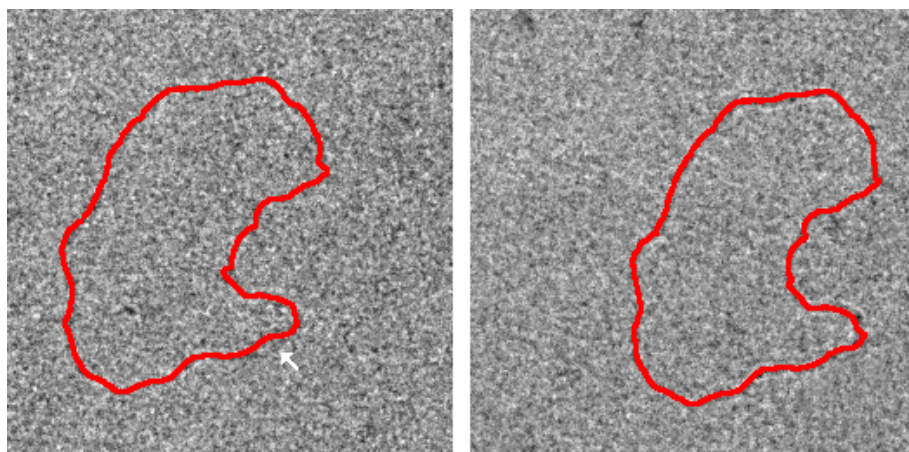


Abbildung 6.20: Projektion des dreidimensionalen Polygonzugs auf die Ausgangsbilder

Abb. 6.20 zeigt die Originalbilder mit und Abb. 6.21 - um genau vergleichen zu können - ohne projizierten Polygonzug. Man erkennt, daß der Polygonzug mit Ausnahme einiger Stellen den (per Augenmaß erkannten) DNA-Ring gut approximiert. Abweichungen um wenige Pixel ergeben sich dadurch, daß die Struktur

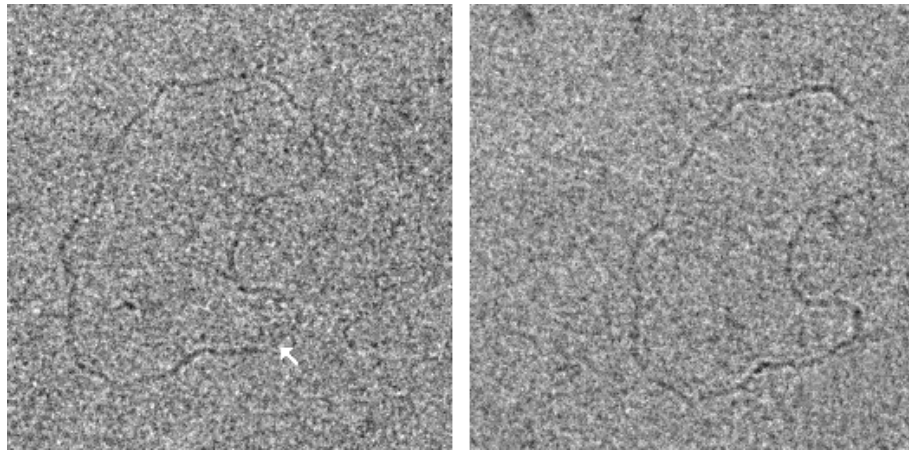


Abbildung 6.21: Ausgangsbilder

in den geglätteten Bildern breiter als in den Originalbildern dargestellt wird und diese Breite beim Lösen der gewöhnlichen Differentialgleichung eine Toleranz für den Verlauf des Polygonzugs darstellt.

An manchen Stellen weicht der gefundene DNA-Ring jedoch deutlich von dem DNA-Ring ab, den das menschliche Auge erkennt. Solche Stellen finden sich vor allem im linken Bild, das im Vergleich zum anderen eine deutlich schlechtere Qualität besitzt. Grund für die Abweichung ist, daß die DNA-Struktur beim Glätten falsch erkannt und hervorgehoben wurde. Eine solche Stelle ist z.B. am unteren rechten Teil der Struktur im linken Bild zu finden (siehe Pfeil). Das menschliche Auge erkennt hier einen eindeutigen Verlauf des DNA-Rings. Betrachtet man den Bereich der Abweichung genauer, so wird klar, daß der Verlauf dort keineswegs eindeutig ist. Das Glättungsverfahren erkennt hier einen anderen, nach menschlichem Augenmaß falschen Verlauf.

6.4 Probleme mit Kreuzungspunkten

Betrachten wir nun noch ein Paar von Stereobildern, bei denen sich die zweidimensionalen Abbilder des DNA-Rings selbst überschneiden. Diese Punkte in 2D hatten wir in 3.1.1 als Kreuzungspunkte bezeichnet.

Abb. 6.22 zeigt ein solches Paar. Die oberen beiden Bilder sind die Originalbilder, die unteren beiden sind das Ergebnis des Glättungsverfahrens mit anschließender Grauwertsegmentierung. Man erkennt in den Originalbildern zwei Kreuzungspunkte des DNA-Rings. Beim Glätten entstehen zusätzlich noch mehrere Verzweigungen in falsch erkannte Strukturen, die, wie wir gesehen haben, bei der

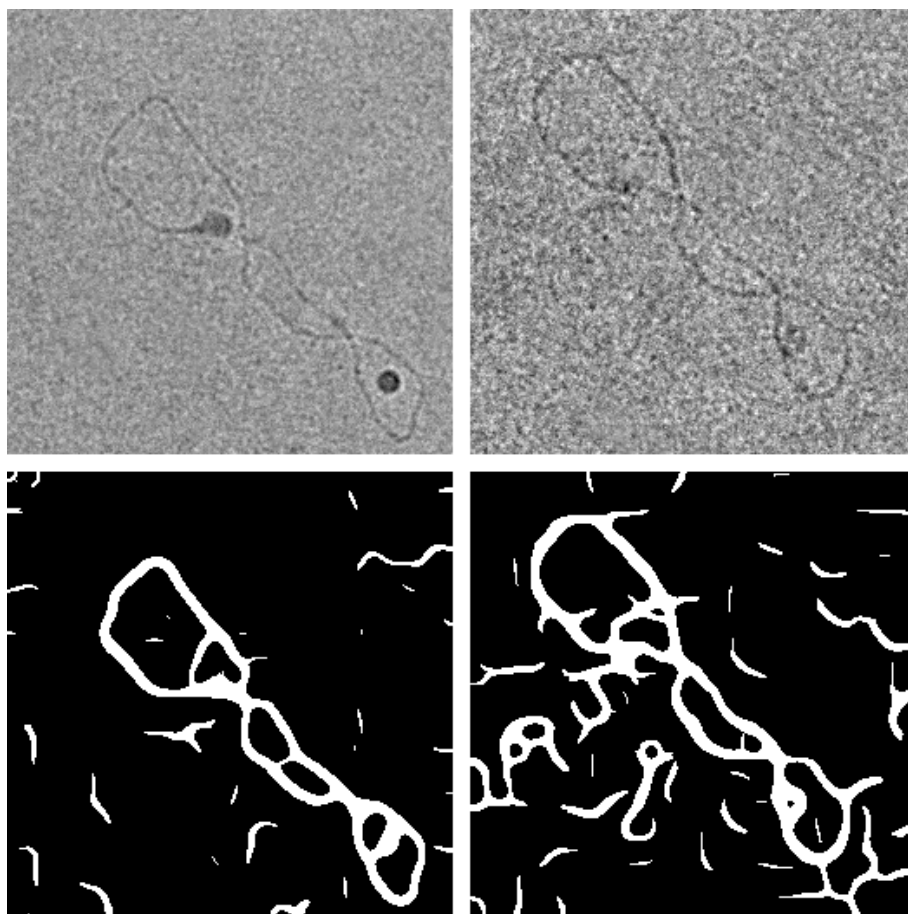


Abbildung 6.22: 2DA-Bilder mit Kreuzungspunkten

Rekonstruktion und dem Tracing kein zu großes Problem darstellen. Uns interessieren im folgenden nur die beiden Kreuzungspunkte.

Diese Kreuzungspunkte existieren nur in der Projektion, im Dreidimensionalen liegen die entsprechenden Teile des Rings voneinander entfernt. Die Tatsache, daß sich beide Projektionen des Rings in etwa an derselben Stelle in den beiden Bildern überschneiden, deutet darauf hin, daß diese Distanz im Vergleich zur Gesamtgröße des Rings sehr klein ist, die Teile des Rings also sehr nahe beieinander liegen. Da DNA-Strukturen im Allgemeinen stark in sich verdrillt sein können, erwarten wir gerade solche geringen Abstände (vgl. Abschnitt 2.1).

Problematisch ist nun, daß solche nahe beieinanderliegenden Teile durch den Diffusionsprozeß zu einer gemeinsamen Struktur verschmelzen können, wie z.B. im obigen Beispiel (Abb. 6.22) zu erkennen ist. Dadurch entstehen nach der 3D-Rekonstruktion auch Kreuzungspunkte im Dreidimensionalen.

Hinzu kommt, daß das von uns verwendete Massenmoment nicht geeignet ist, in

einem Kreuzungspunkt die verschiedenen Richtungen, in die die sich kreuzenden DNA-Stränge verlaufen, exakt zu erkennen. Dieses Problem trat bereits bei den Verzweigungen in falsche Strukturen auf. Es betrifft sowohl den Diffusionsprozeß als auch das Tracing.

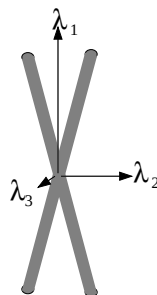


Abbildung 6.23: Modell eines Kreuzungspunkts

Um diese Erkennungsschwierigkeiten zu verdeutlichen, betrachten wir das Modell eines Kreuzungspunktes. Seien g_1, g_2 zwei Geraden in $\mathbb{R}^3$, die sich in einem Punkt p und mit Winkel α kreuzen. Um die beiden Geraden sei Masse stabförmig mit festem Radius angeordnet. Wir setzen voraus, daß $\alpha \neq 0$ gilt (siehe Abb. 6.23). Diese Massenordnung besitzt in p drei Symmetrieachsen, die gleichzeitig die Hauptträgheitsachsen des zugehörigen Trägheitstensors sind. Eine dieser Achsen entspricht der Winkelhalbierenden von α durch p , die zweite liegt in der von g_1, g_2 aufgespannten Ebene orthogonal zur ersten, die dritte liegt orthogonal zu den beiden anderen. Berechnen wir nun in p die Eigenvektoren $v_1 \dots v_3$, so liegen diese parallel zu den Symmetrieachsen. Das Massenmoment ist also nicht geeignet, die Verzweigungsrichtungen in p exakt zu erkennen.

Im Diffusionsmodell liegt die Hauptdiffusionsrichtung parallel zur Symmetrieachse, orthogonal dazu hat die Diffusion, da Kohärenz und Gradient ungleich null sind, einen kleineren Koeffizienten, so daß sich die Struktur durch die Diffusion entlang der Symmetrieachse ausprägen wird. Die sich kreuzenden Strukturen verschwimmen daher einer gemeinsamen parallel zur Symmetrieachse verlaufenden Struktur. Dies gilt analog für die 2D-Diffusion. In den unteren Bildern in Abb. 6.22 ist erkennbar, wie sich die Struktur an Kreuzungspunkten entlang der "falschen" Hauptdiffusionsrichtung ausprägt.

Beim Verfolgen der DNA-Struktur mit Hilfe der gewöhnlichen Differentialgleichung ergibt sich an den Kreuzungspunkten ein ähnliches Problem. Anhand der Kohärenz läßt sich zwar feststellen, ob die Lösungskurve einen Kreuzungspunkt erreicht, wir können jedoch die Verzweigungsrichtungen nicht exakt bestimmen. Als Verzweigungsrichtungen wählen wir in unserem Verfahren $\pm v_i$. Denkbar wäre auch eine geeignete Linearkombination aus den beiden Eigenvektoren. Die Praxis

hat gezeigt, daß die in der Regel vorhandene Unsymmetrie der Umgebung um einen Verzweigungspunkt entscheidet, in welchen der Verzweigungsäste der Massenpunkt hineinläuft.

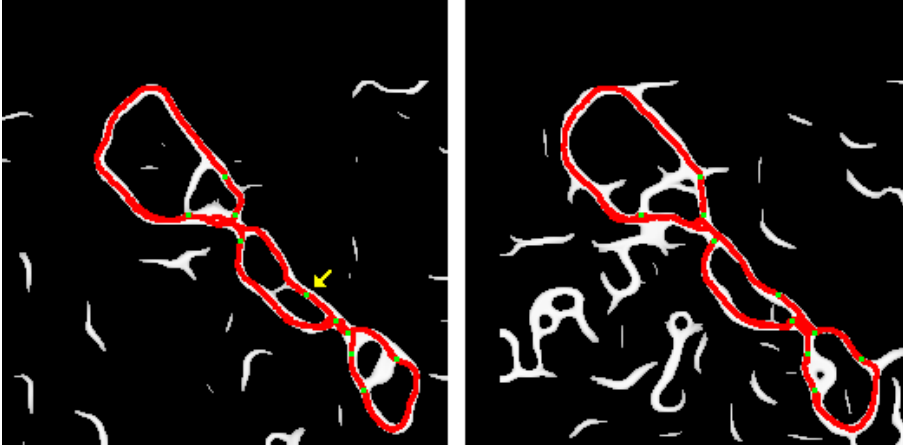


Abbildung 6.24: Projektion des dreidimensionalen Rings auf die vorgeglätteten Bilder

Betrachten wir nun die Approximation des in Abb. 6.22 dargestellten DNA-Rings, die uns das Tracing-Verfahren liefert. Die berechneten Teillösungen, die in Form von Polygonzügen vorliegen, lassen sich in diesem Fall nicht zu einem geschlossenen Polygonzug zusammensetzen. Abb. 6.24 zeigt die Projektionen der verschiedenen Polygonzüge. Die grünen Punkte markieren die Startpunkte der einzelnen Polygonzüge. Der erste Startpunkt ist mit einem Pfeil markiert.

In diesem Beispiel lassen wir die Lösung bereits in X_0 in die Richtungen $\pm v_1$ verzweigen, d.h. wir berechnen die Lösungen der Differentialgleichung zum Startpunkt X_0 und den Startvektoren $r_0 = v_1$ und $r'_0 = -v_1$.

Bei der Berechnung der Teillösungen steigen wir im Baum zunächst möglichst tief ab ("Depth-first"). Wir betrachten nun den Ast, in den zuerst abgestiegen wird und fassen die zugehörigen Polygonzüge zu einem zusammen, den wir als Polygonzug A bezeichnen. Sei r_0 der Startvektor der ersten Teillösung. Die oberen beiden Bilder in Abb. 6.25 zeigen die Projektionen dieses Polygonzugs. Auffallend ist, daß in der linken Projektion am unteren Kreuzungspunkt Teile von A in etwa parallel in der Projektionsebene verlaufen, während man, wenn man die Originalbilder betrachtet, eher erwarten würde, daß sie sich kreuzen sollten.

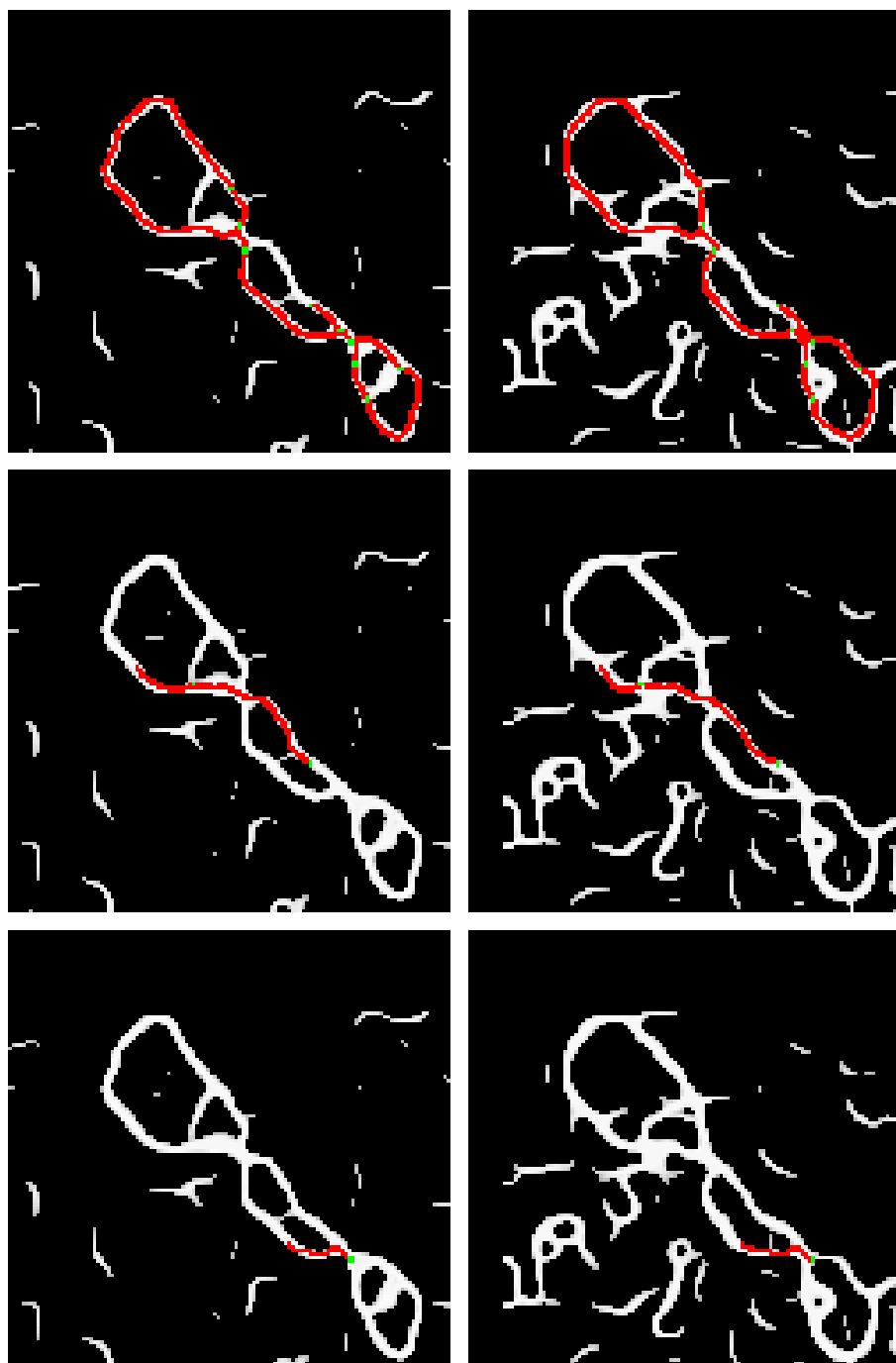


Abbildung 6.25: Polygonzüge A,B,C

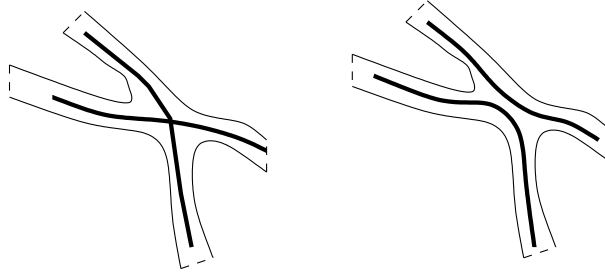


Abbildung 6.26: Mögliche Verläufe am Kreuzungspunkt

Abb. 6.26 zeigt diese beiden Möglichkeiten. Grund für den parallelen Verlauf ist, daß der Eigenvektor v_1 von $M(x)$ wie oben beschrieben diese Richtung vorgibt. Die letzte Teillösung des Astes findet nicht zum ersten Startpunkt zurück, sondern führt in einen bereits durchlaufenen Teil der Struktur, so daß das Verfahren abbricht.

Das Verfahren versucht nun ausgehend von den bisherigen Verzweigungspunkten weitere Teillösungen zu finden. Hier werden zwei weitere Polygonzüge berechnet, die nicht parallel zu einem bereits vorhandenen verlaufen. Der erste, den wir mit B bezeichnen, stellt eine Verbindung der im unteren Kreuzungspunkt parallel verlaufenden Teile von A her (Abb. 6.25, mittlere Reihe).

Den letzten Polygonzug, bezeichnet mit C, erhalten wir als Lösung der Differentialgleichung mit Startpunkt x_0 und Startvektor r'_0 (Abb. 6.25, untere Reihe). Dieser Polygonzug folgt zunächst dem Teil der Struktur, der noch nicht von A und B durchlaufen wurde. Das letzte Stück liegt nahe dem bereits berechneten Polygonzug A. Das Verfahren erkennt, daß dieser Teil der Struktur bereits durchlaufen wurde und bricht ab.



Abbildung 6.27: 3D-Schlauch in verschiedenen Perspektiven

Betrachten wir nun den dreidimensionalen ε -Schlauch um die errechneten Polygonzüge. Abb. 6.27 zeigt verschiedene Ansichten dieses Schlauchs. Deutlich zu erkennen ist, daß der Verlauf des Schlauchs im Bereich der Kreuzungspunkte nicht eindeutig dargestellt wird. Im oberen Bereich sind die Endpunkte der Polygonzüge A und C sichtbar. Das Endstück des Polygonzugs A folgt, wie bereits oben erwähnt, einem falschen, bereits approximierten Teil der Struktur, Polygonzug B schließt die Lücke zwischen dem ersten Startpunkt und dem Polygonzug A.

Im Bereich des unteren Kreuzungspunktes verlaufen zwei Teile des Polygonzugs nahezu parallel, die Lücke zwischen beiden Teilen wird durch den Polygonzug B überbrückt. Wie oben erwähnt, scheint, wenn man die Originalbilder betrachtet, dieser fast parallele Verlauf falsch wiedergegeben zu sein, während das Verbindungsstück die Struktur anscheinend richtig wiedergibt. Entsprechend fehlt ein weiterer Teil, der die übrigen Teile verbindet und in den Projektionen mit B zusammen den Kreuzungspunkt bildet, wie in Abb. 6.26 rechts skizziert.

An dieser Stelle sei nochmal darauf hingewiesen, daß dieses Mehrdeutigkeitsproblem dadurch entsteht, daß bei der Rekonstruktion ein Kreuzungspunkt in $\mathbb{R}^3$ entstanden ist. Erst dadurch ergeben sich die verschiedenen Möglichkeiten, die vier angrenzenden Äste miteinander zu verbinden.

Betrachten wir zum Schluß noch die Projektion des approximierten DNA-Rings auf die Originalbilder. Abb. 6.28 zeigt diese Projektionen, Abb. 6.29 die Originalbilder. Wie im vorherigen Abschnitt gibt die Breite der geglätteten Struktur die maximale Abweichung des Polygonzugs vor.

Der gesuchte DNA-Ring wurde also auch in diesem Fall, abgesehen von den Problemen an den Kreuzungspunkten, gut approximiert.

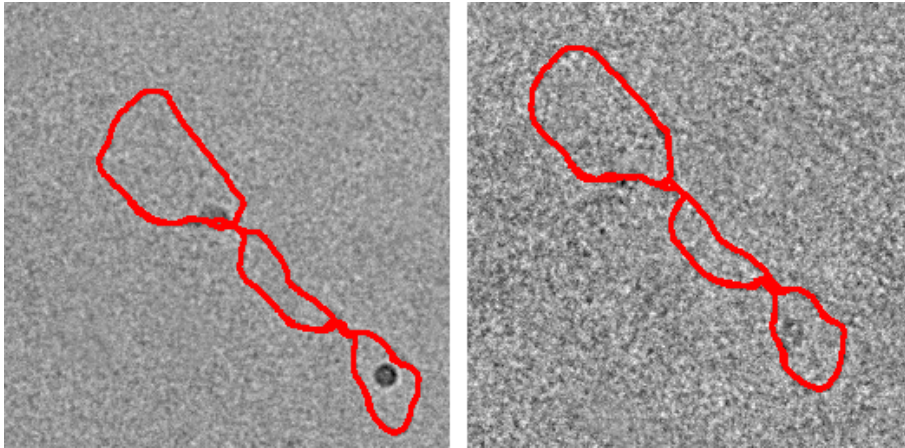


Abbildung 6.28: Projektion des dreidimensionalen Polygonzugs auf die Ausgangsbilder

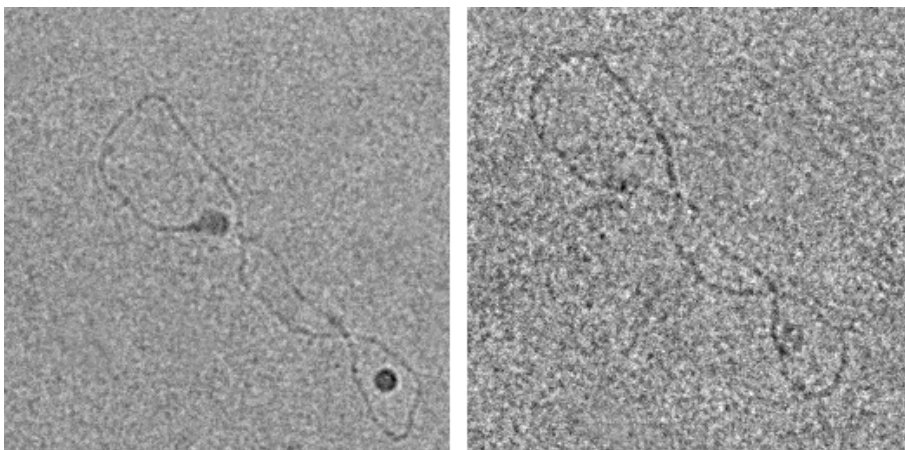


Abbildung 6.29: Ausgangsbilder

Kapitel 7

Resumé

Im vorherigen Kapitel haben wir das vorgestellte Verfahren zur 3D-Rekonstruktion getestet und dabei gute Resultate erzielt.

Zur Vorglättung der zweidimensionalen Bilder haben wir ein Modell einer Anisotropen Diffusion gewählt, das eindimensionale Strukturen erkennt und durch Kontrastschärfung hervorhebt. Wir haben gesehen, daß wir mit Hilfe dieses Modells trotz des starken Rauschens in der Lage sind, die DNA-Strukturen in den vorliegenden Bildern zu erkennen und hervorzuheben. An wenigen Stellen wird aufgrund des Rauschens der Verlauf anders erkannt, als er nach Augenmaß ist. Meiner Meinung nach ist es nicht möglich, diese Erkennung bedeutend zu verbessern. Da der Verlauf der DNA-Ringe in den verrauschten Bildern an vielen Stellen nicht eindeutig ist, wir können nicht erwarten, die DNA-Ringe exakt zu erkennen.

Wir haben das Glättungsverfahren auf Bilder verschiedener DNA-Ringe (siehe Anhang) angewandt. Dabei war eine Anpassung der Parameter nicht notwendig. Lediglich die Parameter der rechten Seite zum Zeitpunkt $t = 0$, die wie in Kapitel 6.1.2 beschrieben einen großen Einfluß auf den Verlauf Erkennung hat, haben wir für jedes Bild neu optimiert, um das Schließen der größeren Lücken zu beschleunigen.

Es hat sich also gezeigt, daß das Verfahren robust ist.

Eine verbesserte Wahl der rechten Seite zum Zeitpunkt $t = 0$ könnte zu einer leichten Verbesserung des Resultats führen, d.h. das weniger zusätzliche Strukturen erkannt werden und sich Lücken des DNA-Rings noch schneller schließen.

Das Tracingverfahren liefert an den Stellen, an denen keine Kreuzungspunkte auftreten, sehr gute Resultate, vorausgesetzt, der Verlauf des DNA-Rings wurde bei der Vorglättung richtig erkannt. An den Kreuzungspunkten ergeben sich einige Schwierigkeiten aufgrund der Nichteindeutigkeit des Problems. Das Tracingverfahren könnte hier noch optimiert werden, indem man andere Kriterien

zur Bestimmung der Verzweigungsrichtungen wählt.

Wir möchten noch einmal betonen, daß das Tracingverfahren nach einer einmaligen Justierung der Parameter in der Lage ist, den gesuchten DNA-Ring automatisch zu isolieren. Lediglich die Wahl des Startpunktes für jedes Bildpaar muß von Hand erfolgen.

Insgesamt liefert das vorgestellte Verfahren also trotz des starken Rauschens in den Stereobildern gute Resultate bei der dreidimensionalen Rekonstruktion der gesuchten DNA-Ringe.

Anhang A

Weitere DNA-Ringe

A.1 DNA-Ring 1

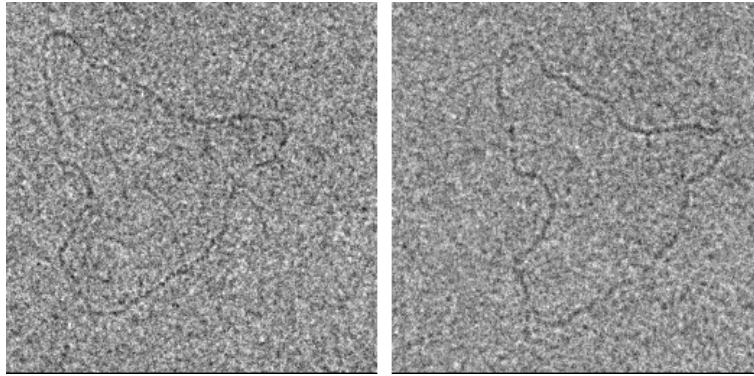


Abbildung A.1: Originalbilder

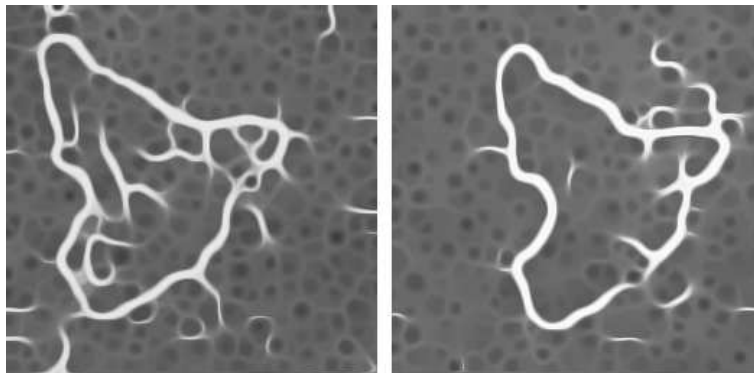


Abbildung A.2: Vorgeglättete Bilder

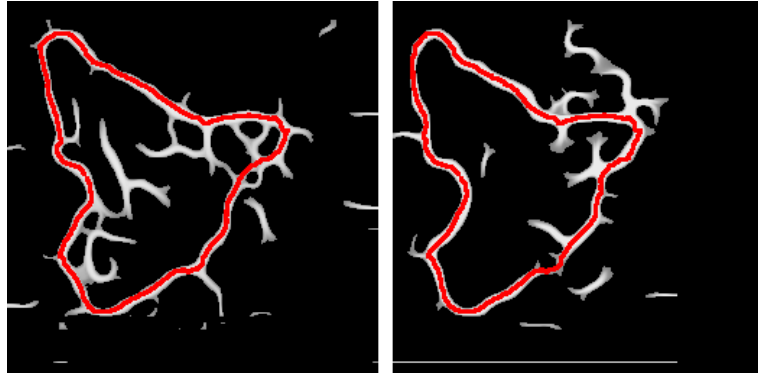


Abbildung A.3: Projektion des 3D-Polygonzugs

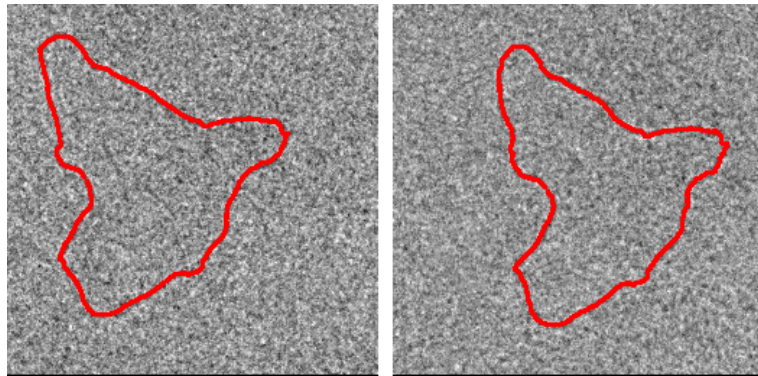


Abbildung A.4: Projektion des 3D-Polygonzugs



Abbildung A.5: Ansichten des 3D-Polygonzugs

A.2 DNA-Ring 2

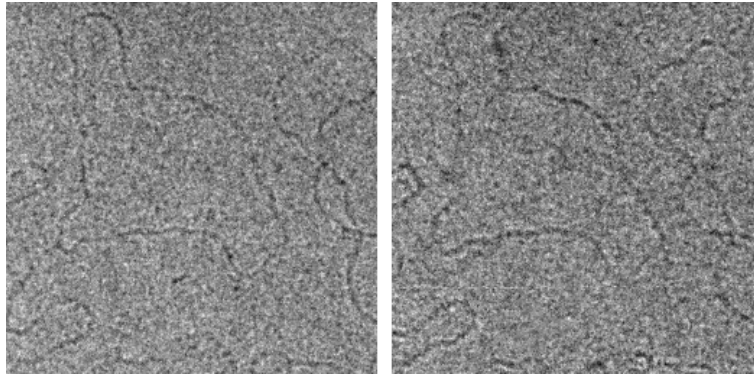


Abbildung A.6: Originalbilder

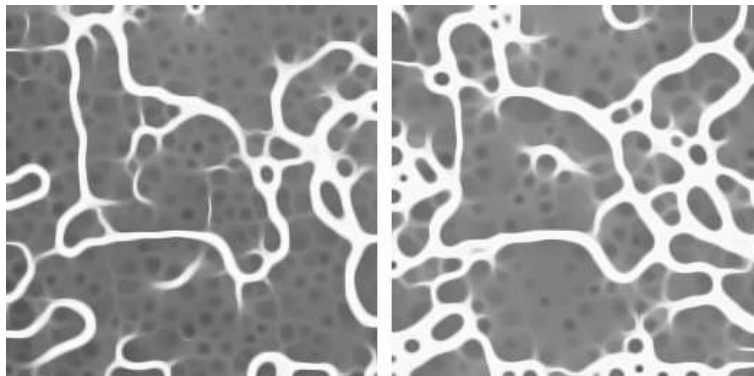


Abbildung A.7: Vorgeglättete Bilder

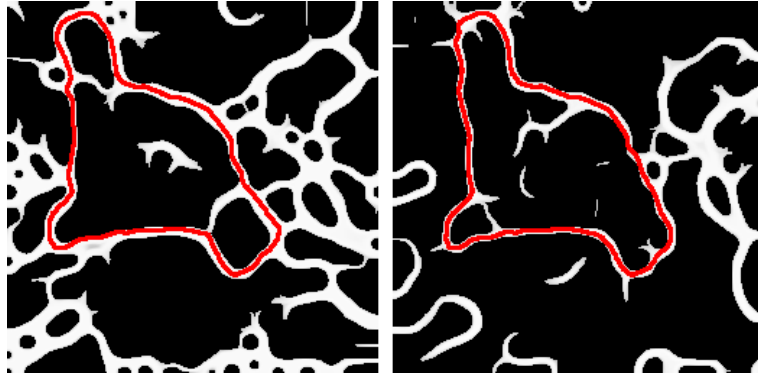


Abbildung A.8: Projektion des 3D-Polygonzugs

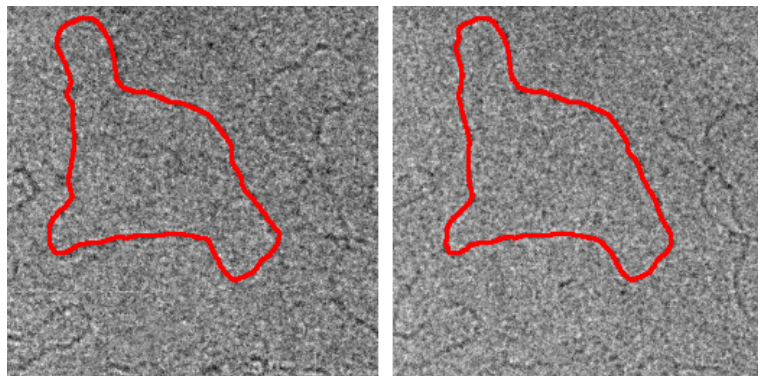


Abbildung A.9: Projektion des 3D-Polygonzugs



Abbildung A.10: Ansichten des 3D-Polygonzugs

A.3 DNA-Ring 3

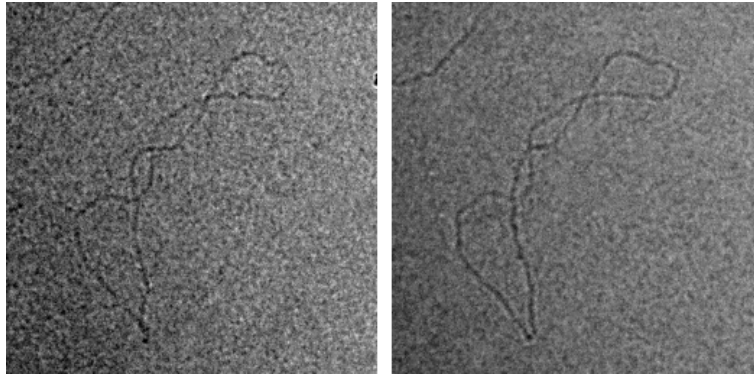


Abbildung A.11: Originalbilder

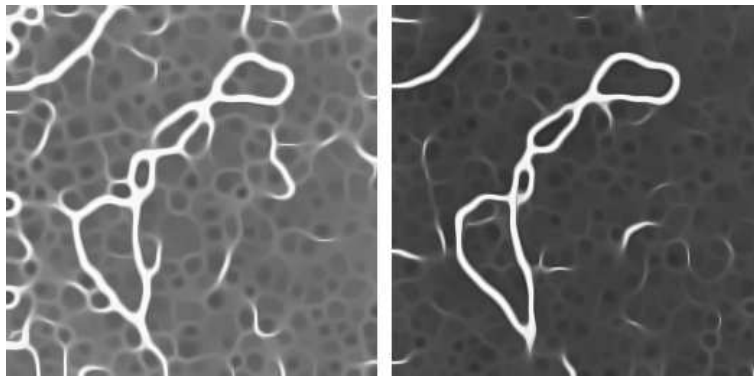


Abbildung A.12: Vorgeglättete Bilder

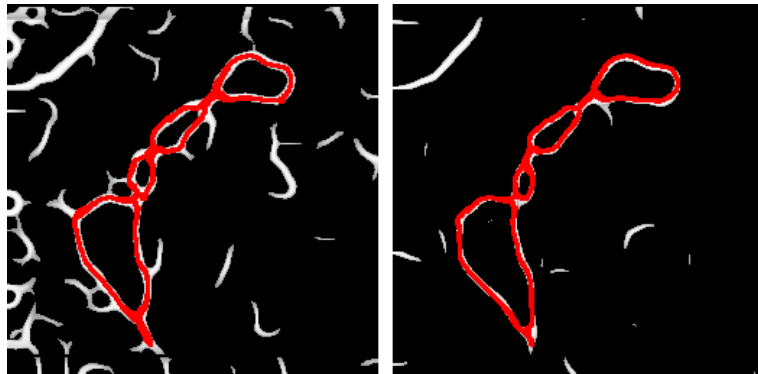


Abbildung A.13: Projektion des 3D-Polygonzugs

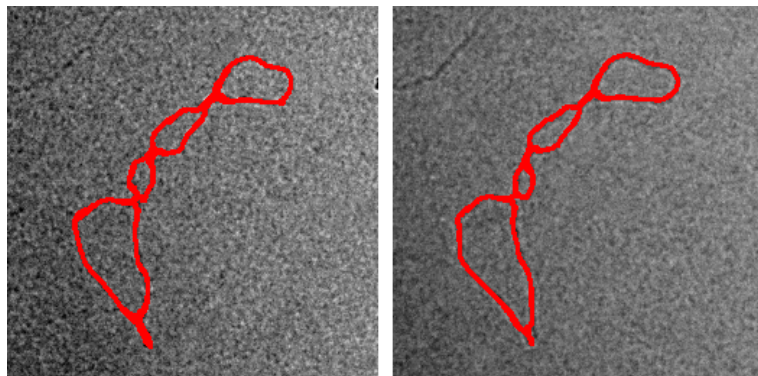


Abbildung A.14: Projektion des 3D-Polygonzugs



Abbildung A.15: Ansichten des 3D-Polygonzugs

A.4 DNA-Ring 4

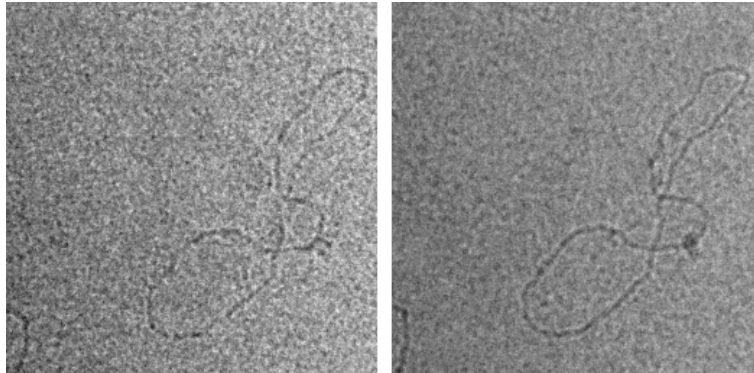


Abbildung A.16: Originalbilder

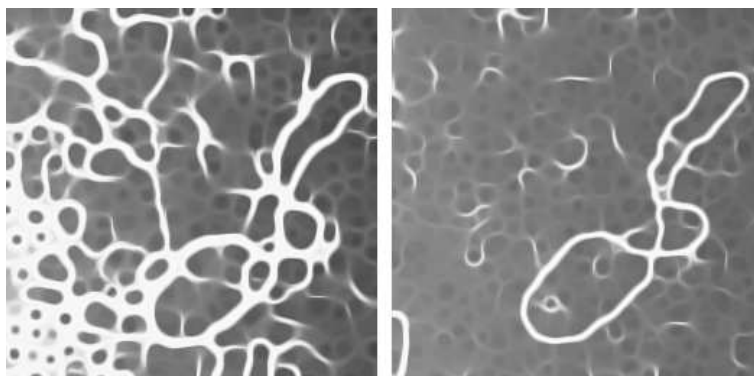


Abbildung A.17: Vorgeglättete Bilder

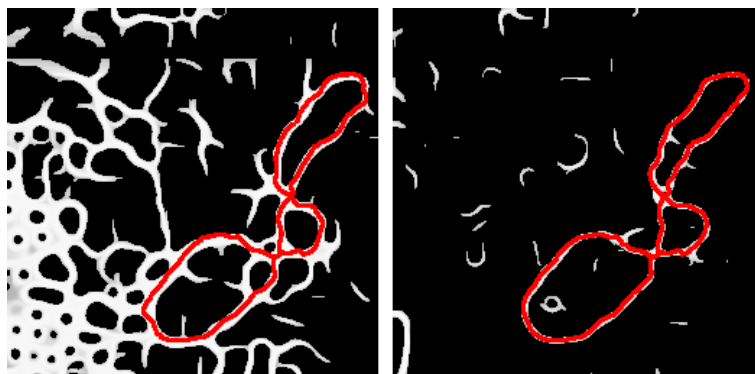


Abbildung A.18: Projektion des 3D-Polygonzugs

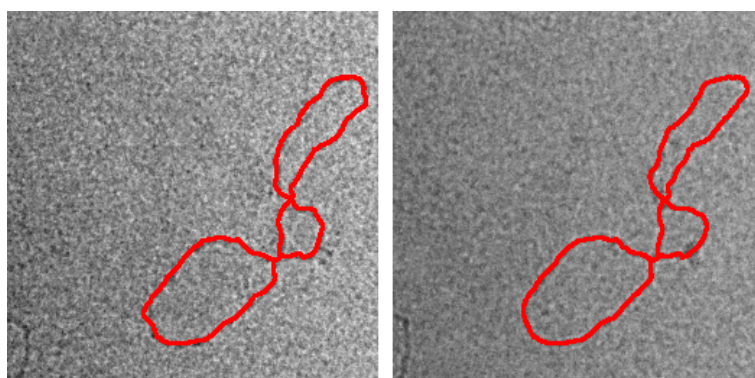


Abbildung A.19: Projektion des 3D-Polygonzugs

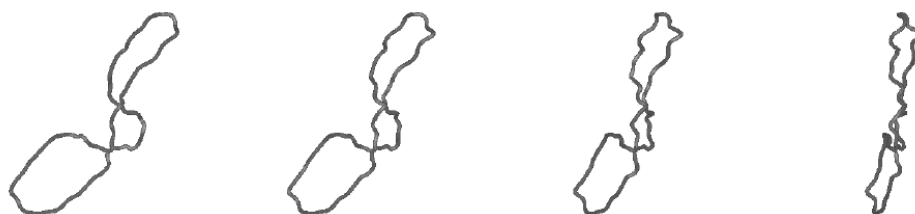


Abbildung A.20: Ansichten des 3D-Polygonzugs

Literaturverzeichnis

- [1] Hans Wilhelm Alt. *Lineare Funktionalanalysis*. Springer Verlag, 3rd edition, 1999.
- [2] Herbert Amann. *Gewöhnliche Differentialgleichungen*. de Gruyter Verlag, 2nd edition, 1995.
- [3] Manfred Auer. Three-dimensional electron cryo-microscopy as a powerful structural tool in molecular medicine. *J. of Molekular Medicine*, 78:919–202, 2000.
- [4] Jan Bednar, Patrick Furrer, Andrzej Stasiak, and Jacques Dubochet. The twist, writhe and overall shape of supercoiled DNA change during counterion-induced transition from a loosely to a tight interwound superhelix. *J.Mol.Biology*, 235:825–847, 1994.
- [5] Rajendra Bhatia. *Matrix Analysis*. Springer Verlag, 1997.
- [6] Dietrich Braess. *Finite Elemente*. Springer Verlag, 1992.
- [7] Bronstein, Semendjajew, Musiol, and Müelig. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, 3rd edition, 1997.
- [8] J.M. Buhmann and T. Fröhlinghaus. Regularizing phase-based stereo. ??, 2000.
- [9] F. Catte, P.-L. Lions, J.-M. Morel, and T.Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM, J. Numerical Analysis*, 29:182–193, February 1992.
- [10] Bernard Couapel. Stereo vision with the use of a virtual plane in space. *Chinese Journal of Electronics*, 4, April 1995.
- [11] J. Dubochet, J. Bednar, P. Furrer, and A. Stasiak. Cryo-electron microscopy of DNA. *J. Nucleic Acids and Molecular Biology*, 8:41–55, 1994.
- [12] Oliver Faugeras and Renaud Keriven. Variational principles, surface evolution, PDE's, level set methods and the stereo problem. *Rapport de recherche no. 3021, INRIA*, 1996.

- [13] H. Gajewski, K. Gröger, and K. Zacharias. *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*. Akademie Verlag Berlin, 1974.
- [14] Gordon Kindlmann, David Weinstein, and Davis Hart. Strategies for direct volume rendering of diffusion tensor fields. *IEEE Transactions on Visualization and Computer Graphics*, 6, april,june 2000.
- [15] O.A. Ladyzenskaja. *Linear and Quasilinear Equations of Parabolic Type*. American mathematical society, 1968.
- [16] E.M. Lifschitz L.D. Landau. *Lehrbuch der Theoretischen Physik, Band I, Mechanik*. Akademie Verlag, 11th edition, 1984.
- [17] H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981.
- [18] Q.T. Luong and O.D. Faugeras. The fundamental matrix: theory, algorithms and stability analysis. *INRIA*, 0.
- [19] Tobias Preußner. Nichtlineare Diffusion in der Bildverarbeitung - Ein adaptives FE-Verfahren. Master's thesis, Universität Bonn, 1999.
- [20] Stoer and Bulirsch. *Numerische Mathematik*, volume 1,2. Springer Lehrbuch, 1994.
- [21] Lubert Stryer. *Biochemie*, pages 90,91. Spektrum Akademischer Verlag, 4th edition, 1996.
- [22] Joachim Weickert. Foundations and applications of nonlinear anisotropic diffusion filtering. *Z.f. Angewandte Mathematik und Mechanik*, 76, 1996.
- [23] Joachim Weickert. *Anisotropic Diffusion in Image Processing*. Teubner Verlag, 1998.