Technische Universität München
Fakultät für Mathematik

Mehrgitterbasierte Beschleunigung von GNC-Verfahren in der Bildsegmentierung

Diplomarbeit von Christian Ludwig

Aufgabensteller:	Prof. Dr. F. Bornemann
Betreuer:	Prof. Dr. F. Bornemann
Abgabetermin:	1. August 2002

Ich erkläre hiermit, dass ich die Diplomarbeit selbständig und nur mit den angegebenen Hilfsmittel angefertigt habe.

München, den 30.07.2002

Inhaltsverzeichnis

1 Motivation

1.1 Mumford-Shah-Funktional

Die Aufgabe bei der Bildsegmentierung ist es, in einem gegebenen verrauschten Bild eine Kantenmenge und ein entrauschtes Idealbild zu finden. Die Ursachen für die Fehler des Eingabebildes sind vielfältig: Lichtbrechung, unterschiedliche Empfindlichkeit der einzelnen Aufnahmesensoren, Quantisierungseffekte, usw. Ein Kantendetektor wird u. a. an seiner Detektions- und Lokalisiationsgüte gemessen. Er soll möglichst viele Kanten im Bild erkennen, während Nicht-Kanten mit geringer Wahrscheinlichkeit fälschlich extrahiert werden sollen. Ferner sollten die erkannten Kanten nahe an der wahren Kante liegen.

Mumford und Shah haben zur Lösung des Segmentierungsproblems in [MS89] die Minimierung von

$$\mathcal{J}(u,K) := \lambda^2 \int_{\Omega\setminus K} \|\nabla u(x)\|^2 \, dx + \alpha \mathcal{H}^1(K) + \int_\Omega |u(x) - g(x)|^2 \, dx \tag{1.1}$$

vorgeschlagen. Dabei sind $\Omega \subset \mathbb{R}^2$ ein durch einen Lipschitz-Rand beschränktes Definitionsgebiet, $g \in L^\infty(\Omega)$ eine Funktion, die den Grauwert bzw. die Lichtintensität des gegebenen Originalbildes beschreibt, $\alpha > 0$ und $\lambda > 0$ zwei Parameter, K eine abgeschlossene, messbare Menge mit endlichem eindimensionalen Hausdorffmaß $\mathcal{H}^1(K)$ und $u \in \mathcal{C}^1(\Omega\setminus K)$.

Die Menge K repräsentiert in diesem Model die Kantenmenge und die Funktion u das entrauschte Bild. Der erste Summand von $\mathcal{J}$ sorgt dafür, dass u außerhalb der Kantenmenge möglichst glatt ist. Durch den zweiten Summanden wird erreicht, dass die Kantenmenge K möglichst klein ist. Der letzte Summand stellt sicher, dass u nicht weit vom Originalbild g abweicht.

In dem Buch [MS95] wurde eine Vielzahl der gängigen Algorithmen zur Kantendetektion untersucht. Die Autoren kamen zu dem sehr überraschenden Ergebnis, dass, obwohl die Verfahren sehr unterschiedlich sind, ihnen allen das Modell von Mumford und Shah (oder leichte Variationen davon) zugrunde liegt. Man kann diese Algorithmen als Versuch interpretieren, dieses Funktional, oder leichte Varianten davon, zu minimieren. In diesem Sinne scheint das Mumford-Shah-Funktional das allgemeine Modell der Bildsegmentierung zu sein.

Die theoretische Frage, die sich nun stellt, ist, ob es überhaupt ein Paar $(\bar{u}, \bar{K})$ gibt, welches (1.1) minimiert. Wenn ja, ob es eindeutig ist und welche Eigenschaften ein solches minimierendes Paar hat. Mumford und Shah haben in [MS89] vermutet, dass es ein $(\bar{u}, \bar{K})$ gibt, bei dem sich $\bar{K}$ aus endlich vielen $\mathcal{C}^1$-Kurven zusammensetzt.

Das Problem (1.1) fällt in die Kategorie der „free discontinuity problems“, für die E. De Giorgi eine schwache Formulierung im Raum SBV vorschlug ([DG88]). Ein $u \in L^1(\Omega)$ liegt genau dann in $SBV(\Omega)$, wenn sich die Ableitung im Sinne der Distributionentheorie Du in

$$Du = \nabla u \cdot \mathcal{L}^n|_\Omega + (u^+ - u^-)\nu \cdot \mathcal{H}^1|_{S_u}$$

zerlegen lässt. Dabei sind $\nabla u \in L^1(\Omega, \mathbb{R}^2)$, $S_u := \{x \in \Omega \,|\, u^-(x) < u^+(x)\}$ die Menge der Sprungstellen von u, ν die Einheitsnormale auf S_u (vgl. Abbildung 1.1) und

$$u^+(x) = \inf\Big\{t \in [-\infty; +\infty] \;\Big|\; \lim_{\rho \to 0^+} \frac{|\{u > t\} \cap B_\rho(x)|}{\rho^2} = 0\Big\}$$

bzw.

$$u^-(x) = \sup\Big\{ t \in [-\infty; +\infty] \;\Big|\; \lim_{\rho \to 0^+} \frac{|\{u < t\} \cap B_\rho(x)|}{\rho^2} = 0 \Big\}$$

der approximative obere bzw. untere Grenzwert von u bei x. Eine genaue Beschreibung des Raumes SBV und seine Eigenschaften findet sich in [Amb89].

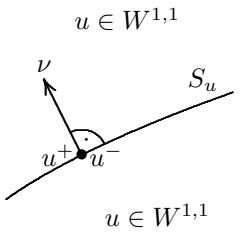

Abbildung 1.1: Veranschaulichung von S_u und ν

Bei der schwachen Formulierung von (1.1) ist dann ein Minimum von

$$\mathcal{J}(u) := \lambda^2 \int_\Omega \|\nabla u\|^2 \, dx + \alpha \mathcal{H}^1(S_u) + \int_\Omega |u(x) - g(x)|^2 \, dx \tag{1.2}$$

mit $u \in SBV(\Omega)$ gesucht.

In [Amb89] wurde gezeigt, dass für ein beliebiges $g \in L^\infty(\Omega)$ mindestens ein Minimum $u \in SBV(\Omega)$ von (1.2) existiert. In [DGCL89] wurde bewiesen, dass für jedes Minimum

1. $u \in L^\infty(\Omega)$ mit $\|u\|_\infty \leq \|g\|_\infty$,
2. $u \in W^{2,p}_{\text{loc}}(\Omega \backslash \overline{S_u})$ für alle $p \in [1; \infty[$ und $\lambda^2 \Delta u = u - g$ in $\Omega \backslash \overline{S_u}$,
3. $\bar{u} = u^+ = u^- \in \mathcal{C}^1(\Omega \backslash \overline{S_u})$ und
4. $\mathcal{H}^1(\Omega \cap \overline{S_u} \backslash S_u) = 0$

gilt.

Der Zusammenhang von Minima von (1.1) und Minima von (1.2) wurde in [DMMS92] untersucht. Dort wurde gezeigt, dass für jedes Minimum $u \in SBV(\Omega)$ von (1.2) das Paar $(u, \overline{S_u})$ ein Minimum von (1.1) ist. Ist umgekehrt (u, K) ein Minimum von (1.1), dann ist u (nach geeigneter Fortsetzung auf $K \cap \Omega$) in $SBV(\Omega)$ und ein Minimum von (1.2) mit $\overline{S_u} \subset K$ und $\mathcal{H}^1(K \backslash S_u) = 0$.

Bonnet hat in [Bon96] lokale Regularitätseigenschaften einer minimalen Kantenmenge K betrachtet und festgestellt, dass jede isolierte Zusammenhangskomponente von K eine Vereinigung von endlich vielen $\mathcal{C}^1$-Kurven ist, die, außer an den Kantenenden, sogar $\mathcal{C}^{1,1}$-Kurven sind.

In [MS95] wurde gezeigt, dass es ein $C = C(\Omega)$ gibt, so dass es für jedes (1.1) minimierendes Paar (u, K) eine Kurve γ gibt, die K enthält, mit $l(\gamma) \leq C\mathcal{H}^1(K \cup \partial\Omega)$, falls K kleinstmöglich ist, also falls es kein abgeschlossenes $\tilde{K} \subset K$, $K \neq \tilde{K}$ gibt mit $J(u, \tilde{K}) \leq J(u, K)$ nach Fortsetzung von u auf $\Omega \backslash \tilde{K}$.

Noch unbeantwortet sind die Fragen, welche Regularitätseigenschaften Kanten an den Enden haben und ob sich jedes minimierende K aus endlich vielen Kurvenstücken zusammensetzt.

1.2 Numerische Behandlung

Für die numerische Behandlung des Mumford-Shah-Funktionals (1.2) wird dieses durch eine Folge von anderen Funktionalen $\mathcal{J}_\tau$ geeignet approximiert. Dabei sollte für die Approximation gelten, dass aus

$$\mathcal{J}_\tau \to \mathcal{J}, \quad u_\tau \to u \quad \text{und} \quad u_\tau = \min_v \mathcal{J}_\tau(v)$$

folgt, dass

$$\mathcal{J}(u) = \min_v \mathcal{J}(v)$$

gilt. In diesem Zusammenhang ist die Γ-Konvergenz das passende Mittel für die Approximation: Ist X ein metrischer Raum, (f_j) eine Folge von Funktionen $f_j\colon X \to \overline{\mathbb{R}}$ und ist $f_\infty\colon X \to \overline{\mathbb{R}}$, so ist f_∞ der Γ-Limes von (f_j), falls für alle $x \in X$ gilt:

1. Für jede Folge (x_j) mit Grenzwert x ist

$$f_\infty(x) \le \liminf_j f_j(x_j).$$

2. Es gibt eine Folge (x_j) mit Grenzwert x, so dass

$$f_\infty(x) \ge \limsup_j f_j(x_j).$$

Eine Einführung in die Theorie der Γ-Konvergenz findet sich z. B. in [DM93].

In [BDM97] wurde gezeigt, dass es unmöglich ist, mit auf H^1 definierten Funktionalen der Form

$$\int_\Omega f_\tau\big(\nabla u(x)\big)\,dx + \int_\Omega |u(x) - g(x)|^2\,dx \tag{1.3}$$

obige gewünschte Konvergenz zu erreichen. In [AT90] wurde dieses Problem von Ambrosio und Tortorelli umgangen, indem sie eine Folge von auf Sobolev-Räumen definierten elliptischen Funktionalen mit einer zusätzlichen Hilfsvariable (Line-Process) einführten. Diese Folge besitzt, im Sinne der Γ-Konvergenz, das Mumford-Shah-Funktional (1.1) als Grenzwert. In [BC94] wird eine Diskretisierung mit stückweise linearen Finiten Elementen beschrieben.

Chambolle und Dal Maso haben in [CDM99] eine Approximation des Mumford-Shah-Funktionals in der Form (1.3) vorgeschlagen, wobei dort aber zusätzliche Restriktionen an den Funktionenraum, auf dem diese Funktionale definiert sind, gestellt wurden, um eine Approximation im Sinne der Γ-Konvergenz zu erreichen. Daher sind dort spezielle Triangulierungen erforderlich, so dass bei der Implementierung (vgl. [BC00]) Algorithmen zur adaptiven Gitterverfeinerung zum Einsatz kommen. Die Approximation findet hier durch sukzessive Verfeinerung statt.

Der in dieser Diplomarbeit behandelte Algorithmus von Prof. F. Bornemann hat folgende Eigenschaften (vgl. [Bor00]): Hier werden auf Sobolev-Räumen definierte Funktionale J_τ mit einer zusätzlichen Hilfsvariable (Line-Process) betrachtet. Dabei wird mit Finiten Elementen diskretisiert. Im Gegensatz zu [CDM99] sind der Approximationsparameter τ und der Parameter h für die Feinheit der Vergitterung entkoppelt, ähnlich wie in [BC94]. Dadurch ist es möglich, den Algorithmus für ein festes h zu verwenden. Diese Situation tritt z. B. ein, wenn man in einem digitalen Bild Kanten sucht. Dort ist eine Verfeinerung nicht möglich, da zwischen zwei Pixeln keine zusätzliche Grauwertinformation vorliegt. In

dieser Arbeit werden auch speziell die Effekte, die aus uniformen Triangulierungen bzw. Vergitterungen entstehen, näher untersucht. Das Mumford-Shah-Funktional wird durch eine **g**raduated **n**on-**c**onvexity (GNC) Homotopie nach einer Idee von Blake und Zisserman (vgl. [BZ87]) approximiert. Jede einzelne Funktion J_τ der verwendeten Homotopie hat, im Grenzfall $h \to 0$, eine anisotrope Version des Mumford-Shah-Funktionals als Γ-Grenzwert. Um die Minimierung der J_τ zu vereinfachen, wird jedes J_τ mit Hilfe der Half-Quadratic Regularization (HQR) (vgl. [GY95] und [CBFAB97]) in ein J_τ^* überführt, welches dann mit einem ADI-Verfahren minimiert wird. Dabei ist die Minimierung bzgl. der einen Variablen ein explizit lösbares konvexes Problem und die Minimierung bzgl. der anderen Variablen ein linear elliptisches Problem. Letzteres wird mit algebraischen Mehrgitterverfahren gelöst.

1.3 Überblick

Diese Diplomarbeit ist wie folgt aufgebaut:

In Kapitel 2 wird das behandelte Problem dargestellt und Notation eingeführt. In Kapitel 3 wird die Idee der GNC-Homotopie behandelt. Außerdem wird am Ende dieses Kapitels untersucht, welche anisotropen Effekte durch eine uniforme Vergitterung entstehen. In Kapitel 4 wird die Legendre-Fenchel-Transformation eingeführt und auf das hier betrachtete Funktional angewendet. Anschließend wird bei der daraus entstehenden HQR-Iteration eine Konvergenzanalyse durchgeführt. In Kapitel 5 wird eine Implementierung mit bilinearen Elementen beschrieben. Dabei wird auch auf die Steuerung der verschiedenen Iterationen eingegangen. Das Kaptiel 6 ist eine Beschreibung der algebraischen Mehrgitterverfahren, die innerhalb der HQR-Iteration zur Lösung von linearen elliptischen Problemen zum Einsatz kommen. Zum Abschluss finden sich in Kapitel 7 einige numerische Beispiele.

Nur im Abschnitt 3.5 und bei der Implementierung in Kapitel 5 wird mit uniformen Gittern gearbeitet. Die Aussagen und Ergebnisse der anderen Kapitel bzw. Abschnitte gelten für eine beliebige Vergitterung. Insbesondere der schnelle AMG-Löser benötigt keine uniforme Vergitterung.

2 Problemstellung

2.1 Das diskrete Mumford-Shah-Funktional

Es sei $\Omega \subset \mathbb{R}^2$ ein durch einen Lipschitz-Rand beschränktes Definitionsgebiet. Zum kontinuierlichen Mumford-Shah-Funktional

$$\mathcal{J}(u, K) := \lambda^2 \int_{\Omega \setminus K} \|\nabla u(x)\|^2 \, dx + \alpha \mathcal{H}^1(K) + \int_\Omega |u(x) - g(x)|^2 \, dx, \tag{2.1}$$

wie es in Abschnitt 1.1 vorgestellt wurde, wird in dieser Diplomarbeit die Minimumsuche für das diskretes Mumford-Shah-Funktional

$$J^{(2)}(u, v) := \lambda^2 \sum_{t \in \mathcal{T}} v_t u^* A_t u + \alpha \sum_{t \in \mathcal{T}} (1 - v_t) h_t + (u - g)^* M (u - g) \tag{2.2}$$

behandelt. Dabei ist $\mathcal{T}$ eine Vergitterung von Ω mit Finiten Elementen. Es seien $\Phi_1, \ldots, \Phi_d$ die zugehörigen Basisfunktionen. Die Vektoren u und g aus dem $\mathbb{R}^d$ sind die Koordinatendarstellungen der jeweiligen Funktionen bzgl. der Basis $\Phi_1, \ldots, \Phi_d$.

Die Kantenmenge wird durch einen Vektor $v \in \mathbb{R}^{|\mathcal{T}|}$ modelliert. Für jedes Element $t \in \mathcal{T}$ der Vergitterung ist $v_t \in \{0, 1\}$ ein Flag, welches angibt, ob durch dieses Element eine Kante verläuft ($v_t = 0$) oder nicht ($v_t = 1$). In Abbildung 2.1 wird dies an einer Skizze verdeutlicht. Geman und Geman haben dafür in [GG84] den Namen „Line-Process“ eingeführt.

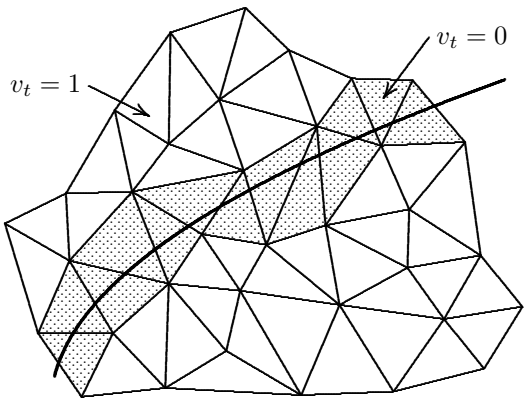

Abbildung 2.1: Modellierung der Kantenmenge durch Flags (Line-Process)

Die Massenmatrix $M \in \mathbb{R}^{d \times d}$ ist symmetrisch und positiv definit und die lokalen Steifigkeitsmatrizen $A_t \in \mathbb{R}^{d \times d}$ sind symmetrisch und positiv semidefinit für alle $t \in \mathcal{T}$. Mit h_t wird für jedes $t \in \mathcal{T}$ die „Länge“ des Elements t bezeichnet, wenn es zu einer Kante beiträgt. λ und α sind Parameter zur Gewichtung der einzelnen Summanden.

Anmerkung
Die (2) *im Superscript in (2.2) deutet an, dass das Funktional von zwei Variablen (u und v) abhängt. Im Laufe der Arbeit wird* $J^{(2)}$ *umgeformt und z. B. in Kapitel 3 die Variable v*

eliminiert. Zur besseren Unterscheidbarkeit trägt dann das dortige Funktional eine (1) *im Superscript.*

Ferner werden noch die Abkürzungen

$$A := \sum_{t \in \mathcal{T}} A_t \tag{2.3}$$

und

$$A_v := \sum_{t \in \mathcal{T}} v_t A_t \tag{2.4}$$

benötigt.

2.2 Interpretation der Parameter

Für die Parameter α und λ in (2.1) bzw. in (2.2) hat man zunächst keine Anhaltspunkte, wie man sie wählen soll. Die Bedeutung dieser Parameter wurde von Blake und Zisserman in [BZ87] untersucht, indem für verschiedene Testdatensätze g (einfacher Sprung, doppelter Sprung, Rampe, usw.) die Funktionalwerte für ein Bild u_1 ohne Kanten und für ein Bild u_2 mit Kanten an den Sprungstellen verglichen wurden.

Für λ gibt es zwei Interpretationen. Zunächst handelt es sich dabei um eine charakteristische Länge, die die Größe des Bereichs beschreibt, in dem geglättet wird, sofern dort keine Kanten liegen. Andererseits ist λ auch eine charakteristische Distanz, die angibt, ab wann zwei isolierte Einzelkanten zu einer Doppelkante verschmelzen.

Die Untersuchung eines isolierten Sprungs hat weiter ergeben, dass $h_0 := \sqrt{2\alpha/\lambda}$ die Kontrastschwelle ist, ab der eine isolierte Kante detektiert wird. Deshalb wird h_0 auch „sensitivity“ genannt. Zwei Sprünge mit Abstand a ($a \ll \lambda$) werden als Doppelkante erkannt, wenn der Kontrastsprung größer als $h_0\sqrt{\lambda/a} = \sqrt{2\alpha/a}$ ist. In der Abbildung 2.2 sind die Ergebnisse von Blake und Zisserman für Sprünge zusammengefasst.

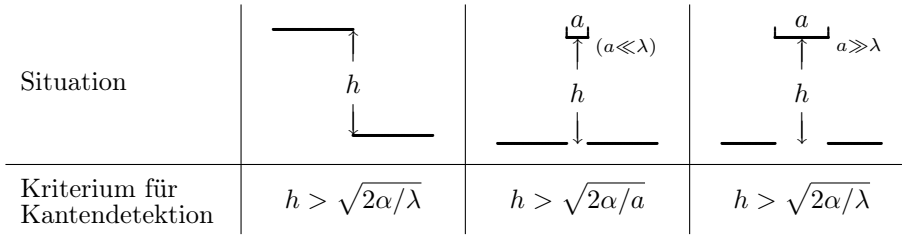

Abbildung 2.2: Interpretation der Parameter λ und α

Die Auswirkungen bei der Wahl von verschiedenen (λ, h_0)-Paaren wird in Abbildung 2.3 dargestellt. In Abbildung 2.4 sind die jeweils gefundenen Kanten zu sehen. Die Bezugslängen sind immer Pixel.

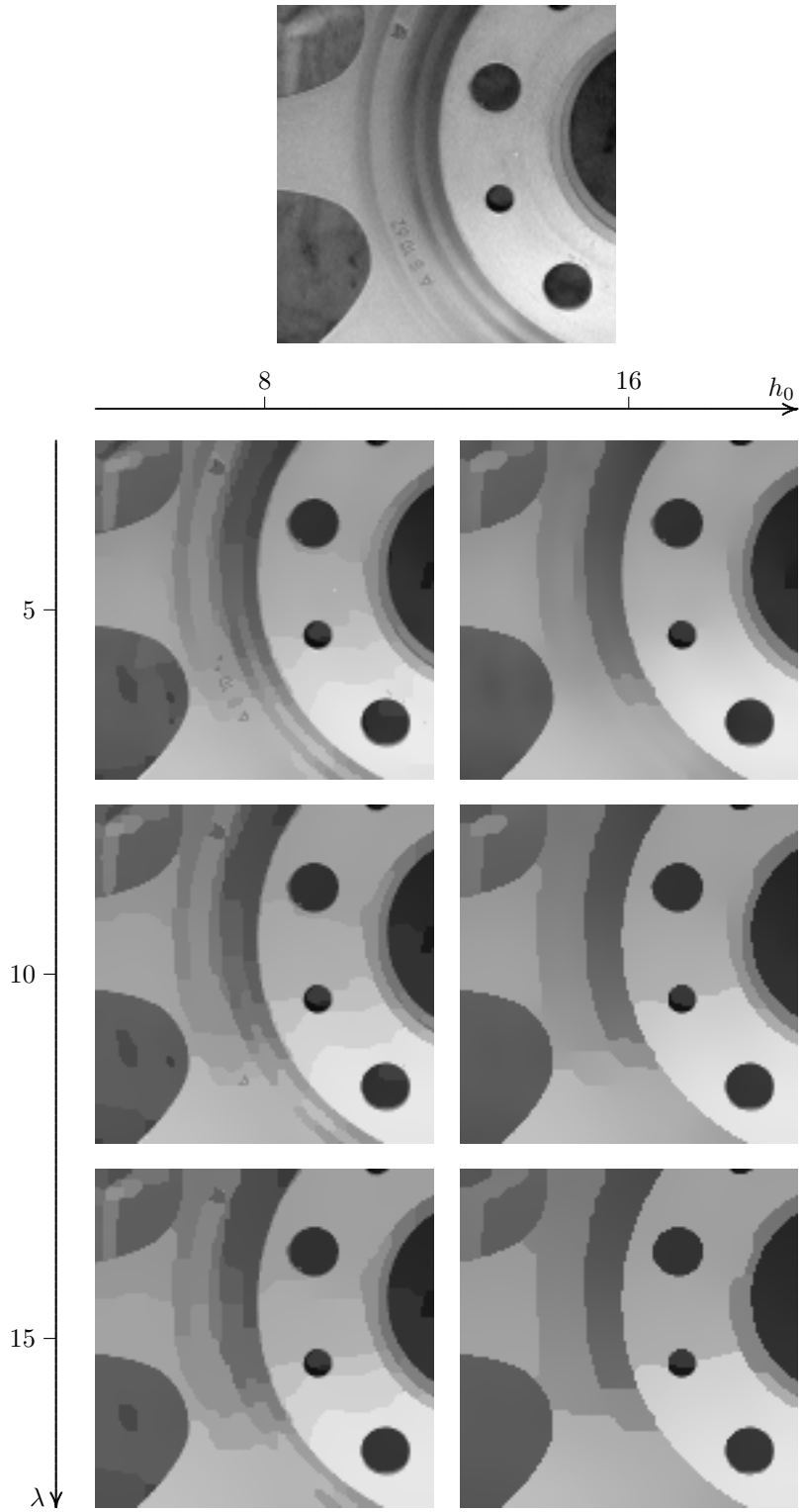

Abbildung 2.3: Testbeispiel: Autofelge

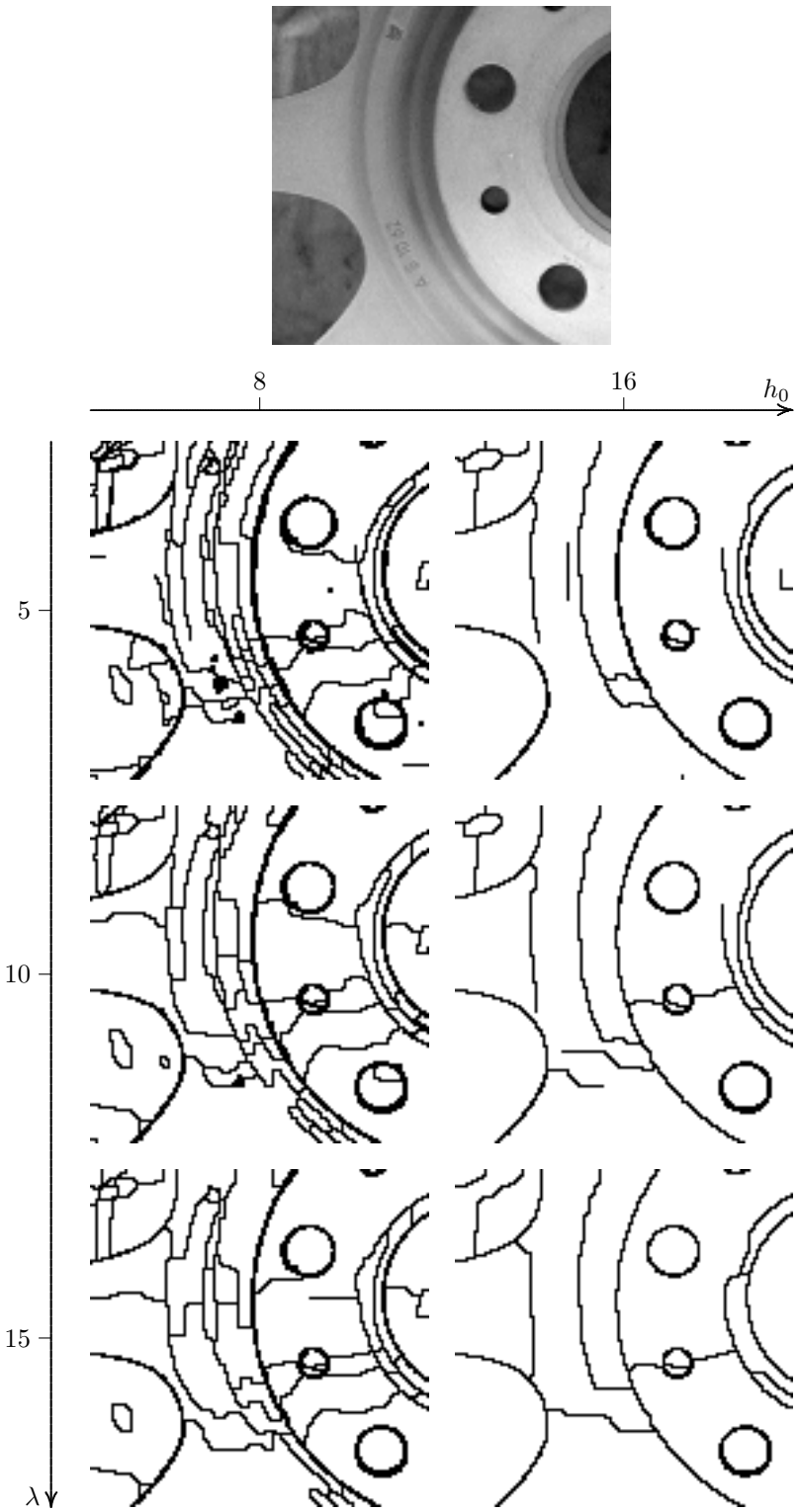

Abbildung 2.4: Testbeispiel: Autofelge (gefundene Kanten)

3 Graduated Non-Convexity (GNC)

Blake und Zisserman haben in ihrem Buch [BZ87] ein **g**raduated **n**on-**c**onvexity-Verfahren (GNC-Verfahren) zur Bildsegmentierung vorgestellt. Die dort beschriebene Idee zur Suche eines Minimums des diskreten, nicht konvexen Mumford-Shah-Funktionals besteht darin, eine Homotopie von einem konvexen zu diesem nicht konvexen Funktional zu verwenden, um so das Minimum des konvexen Funktionals „entlang“ der Homotopie zu verfolgen.

Dazu wird zunächst in Abschnitt 3.1 das hier behandelte Problem umgeformt, indem der Line-Process eliminiert wird. In 3.2 wird dann gezeigt, wie sich dieses umgeformte Problem durch eine ganze Familie anderer Probleme annähern lässt. Im Anschluss wird in Abschnitt 3.3 genauer untersucht, wie eine GNC-Homotopie konstruiert werden kann. Eine hinreichende Bedingung, wann ein solches GNC-Verfahren ein globales Minimum findet, wird in 3.4 gegeben. Im Abschnitt 3.5 wird für uniforme Vergitterungen gezeigt, dass durch jedes Funktional der konstruierten Homotopie bei immer feinerer Vergitterung im Sinne der Γ-Konvergenz „anisotrope“ Mumford-Shah-Funktionale angenähert werden. Für verschiedene Vergitterungen werden die Anisotropien betrachtet.

3.1 Elimination des Line-Process

Da im Funktional

$$\begin{aligned} J^{(2)}(u,v) &= \lambda^2 \sum_{t\in\mathcal{T}} v_t u^* A_t u + \alpha \sum_{t\in\mathcal{T}} (1-v_t)h_t + (u-g)^* M (u-g) \\ &= \sum_{t\in\mathcal{T}} \big[\lambda^2 v_t u^* A_t u + \alpha(1-v_t)h_t\big] + \underbrace{(u-g)^* M (u-g)}_{\geq 0} = \min_{u,v}! \end{aligned}$$

alle Summanden nicht negativ sind und $v_t \in \{0,1\}$ für alle $t \in \mathcal{T}$, kann die Minimierung über v durch

$$\min_v \sum_{t\in\mathcal{T}} \big[\lambda^2 v_t u^* A_t u + \alpha(1-v_t)h_t\big] = \sum_{t\in\mathcal{T}} \min_{v_t\in\{0,1\}} h_t \big[v_t \lambda^2 h_t^{-1} u^* A_t u + \alpha(1-v_t)\big] =$$

$$= \sum_{t\in\mathcal{T}} h_t \min(\alpha, \lambda^2 h_t^{-1} u^* A_t u) = \sum_{t\in\mathcal{T}} h_t f_0(\lambda^2 h_t^{-1} u^* A_t u)$$

eliminiert werden. Dabei ist f_0 durch

$$f_0(x) := \min(\alpha, x) \tag{3.1}$$

definiert (vgl. Abbildung 3.1).
Setzt man nun

$$J_0^{(1)}(u) := \sum_{t\in\mathcal{T}} h_t f_0(h_t^{-1}\lambda^2 u^* A_t u) + (u-g)^* M (u-g), \tag{3.2}$$

so kann man anstatt $J^{(2)}$ auch $J_0^{(1)}$ minimieren. Ist ein u gegeben, so ist jederzeit mittels

$$v_t := f_0'(h_t^{-1}\lambda^2 u^* A_t u) \qquad \text{für alle } t \in \mathcal{T} \tag{3.3}$$

ein zu u bestmögliches v konstruierbar. Dabei kann $f_0'(\alpha) \in \{0,1\}$ beliebig gewählt werden.

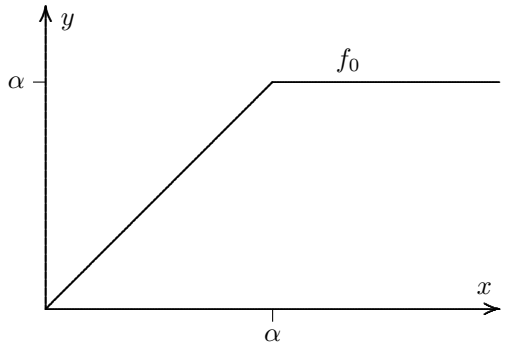

Abbildung 3.1: Graph von f_0: $y = f_0(x)$

3.2 Fuzzy-Kanten

Die durch (3.1) definierte Funktion f_0 bestimmt in (3.2) in Abhängigkeit vom Verhalten des Gradiententerms, ab wann ein $t \in \mathcal{T}$ als Kantenbestandteil gesehen wird. Dieser Übergang, der bei f_0 im Punkt $(\alpha, f_0(\alpha))$ stattfindet, kann auf einen ganzen Bereich erweitert werden. Dazu wird f_0 durch andere Funktionen f ersetzt, die die wesentlichen Eigenschaften von f_0 haben. Ein solches f sollte folgende Eigenschaften besitzen

1. $f \in \mathcal{C}^1\big([0;\infty[\big)$ und konkav,
2. $f(x) \sim x$ für $x \to 0$,
3. $f(x) \sim \alpha$ für $x \to \infty$ und
4. $f(x) \leq f_0(x)$ für alle $x \geq 0$.

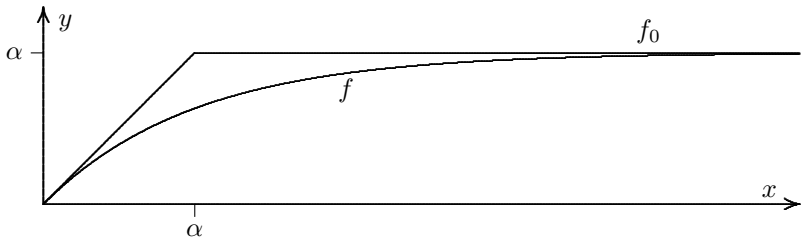

Abbildung 3.2: Beispiel einer Ersatzfunktion f für f_0

Setzt man nun analog zu (3.3) wieder

$$v_t := f'(h_t^{-1}\lambda^2 u^* A_t u) \qquad \text{für alle } t \in \mathcal{T}, \tag{3.4}$$

so ist dann $v_t \in [0;1]$ und es entstehen damit „fuzzy-Kanten“. Ein solches v heißt dann auch generalized Line-Process.
Welche Vorteile solche verallgemeinerten Kanten mit sich bringen, zeigt der nächste Abschnitt.

3.3 Homotopie: vom konvexen zum stark nicht konvexen Funktional

Die in Abschnitt 3.2 beschriebenen Freiheiten bei der Wahl von Ersatzfunktionen f für f_0 verwendet man, um eine stetige Homotopie $\tau \to f_\tau$ von Funktionen zu konstruieren, die folgende Eigenschaften hat:

1. Es gibt ein $\tau_0 > 0$, so dass f_{τ_0} konvex ist.
2. Es gilt: $\|f_\tau - f_0\|_\infty \to 0$ für $\tau \to 0$.

Nun bietet sich folgendes heuristische Verfahren an. Man sucht beim Funktional

$$J_\tau^{(1)}(u) := \sum_{t \in \mathcal{T}} h_t f_\tau(\lambda^2 h_t^{-1} u^* A_t u) + (u-g)^* M (u-g), \tag{3.5}$$

für $\tau = \tau_0$ das globale Minimum, lässt dann $\tau \to 0$ laufen und versucht, das Minimum zu „verfolgen“ (Minimum Tracking).

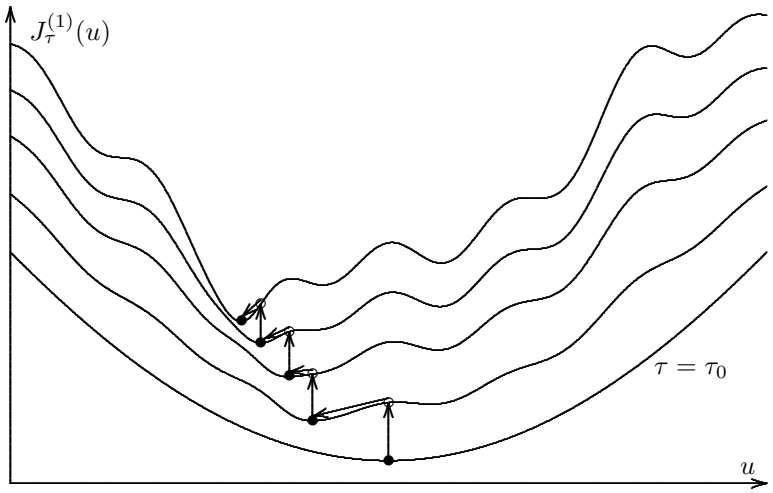

Abbildung 3.3: Beispiel für Minimum Tracking

Da die Homotopie $\tau \to f_\tau$ bei einem konvexen Funktional f_{τ_0} startet und beim stark nicht konvexen Funktional f_0 endet, heißt dieses Verfahren auch **g**raduated **n**on-**c**onvexity-Verfahren (GNC-Verfahren).
Für die Konstruktion der GNC-Homotopie ist folgender Satz hilfreich.

Satz 3.1 (Bornemann, 2000)
Erfüllt ein $f_\tau \in \mathcal{C}^2$ zusätzlich zu den Eigenschaften 1 bis 4 aus Abschnitt 3.2 die Differenzialungleichung

$$2x f_\tau''(x) + f_\tau'(x) \geq -\frac{1}{\tau}, \tag{3.6}$$

dann ist $J_\tau^{(1)}$ konvex für $\tau > \lambda^2 \varrho(M^{-1}A)$.

Beweis

Bei der Differenziation von

$$J_\tau^{(1)}(u) = \underbrace{\sum_{t\in\mathcal{T}} h_t f_\tau(\lambda^2 h_t^{-1} u^* A_t u)}_{=:E_\tau(u)} + \underbrace{(u-g)^* M (u-g)}_{=:D(u)}$$

ergibt sich

$$\nabla E_\tau(u) = \sum_{t\in\mathcal{T}} 2\lambda^2 f_\tau'(\lambda^2 h_t^{-1} u^* A_t u) A_t u$$

und

$$\operatorname{Hess} E_\tau(u) = \sum_{t\in\mathcal{T}} 4\lambda^4 h_t^{-1} f_\tau''(\lambda^2 h_t^{-1} u^* A_t u) A_t u (A_t u)^* + \sum_{t\in\mathcal{T}} 2\lambda^2 f_\tau'(\lambda^2 h_t^{-1} u^* A_t u) A_t.$$

Da A_t symmetrisch positiv semidefinit für alle $t \in \mathcal{T}$ ist, gibt es W_t, so dass $A_t = W_t^* W_t$. Damit gilt

$$v^*(A_t u)(A_t u)^* v = (v^* A_t u)^2 = \langle W_t v, W_t u\rangle^2 \le \langle W_t v, W_t v\rangle \langle W_t u, W_t u\rangle = v^* A_t v \cdot u^* A_t u.$$

Weil $f_\tau'' \le 0$ und $f_\tau' \ge 0$, gilt die Abschätzung

$$v^*[\operatorname{Hess} E_\tau(u)]v \ge$$

$$\ge \sum_{t\in\mathcal{T}} 4\lambda^4 h_t^{-1} f_\tau''(\lambda^2 h_t^{-1} u^* A_t u) u^* A_t u \cdot v^* A_t v + \sum_{t\in\mathcal{T}} 2\lambda^2 f_\tau'(\lambda^2 h_t^{-1} u^* A_t u) v^* A_t v =$$

$$= 2\lambda^2 \sum_{t\in\mathcal{T}} \big\{ 2(\lambda^2 h_t^{-1} u^* A_t u) f_\tau''(\lambda^2 h_t^{-1} u^* A_t u) + f_\tau'(\lambda^2 h_t^{-1} u^* A_t u) \big\} v^* A_t v \overset{(3.6)}{\ge}$$

$$= 2\lambda^2 \sum_{t\in\mathcal{T}} \Big(-\frac{1}{\tau} v^* A_t v \Big) = -\frac{2\lambda^2}{\tau} v^* A v$$

mit $A = \sum_{t\in\mathcal{T}} A_t$. Daher ergibt sich nun mit $\operatorname{Hess} D(u) = 2M$

$$v^*[\operatorname{Hess} J_\tau^{(1)}(u)]v \ge -\frac{2\lambda^2}{\tau} v^* A v + 2 v^* M v.$$

Damit wird nun $J_\tau^{(1)}$ konvex, falls für alle $v \in \mathbb{R}^N \backslash \{0\}$

$$v^* A v < \frac{\tau}{\lambda^2} v^* M v \qquad \Leftrightarrow \qquad \varrho(M^{-1}A) < \frac{\tau}{\lambda^2} \qquad \Leftrightarrow \qquad \tau > \lambda^2 \varrho(M^{-1}A)$$

gilt. □

Die Existenz einer solchen Homotopie mit allen Eigenschaften aus Satz 3.1 wird im Abschnitt 5.3 gezeigt; dort werden die f_τ explizit konstruiert.

3.4 Stetige Abhängigkeit des Minimums vom Homotopieparameter

Wenn keine weiteren Voraussetzungen gemacht werden, ist natürlich nicht sichergestellt, dass das in Abschnitt 3.3 vorgestellte GNC-Verfahren stets ein globales Minimum von $J_0^{(1)}$ findet. Folgender Satz gibt hinreichende Bendingungen an, die genau dies sicherstellen.

Satz 3.2 (Bornemann, 2000)
Ist $\Omega \in \mathbb{R}^d$ kompakt, die Abbildung $[0;1] \to \big(\mathcal{C}(\Omega,\mathbb{R}), \|\cdot\|_\infty\big)$, $\tau \mapsto J_\tau$ stetig und gibt es für jedes $\tau \in [0;1]$ genau ein $u(\tau) \in \Omega$ mit $J_\tau(u(\tau)) = \min\{J_\tau(v) \,|\, v \in \Omega\}$, dann ist $[0;1] \to \Omega$, $\tau \mapsto u(\tau)$ stetig.

Beweis

Es seien die Voraussetzungen des Satzes erfüllt. Angenommen, $\tau \mapsto u(\tau)$ ist bei $\tau_0 \in [0;1]$ unstetig. Dann gibt es eine konvergente Folge $\big(\tau_n\big)_{n\in\mathbb{N}}$ mit $\tau_n \to \tau_0$ für $n \to \infty$, so dass $\big(u(\tau_n)\big)_{n\in\mathbb{N}}$ nicht gegen $u(\tau_0)$ konvergiert.
Da $u(\tau_n) \in \Omega$ für alle $n \in \mathbb{N}$ und Ω kompakt ist, besitzt $\big(u(\tau_n)\big)_{n\in\mathbb{N}}$ eine konvergente Teilfolge; o. E. ist $\big(u(\tau_n)\big)_{n\in\mathbb{N}}$ selbst konvergent mit

$$\Omega \ni u_* := \lim_{n\to\infty} u(\tau_n) \neq u(\tau_0).$$

Da J_{τ_0} das eindeutige Minimum $u(\tau_0)$ besitzt, gilt

$$J_{\tau_0}(u_*) > J_{\tau_0}(u(\tau_0)). \qquad (*)$$

Weil jedes $u(\tau_n)$ das eindeutige Minimum von J_{τ_n} für alle $n \in \mathbb{N}$ ist, gilt auch

$$J_{\tau_n}(u(\tau_n)) < J_{\tau_n}(u(\tau_0)) \qquad \text{für alle } n \in \mathbb{N}.$$

Der Grenzübergang $n \to \infty$ auf beiden Seiten liefert mit

$$|J_{\tau_n}(u(\tau_0)) - J_{\tau_0}(u(\tau_0))| \leq \|J_{\tau_n} - J_{\tau_0}\|_\infty \to 0 \quad \text{für } n \to \infty$$

und

$$|J_{\tau_n}(u(\tau_n)) - J_{\tau_0}(u_*)| \leq \|J_{\tau_n} - J_{\tau_0}\|_\infty + |J_{\tau_0}(u(\tau_n)) - J_{\tau_0}(u_*)| \to 0 \qquad \text{für } n \to \infty$$

schließlich

$$J_{\tau_0}(u_*) \leq J_{\tau_0}(u(\tau_0)),$$

was einen Widerspruch zu $(*)$ darstellt. □

3.5 Approximation von anisotropen Mumford-Shah-Funktionalen

In diesem Abschnitt wird gezeigt, dass jedes Funktional $J_\tau^{(1)}$ aus der oben besprochenen Homotopie bei uniformer Vergitterung als Γ-Limes eine anisotrope Version

$$\mathcal{J}_A(u) = \lambda^2 \int_\Omega \|\nabla u(x)\|^2 \, dx + \alpha \int_{S_u} \phi(\nu_u) \, d\mathcal{H}^1 + \int_\Omega |u(x) - g(x)|^2 \, dx$$

des Mumford-Shah-Funktionals besitzt. Für verschiedene Vergitterungen werden die entsprechenden ϕ angegeben.
Es werden Gitter untersucht, welche durch endlich viele Einheitsvektoren $e_1, \ldots, e_k \in \mathbb{R}^2$ ($k \geq 2$) aufgebaut werden, wobei $e_i \nparallel e_j$ für alle $i \neq j$ gelte. Ein Gitter wird durch Geradenscharen erzeugt, indem für alle $1 \leq j \leq k$ eine Geradenschar mit parallelen Geraden in Richtung von e_j im Abstand von h_j betrachtet werden (vgl. Abbildung 3.4).

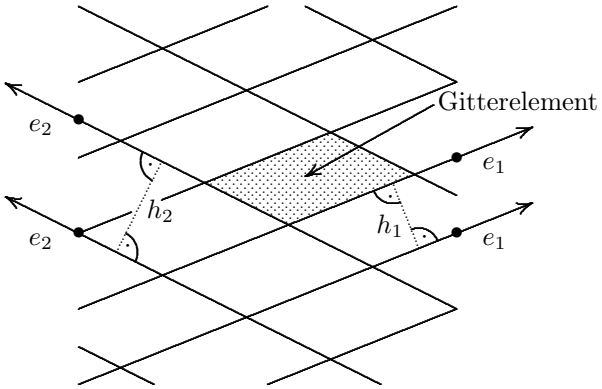

Abbildung 3.4: Uniformes Gitter (erzeugt durch e_1 und e_2)

Eine Strecke S der Länge 1 mit Normale ν ($\|\nu\| = 1$) schneidet die zum Vektor e_j gehörenden parallelen Geraden in asymptotisch

$$\frac{|\langle e_j, \nu\rangle|}{h_j}$$

vielen Punkten, denn in der Abbildung 3.5 gilt $x = \cos\alpha = \langle e_j, \nu\rangle$ und die asymptotische Anzahl der Schnittpunkte ist x/h_j.

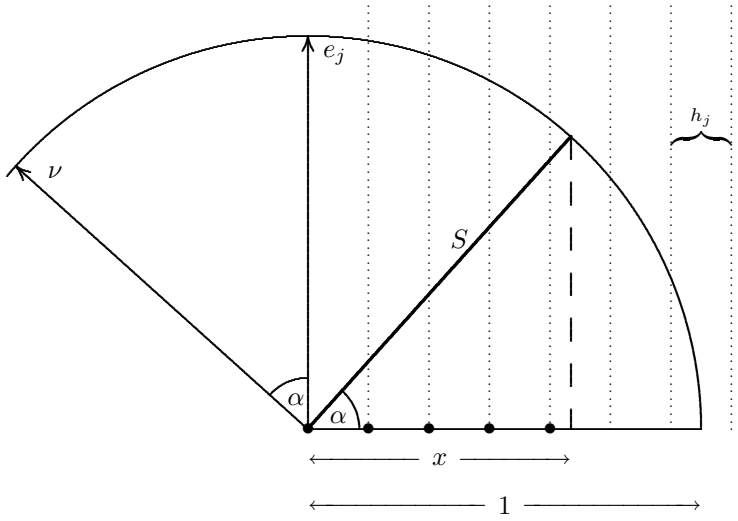

Abbildung 3.5: Asymptotische Anzahl der Schnittpunkte der Strecke S

Damit werden insgesamt asymptotisch

$$\sum_{j=1}^{k} \frac{|\langle e_j, \nu\rangle|}{h_j}$$

viele Gitterelemente von der Strecke S geschnitten. Jedes Element habe die Fläche F und h bezeichne den Diskretisierungsparameter, der durch

$$F = \rho h^2 \qquad \text{und} \qquad h_j = \sigma_j h$$

mittels ρ und σ_j mit der Fläche F und den Abständen h_j verknüpft sei. Wird nun jedem Gitterelement die Länge F/h zugeordnet, so erhält man die Längenfunktion

$$\varphi(\nu) = \lim_{h\to 0} \frac{F}{h} \sum_{j=1}^{k} \frac{|\langle e_j, \nu\rangle|}{h_j} = \rho \sum_{j=1}^{k} \frac{|\langle e_j, \nu\rangle|}{\sigma_i},$$

die die (für $h \to 0$ asymptotische) im Gitter gemessene Länge einer Strecke der Länge 1 mit Normale ν angibt. Die Funktion φ ist, auf dem $\mathbb{R}^2$ betrachtet, eine Norm, die mit der Matrix

$$\Phi = \rho \begin{pmatrix} \sigma_1^{-1} e_1^* \\ \vdots \\ \sigma_k^{-1} e_k^* \end{pmatrix} \in \mathbb{R}^{k\times 2}$$

alternativ auch durch

$$\varphi(v) = \|\Phi v\|_1, \qquad v \in \mathbb{R}^2$$

dargestellt wird.
Diese Überlegungen erlauben systematische Untersuchungen der von uniformen Gittern erzeugten Anisotropien ohne Fallunterscheidungen, wie sie in [Neg99] vorkommen.
Es werden nun die Auswirkungen von verschiedenen Vergitterungen (vgl. Abbildung 3.6) im Einzelnen betrachtet. Dazu sei

$$M := \max_{\|v\|_2=1} \varphi(v) \qquad \text{bzw.} \qquad m := \min_{\|v\|_2=1} \varphi(v).$$

Je größer das Aspektverhältnis

$$a := \frac{M}{m}$$

ist, desto größer ist die Anisotropie. Da Φ vollen Rang hat, gilt

$$M = \max_{\|v\|_2=1} \|\Phi v\|_1 = \max_{v\neq 0} \frac{\|\Phi v\|_1}{\|v\|_2} = \frac{1}{\min\limits_{\Phi v\neq 0} \frac{\|v\|_2}{\|\Phi v\|_1}} = \frac{1}{\min\limits_{\|\Phi v\|_1=1} \|v\|_2} \qquad \text{und}$$

$$m = \min_{\|v\|_2=1} \|\Phi v\|_1 = \min_{v\neq 0} \frac{\|\Phi v\|_1}{\|v\|_2} = \frac{1}{\max\limits_{\Phi v\neq 0} \frac{\|v\|_2}{\|\Phi v\|_1}} = \frac{1}{\max\limits_{\|\Phi v\|_1=1} \|v\|_2},$$

weshalb in den nächsten Unterabschnitten stets die Menge $\{v \in \mathbb{R}^2 \,|\, \varphi(v) = 1\}$ zusammen mit dem euklidschen Einheitskreis dargestellt wird. Dabei ist h stets die maximale Kantenlänge eines Gitterelements, welche mit $h = 1$ normiert wird.

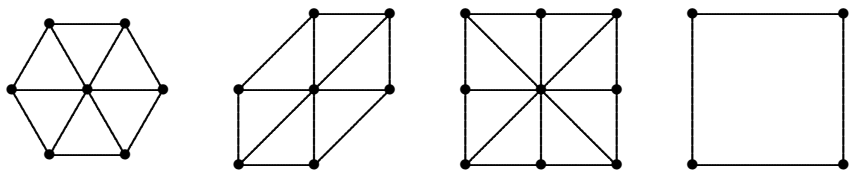

Abbildung 3.6: Betrachtete uniforme Vergitterungen

3.5.1 Hexagonale Triangulierung

Bei der hexagonalen Triangulierung (vgl. Abbildung 3.6) ist

$$e_1 = (1,0)^*,\ e_2 = \left(\frac{1}{2}, \frac{1}{2}\sqrt{3}\right)^* \qquad \text{und} \qquad e_3 = \left(-\frac{1}{2}, \frac{1}{2}\sqrt{3}\right)^*$$

mit

$$\sigma_1 = \sigma_2 = \sigma_3 = \frac{1}{2}\sqrt{3} \quad \text{und} \quad \rho = \frac{1}{4}\sqrt{3}.$$

Die Abbildung 3.7 zeigt die zugehörige Einheitssphäre.

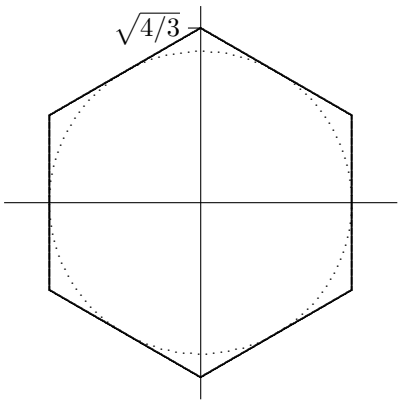

Abbildung 3.7: Einheitssphäre $\varphi(v) = 1$ (hexagonale Triangulierung)

Daraus ist

$$M = 1, \qquad m = \frac{1}{2}\sqrt{3} \qquad \text{und damit} \qquad a \approx 1{,}15$$

ersichtlich.

3.5.2 Vergitterung mit Courant-Element

Beim Courant-Element (vgl. Abbildung 3.6) ist

$$e_1 = (1,0)^*,\ e_2 = (0,1)^* \qquad \text{und} \qquad e_3 = \left(\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}\right)^*$$

mit

$$\sigma_1 = \sigma_2 = \frac{1}{\sqrt{2}},\ \sigma_3 = \frac{1}{2} \quad \text{und} \quad \rho = \frac{1}{4}.$$

Abbildung 3.8 zeigt wieder die Einheitssphäre.
Hier ist nun

$$M = 1, \qquad m = \frac{1}{2} \qquad \text{und damit} \qquad a = 2.$$

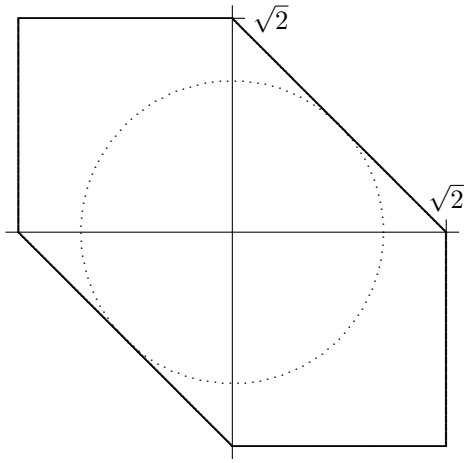

Abbildung 3.8: Einheitssphäre $\varphi(v) = 1$ (Courant-Element)

3.5.3 Sternförmige Vergitterung

Bei dieser Vergitterung (vgl. Abbildung 3.6) ist

$$e_1 = (1,0)^*,\ e_2 = (0,1)^*,\ e_3 = \left(\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}\right)^* \qquad \text{und} \qquad e_4 = \left(\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2}\right)^*$$

mit

$$\sigma_1 = \sigma_2 = \frac{1}{2}\sqrt{2},\ \sigma_3 = \sigma_4 = 1 \quad \text{und} \quad \rho = \frac{1}{4}.$$

In Abbildung 3.9 ist wieder die Einheitssphäre zu sehen.

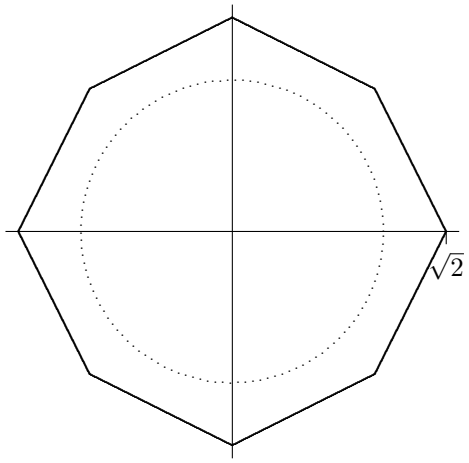

Abbildung 3.9: Einheitskugel $\varphi(v) = 1$ (Sternförmige Vergitterung)

Mit

$$M = \frac{1}{4}\sqrt{10}, \qquad m = \frac{1}{2}\sqrt{2}, \qquad \text{ergibt sich} \qquad a \approx 1{,}12.$$

3.5.4 Quadratische Vergitterung

Hier ist (vgl. Abbildung 3.6)

$$e_1 = (1,0)^*, \qquad \text{und} \qquad e_2 = (0,1)^*$$

mit

$$\sigma_1 = \sigma_2 = 1 \quad \text{und} \quad \rho = 1.$$

In Abbildung 3.10 ist wieder die Einheitssphäre zu sehen.

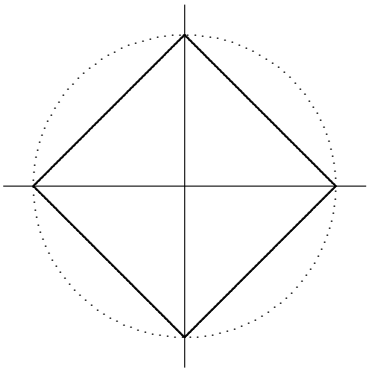

Abbildung 3.10: Einheitskugel $\varphi(v) = 1$ (quadratisches Element)

Hier ist abzulesen, dass

$$M = \sqrt{2}, \qquad m = 1 \qquad \text{und damit} \qquad a \approx 1{,}41.$$

3.5.5 Folgerungen

Nach [Neg99] führt das Modell

$$\sum_{t\in\mathcal{T}} h_t f_\tau(h_t^{-1}\lambda^2 u^* A_t u) = \sum_{t\in\mathcal{T}} \frac{F}{h} f_\tau\left(\frac{h}{F}\lambda^2 u^* A_t u\right)$$

im Γ-Limes auf den Längenterm

$$\alpha \int_{S_u} \varphi(v)\, d\mathcal{H}^1,$$

der sich mit Hilfe von m und M durch

$$\alpha m \mathcal{H}^1(S_u) \le \alpha \int_{S_u} \varphi(v)\, d\mathcal{H}^1 \le \alpha M \mathcal{H}^1(S_u)$$

abschätzen lässt. Will man beim Mumford-Shah-Funktional einen symmetrischen Fehler haben, so sollte

$$\hat{\alpha} = \frac{\alpha}{\sqrt{Mm}}$$

verwendet werden, was zur Abschätzung

$$\alpha\sqrt{\frac{m}{M}}\mathcal{H}^1(S_u) \le \hat{\alpha} \int_{S_u} \varphi(v)\, d\mathcal{H}^1 \le \alpha\sqrt{\frac{M}{m}}\mathcal{H}^1(S_u)$$

führt.
Damit ergeben sich für die oben besprochenen Vergitterungen folgende symmetrische Längenabweichungen:

Vergitterung	$\sqrt{M/m}$	Fehler
hexagonal	1,075	$\pm$ 7,5 %
Courant	1,414	$\pm$41,4 %
sternförmig	1,057	$\pm$ 5,7 %
quadratisch	1,189	$\pm$18,9 %

Das wichtige Ergebnis dieses Abschnitts ist, dass alle Funktionale $J_\tau^{(1)}$ ($\tau > 0$) im Grenzfall ($h \to 0$) denselben Γ-Limes, nämlich eine anisotrope Version des Mumford-Shah-Funktionals, besitzen. Damit approximieren alle Funktionale der Homotopie dieselbe kontinuierliche Situation. Falls das kontinuierliche Problem eine eindeutige Lösung besitzt, so geben obige Ergebnisse Anlass zu hoffen, dass, bei geeigneter Diskretisierung, diese Eindeutigkeit ins Diskrete vererbt werden kann und dadurch die Lösung mittels der Homotopie schnell gefunden werden kann.

4 Half-Quadratic Regularization (HQR)

Die in Abschnitt 3.3 in Gleichung (3.5) betrachteten Funktionale $J_\tau^{(1)}$ der GNC-Homotopie sind für kleine τ i. Allg. nicht konvex und besitzen viele lokale Minima. Um diese Funktionale leichter minimieren zu können, werden sie regularisiert. Dazu wird in Abschnitt 4.1 die Legendre-Fenchel-Transformation eingeführt und im Abschnitt 4.2 auf die Funktionen f_τ angewendet. Dadurch entstehen neue Funktionale J_τ^* in zwei Variablen. Im Abschnitt 4.3 wird für die Minimierung der J_τ^* eine Iteration angegeben, deren Konvergenz in Abschnitt 4.4 untersucht wird.

4.1 Legendre-Fenchel-Transformation

Definition 4.1
Es sei X ein normierter Raum und $f\colon X \to \mathbb{R} \cup \{-\infty, +\infty\}$ eine Funktion auf X. Als Legendre-Fenchel-Transformierte oder auch Young-Fenchel-Transformierte der Funktion f bezeichnet man die durch

$$f^*(x') = \sup_{x \in X} \big(x'(x) - f(x)\big)$$

auf X' definierte Funktion. f^* wird auch als die zu f konjugierte Funktion bezeichnet.

Wie man sofort erkennt, gilt für $x'_1, x'_2 \in X'$ und $0 \le \lambda \le 1$

$$\begin{aligned} f^*\big(\lambda x'_1 + (1-\lambda)x'_2\big) &= \sup_{x \in X} \big(\lambda x'_1(x) + (1-\lambda)x'_2(x) - f(x)\big) \le \\ &\le \lambda \sup_{x \in X} \big(x'_1(x) - f(x)\big) + (1-\lambda) \sup_{x \in X} \big(x'_2(x) - f(x)\big) = \\ &= \lambda f^*(x'_1) + (1-\lambda) f^*(x'_2), \end{aligned}$$

womit also f^* konvex ist.
X sei ab jetzt immer ein Hilbertraum. Man sucht zu gegebenem $x^* \in X$ das kleinste η, so dass die Hyperebene $\langle x^*, x\rangle - \eta$ niemals oberhalb von f liegt, also

$$\langle x^*, x\rangle - \eta \le f(x) \qquad \text{für alle } x \in X$$

gilt. Dieses η wird gerade durch

$$\eta = \sup_{x \in X} \big(\langle x^*, x\rangle - f(x)\big) = f^*(x^*)$$

festgelegt.
Die Hyperebene $\{(x, y) \,|\, y = \langle x^*, x\rangle - f^*(x^*)\}$ ist also Stützhyperebene an epi f.

Aus der Definition unmittelbar ersichtlich ist, dass $f^{**} \le f$.
Der nachfolgende Satz gibt Aufschluss, wann $f^{**} = f$ gilt.

Satz 4.2 (Fenchel-Moreau)
Für eine Funktion $f\colon X \to \mathbb{R} \cup \{+\infty\}$ gilt $f^{**} = f$ genau dann, wenn f konvex und abgeschlossen ist.

Einen Beweis findet man z. B. in [Roc97] oder in [ET76].

Im Folgenden wird die Legendre-Fenchel-Transformation nur für den Fall $X = \mathbb{R}$ benötigt.

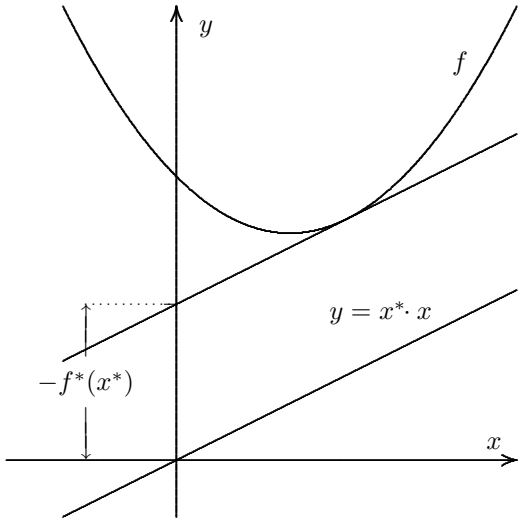

Abbildung 4.1: Geometrische Veranschaulichung der Legendre-Fenchel-Transformation

4.2 Anwendung auf das zu minimierende Funktional

Die Anwendung der Legendre-Fenchel-Transformation auf eine Funktion f mit den Eigenschaften 1 bis 4 aus Abschnitt 3.2 macht es möglich, f mit Hilfe der konjugierten Funktion darzustellen:

Satz 4.3
Sei $f\colon [0;\infty[\to \mathbb{R}$ eine Funktion mit den Eigenschaften 1 bis 4 aus Abschnitt 3.2, dann gilt für $x \geq 0$

$$f(x) = \min_{v\in[0;1]} \big(xv + \Psi(v)\big) \tag{4.1}$$

mit $\Psi(v) = (-f)^*(-v)$. Dabei wird das Minimum bei $v = f'(x)$ angenommen.

Beweis

Die durch

$$g(x) := \begin{cases} +\infty & \text{für } x < 0 \\ -f(x) & \text{für } x \geq 0 \end{cases}$$

definierte Funktion ist konvex und abgeschlossen. Es gilt

$$g^*(v) = \sup_{x\in\mathbb{R}} \big(vx - g(x)\big) = \sup_{x\geq 0} \big(vx - g(x)\big).$$

Für $v > 0$ ist wegen $g(x) \leq 0$ für $x \geq 0$

$$g^*(v) = \sup_{x\geq 0} \big(xv - g(x)\big) = +\infty. \tag{$*$}$$

Für $v \leq -1$ ist

$$g^*(v) = \sup_{x\geq 0} \big(vx - g(x)\big) \leq \sup_{x\geq 0} (vx + x) \leq 0 \qquad \text{und}$$
$$g^*(v) = \sup_{x\geq 0} \big(vx - g(x)\big) \geq v \cdot 0 - g(0) = 0,$$

also

$$g^*(v) = 0 \qquad (**)$$

Mit Satz 4.2 ist $g^{**} = g$ und somit

$$g(x) = g^{**}(x) = \sup_{v\in\mathbb{R}}\big(xv - g^*(v)\big) \overset{(*)}{=} \sup_{v\le 0}\big(xv - g^*(v)\big).$$

Weil für $x \ge 0$

$$\sup_{v<-1}\big(xv - g^*(v)\big) \overset{(**)}{=} \sup_{v<-1}(xv) \le -x \overset{(**)}{=} x\cdot(-1) - g^*(-1),$$

ist nun

$$\begin{aligned} -f(x) = g(x) = g^{**}(x) &= \max_{v\in[-1;0]}\big(xv - g^*(v)\big) = \max_{v\in[0;1]}\big(-xv - g^*(-v)\big) = \\ &= -\min_{v\in[0;1]}\big(xv + \Psi(v)\big), \end{aligned}$$

und f hat die Darstellung

$$f(x) = \min_{v\in[0;1]}\big(xv + \Psi(v)\big).$$

Nun zur Feststellung, wo das Minimum angenommen wird.
Sei dazu $x_0 \ge 0$ fest und $v_0 := g'(x_0)$ (für $x_0 = 0$ ist natürlich der rechtsseitige Grenzwert gemeint). Für g gilt (vgl. auch Abbildung 4.1)

$$g^*(v_0) = x_0v_0 - g(x_0),$$

also

$$\begin{aligned} g^*(v_0) - x_0v_0 &= -g(x_0) \\ g^*(v_0) + x_0(v - v_0) &= x_0v - g(x_0) && \forall v\in\mathbb{R} \\ g^*(v_0) + x_0(v - v_0) &\le \sup_{x\in\mathbb{R}}\big(xv - g(x)\big) = g^*(v) && \forall v\in\mathbb{R} \\ g^*(-v) &\ge g^*(v_0) - x_0(v + v_0) && \forall v\in\mathbb{R} \\ x_0v + g^*(-v) &\ge x_0(-v_0) + g^*\big[-(-v_0)\big] && \forall v\in\mathbb{R}. \end{aligned}$$

Damit wird das Minimum von $x_0v + g^*(-v)$ bei

$$v = -v_0 = -g'(x_0) = f'(x_0)$$

angenommen. □

Zusammenfassend kann man nun feststellen, dass das Problem

$$J^{(2)}(u,v) = \lambda^2\sum_{t\in\mathcal{T}} v_t u^* A_t u + \alpha\sum_{t\in\mathcal{T}} h_t(1-v_t) + (u-g)^*M(u-g) = \min_{u,v}!$$

durch die Elimination des Line-Process in Abschnitt 3.1 umgeformt wurde zu

$$J_0^{(1)}(u) = \sum_{t\in\mathcal{T}} h_t f_0(\lambda^2 h_t^{-1} u^* A_t u) + (u-g)^*M(u-g) = \min_{u}!$$

Nachdem f_0 durch eine Folge f_τ ersetzt wurde, gilt nun mit Satz 4.3

$$J_\tau^{(1)}(u) = \sum_{t\in\mathcal{T}} h_t f_\tau(\lambda^2 h_t^{-1} u^* A_t u) + (u-g)^* M(u-g) = \min_u!$$
$$\Updownarrow (4.1)$$
$$\sum_{t\in\mathcal{T}} h_t \min_{v_t\in[0;1]} \left(\lambda^2 h_t^{-1} u^* A_t u \cdot v_t + \Psi_\tau(v_t)\right) + (u-g)^* M(u-g) = \min_u!$$

Also wird im Folgenden das Minimierungsproblem

$$J_\tau^*(u,v) := \lambda^2 \sum_{t\in\mathcal{T}} v_t u^* A_t u + \sum_{t\in\mathcal{T}} h_t \Psi_\tau(v_t) + (u-g)^* M(u-g) = \min_{u,v}! \tag{4.2}$$

behandelt.
Betrachtet man für ein fest vorgegebenes $v \in [0;1]^{|\mathcal{T}|}$ das Funktional $u \mapsto J_\tau^*(u,v)$, so handelt es sich um ein quadratisches Funktional. Daher wird die Anwendung der Legendre-Fenchel-Transformation in diesem Zusammenhang auch „**H**alf-**Q**uadratic **R**egularization“ (HQR) genannt (vgl. [GY95] oder auch [CBFAB97]).
Betrachtet man dagegen für ein fest vorgegebenes $u \in \mathbb{R}^d$ das Funktional $v \mapsto J_\tau^*(u,v)$, so handelt es sich um ein konvexes Funktional, da Ψ_τ konvex ist.

4.3 HQR-Iteration

Um nun das Funktional (4.2)

$$J_\tau^*(u,v) = \lambda^2 \sum_{t\in\mathcal{T}} v_t u^* A_t u + \sum_{t\in\mathcal{T}} h_t \Psi_\tau(v_t) + (u-g)^* M(u-g)$$

bei einem fest vorgegebenen τ zu minimieren, wird alternierend minimiert, also in folgenden zwei Schritten vorgegangen:

1. u wird festgehalten und bezüglich v wird minimiert.
2. v wird festgehalten und bezüglich u wird minimiert.

Es wird also zu einem vorgegebenen Startwert u^0 die Iteration

$$v^{n+1} = \operatorname*{argmin}_v J_\tau^*(u^n, v) \tag{4.3}$$

und

$$u^{n+1} = \operatorname*{argmin}_u J_\tau^*(u, v^n) \tag{4.4}$$

für $n = 0, 1, 2, \ldots$ ausgeführt. Im nächsten Abschnitt wird das Konvergenzverhalten dieser Iteration untersucht.

4.4 Konvergenzanalyse der HQR-Iteration

Für die Konvergenzuntersuchung ist folgender Satz hilfreich.

Satz 4.4
Ist $(X, \|\cdot\|)$ ein endlichdimensionaler normierter Raum, (x_n) eine Folge in X mit einem isolierten Häufungspunkt $\bar{x} \in X$ und gilt

$$\|x_{n+1} - x_n\| \to 0 \qquad \text{für } n \to \infty, \tag{4.5}$$

dann konvergiert (x_n) gegen $\bar{x}$.

Beweis

Angenommen, (x_n) hat einen weiteren Häufungspunkt $\hat{x} \in X$ mit $\hat{x} \neq \bar{x}$. Da $\bar{x}$ isoliert ist, gibt es ein $\varepsilon > 0$, so dass $\bar{x}$ der einzige Häufungspunkt in $U := \{x \in X \,|\, \|x - \bar{x}\| \leq 3\varepsilon\}$ ist. Ferner sei $R := \{x \in X \,|\, \varepsilon \leq \|x - \bar{x}\| \leq 2\varepsilon\}$ (vgl. Abbildung 4.2).

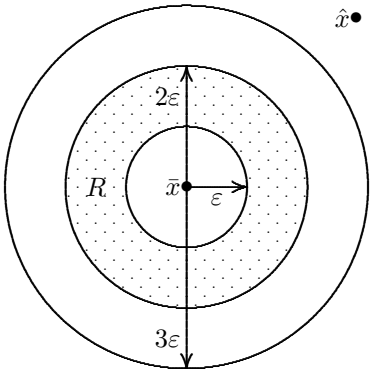

Abbildung 4.2: Skizze zum Beweis von Satz 4.4

Wegen (4.5) gibt es ein $N \in \mathbb{N}$, so dass für alle $n > N$ gilt: $\|x_{n+1} - x_n\| < \varepsilon$.
Da $\bar{x}$ Häufungspunkt ist, gibt es eine Teilfolge $(x_{\bar{n}_1}, x_{\bar{n}_2}, \ldots)$ mit $x_{\bar{n}_k} \to \bar{x}$ für $k \to \infty$. Da auch $\hat{x}$ Häufungspunkt ist, gibt es eine Teilfolge $(x_{\hat{n}_1}, x_{\hat{n}_2}, \ldots)$ mit $x_{\hat{n}_k} \to \hat{x}$ für $k \to \infty$.
Dabei gelte o. E. $N < \bar{n}_k < \hat{n}_k$, $\|x_{\bar{n}_k} - \bar{x}\| < \varepsilon$ und $\|x_{\hat{n}_k} - \hat{x}\| < \varepsilon$ für alle $k \in \mathbb{N}$.
Setzt man nun

$$m_k := \max\{n \,|\, \bar{n}_k \leq n \leq \hat{n}_k \quad \text{und} \quad \|x_n - \bar{x}\| < \varepsilon\} \qquad \text{und} \qquad \tilde{n}_k := m_k + 1$$

dann gilt $\bar{n}_k \leq m_k < \hat{n}_k$, da $\|x_{\hat{n}_k} - \bar{x}\| \geq \|\hat{x} - \bar{x}\| - \|x_{\hat{n}_k} - \hat{x}\| \geq 3\varepsilon - \varepsilon = 2\varepsilon$ ist. Somit ist $\bar{n}_k < \tilde{n}_k \leq \hat{n}_k$ und wegen der Wahl von m_k muss gelten

$$\varepsilon \leq \|x_{\tilde{n}_k} - \bar{x}\|. \tag{$*$}$$

Außerdem gilt wegen (4.5)

$$\|x_{\tilde{n}_k} - \bar{x}\| \leq \|x_{\tilde{n}_k} - x_{\tilde{n}_k - 1}\| + \|x_{m_k} - \bar{x}\| \leq \varepsilon + \varepsilon = 2\varepsilon. \tag{$**$}$$

Aus $(*)$ und $(**)$ folgt, dass $x_{\tilde{n}_k} \in R$ für jedes $k \in \mathbb{N}$ ist. Damit ist $(x_{\tilde{n}_k})$ eine Folge im Kompaktum R und besitzt demnach einen Häufungspunkt $\tilde{x} \in R \subset U$. Damit hat (x_n) einen weiteren Häufungspunkt in U, was im Widerspruch zur Wahl von ε steht. □

Zum Beweis der Konvergenz der HQR-Iteration für ein beliebiges, aber festes $\tau > 0$ wird von folgenden Voraussetzungen ausgegangen:

1. Die Funktion f_τ im Funktional $J_\tau^{(1)}$ (vgl. Gleichung (3.5)) erfüllt die Eigenschaften 1 bis 4 aus Abschnitt 3.2.

2. Die zu f_τ gehörende Funktion $\Psi_\tau(v) := (-f_\tau)^*(-v)$ (vgl. Satz 4.3) ist auf dem Intervall $[0;1]$ zweimal stetig diffbar und es gilt $\Psi_\tau''(v) \geq c_\tau$ für alle $v \in [0;1]$ mit einem $c_\tau > 0$.

3. Die stationären Punkte des Funktionals J_τ^* sind isoliert.

Anmerkung
Die in Abschnitt 5.3 konstruierten Homotopie-Funktionen f_τ erfüllen alle die Voraussetzungen 1 und 2.

Satz 4.5
Es seien die oben genannten Voraussetzungen 1 bis 3 erfüllt. Ferner bezeichne (u^n, v^n) $(n \in \mathbb{N})$ die Folge, welche von der HQR-Iteration aus Abschnitt 4.3, bei vorgegebenem Startwert $u^0 \in \mathbb{R}^d$ generiert wird. Dann gilt:

1. *(u^n, v^n) konvergiert für jeden Startwert $u^0 \in \mathbb{R}^d$ gegen einen stationären Punkt $(\bar{u}, \bar{v})$ von J_τ^*. Dabei ist die Folge $\left(J_\tau^*(u^n, v^n)\right)$ monoton fallend.*

2. *Ist $(\hat{u}, \hat{v})$ das globale Minimum von J_τ^* und ist der Startwert u^0 hinreichend nahe an $\hat{u}$, so konvergiert (u^n, v^n) gegen $(\hat{u}, \hat{v})$.*

Beweis

Für den Beweis sei $J^n := J_\tau^*(u^n, v^{n+1})$. Der Beweis für den Punkt 1 läuft in mehreren Schritten ab.

1. Schritt: (J^n) ist monoton fallend und konvergent:
Wegen

$$J^n = J_\tau^*(u^n, v^{n+1}) \overset{(4.3)}{\leq} J_\tau^*(u^n, v^n) \overset{(4.4)}{\leq} J_\tau^*(u^{n-1}, v^n) = J^{n-1},$$

ist (J^n) monoton fallend. Außerdem ist (J^n) durch 0 nach unten beschränkt, also konvergent. Sei etwa $\bar{J} := \lim_{n\to\infty} J^n$. Natürlich ist aus dem gleichen Grund auch die Folge $J_\tau^*(u^n, v^n)$ monoton fallend.

2. Schritt: Es gilt: $\left\|v^{n+1} - v^n\right\|_2 \to 0$ für $n \to \infty$:
Es gilt

$$J_\tau^*(u^n, v^n) - J_\tau^*(u^n, v^{n+1}) =$$
$$= \lambda^2 \sum_{t\in\mathcal{T}} \left(v_t^n (u^n)^* A_t u^n + \frac{h_t}{\lambda^2} \Psi_\tau(v_t^n) \right) - \lambda^2 \sum_{t\in\mathcal{T}} \left(v_t^{n+1} (u^n)^* A_t u^n + \frac{h_t}{\lambda^2} \Psi_\tau(v_t^{n+1}) \right).$$

Setze für $t \in \mathcal{T}$ und $w \in \mathbb{R}$

$$g_{n,t}(w) := w \cdot (u^n)^* A_t u^n + \Psi_\tau(w) \frac{h_t}{\lambda^2}, \quad \text{dann ist}$$
$$g'_{n,t}(w) = (u^n)^* A_t u^n + \Psi'_\tau(w) \frac{h_t}{\lambda^2} \quad \text{und}$$
$$g''_{n,t}(w) = \Psi''_\tau(w) \frac{h_t}{\lambda^2}.$$

Damit gilt

$$J_\tau^*(u^n, v^n) - J_\tau^*(u^n, v^{n+1}) = \lambda^2 \sum_{t\in\mathcal{T}} \big[g_{n,t}(v_t^n) - g_{n,t}(v_t^{n+1})\big].$$

Setzt man nun die Taylor-Entwicklung von $g_{n,t}$

$$g_{n,t}(v_t^n) = g_{n,t}(v_t^{n+1}) + g'_{n,t}(v_t^{n+1})(v_t^n - v_t^{n+1}) + \frac{1}{2} g''_{n,t}(c_t^n)(v_t^n - v_t^{n+1})^2,$$

mit c_t^n zwischen v_t^{n+1} und v_t^n, ein, so erhält man

$$J_\tau^*(u^n, v^n) - J_\tau^*(u^n, v^{n+1}) = \lambda^2 \sum_{t\in\mathcal{T}} (v_t^n - v_t^{n+1}) g'_{n,t}(v_t^{n+1}) + \frac{\lambda^2}{2} \sum_{t\in\mathcal{T}} (v_t^n - v_t^{n+1})^2 g''_{n,t}(c_t^n).$$

Aus

$$J_\tau^*(u^n, v^{n+1})) \le J_\tau^*(u^n, v) \quad \forall v$$

und

$$\frac{\partial J_\tau^*}{\partial v_t}(u^n, v) = \lambda^2 (u^n)^* A_t u^n + h_t \Psi'_\tau(v_t) = \lambda^2 g'_{n,t}(v_t)$$

folgt

$$0 = \frac{\partial J_\tau^*}{\partial v_t}(u^n, v^{n+1}) = \lambda^2 g'_{n,t}(v_t^{n+1}),$$

also $g'_{n,t}(v_t^{n+1}) = 0$. Damit gilt nun

$$J_\tau^*(u^n, v^n) - J_\tau^*(u^n, v^{n+1}) = \frac{\lambda^2}{2} \sum_{t\in\mathcal{T}} (v_t^n - v_t^{n+1})^2 g''_{n,t}(c_t^n).$$

Mit

$$g''_{n,t}(w) \ge c_\tau \frac{h_t}{\lambda^2} \qquad \text{für alle } w \in [0;1]$$

ergibt sich dann

$$\begin{aligned} J^{n-1} - J^n &= J_\tau^*(u^{n-1}, v^n) - J_\tau^*(u^n, v^{n+1}) = \\ &= \big[J_\tau^*(u^{n-1}, v^n) - J_\tau^*(u^n, v^n)\big] + \big[J_\tau^*(u^n, v^n) - J_\tau^*(u^n, v^{n+1})\big] \ge \\ &\ge 0 + \frac{\lambda^2}{2} \sum_{t\in\mathcal{T}} (v_t^n - v_t^{n+1})^2 g''_{n,t}(c_t^n) \ge \frac{h_t}{2} c_\tau \sum_{t\in\mathcal{T}} (v_t^n - v_t^{n+1})^2 = \\ &= \frac{h_t}{2} c_\tau \big\| v^{n+1} - v^n \big\|_2^2 \ge 0. \end{aligned}$$

Weil (J^n) nach Schritt 1 konvergiert, also für $n \to \infty$ die linke Seite gegen 0 geht, muss auch $\big\|v^n - v^{n+1}\big\|_2 \to 0$ für $n \to \infty$ gelten.

3. Schritt: Die Folge (u^n, v^n) ist beschränkt:
Mit Hilfe von Satz 4.3 lässt sich die Minimierung in (4.3) explizit durchführen und es gilt

$$\lambda^2 \sum_{t\in\mathcal{T}} v_t u^* A_t u + \sum_{t\in\mathcal{T}} h_t \Psi_\tau(v_t) = \sum_{t\in\mathcal{T}} h_t \big[(\lambda^2 h_t^{-1} u^* A_t u) v_t + \Psi_\tau(v_t)\big] = \min_v !$$

$$\Updownarrow$$

$$(\lambda^2 h_t^{-1} u^* A_t u) v_t + \Psi_\tau(v_t) = \min_{v_t} ! \qquad \text{für alle } t \in \mathcal{T}$$

$$\Updownarrow \quad \text{(Satz 4.3)}$$

$$v_t = f'_\tau(\lambda^2 h_t^{-1} u^* A_t u) \qquad \text{für alle } t \in \mathcal{T}.$$

Wegen den Eigenschaften 1 bis 4 aus Abschnitt 3.2 ist $f'_\tau(w) \in [0;1]$ und damit ist

$$\|v^n\|_\infty \leq 1 \qquad \text{für alle } n.$$

Es gilt

$$J^*_\tau(u^n, v^{n+1}) - \underbrace{\lambda^2 \sum_{t\in\mathcal{T}} v_t^{n+1}(u^n)^* A_t u^n}_{\geq 0} - \underbrace{\sum_{t\in\mathcal{T}} h_t \Psi_\tau(v_t^{n+1})}_{\geq 0} = (u-g)^* M (u-g),$$

und mit der Monotonie aus Schritt 1

$$(u-g)^* M (u-g) \leq J^*_\tau(u^n, v^{n+1}) = J^n \leq J^0.$$

Mit der Norm $\|x\|^2_M := x^* M x$ ist

$$\|u\|_M = \|u - g + g\|_M \leq \|u-g\|_M + \|g\|_M \leq \sqrt{J_0} + \|g\|_M,$$

und damit ist auch (u^n) beschränkt.

4. Schritt: Die Folge (u^n, v^n) hat mindestens einen Häufungspunkt:
Nach Schritt 3 ist (u^n, v^n) eine Folge in einer kompakten Menge und besitzt damit eine konvergente Teilfolge. Also hat (u^n, v^n) mindestens einen Häufungspunkt.

5. Schritt: Jeder Häufungspunkt $(\bar{u}, \bar{v})$ der Folge (u^n, v^n) ist ein stationärer Punkt von J^*_τ:
Sei dazu (u^{n_k}, v^{n_k}) eine Teilfolge von (u^n, v^n) mit $u^{n_k} \to \bar{u}$ und $v^{n_k} \to \bar{v}$ für $k \to \infty$.
Wegen (4.4) gilt

$$\nabla_u J^*_\tau(u^{n_k}, v^{n_k}) = 0 \qquad \text{für alle } k \in \mathbb{N}.$$

Der Grenzübergang $k \to \infty$ liefert dann

$$\nabla_u J^*_\tau(\bar{u}, \bar{v}) = 0.$$

Wegen (4.3) gilt

$$\nabla_v J^*_\tau(u^{n_k}, v^{n_k+1}) = 0 \qquad \text{für alle } k \in \mathbb{N}.$$

Damit gilt

$$\nabla_v J^*_\tau(u^{n_k}, v^{n_k}) = \operatorname{Hess}_v J^*_\tau(u^{n_k}, v^{n_k+1})(v^{n_k} + v^{n_k+1}) + O\big(\big\|v^{n_k} - v^{n_k+1}\big\|_2^2\big),$$

und der Grenzübergang $k \to \infty$, zusammen mit Schritt 2, liefert

$$\nabla_v J^*_\tau(\bar{u}, \bar{v}) = 0,$$

womit also $(\bar{u}, \bar{v})$ stationär ist.

6. Schritt: Die Folge (u^n, v^n) konvergiert gegen $(\bar{u}, \bar{v})$:
Die Folge (v^n) erfüllt die Voraussetzungen des Satzes 4.4. Damit konvergiert (v^n) gegen $\bar{v}$. Mit (v^n) konvergiert auch (u^n). Also gilt $(u^n, v^n) \to (\bar{u}, \bar{v})$ für $n \to \infty$.

Nun noch zur Behauptung im Punkt 2.
Es gibt ein Kompaktum $U \times V$, so dass $(\hat{u}, \hat{v})$ im Inneren von $U \times V$ liegt und $(\hat{u}, \hat{v})$ der einzige stationäre Punkt in $U \times V$ ist. Da $(\hat{u}, \hat{v})$ das globale Minimum ist, gilt zusätzlich, nach eventueller Verkleinerung von U und V, für jedes $(\tilde{u}, \tilde{v}) \notin U \times V$

$$J^*_\tau(u, v) < J^*_\tau(\tilde{u}, \tilde{v}) \qquad \text{für alle } (u, v) \in U \times V. \qquad (*)$$

Wird nun die HQR-Iteration mit einem u^0 in U gestartet, so ist wegen $(*)$ $v^1 \in V$. Da die Folge $(J_\tau^*(u^n, v^n))$ monoton fallend ist, gilt wieder wegen $(*)$

$$(u^n, v^n) \in U \times V$$

für alle $n \in \mathbb{N}$. Wegen Punkt 1 konvergiert folglich (u^n, v^n) gegen $(\hat{u}, \hat{v})$. □

Anmerkung
Es ist noch ungeklärt, ob der Punkt 2 des Satzes 4.5 auch für lokale Minima $(\hat{u}, \hat{v})$ gültig bleibt. Schwierigkeiten bereiten hier die $(\hat{u}, \hat{v})$, die am Rand des Definitonsgebiets von J_τ^* liegen.

5 Implementierung

In diesem Kapitel wird eine Implementierung des GNC-Algorithmus zur Minimierung von (4.2) vorgestellt. Nachdem im Abschnitt 5.1 die Voraussetzungen und die Ausgangssituation für den hier behandelten Algorithmus dargestellt wurden, wird im folgenden Abschnitt 5.2 das grobe Grundgerüst vorgestellt. In den weiteren Abschnitten 5.3, 5.4 und 5.5 werden dann die Iterationen getrennt im Detail behandelt. Zum Abschluss werden in 5.6 die Abbruchbedingungen der einzelnen Iterationen besprochen.

5.1 Ausgangssituation

Der im Rahmen dieser Diplomarbeit implementierte Algorithmus geht von folgender Situation aus. Gegeben ist ein Bild bestehend aus Pixeln in m Zeilen und n Spalten. Jeder Pixel hat einen Grauwert. Als Vergitterung werden bilineare Elemente verwendet (vgl. Abbildung 5.1). Also sind $N := mn$ Unbekannte gesucht und die Anzahl der Elemente in der Vergitterung beträgt $|\mathcal{T}| = (m-1)(n-1)$.

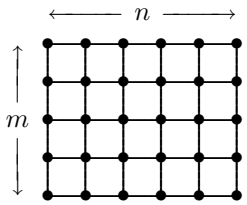

Abbildung 5.1: Vergitterung mit bilinearen Elementen

Die Pixel bzw. Elemente der Vergitterung werden zeilenweise nummeriert. Als Basis wird die Knotenbasis

$$\Phi_{ij}(x,y) = \Phi(x - x_j, y - y_i) \quad \text{mit} \quad \Phi(x,y) = \begin{cases} (1-|x|)(1-|y|) & \text{für } -1 \le x, y \le 1 \\ 0 & \text{sonst} \end{cases}$$

verwendet. Dabei ist $i = 1, \dots m$, $j = 1, \dots, n$ und die (x_j, y_i) bezeichnen die Koordinaten des Pixels mit der Nummer $i \cdot j$. Die Pixel haben einen normierten Abstand von $h = 1$. Nun kann man A_t (für $t \in \mathcal{T}$) bestimmen. Dazu betrachtet man zunächst das Einheitsquadrat und die bilineare Funktion

$$B = a\Phi_{00} + b\Phi_{10} + c\Phi_{11} + d\Phi_{01},$$

die an den vier Ecken die Werte a, b, c bzw. d annimmt. Dann ist

$$\|\nabla B(x,y)\|_2^2 = \big[(1-y)(b-a) + y(c-d)\big]^2 + \big[(1-x)(d-a) + x(c-d)\big]^2$$

und

$$\int_0^1 \int_0^1 \|\nabla B(x,y)\|_2^2 \, dy\, dx = \frac{1}{6}\big[4(a^2+b^2+c^2+d^2) - 4(ac+bd) - 2(ab+ad+bc+cd)\big].$$

Setzt man

$$\tilde{A} = \frac{1}{6}\begin{pmatrix} 4 & -1 & -2 & -1 \\ -1 & 4 & -1 & -2 \\ -2 & -1 & 4 & -1 \\ -1 & -2 & -1 & 4 \end{pmatrix} \in \mathbb{R}^{4\times 4},$$

so gilt hier

$$(a,b,c,d)\tilde{A}\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \int_0^1 \int_0^1 \|\nabla B(x,y)\|_2^2 \, dy\, dx.$$

Übertragen auf ein beliebiges $t \in \mathcal{T}$, dessen vier Pixel die Nummern i_1, i_2, i_3 und i_4 haben, gilt

$$(e_{i_1}, e_{i_2}, e_{i_3}, e_{i_4})^* A_t (e_{i_1}, e_{i_2}, e_{i_3}, e_{i_4}) = \tilde{A},$$

wobei $e_j \in \mathbb{R}^N$ der j-te Einheitsvektor ist. An allen anderen Positionen von A_t sind die Einträge 0.
Wie bei derartigen diskreten Problemen üblich, wird eine gelumpte Massenmatrix

$$M = \operatorname{diag}(h^2, \ldots, h^2) = I \in \mathbb{R}^{N\times N}$$

verwendet.
Mit den Abkürzungen (2.3) bzw. (2.4)

$$A = \sum_{t\in\mathcal{T}} A_t \qquad \text{und} \qquad A_v = \sum_{t\in\mathcal{T}} v_t A_t$$

befinden sich folglich alle Nichtnulleinträge dieser zwei Matrizen auf der Diagonalen und den Nebendiagonalen 1, $n-1$, n und $n+1$. Alle zwei Matrizen haben also die in Abbildung 5.2 angedeutete Besetzungsstruktur.

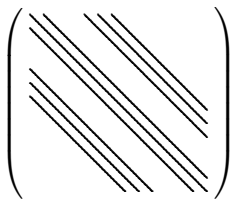

Abbildung 5.2: Besetzungsstruktur von A und A_v

Es sei nocheinmal darauf hingewiesen, dass die hier vorgestellte Implementierung nicht auf diese Spezialstruktur angewiesen ist.

5.2 Grundgerüst

Der implementierte Algorithmus hat ein Grundgerüst, bestehend aus drei ineinander verschachtelten Iterationen. Abbildung 5.3 zeigt dieses Gerüst.
In den nachfolgenden Abschnitten wird jede Iteration einzeln besprochen.

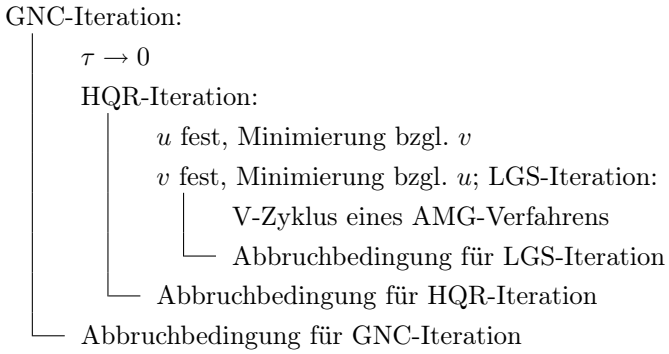

Abbildung 5.3: Grundgerüst des implementierten Algorithmus

5.3 GNC-Iteration

5.3.1 Konstruktion der GNC-Homotopie

Für die Konstruktion der f_τ wird der Satz 3.1 verwendet. Die Lösungen der Differenzialgleichung

$$2xy''(x) + y'(x) = -\frac{1}{\tau}$$

haben die Form

$$y(x) = c_1\sqrt{x} - \frac{x}{\tau} + c_2.$$

Die Abbildung 5.4 zeigt den Bereich $]x_0; x_1[$, in dem die f_τ diese Differenzialgleichung erfüllen sollen.

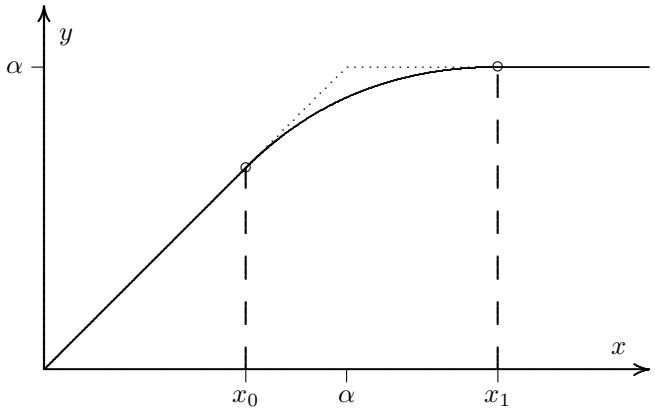

Abbildung 5.4: GNC-Homotopie: Konstruktion der f_τ

Dabei ist $x_0 = \alpha/(1+\tau)$ und $x_1 = \alpha(1+\tau)$. Für die $x \notin]x_0; x_1[$ wird, wie auch in der Abbildung 5.4 angedeutet, $f_\tau(x) := f_0(x)$ gesetzt. Die Konstanten c_1 und c_2 werden so bestimmt, dass

$$f_\tau(x_0) = c_1\sqrt{x_0} - \frac{x_0}{\tau} + c_2 = x_0 \qquad \text{und}$$

$$f_\tau(x_1) = c_1\sqrt{x_1} - \frac{x_1}{\tau} + c_2 = \alpha,$$

um die Stetigkeit von f_τ sicherzustellen. Insgesamt ist dann f_τ durch

$$f_\tau(x) := \begin{cases} x & \text{für } 0 \le x \le x_0 \\ -\dfrac{x}{\tau} - \dfrac{\alpha}{\tau} + \dfrac{2}{\tau}\sqrt{x_1 x} & \text{für } x_0 < x < x_1 \\ \alpha & \text{für } x \ge x_1 \end{cases} \tag{5.1}$$

festgelegt. Man rechnet leicht nach, dass $f_\tau \in \mathcal{C}^1$ für alle $\tau > 0$ und dass f_τ, außer bei $x \in \{x_0, x_1\}$, zweimal stetig differenzierbar ist. Somit gilt

$$2x f_\tau''(x) + f_\tau'(x) = \left\{ \begin{matrix} 1 & \text{für } 0 < x < x_0 \\ -1/\tau & \text{für } x_0 < x < x_1 \\ 0 & \text{für } x_1 < x \end{matrix} \right\} \ge -\frac{1}{\tau}.$$

Da

$$f_\tau'(x) = \frac{1}{\tau}\left(\sqrt{\frac{x_1}{x}} - 1\right) > 0 \qquad \text{für } x_0 < x < x_1, \tag{5.2}$$

ist f_τ in $]x_0; x_1[$ monoton wachsend. Weil ferner

$$f_\tau''(x) = -\frac{1}{2\tau}\sqrt{x_1} \cdot x^{-3/2} < 0 \qquad \text{für } x_0 < x < x_1,$$

ist f_τ insgesamt konkav und deshalb gilt

$$f_\tau(x) \le f_0(x) \qquad \text{für alle } \tau \ge 0.$$

Außerdem ist $\|f_\tau - f_0\|_\infty \to 0$ für $\tau \to 0$.
Folglich sind die Eigenschaften 1 bis 4 aus Abschnitt 3.2 und damit auch die Voraussetzungen des Satzes 3.1 erfüllt und $\tau \to f_\tau$ ist für die durch Gleichung (5.1) festgelegten f_τ eine geeignete Homotopie für das GNC-Verfahren.

5.3.2 Berechnung der konjugierten Funktionen

Die Legendre-Fenchel-Transformierten der im Unterabschnitt 5.3.1 konstruierten f_τ haben dann folgendes Aussehen:

Satz 5.1
Sei f_τ für alle $\tau > 0$ wie in (5.1) definiert. Dann gilt

$$\Psi_\tau(v) = \begin{cases} +\infty & \textit{für } v < 0 \\ \alpha\dfrac{1-v}{1+\tau v} & \textit{für } 0 \le v \le 1 \\ 0 & \textit{für } 1 < v \end{cases} \tag{5.3}$$

mit $\Psi_\tau(v) := (-f_\tau)^*(-v)$.

Beweis

Ist $v < 0$, so gilt

$$\Psi_\tau(v) = (-f_\tau)^*(-v) = \sup_{x \in \mathbb{R}} \big(f_\tau(x) - vx\big) = +\infty.$$

Für $v = 0$ ist

$$\Psi_\tau(v) = \sup_{x \in \mathbb{R}} f_\tau(x) = \alpha = \alpha\frac{1-v}{1+\tau v}.$$

Für $0 < v < 1$ gibt es genau ein $x_0 < x < x_1$ mit $v = f'_\tau(x)$. Wegen (5.2) gilt

$$1 + \tau v = \sqrt{\frac{x_1}{x}}. \qquad (*)$$

und damit

$$\begin{aligned}
\Psi_\tau(v) = (-f_\tau)^*(-v) &= x(-v) + f_\tau(x) = f_\tau(x) - x f'_\tau(x) = \\
&= \frac{2}{\tau}\sqrt{x_1 x} - \frac{x+\alpha}{\tau} - \frac{x}{\tau}\left(\sqrt{\frac{x_1}{x}} - 1\right) = \frac{1}{\tau}(\sqrt{x_1 x} - \alpha) = \\
&= \frac{1}{\tau}\left(x_1\sqrt{\frac{x}{x_1}} - \alpha\right) \overset{(*)}{=} \frac{1}{\tau}\left(\frac{x_1}{1+\tau v} - \alpha\right) = \\
&= \frac{1}{\tau}\frac{\alpha(1+\tau) - \alpha(1+\tau v)}{1+\tau v} = \alpha\frac{1-v}{1+\tau v}.
\end{aligned}$$

Ist $v \geq 1$, dann gilt sowohl

$$\Psi_\tau(v) = \sup_{x\in\mathbb{R}}\big(f_\tau(x) - vx\big) \geq f_\tau(0) - v\cdot 0 = f_\tau(0) = 0,$$

also auch

$$\Psi_\tau(v) = \sup_{x\in\mathbb{R}}\big(f_\tau(x) - vx\big) \leq \sup_{x\in\mathbb{R}}\big(f_\tau(x) - x\big) = 0,$$

und somit $\Psi_\tau(v) = 0$. □

Damit sind die Ψ_τ für alle $\tau > 0$ im Intervall $[0;1]$ zweimal stetig differenzierbar mit

$$\Psi''_\tau(v) = 2\alpha\tau\frac{1+\tau}{(1+\tau v)^3} \geq \frac{2\alpha\tau}{(1+\tau)^2} =: c_\tau > 0,$$

womit die Ψ_τ also auch die Voraussetzung 2 aus Abschnitt 4.4 erfüllen.
Die so konstruierte Homotopie $\tau \to f_\tau$ hat noch eine schöne Eigenschaft. Es gilt nämlich

$$\lim_{\tau\to 0+}\Psi_\tau(v) = \lim_{\tau\to 0+}\alpha\frac{1-v}{1+\tau v} = \alpha(1-v) = (-f_0)^*(-v) =: \Psi_0(v).$$

Diese Konvergenz ist monoton. Daher konvergiert auch $f_\tau(x)$ monoton gegen $f_0(x)$.

5.4 HQR-Iteration

In diesem Abschnitt wird die Implementierung der beiden HQR-Schritte, wie sie in Abschnitt 4.3 vorgestellt wurden, angegeben.
Wie schon im Schritt 3 des Beweises von Satz 4.5 gezeigt, lässt sich mit Hilfe von Satz 4.3 die Minimierung (4.3) leicht ausführen, denn

$$\begin{gathered}
\lambda^2\sum_{t\in\mathcal{T}} v_t u^* A_t u + \sum_{t\in\mathcal{T}} h_t\Psi_\tau(v_t) = \sum_{t\in\mathcal{T}} h_t\big[(\lambda^2 h_t^{-1} u^* A_t u)v_t + \Psi_\tau(v_t)\big] = \min_v! \\
\Updownarrow \\
(\lambda^2 h_t^{-1} u^* A_t u)v_t + \Psi_\tau(v_t) = \min_{v_t}! \qquad \text{für alle } t\in\mathcal{T} \\
\Updownarrow \quad \text{(Satz 4.3)} \\
v_t = f'_\tau(\lambda^2 h_t^{-1} u^* A_t u) \qquad \text{für alle } t\in\mathcal{T}.
\end{gathered}$$

Damit ist die Minimierung von v bei festem u schon erledigt.
Bei (4.4) handelt es sich um ein lineares elliptisches Problem. Es gilt

$$\begin{gathered} J_\tau^*(u,v) = \min_u ! \\ \Updownarrow \\ \lambda^2 \sum_{t\in\mathcal{T}} v_t u^* A_t u + (u-g)^* M (u-g) = \min_u ! \\ \Updownarrow \\ \lambda^2 u^* A_v u + (u-g)^* M (u-g) = \min_u ! \\ \Updownarrow \\ \lambda^2 \cdot 2A_v u + 2Mu - 2Mg = 0 \\ \Updownarrow \\ \underbrace{(\lambda^2 A_v + M)}_{=:L} u = \underbrace{Mg}_{=:f}, \end{gathered}$$

wobei $A_v := \sum_{t\in\mathcal{T}} v_t A_t$ ist.
Wie dieses lineare Gleichungssystem

$$Lu = f \tag{5.4}$$

gelöst wird, wird in Abschnitt 5.5 behandelt.

5.5 LGS-Iteration

Betrachtet man, bei vorgegebenem v, den Ursprung des Gleichungssystems (5.4), nämlich

$$\lambda^2 \sum_{t\in\mathcal{T}} v_t u^* A_t u + (u-g)^* M (u-g) = \min_u !,$$

so erkennt man, dass dies die Diskretisierung des kontinuierlichen, elliptischen Problems

$$-\operatorname{div}\big(v(x)\nabla u(x)\big) + u = g \tag{5.5}$$

ist. Dabei ist die Koeffizientenfunktion v unstetig mit großen Sprüngen. Trennt also eine geschlossene Kante, wie in Abbildung 5.5 gezeigt, ein Gebiet Ω_1 vom Rest Ω_2 ab, so beschreibt (5.5) zwei unabhängige Probleme mit Neumann Randbedingungen. Es wird (bei realistischen Daten) sehr häufig passieren, dass auf diese Art und Weise Inseln entstehen, die auch eine unabhängige numerische Behandlung erfordern. Da am Anfang nicht klar ist, wo diese Inseln entstehen (es ist ja gerade die Aufgabe der Segmentierung solche Inseln zu finden), wird ein Löser benötigt, der völlig automatisch solche Situationen erkennt.

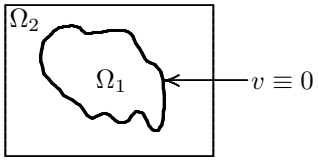

Abbildung 5.5: Unabhängige Neumann-Probleme durch geschlossene Kanten

Außerdem muss natürlich, wegen der großen Dimension des Gleichungssystems, eine iterative Methode gewählt werden. Daher bieten sich Mehrgitterverfahren an. Um die nötige Flexibilität für oben genannte Inseln zu garantieren, kommen algebraische Mehrgitterverfahren (AMG-Verfahren) zum Einsatz. In dieser Diplomarbeit wurde ein von Ruge und Stüben entwickeltes AMG-Verfahren verwendet. In Kapitel 6 wird dieses Verfahren (und eine Erweiterung) vorgestellt. Die Verwendung solcher Verfahren hat noch den weiteren Vorteil, dass diese Methoden keine geometrischen Voraussetzungen an die Vergitterungen stellen. Dies ist der Grund, warum die hier vorgestellte Implementierung auch für andere Vergitterungen verwendbar ist.

5.6 Steuerung

Nachdem nun alle Module des Algorithmus behandelt wurden, fehlen zu einer Implementierung noch die geeigneten Abbruchbedingungen. Diese werden jetzt besprochen.

5.6.1 Abbruch der LGS-Iteration

Der in Abschnitt 5.5 dargestellte iterative Löser wird nach einer fest vorgegebenen Anzahl von Iterationen abgebrochen. Daher ist eine wichtige Forderung an diesen Löser, dass für die Iterierten u^k

$$J_\tau^*(u^{k+1}, v) \leq J_\tau^*(u^k, v)$$

für beliebige v gilt, also Monotonie bzgl. J_τ^* vorliegt. Dies wird in Abschnitt 6.3 für eine große Klasse von AMG-Lösern bewiesen.

5.6.2 Abbruch der HQR-Iteration

Das Abbruchkriterium für die HQR-Iteration aus Abschnitt 5.4 ist nicht mehr so einfach. In der Abbildung 5.6 ist der typische Verlauf $J_\tau^*(u^0, v^1), J_\tau^*(u^1, v^1), J_\tau^*(u^1, v^2), \ldots$ aufgetragen.

Abbildung 5.6: Typischer Energieverlauf während der HQR-Iteration

Im Kontext der GNC-Homotopie muss hier sicherlich nicht solange iteriert werden, bis ein Minimum auf sehr hohe Genauigkeit erreicht wird, denn das Funktion J_τ^* wird sich beim nächsten GNC-Schritt wieder verändern. Es muss aber sichergestellt sein, dass die letzte Iterierte im „richtigen Potentialtopf“ sitzt (vgl. Abbildung 5.7) und somit das globale Minimum durch die GNC-Iteration verfolgt werden kann.

Deshalb wird beim Start der HQR-Iteration der Energieunterschied

$$\Delta_0 := J_\tau^*(u^0, v^1) - J_\tau^*(u^1, v^1)$$

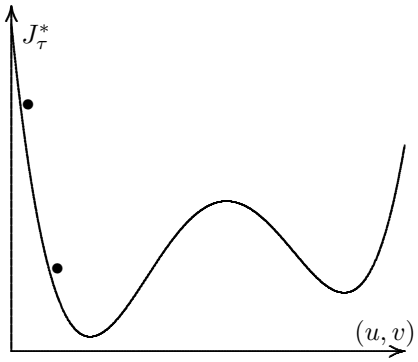

Abbildung 5.7: HQR-Iteration bis letzte Iterierte nahe am Minimum

gespeichert und während der Iteration die letzten drei Energieunterschiede

$$\begin{aligned}\Delta_1^k &:= J_\tau^*(u^{k-1}, v^k) - J_\tau^*(u^k, v^k)\\ \Delta_2^k &:= J_\tau^*(u^{k-1}, v^{k-1}) - J_\tau^*(u^{k-1}, v^k)\\ \Delta_3^k &:= J_\tau^*(u^{k-2}, v^{k-1}) - J_\tau^*(u^{k-1}, v^{k-1})\end{aligned}$$

verglichen. Gilt mit einem vorgebenen $0 < \eta < 1$

$$\max(\Delta_1^k, \Delta_2^k, \Delta_3^k) \leq \eta\Delta_1,$$

so wird davon ausgegangen, dass sich die Iteration in einem flachen Bereich befindet und abgebrochen. Um zu verhindern, dass zu viele Iterationen gemacht werden, falls die HQR-Iteration schon in einem sehr flachen Bereich gestartet wurde, wird zusätzlich nach jeder Iteration getestet, ob für ein fest vorgegebenes Δu

$$\left\|u^k - u^{k-1}\right\|_\infty \leq \Delta u$$

gilt. Falls dies erfüllt ist, wird ebenfalls die Iteration beendet.

5.6.3 Steuerung der GNC-Iteration

Die Aufgabe der GNC-Steuerung ist es, eine rasch fallende Folge (τ) von Homotopieparametern zu ermitteln, so dass es aber immer noch möglich ist, bei jedem Übergang das globale Minimum mit der HQR-Iteration zu verfolgen. Zur Beschreibung dieser Steuerung wird von folgender Situation ausgegangen.

Zu einem Homotopieparameter $\tau_0 > 0$ gibt es ein Paar (u^0, v^0) als Approximation für das Minimum. Ferner sei eine „Schrittweite“ $\Delta\tau > 0$ gegeben. Dann werden

$$\tau_1 := \frac{\tau_0}{1 + \frac{1}{2}\Delta\tau} \qquad \text{und}$$

$$\tau_2 := \frac{\tau_0}{1 + \Delta\tau}$$

definiert. Es gilt offensichtlich $\tau_0 > \tau_1 > \tau_2$.

Nun wird die HQR-Iteration für τ_1 mit dem Startwert u^0 gestartet. Nach Abschluss dieser Iteration erhält man ein Paar (u^{01}, v^{01}) als Approximation für das Minimum von $J_{\tau_1}^*$. Im

Anschluss wird die HQR-Iteration erneut für τ_2 mit dem Startwert u^{01} gestartet und man erhält somit ein Paar (u^{12}, v^{12}) als Approximation des Minimums von $J^*_{\tau_2}$. Im letzten Teilschritt erhält man aus der HQR-Iteration für τ_2 mit dem Startwert u^0 ein Paar (u^{02}, v^{02}) als weitere Approximation für das Minimum von $J^*_{\tau_2}$. In der Abbildung 5.8 sind die eben beschriebenen Teilschritte tabellarisch dargestellt.

	Startwert	τ	HQR liefert
1. Teilschritt	u^0	τ_1	(u^{01}, v^{01})
2. Teilschritt	u^1	τ_2	(u^{12}, v^{12})
3. Teilschritt	u^0	τ_2	(u^{02}, v^{02})

Abbildung 5.8: Teilschritte der GNC-Iteration

Um nun die Güte des GNC-Schrittes $\tau_0 \to \tau_2$ zu messen, werden die Paare (u^{02}, v^{02}) und (u^{12}, v^{12}) verglichen. Da das globale Minimum nach Satz 3.2 stetig von τ abhängt, ist die Abweichung zwischen den beiden Paaren bei hinreichen kleinem $\Delta\tau$ sehr klein. Wie diese Abweichung gemessen wird, wird weiter unten erklärt. Es wird nun so vorgegangen:

1. Ist die Abweichung annehmbar klein, so wird aus den beiden Paaren (u^{02}, v^{02}) und (u^{12}, v^{12}) das Paar $(\bar{u}, \bar{v})$ ermittelt mit

   $$J^*_{\tau_2}(\bar{u}, \bar{v}) = \min\big(J^*_{\tau_2}(u^{02}, v^{02}), J^*_{\tau_2}(u^{12}, v^{12})\big).$$

   Der Schritt $\tau_0 \to \tau_2$ wird akzeptiert und $(\bar{u}, \bar{v})$ ist die Approximation für das Minimum von $J^*_{\tau_2}$. Jetzt kann die GNC-Iteration erneut gestartet werden. War die Abweichung sehr klein, so wird die Schrittweite durch $\overline{\Delta\tau} = 2\Delta\tau$ erhöht.

2. Ist die Abweichung unzumutbar groß, so wird die Schrittweite durch $\overline{\Delta\tau} = \Delta\tau/2$ halbiert und die GNC-Iteration mit dem Startwert $(\bar{u}, \bar{v}) = (u^0, v^0)$ erneut gestartet. Hier kann aber wegen

   $$\bar{\tau}_2 = \frac{\tau_0}{1 + \overline{\Delta\tau}} = \frac{\tau_0}{1 + \frac{1}{2}\Delta\tau} = \tau_1$$

   der dortige 3. Teilschritt eingespart werden. Er entspricht dem schon berechneten Schritt $\tau_0 \to \tau_1$. So wird bei einer Ablehnung eines GNC-Schrittes wenigstens einer der drei Schritte wiederverwertet. In Abbildung 5.9 ist die Lage der verschiedenen Parameter dargestellt.

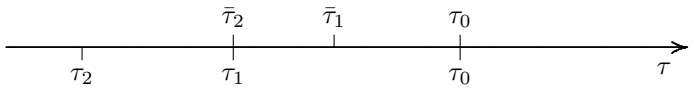

Abbildung 5.9: Lage der Homotopieparameter τ_i bei der GNC-Iteration

Wie oben angekündigt, kommt jetzt noch die Frage, wie man die Abweichung zwischen den Paaren (u^{12}, v^{12}) und (u^{02}, v^{02}) problemangepasst misst. Dazu errinnert man sich daran, dass ein jedes solches Paar ein Bild mit Kanten beschreibt. Man soll also feststellen, wie sich zwei gegebene Bilder unterscheiden. Dazu geht man zum „Differenzbild“ $(u^d, v^d) := (u^{12} - u^{02}, e - |v^{12} - v^{02}|)$ über. Dabei ist $e = (1, \ldots, 1)^T$ und $u^{12} - u^{02}$ offensichtlich die Grauwertunterschiede. $e - |v^{12} - v^{02}|$ entspricht der Kantenmenge, die nicht beide

Bilder gemeinsam besitzen. Im Idealfall würde so das „Nullbild“ entstehen. Als Maß für die Abweichung wird nun die Energie

$$A := J_\tau^*(u^d, v^d) = \lambda^2 \sum_{t \in \mathcal{T}} v_t^d (u^d)^* A_t u^d + \sum_{t \in \mathcal{T}} h_t \Psi_\tau(v_t^d) + (u^d)^* M u^d$$

des Differenzbildes verwendet.

Man benötigt also für die GNC-Steuerung zwei Parameter $A_{\min}$ und $A_{\max}$, um zu entscheiden, ob die Abweichung unzumutbar groß ($A > A_{\max}$), annehmbar ($A \leq A_{\max}$) bzw. sehr klein ($A \leq A_{\min}$) ist.

Der komplette Algorithmus kann beendet werden, wenn ein akzeptierter GNC-Schritt nur noch sichere Kanten bzw. Nichtkanten enthält, d. h. wenn $v \in \{0, 1\}^{|\mathcal{T}|}$ gilt.

6 Algebraische Mehrgitterverfahren (AMG-Verfahren)

In diesem Kapitel werden algebraische Mehrgitterverfahren beschrieben, die als Löser für die linearen Gleichungssysteme aus Abschnitt 5.5 verwendet werden. In Abschnitt 6.1 werden grundlegende Begriffe und Notation eingeführt. In 6.2 folgt dann die Theorie der Zweigittermethoden. Danach wird in 6.3 die Energiemonotonie der Verfahren, die innerhalb der HQR-Iteration von entscheidender Bedeutung ist, untersucht. Für die Konstruktion von AMG-Verfahren spielen algebraisch glatte Fehler und deren Eigenschaften eine entscheidende Rolle. Sie werden in 6.4 behandelt. Die Abschnitte 6.5 und 6.6 zeigen die Konstruktion der Grobgitter und der Interpolationsoperatoren, wie sie von Ruge und Stüben in [RS86] bzw. [Stü99] für M-Matrizen eingeführt wurden. In Abschnitt 6.7 wird nach der Idee von Chang, Wong und Fu gezeigt, wie man Interpolationsoperatoren konstruieren kann, die auch für Matrizen, die keine M-Matrizen sind, effiziente AMG-Verfahren liefern.

6.1 Problemstellung, Begriffe und Notation

Gegeben sind eine symmetrische, positiv definite Matrix $A_h \in \mathbb{R}^{n\times n}$ und eine rechte Seite $f^h \in \mathbb{R}^n$. Gesucht ist ein $u^h \in \mathbb{R}^n$, so dass gilt

$$A_h u^h = f^h. \tag{6.1}$$

Man setzt $\Omega^h := \{1, 2, \ldots, n\}$ und betrachtet A_h als Knoteninzidenzmatrix auf der Knotenmenge Ω^h, welche auch als Feingitter bezeichnet wird. Ein $\Omega^H \neq \Omega^h$ heißt Grobgitter, falls $\Omega^H \subset \Omega^h$.
Ein Punkt $i \in \Omega^h$ heißt direkt verbunden oder direkt verkoppelt mit $j \in \Omega^h$, falls $a^h_{ij} \neq 0$. Durch

$$N^h_i := \{j \in \Omega^h \,|\, j \neq i,\ a^h_{ij} \neq 0\}$$

wird dementsprechend die Nachbarschaft eines Punktes $i \in \Omega^h$ festgelegt. Damit kann man nun (6.1) auch als

$$a^h_{ii} u^h_i + \sum_{j\in N^h_i} a^h_{ij} u^h_j = f^h_i \qquad \text{für } i \in \Omega^h \tag{6.2}$$

schreiben.
Die grundlegende Idee bei einem algebraischen Zweigitterverfahren ist, ein „günstiges“ Grobgitter Ω^H zu konstruieren und darauf eine „geeignete“ Grobgittergleichung

$$A_H u^H = f^H$$

zu lösen, um für u^h mit Hilfe von u^H eine bessere Näherung zu finden. Durch das Grobgitter Ω^H kann das Feingitter Ω^h durch

$$\Omega^h = C^h \cup F^h, \qquad \Omega^H = C^h$$

disjunkt zerlegt werden. Diese Zerlegung wird auch C/F-Splitting genannt. Die Indizes in C^h heißen Grobgitterindizes (bzw. Grobgitterpunkte), die Indizes in F^h heißen Feingitterindizes (bzw. Feingitterpunkte). Durch Umordnung kann erreicht werden, dass bei allen beteiligten Vektoren und Matrizen die Feingitterindizes am Anfang stehen, so dass die Gleichung (6.1) auch als

$$\begin{pmatrix} A_{FF} & A_{FC} \\ A_{CF} & A_{CC} \end{pmatrix} \begin{pmatrix} u_F \\ u_C \end{pmatrix} = \begin{pmatrix} f_F \\ f_C \end{pmatrix}$$

formuliert werden kann.
Ferner werden zwei Operatoren I_H^h und I_h^H benötigt, die den Übergang vom Grobgitter zum Feingitter bzw. vom Feingitter zum Grobgitter beschreiben. I_H^h heißt Prolongation und I_h^H Restriktion. Beide zusammen bilden die Interpolation. Vereinfachend wird hier gleich angenommen, dass

$$I_h^H = (I_H^h)^* \qquad \text{und} \qquad u^h = \begin{pmatrix} u_F \\ u_C \end{pmatrix} = \begin{pmatrix} I_{FC} \\ I_{CC} \end{pmatrix} u_C = \begin{pmatrix} I_{FC} \\ I_{CC} \end{pmatrix} u^H,$$

wobei I_{CC} die Identität ist. Folglich hat jede hier betrachtete Interpolation die Form

$$e_i^h = (I_H^h e^H)_i = \begin{cases} e_i^H & \text{falls } i \in C^h \\ \sum_{k \in P_i^h} w_{ik}^h e_k^H & \text{falls } i \in F^h, \end{cases} \tag{6.3}$$

wobei $P_i^h \subset C^h$. Damit hat I_H^h automatisch vollen Rang.
Als Grobgitteroperator A_H wird der Galerkin-Operator

$$A_H := I_h^H A_h I_H^h \tag{6.4}$$

verwendet, der mit diesen Voraussetzungen dann auch wieder symmetrisch und positiv definit ist.
Außerdem benötigt ein AMG-Verfahren noch einen Glätter S_h, der durch

$$u^h \to \bar{u}^h, \qquad \bar{u}^h = S_h u^h + (I_h - S_h) A_h^{-1} f^h$$

beschrieben wird. Bezeichnet u_*^h die exakte Lösung von (6.1), dann ist

$$e^h = u_*^h - u^h$$

der Fehler einer Näherung u^h. Da nun

$$\begin{aligned} \bar{u}^h - u_*^h &= S_h u^h + (I_h - S_h) u_*^h - u_*^h \\ -\bar{e}^h &= S_h (u^h - u_*^h) = -S_h e^h, \end{aligned}$$

hat der Glätter S_h bezüglich der Fehler die Gestalt

$$e^h \to \bar{e}^h, \qquad \bar{e}^h = S_h e^h.$$

Sei nun u_{old}^h eine Approximation für u_*^h. Bei einem Grobgitterkorrekturschritt wird zur Verbesserung von u_{old}^h

$$u_{\text{new}}^h := u_{\text{old}}^h + I_H^h e^H \tag{6.5}$$

gesetzt, wobei e^H die Lösung der Gleichung

$$A_H e^H = I_h^H r_{\text{old}}^h = I_h^H (f^h - A_h u_{\text{old}}^h) \tag{6.6}$$

ist. Weil

$$u^h_{\text{new}} - u^h_* = u^h_{\text{old}} - u^h_* + I^h_H A_H^{-1} I^H_h (A_h u^h_* - A_h u^h_{\text{old}})$$
$$e^h_{\text{new}} = e^h_{\text{old}} + I^h_H A_H^{-1} I^H_h A_h e^h_{\text{old}},$$

lässt sich ein Zweigitterkorrekturschritt bzgl. der Fehler als

$$e^h_{\text{new}} = K_{h,H} e^h_{\text{old}}$$

mit

$$K_{h,H} = I_h + I^h_H A_H^{-1} I^H_h A_h \tag{6.7}$$

schreiben. $K_{h,H}$ heißt Grobgitterkorrektur. Wird vor dem Grobgitterkorrekturschritt ν_1 mal geglättet und danach ν_2 mal, so wird die Zweigitteriteration $M_{h,H}$ des AMG-Verfahrens durch

$$M_{h,H} = S_h^{\nu_2} K_{h,H} S_h^{\nu_1} \tag{6.8}$$

beschrieben.
Es bezeichne in diesem Kapitel $\langle \cdot, \cdot \rangle$ das euklidsche Skalarprodukt. Außerdem sei D_h die Diagonalmatrix mit der Diagonaleinträgen von A_h. Ferner werden noch die drei weiteren Skalarprodukte

$$\langle u, v \rangle_0 := \langle D_h u, v \rangle, \qquad \langle u, v \rangle_1 := \langle A_h u, v \rangle, \qquad \langle u, v \rangle_2 := \left\langle D_h^{-1} A_h u, A_h v \right\rangle$$

benötigt.
Durch rekursive Anwendung des soeben beschriebenen Zweigitter-Verfahrens auf die jeweiligen Grobgittergleichungen kann sofort ein Mehrgitterverfahren mit mehr als zwei Gittern programmiert werden. Auch die Zweigitterkonvergenzaussagen übertragen sich mit einer schwachen Zusatzvoraussetzung, die hier immer erfüllt ist, auf den Mehrgitterfall (vgl. [Hac85], Kapitel 7).
Anders als bei einem geometrischen Mehrgitterverfahren, bei dem die Grobgitter und die Interpolation aus geometrischen Überlegungen gegeben sind und passende Glätter konstruiert werden, ist bei einem AMG-Verfahren der Glätter fest vorgegeben und man versucht, das C/F-Splitting und I_{FC} geeignet zu konstruieren.

6.2 Theorie der Zweigittermethoden

Die Grobgitterkorrektur hat folgende Eigenschaften.

Satz 6.1
Für die Grobgitterkorrektur $K_{h,H}$ aus (6.7) gilt

1. $K_{h,H} I^h_H = 0$,
2. $K^2_{h,H} = K_{h,H}$ *und*
3. $I^H_h A_h K_{h,H} = 0$.

Insbesondere gilt

$$\mathcal{N}(K_{h,H}) = \mathcal{R}(I^h_H) \qquad \textit{und} \qquad \mathcal{R}(K_{h,H}) = \mathcal{N}(I^H_h A_h). \tag{6.9}$$

Beweis

Wegen

$$K_{h,H} I_H^h = (I_h - I_H^h A_H^{-1} I_h^H A_h) I_H^h = I_h - I_H^h A_H^{-1} A_H = 0$$

gilt Punkt 1 und $\mathcal{R}(I_H^h) \subset \mathcal{N}(K_{h,H})$. Wegen (6.7) ist $\mathcal{N}(K_{h,H}) \subset \mathcal{R}(I_H^h)$, womit dann $\mathcal{N}(K_{h,H}) = \mathcal{R}(I_H^h)$ gezeigt ist.
Die Rechnung

$$\begin{aligned} K_{h,H}^2 &= I_h - 2 I_H^h A_H^{-1} I_h^H A_h + I_H^h A_H^{-1} I_h^H A_h I_H^h A_H^{-1} I_h^H A_h = \\ &= I_h - I_H^h A_H^{-1} I_h^H A_h = K_{h,H} \end{aligned}$$

zeigt Punkt 2.
Wegen

$$I_h^H A_h K_{h,H} = I_h^H A_h (I_h - I_H^h A_H^{-1} I_h^H A_h) = I_h^H A_h - A_H A_H^{-1} I_h^H A_h = 0$$

gilt Punkt 3 und $\mathcal{R}(K_{h,H}) \subset \mathcal{N}(I_h^H A_h)$. Wegen (6.7) ist $\mathcal{N}(I_h^H A_h) \subset \mathcal{R}(K_{h,H})$, womit auch $\mathcal{R}(K_{h,H}) = \mathcal{N}(I_h^H A_h)$ gezeigt ist. □

Desweiteren wird noch folgender Hilfssatz benötigt.

Hilfssatz 6.2
Sei X ein Hilbertraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und $\|\cdot\|$ die zugehörige Norm. Ist $Q \in L(X)$ symmetrisch bzgl. $\langle \cdot, \cdot \rangle$ und $Q^2 = Q$, dann ist Q orthogonaler Projektor und es gilt

1. $\mathcal{R}(Q) \perp \mathcal{R}(I - Q)$,
2. *aus* $u \in \mathcal{R}(Q)$ *und* $v \in \mathcal{R}(I - Q)$ *folgt:* $\|u + v\|^2 = \|u\|^2 + \|v\|^2$,
3. $\|Q\| = 1$ *oder* $Q = 0$ *und*
4. *für alle* $u \in X$ *gilt:* $\|Qu\| = \min\limits_{v \in \mathcal{R}(I-Q)} \|u - v\|$.

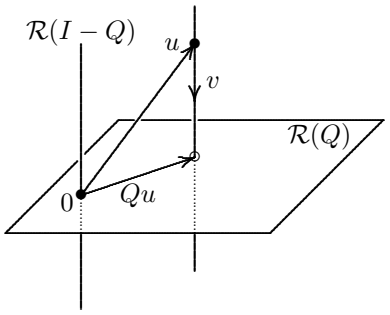

Abbildung 6.1: Skizze zu Satz 6.2

Beweis

Für $Q = 0$ sind alle Aussagen trivial. Sei also $Q \neq 0$.

$$\langle Qu, (I - Q)v \rangle = \langle u, Q(I - Q)v \rangle = \langle u, Qv - Qv \rangle = 0$$

zeigt Punkt 1, woraus sofort Punkt 2 folgt.
Da $Q \neq 0$ und $Q^2 = Q$, ist $\|Q\| \geq 1$. Mit der Zerlegung $u = Qu + (I - Q)u$ gilt

$$\|Q\|^2 = \sup_{u \neq 0} \frac{\|Qu\|^2}{\|u\|^2} = \sup_{u \neq 0} \frac{\|Qu\|^2}{\|Qu\|^2 + \|(I-Q)u\|^2} \leq 1,$$

womit auch Punkt 3 gezeigt ist.
Und

$$\begin{aligned} \min_{v \in \mathcal{R}(I-Q)} \|u - v\|^2 &= \min_{v \in \mathcal{R}(I-Q)} \big\|Qu + \underbrace{(I-Q)u - v}_{\in \mathcal{R}(I-Q)}\big\|^2 = \\ &= \min_{v \in \mathcal{R}(I-Q)} \big(\|Qu\|^2 + \|(I-Q)u - v\|^2\big) = \\ &= \|Qu\|^2 + \min_{v \in \mathcal{R}(I-Q)} \|(I-Q)u - v\|^2 = \|Qu\|^2 \end{aligned}$$

zeigt Punkt 4. □

Mit der Beobachtung

$$\begin{aligned} (A_h K_{h,H})^* = K_{h,H}^* A_h^* &= (I_h - I_H^h A_H^{-1} I_h^H A_h)^* A_h = (I_h - A_h I_H^h A_H^{-1} I_h^H) A_h \\ &= A_h (I_h - I_H^h A_H^{-1} I_h^H A_h) = A_h K_{h,H} \end{aligned}$$

folgt, dass

$$\langle K_{h,H} u, v \rangle_1 = \langle A_h K_{h,H} u, v \rangle = \langle u, A_h K_{h,H} v \rangle = \langle u, K_{h,H} v \rangle_1 ,$$

also, dass $K_{h,H}$ bzgl. des Energieskalarprodukts $\langle \cdot, \cdot \rangle_1$ symmetrisch ist. Wegen Punkt 2 in Satz 6.1 lässt sich Hilfssatz 6.2 anwenden und man erhält mit

$$\mathcal{R}(I_h - K_{h,H}) = \mathcal{R}(I_H^h)$$

den folgenden Satz.

Satz 6.3
Für die Grobgitterkorrektur $K_{h,H}$ gilt

1. $\mathcal{R}(K_{h,H}) \perp_1 \mathcal{R}(I_H^h)$,
2. *aus* $u \in \mathcal{R}(K_{h,H})$ *und* $v \in \mathcal{R}(I_H^h)$ *folgt:* $\|u + v\|_1^2 = \|u\|_1^2 + \|v\|_1^2$,
3. $\|K_{h,H}\|_1 = 1$ *und*
4. *für alle* e^h *gilt:* $\big\|K_{h,H} e^h\big\|_1 = \min_{e^H} \big\|e^h - I_H^h e^H\big\|_1$.

Der Punkt 4 in Satz 6.3 wird auch als Variationsprinzip bezeichnet: Die Grobgitterkorrektur $K_{h,H}$ minimiert unter allen möglichen Grobgitterfehlern e^H die $\|\cdot\|_1$-Norm.
Betrachtet man die Zweigitteriteration $M_{h,H}$ aus (6.8), so kann man aus dieser Darstellung zwei Richtlinien ablesen, die eine effektive Fehlerminimierung garantieren. Gilt nämlich $K_{h,H} S_h e^h \approx 0$ für möglichst viele e^h, also im Idealfall

$$\mathcal{R}(S_h) \subset \mathcal{N}(K_{h,H}) \overset{(6.9)}{=} \mathcal{R}(I_H^h), \tag{6.10}$$

oder auch $S_h K_{h,H} e^h \approx 0$, also hier im Idealfall

$$\mathcal{N}(I_h^H A_h) \overset{(6.9)}{=} \mathcal{R}(K_{h,H}) \subset \mathcal{N}(S_h), \tag{6.11}$$

so ist $M_{h,H} = 0$.

Anmerkung
Nimmt man

$$\hat{S}_h = \begin{pmatrix} 0 & -A_{FF}^{-1} A_{FC} \\ 0 & I_{CC} \end{pmatrix}$$

als „Glätter“ und wählt man

$$\hat{I}_{FC} = -A_{FF}^{-1} A_{FC},$$

so sind beide Richtlinien (6.10) und (6.11) erfüllt, falls das C/F-Splitting so gewählt wird, dass A_{FF} invertierbar ist. In diesem Fall ist dann

$$A_H = A_{CC} - A_{CF} A_{FF}^{-1} A_{FC}.$$

Dies ist aber ein unpraktikabler Grenzfall, da die Inversion von A_{FF} (bis auf wenige Ausnahmen) viel zu teuer ist.

Um diese beiden Richtlinien (näherungsweise) umsetzen zu können, muss man zunächst den Glätter S_h und die Räume $\mathcal{R}(S_h)$ und $\mathcal{N}(S_h)$ näher untersuchen. Dies geschieht in Abschnitt 6.4. Zunächst wird aber im folgenden Abschnitt eine Eigenschaft untersucht, die für die Verwendung des AMG-Lösers innerhalb des GNC-Algorithmus von großer Bedeutung ist.

6.3 Monotone AMG-Verfahren

Beim GNC-Algorithmus aus Kapitel 5 wird der AMG-Löser innerhalb der HQR-Iteration verwendet, um das Minimum eines Funktionals der Form

$$J(v) = \frac{1}{2} v^* A v - f^* v \tag{6.12}$$

zu finden (vgl. Abschnitt 5.4). Klar ist, dass

$$J(u) = \min_v J(v) \qquad \Leftrightarrow \qquad Au = f.$$

Anmerkung
Solange keine Mehrdeutigkeiten entstehen, wird in diesem Abschnitt der Super- bzw. Subscript h weggelassen.

Nun iteriert aber das AMG-Verfahren nicht bis Konvergenz eintritt, sondern wird nach einer fest vorgegebenen Anzahl von Iterationen abgegrochen. Dabei ist es also wichtig zu wissen, ob bei jedem Zyklus wirklich eine Verbesserung eintritt, d. h. ob jede neu berechnete Näherung in (6.12) einen kleineren Funktionalwert liefert. Ein AMG-Verfahren ist sicherlich bzgl. J monoton, wenn sein Grobgitterkorrekturschritt und sein Glätter monoton sind. Durch die Wahl eines Gauss-Seidel-Glätters ist die Monotonie im Glätter sichergestellt. Über die Monotonie des Grobgitterkorrekturschritts gibt folgender Satz Auskunft.

Satz 6.4
Jeder nach Abschnitt 6.1 konstruierte Grobgitterkorrekturschritt $K_{h,H}$ ist bzgl. J aus (6.12) monoton.

Beweis

Sei ein u_{old} gegeben. Nach der Gleichung (6.5) ist dann die neu berechnete Approximation

$$u_{\text{new}} = u_{\text{old}} + \underbrace{I_H^h e^H}_{=:\Delta u},$$

wobei e^H die Grobgitterkorrektur aus der Gleichung (6.6) ist. Jetzt wird nicht nur

$$J(u_{\text{new}}) \leq J(u_{\text{old}})$$

gezeigt, sondern sogar

$$J(u_{\text{new}}) = \min_t J(u_{\text{old}} + t\Delta u).$$

Dazu sei $\Delta u \neq 0$, sonst ist die Behauptung trivialerweise erfüllt. Da

$$\frac{\partial}{\partial t} J(u_{\text{old}} + t\Delta u) = t(\Delta u)^* A\Delta u + u_{\text{old}}^* A\Delta u - f^*\Delta u,$$

wird das Minimum bei

$$t = \alpha := \frac{f^*\Delta u - u_{\text{old}}^* A\Delta u}{(\Delta u)^* A\Delta u}$$

angenommen. Es wird gezeigt, dass $\alpha = 1$ ist.
Mit $e_{\text{old}} = u_* - u_{\text{old}}$ ist

$$u_{\text{old}}^* A\Delta u = \langle Au_{\text{old}}, \Delta u\rangle = \langle Au_*, \Delta u\rangle - \langle Ae_{\text{old}}, \Delta u\rangle = \langle f, \Delta u\rangle - \langle e_{\text{old}}, \Delta u\rangle_1 ,$$

also

$$\alpha = \frac{\langle e_{\text{old}}, \Delta u\rangle_1}{\langle \Delta u, \Delta u\rangle_1}.$$

Wegen $\Delta u = u_{\text{new}} - u_{\text{old}} = u_* - e_{\text{new}} - u_* + e_{\text{old}} = e_{\text{old}} - e_{\text{new}}$ gilt weiter

$$\alpha = \frac{\langle \Delta u, \Delta u\rangle_1 + \langle e_{\text{new}}, \Delta u\rangle_1}{\langle \Delta u, \Delta u\rangle_1}.$$

Da $\Delta u \in \mathcal{R}(I_H^h)$, gilt nach Satz 6.3

$$e_{\text{new}} = K_{h,H} e_{\text{old}} \perp_1 \Delta u,$$

folglich $\langle e_{\text{new}}, \Delta u\rangle_1 = 0$ und $\alpha = 1$. □

6.4 Algebraisch glatte Fehler

Bei geometrischen Mehrgitterverfahren sind von Anfang an Grobgitter festgelegt und der Begriff „glatt“ hat die natürliche Bedeutung: Ein Fehler e^h ist glatt, wenn er keine hochfrequenten Anteile enthält, wenn er sich also sehr gut auf dem Grobgitter durch ein e^H darstellen lässt. Da ein Jacobi- oder Gauss-Seidel-Verfahren hochfrequente Anteile stark dämpft, haben sie den Namen „Glätter“ bekommen.
Bei einem algebraischen Mehrgitterverfahren gibt es zunächst keine vordefinierten Grobgitter. Am Anfang ist nur der „Glätter“ vorgegeben. Damit nun dieser seinen Namen zurecht trägt, wird festgelegt: Alles, was der Glätter (fast) unbearbeitet lässt, ist algebraisch glatt. Also

$$e^h \text{ ist algebraisch glatt} \quad \Leftrightarrow \quad S_h e^h \approx e^h.$$

Man beachte, dass algebraisch glatte Fehler nicht geometrisch glatt sein müssen.
Als nächstes folgt eine Charakterisierung von algebraisch glatten Fehlern. Da keine Grobgittergrößen vorkommen, wird ab hier der Super- und Subscript h weggelassen. Als Glätter wird dazu zunächst ein relaxiertes Jacobi-Verfahren

$$u \to \bar{u}, \qquad \bar{u} = u + \omega D^{-1}(f - Au)$$

betrachtet. Bezüglich der Fehler hat dieser Glätter die Darstellung

$$\bar{e} = e - \omega D^{-1} A(u_* - u) = Se$$

mit

$$S = I - \omega D^{-1} A.$$

Nun gilt

$$e \text{ alg. glatt} \quad \Leftrightarrow \quad Se \approx e \quad \Leftrightarrow \quad e - \omega D^{-1} Ae \approx e.$$

Dies ist genau dann der Fall, wenn e Eigenvektor von $D^{-1}A$ zu einem Eigenwert λ mit $|\lambda| \ll 1$ ist. Gleiches lässt sich auch für ein Gauss-Seidel-Verfahren zeigen.
Ist nun (λ, e) ein Eigenpaar von $D^{-1}A$, dann gilt

$$\|e\|_2^2 = \left\langle D^{-1}Ae, Ae \right\rangle = \lambda \langle e, Ae \rangle = \lambda \|e\|_1^2 \qquad \text{und}$$
$$\|e\|_1^2 = \langle Ae, e \rangle = \left\langle DD^{-1}Ae, e \right\rangle = \lambda \langle e, De \rangle = \lambda \|e\|_0^2 .$$

Ist nun $|\lambda| \ll 1$, dann gilt

$$\|e\|_2 \ll \|e\|_1 \ll \|e\|_0 .$$

Für algebraisch nicht glatte Fehler gilt hingegen

$$\|e\|_2 \approx \|e\|_1 \approx \|e\|_0 .$$

Damit ist eine Charakterisierung für algebraisch glatte Fehler gefunden:

$$e \text{ algebebraisch glatt} \qquad \Leftrightarrow \qquad \|e\|_2 \ll \|e\|_1 \ll \|e\|_0 . \tag{6.13}$$

Aus diesem Kriterium lässt sich leider aber noch nicht ablesen, wie die Interpolation I_H^h konstruiert werden muss, damit in $\mathcal{R}(I_H^h)$ möglichst alle algebraisch glatten Fehler liegen. Um dafür ein Kriterium zu finden, betrachtet man zu einem alg. glatten Fehler e und dessen Residuum $r := Ae$ die Bedingung (6.13) genauer. Aus $\|e\|_2 \ll \|e\|_1$ oder auch

$$\left\langle D^{-1}Ae, Ae \right\rangle \ll \langle Ae, e \rangle \qquad \Leftrightarrow \qquad \left\langle D^{-1}r, r \right\rangle \ll \langle r, e \rangle$$

folgt, dass algebraisch glatte Fehler kleine Residuen haben.
Dies kann man beispielsweise auch am Gauss-Seidel-Verfahren

$$\bar{u}_i = \frac{1}{a_{ii}}\Big(f_i - \sum_{j\neq i} a_{ij}u_j\Big) = \frac{1}{a_{ii}}\Big(a_{ii}u_i + f_i - \sum_{i} a_{ij}u_j\Big) = u_i + \frac{r_i}{a_{ii}}$$

direkt erkennen. In Fehlern lautet es

$$\bar{e}_i = e_i + \frac{r_i}{a_{ii}}.$$

Ist nun $\bar{e}_i \approx e_i$, so muss

$$|r_i| \ll |a_{ii}| \cdot |e_i|$$

sein.
Also gilt für algebraisch glatte Fehler e

$$Ae = r \approx 0$$

bzw.

$$A_{FF}e_F + A_{FC}e_C \approx 0. \tag{6.14}$$

Verwendet man die zweite Ungleichung $\|e\|_1 \ll \|e\|_0$ in (6.13), so erhält man eine weitere Interpretation von $r \approx 0$:

$$\langle Ae, e\rangle \ll \langle e, De\rangle. \tag{6.15}$$

Es folgt nun eine Möglichkeit, die Glättungseigenschaften eines Glätters quantitativ zu beschreiben.

Definition 6.5
Ein Glätter S erfüllt bzgl. einer (symmetrisch positiv definiten) Matrix A die Glättungseigenschaft mit $\sigma > 0$, falls

$$\|Se\|_1^2 \leq \|e\|_1^2 - \sigma\|e\|_2^2 \tag{6.16}$$

für alle e erfüllt ist. Man sagt, der Glätter S erfüllt die Glättungseigenschaft gleichmäßig für alle $A \in \mathcal{A}$ mit einem $\sigma > 0$, wenn (6.16) für alle Matrizen $A \in \mathcal{A}$ erfüllt ist.

In [RS86] wird gezeigt, dass der Gauss-Seidel-Glätter die Glättungseigenschaft (6.16) mit $\sigma = \frac{1}{4}$ gleichmäßig in der Menge der M-Matrizen erfüllt.

Anmerkung
Eine symmetrisch positiv definite Matrix A heißt M-Matrix, falls A auf der Diagonalen nur positive Einträge und auf allen anderen Positionen keine positiven Einträge besitzt.

6.5 Konstruktion des Grobgitters

Ziel dieses Abschnitts ist es, eine Methode zur automatischen Einteilung

$$\Omega = C \cup F \qquad \text{(C/F-Splitting)}$$

anzugeben. Dabei wird davon ausgegangen, dass die symmetrisch positive Matrix A eine M-Matrix ist. Dieser Abschnitt ist eine Zusammenfassung der wesentlichen Ideen zur Konstruktion des Grobgitters, wie sie in [Stü99] beschrieben werden.

Anmerkung
Solange keine Mehrdeutigkeiten entstehen, wird auch in diesem Abschnitt der Super- bzw. Subscript h weggelassen.

An der Interpolationsformel (6.3) wird deutlich, dass für eine effiziente Interpolation sichergestellt sein muss, dass jeder F-Punkt i ($i \in F$) eine ausreichend starke Kopplung zu C-Punkten haben muss. Es werden nur direkte Kopplungen betrachtet, also wird ab hier immer

$$P_i \subset C \cap N_i \qquad \text{für } i \in F$$

gelten. Erfahrungsgemäß liefert eine Interpolation dann gute Ergebnisse, wenn jede F-Variable von ausreichend vielen C-Variablen umgeben (eingekreist) ist.
Auf der anderen Seite will man natürlich die Menge C so klein wie möglich halten, damit das Grobgitterproblem erheblich kleiner wird und schneller behandelt werden kann.

Anmerkung
Ein Punkt i, der von Anfang an keine Nachbarn besitzt, wird sofort F-Punkt und muss nicht interpoliert werden. Die Matrix A hat dann in der Zeile i nur die Diagonale besetzt und die i-te Gleichung lautet

$$a_{ii} u_i = f_i.$$

Eine solche Gleichung löst ein Glätter wie Gauss-Seidel in einer Iteration exakt. Solche Punkte seien aus den folgenden Überlegungen ausgenommen.

Stüben hat eine Methode für das C/F-Splitting angegeben, die davon ausgeht, dass die betrachtete Matrix keine großen positiven Einträge außerhalb der Diagonalen aufweist. Die grundlegende Idee ist, die Punkte i C-Variablen werden zu lassen, die viele noch nicht zugeteilte Punkte oder viele F-Punkte als Nachbarn haben, die stark von i abhängen.

Definition 6.6
Ein Punkt $i \in \Omega$ heißt stark negativ gekoppelt an ein $j \in N_i$ (in Zeichen $i \bullet\!\!-\!\!\circ\, j$), falls

$$-a_{ij} \geq \varepsilon_{\text{str}} \max_{a_{ik}<0} |a_{ik}|$$

gilt. Dabei ist $0 < \varepsilon_{\text{str}} < 1$ ein fest vorgegebener Parameter.

Ferner bezeichne für ein $i \in \Omega$

$$S_i := \{j \in N_i \,|\, i \bullet\!\!-\!\!\circ\, j\}$$

die Menge aller Punkte $j \in N_i$, an die der Punkt i stark gekoppelt ist. Man beachte, dass die Relation $\bullet\!\!-\!\!\circ$ i. Allg. nicht symmetrisch ist. Deshalb bezeichne für ein $j \in \Omega$

$$S_j^T := \{i \in N_j \,|\, j \in S_i\} = \{i \in N_j \,|\, i \bullet\!\!-\!\!\circ\, j\} = \{i \in N_j \,|\, j \circ\!\!-\!\!\bullet\, i\}$$

die Menge aller Nachbarn i von j, welche stark negativ an j gekoppelt sind.
Bei der Erklärung des Zuteilunsprozesses für das C/F-Splitting bezeichne im Folgenden $U \subset \Omega$ die Menge der bisher noch nicht zugeteilten Punkte. Analog bezeichne C und F die Menge der bisher festgelegten C- bzw. F-Variablen. Am Anfang des Zuteilungsprozesses ist demnach $U = \Omega$, $C = F = \varnothing$. Für jeden Punkt $i \in U$ wird ein „Maß“ $\lambda_i \in \mathbb{N}_0$ durch

$$\lambda_i := |S_i^T \cap U| + 2|S_i^T \cap F|$$

festgelegt, welches die Dringlichkeit angibt, mit der dieser Punkt C-Variable werden soll. Der erste Summand misst dabei, wie viele noch nicht zugeteilte Punkte stark negativ an den Punkt i gekoppelt sind. Der zweite Summand misst, wie viele schon bekannte F-Variablen stark negativ an den Punkt i gekoppelt sind.
Die Idee ist, einen Punkt aus U mit höchster Dringlichkeit zur C-Variable zu machen und alle Punkte, die stark an ihn negativ gekoppelt sind, F-Variablen werden zu lassen. Mit U und F verändern sich auch die λ_i. Diese Veränderungen können inkrementell berechnet werden.
Damit sieht das C/F-Splitting so aus:

1. $C \leftarrow \varnothing$; $F \leftarrow \varnothing$; $U \leftarrow \Omega$; $\lambda_i := |S_i^T| \quad \forall i \in U$
2. wähle $i \in U$, so dass $\lambda_i = \max_{k \in U} \lambda_k$; $C \leftarrow C \cup \{i\}$; $U \leftarrow U \backslash \{i\}$
3. $\forall j \in S_i^T \cap U$:
   (a) $F \leftarrow F \cup \{j\}$; $U \leftarrow U \backslash \{j\}$
   (b) $\forall l \in S_j \cap U : \ \lambda_l \leftarrow \lambda_l + 1$
4. $\forall j \in S_i \cap U : \ \lambda_j \leftarrow \lambda_j - 1$
5. falls $U \neq \varnothing$, gehe zu 2.

6.6 Konstruktion der Interpolation

Jetzt geht es darum, den Interpolationsoperator I_H^h bzw. die w_{ik} aus der Formel (6.3) so zu bestimmen, dass die beiden Richtlinien (6.10) und (6.11) möglichst gut erfüllt werden. In diesem Abschnitt wird die Konstruktion der Interpolation nach [RS86] bzw. [Stü99] beschrieben.
Es wird zunächst (6.11) betrachtet und untersucht, wann diese Bedingung (näherungsweise) erfüllt ist. Für die Gültigkeit von (6.11) müssen Fehler nach der Grobgitterkorrektur K gut geglättet werden können. Diese Fehler dürfen also nicht algebraisch glatt sein. Damit darf $\|Ke\|_2$ relativ zu $\|Ke\|_1$ nicht zu klein werden. Es gilt folgender Satz:

Satz 6.7
Erfüllt ein Glätter S die Glättungseigenschaft (6.16) mit einem $\sigma > 0$ und wird das C/F-Splitting zusammen mit dem Interpolationsoperator so konstruiert, dass

$$\|Ke\|_1^2 \leq \tau \|Ke\|_2^2 \tag{6.17}$$

mit einem $\tau > 0$ unabhängig von e gilt, dann ist $\tau \geq \sigma$ und $\|SK\|_1 \leq \sqrt{1 - \sigma/\tau}$.

Beweis

Die Abschätzung

$$\|SKe\|_1^2 \overset{(6.16)}{\leq} \|KE\|_1^2 - \sigma \|Ke\|_2^2 \overset{(6.17)}{\leq} \left(1 - \frac{\sigma}{\tau}\right) \|Ke\|_1^2 \leq \left(1 - \frac{\sigma}{\tau}\right) \|e\|_1^2$$

zeigt die Behauptung. □

Der folgende Satz liefert eine Möglichkeit, die Bedingung (6.17) direkt an der Interpolation zu überprüfen. Dabei ist $\|\cdot\|_{0,F}$ die zugehörige Norm zu $\langle u_F, v_F \rangle_{0,F} := \langle D_{FF} u_F, v_F \rangle$.

Satz 6.8
Wird das C/F-Splitting und die Interpolation I_{FC} so konstruiert, dass für ein $\tau > 0$

$$\|e_F - I_{FC} e_C\|_{0,F}^2 \leq \tau \|e\|_1^2 \qquad \forall e \tag{6.18}$$

gilt, dann ist (6.17) (mit demselben τ) erfüllt.

Beweis

Sei $e \in \mathcal{R}(K)$. Wegen des Punktes 3 im Satz 6.1 gilt $I_h^H A_h e = 0$ und daher

$$\begin{aligned} \|e\|_1^2 = \langle Ae, e\rangle = \left\langle Ae, e - I_H^h e^H \right\rangle = \left\langle D^{-\frac{1}{2}} Ae, D^{\frac{1}{2}}(e - I_H^h e^H)\right\rangle = \\ \leq \left\| D^{-\frac{1}{2}} Ae \right\| \cdot \left\| D^{\frac{1}{2}}(e - I_H^h e^H)\right\| = \|e\|_2 \cdot \left\| e - I - H^h e^H \right\|_0 \end{aligned}$$

für alle e^H. Ist $e = (e_F, e_C)^*$ und setzt man $e^H = e_C$, dann ist

$$\begin{aligned} \|e\|_1^2 \leq \|e\|_2 \cdot \left\| \begin{pmatrix} e_F \\ e_C \end{pmatrix} - \begin{pmatrix} I_{FC} \\ I_{CC} \end{pmatrix} e_C \right\|_0 = \\ = \|e\|_2 \cdot \|e_F - I_{FC} e_C\|_{0,F} \overset{(6.18)}{\leq} \|e\|_2 \sqrt{\tau} \|e\|_1 \end{aligned}$$

und daher

$$\|e\|_1^2 \leq \tau \|e\|_2^2 ,$$

womit die Gleichung (6.17) gezeigt ist. □

Nun wird die Richtlinie (6.10) benutzt, um an eine Interpolationsformel zu gelangen. Von dieser wird anschließend gezeigt, dass sie auch (6.18) erfüllt. Aus der Inklusion (6.10) wird deutlich, dass die Interpolation I_H^h so konstruiert werden muss, dass möglichst algebraisch glatte Fehler in $\mathcal{R}(I_H^h)$ liegen. Nun wurde in Abschnitt 6.4 gezeigt, dass algebraisch glatte Fehler kleine Residuen haben. Deshalb wird zur Konstruktion der Interpolationsformel von der Beziehung (6.14) ausgegangen. Man versucht also die w_{ik} aus (6.3) so zu bestimmen, dass

$$a_{ii} e_i + \sum_{j \in N_i} a_{ij} e_j = 0 \tag{6.19}$$

für alle $i \in F$ gilt. Mit der in Abschnitt 6.5 verwendeten Methode des C/F-Splittings gilt $\varnothing \neq P_i \subset C \cap N_i$ $(i \in F)$ und

$$\frac{1}{\sum_{j \in N_i} a_{ij}} \sum_{j \in N_i} a_{ij} e_j \approx \frac{1}{\sum_{k \in P_i} a_{ik}} \sum_{k \in P_i} a_{ik} e_k$$

ist umso besser erfüllt, je mehr Nachbarn von i in P_i sind. Setzt man diese Näherung in (6.19) ein, so hat man

$$a_{ii} e_i + \underbrace{\frac{\sum_{j \in N_i} a_{ij}}{\sum_{k \in P_i} a_{ik}}}_{=:\alpha_i} \sum_{k \in P_i} a_{ik} e_k = 0.$$

Daraus ergibt sich

$$e_i = \sum_{k \in P_i} -\frac{\alpha_i a_{ik}}{a_{ii}} e_k.$$

Ein Vergleich mit Formel (6.3) zeigt, dass damit eine Interpolationsformel mit

$$w_{ik} = -\frac{\alpha_i a_{ik}}{a_{ii}} \qquad \text{und} \qquad \alpha_i = \frac{\sum_{j \in N_i} a_{ij}}{\sum_{k \in P_i} a_{ik}} \tag{6.20}$$

gefunden wurde. Für diese Interpolation gilt ferner

$$a_{ii}\big(1 - \sum_{k \in P_i} w_{ik}\big) = a_{ii} + \alpha_i \sum_{k \in P_i} a_{ik} = a_{ii} + \sum_{j \in N_i} a_{ij} =: s_i, \tag{6.21}$$

woraus man schließen kann, dass $\sum_{k \in P_i} w_{ik} = 1$, falls $s_i = 0$, also konstante Funktionen exakt interpoliert werden, falls $s_i = 0$. Der nächste Satz zeigt, dass bei passendem C/F-Splitting auch (6.18) erfüllt ist.

Satz 6.9
Ist A eine M-Matrix mit $s_i = \sum_{j \in N_i} a_{ij} \geq 0$ und gibt es für das C/F-Splitting ein $\tau \geq 1$, so dass für alle $i \in F$

$$\sum_{k \in P_i} |a_{ik}| \geq \frac{1}{\tau} \sum_{j \in N_i} |a_{ij}| \tag{6.22}$$

gilt mit $\varnothing \neq P_i \subset C \cap N_i$, dann erfüllt die Interpolationsformel (6.3) mit den Gewichten aus (6.20) die Bedingung (6.18).

Ein Beweis findet sich in [Stü99].

In Abbildung 6.2 wurden das C/F-Splitting und der Interpolationsoperator an einem Beispiel graphisch dargestellt. Die rot markierten Pixel sind die Grobgitterpunkte, die grün markierten Pixel die Feingitterpunkte. Die blauen Verbindungslinien geben an, von welchen Grobgitterpunkten der Interpolationswert bei einem Feingitterpunkt abhängt. In der Abbildung 6.2 ist links der erste Level und rechts der zweite Level zu sehen.

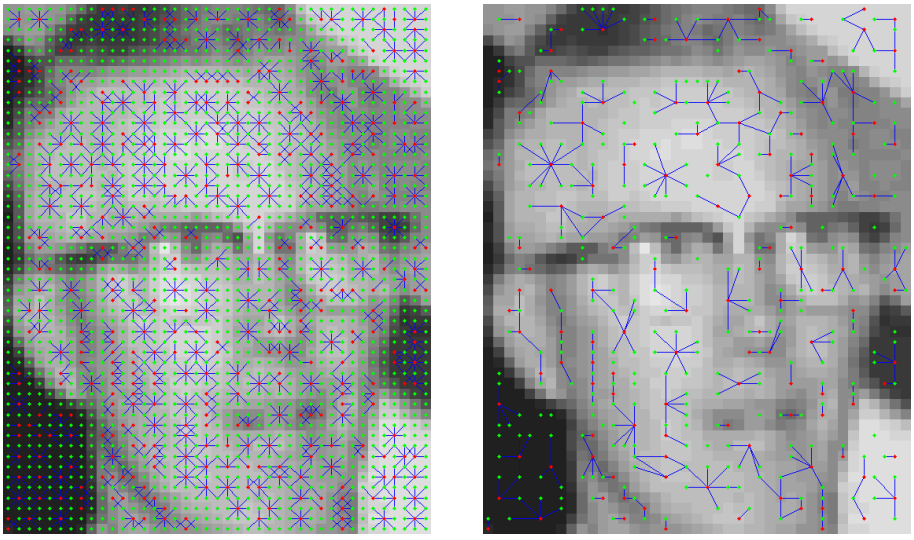

Abbildung 6.2: Graphische Darstellung des C/F-Splittings mit Interpolationsoperator

6.7 Erweiterung auf allgemeinere spd-Matrizen

Die Konstruktion des Grobgitters in Abschnitt 6.5 und die Interpolationsformel mit den Gewichten aus (6.20) in Abschnitt 6.6 sind nur geeignet, falls A eine M-Matrix ist, oder falls A nur wenige, relativ kleine postive Einträge außerhalb der Diagonalen aufweist. Nun wird das C/F-Splitting verallgemeinert und eine andere Idee von Chang, Wong und Fu aus [CWF96] vorgestellt, die bei einer größeren Klasse von Matrizen zu effizienteren AMG-Verfahren führt.
Zunächst werden jetzt alle starken Kopplungen (nicht nur stark negative Kopplungen) berücksichtigt.

Definition 6.10
Ein Punkt $i \in \Omega$ heißt stark gekoppelt an ein $j \in N_i$ (in Zeichen $i \bullet\!\!-\!\!\circ\, j$), falls

$$|a_{ij}| \geq \varepsilon_{\text{str}} \max_{k \neq i} |a_{ik}|$$

gilt. Dabei ist $0 < \varepsilon_{\text{str}} < 1$ ein fest vorgegebener Parameter.

Alle Größen in Abschnitt 6.5 werden nun analog definiert, indem „$\bullet\!\!\bar{-}\!\!\circ$“ durch „$\bullet\!\!-\!\!\circ$“ ersetzt wird. Dann kann die Methode des C/F-Splittings direkt übernommen werden.
Für ein $i \in F$ werden die F-Nachbarschaftspunkte $D_i := N_i \backslash C$ eingeteilt in $D_i^s := D_i \cap S_i$, den stark gekoppelten F-Nachbarn und den Rest $D_i^w := D_i \backslash D_i^s$. Dann kann man für alle $i \in F$ die Gleichung (6.19) in der Form

$$a_{ii} e_i + \sum_{k \in N_i \cap C} a_{ik} e_k + \sum_{j \in D_i^s} a_{ij} e_j + \sum_{j \in D_i^w} a_{ij} e_j = 0 \tag{6.23}$$

schreiben.
Sei nun ein $i \in F$ fest, dann wird in (6.23) jedes e_j ($j \in D_i$) durch eine Kombination von e_k ($k \in C \cap N_i$) approximiert. Auf diese Weise gelangt man zu einer Interpolationsformel der Form (6.3).
Um die Approximation herzuleiten, betrachtet man die Ungleichung (6.15), die von algebraisch glatten Fehlern erfüllt wird. Mit der Umformung

$$\begin{aligned}
\langle e, Ae \rangle &= \sum_{i,j \in \Omega} a_{ij} e_i e_j = \sum_{i \in \Omega} a_{ii} e_i^2 + \sum_{\substack{i,j \in \Omega \\ i \neq j}} a_{ij} e_i e_j = \\
&= \sum_{i \in \Omega} a_{ii} e_i^2 - \sum_{\substack{i,j \in \Omega \\ i \neq j}} |a_{ij}| e_i^2 + \frac{1}{2} \sum_{\substack{i,j \in \Omega \\ i \neq j}} |a_{ij}| e_i^2 + \frac{1}{2} \sum_{\substack{i,j \in \Omega \\ i \neq j}} |a_{ij}| e_j^2 + \sum_{\substack{i,j \in \Omega \\ i \neq j}} a_{ij} e_i e_j = \\
&= \sum_{i \in \Omega} a_{ii} e_i^2 - \sum_{\substack{i,j \in \Omega \\ i \neq j}} |a_{ij}| e_i^2 + \frac{1}{2} \sum_{\substack{i,j \in \Omega \\ i \neq j}} |a_{ij}| \left[e_i^2 + e_j^2 + 2 \operatorname{sign}(a_{ij}) e_i e_j \right] = \\
&= \sum_{i \in \Omega} a_{ii} e_i^2 - \sum_{\substack{i,j \in \Omega \\ i \neq j}} |a_{ij}| e_i^2 + \frac{1}{2} \sum_{\substack{i,j \in \Omega \\ i \neq j}} |a_{ij}| \left[e_i + \operatorname{sign}(a_{ij}) e_j \right]^2
\end{aligned}$$

sieht man, dass (6.15) erfüllt ist, falls

$$\frac{1}{2} \sum_{\substack{i,j \in \Omega \\ i \neq j}} |a_{ij}| \left[e_i + \operatorname{sign}(a_{ij}) e_j \right]^2 \approx 0 \tag{6.24}$$

gilt. Daran kann man sehr schön die Eigenschaften von algebraisch glatten Fehlern ablesen: Betrachtet man ein $i \in \Omega$ und ein $j \in N_i$ und ist $a_{ij} \ll 0$, so wird der zu i und j gehörige Summand in (6.24) klein, wenn $e_i \approx e_j$. Dies begründet die folgende Aussage.

> Entlang starker negativer Kopplungen ist ein algebraisch glatter Fehler auch geometrisch glatt.

Ist hingegen $a_{ij} \gg 0$, so wird der zugehörige Summand klein, falls $e_i \approx -e_j$ und man sagt:

> Entlang starker positiver Kopplungen führt ein algebraisch glatter Fehler (geometrische) Oszillationen aus.

Die beiden Aussagen und die geometrische Vermutung, dass für ein $i \in \Omega$ ein $j \in N_i$ umso näher an i liegt, je größer $|a_{ij}|$ ist, werden verwendet, um eine Interpolationsformel zu erzeugen.
Dazu bezeichnet

$$S_{ij} := \{k \in C \cap N_i \,|\, a_{jk} \neq 0\}$$

die Menge aller C-Nachbarn von i, die einen Einfluss auf j haben. Ferner sei $l_{ij} := |S_{ij}|$. Für $i \in F$ und $j \in F \cap N_i$ wird mit Hilfe von

$$\xi_{ij} := \frac{-\sum_{k \in S_{ij}} a_{jk}}{\sum_{k \in S_{ij}} |a_{jk}|}$$

entschieden, wie sich ein algebraisch glatter Fehler zwischen i und j verhält. Er wird als geometrisch glatt angenommen, wenn $\xi_{ij} \geq \frac{1}{2}$ und $a_{ij} < 0$. Wegen den geometrischen Vermutungen ist

$$N := \frac{1}{l_{ij}} \sum_{k \in S_{ij}} |a_{jk}|$$

die durchschnittliche Nähe (Abstand) von j zu S_{ij}. Damit kann man mit

$$\eta_{ij} := \frac{|a_{ji}|}{N} = \frac{|a_{ji}| l_{ij}}{\sum_{k \in S_{ij}} |a_{jk}|}$$

die Lage von i, j und S_{ij} beschreiben:

$\eta_{ij} \gg 1$	j liegt nahe an i, aber weit weg von S_{ij}
$\eta_{ij} \ll 1$	j liegt weit weg von i, aber nahe bei S_{ij}

Durch

$$g_{jk} = \frac{|a_{jk}|}{\sum_{l \in S_{ij}} |a_{jl}|}$$

wird der Einfluss von $k \in S_{ij}$ auf j quantifiziert.
Aufgrund der geometrischen Vorüberlegungen sieht nun die Approximation eines $j \in D_i^w$ folgendermaßen aus:

$$e_j = \begin{cases} e_i & \text{falls } l_{ij} = 0 \text{ und } a_{ij} < 0 \qquad \text{(Fall I)} \\ -e_i & \text{falls } l_{ij} = 0 \text{ und } a_{ij} > 0 \qquad \text{(Fall II)} \\ 2\sum_{k \in S_{ij}} g_{jk} e_k - e_i & \text{falls } l_{ij} > 0,\ \xi_{ij} \geq \frac{1}{2} \text{ und } a_{ij} < 0 \qquad \text{(Fall III)} \\ \sum_{k \in S_{ij}} g_{jk} e_k & \text{sonst} \qquad \text{(Fall IV)} \end{cases} \tag{6.25}$$

Für ein $j \in D_i^s$ wird bzgl. der Lage von i, j und S_{ij} weiter differenziert:

$$e_j = \begin{cases} 2\sum_{k\in S_{ij}} g_{jk}e_k - e_i & \text{falls } \eta_{ij} < \frac{3}{4},\ \xi_{ij} \geq \frac{1}{2} \text{ und } a_{ij} < 0 \quad \text{(Fall V)} \\ \frac{1}{2}(\sum_{k\in S_{ij}} g_{jk}e_k + e_i) & \text{falls } \eta_{ij} > 2,\ \xi_{ij} \geq \frac{1}{2} \text{ und } a_{ij} < 0 \quad \text{(Fall VI)} \\ \sum_{k\in S_{ij}} g_{jk}e_k & \text{sonst} \quad \text{(Fall VII)} \end{cases} \tag{6.26}$$

Teilt man nun für das feste $i \in F$ die Menge D_i in die vier Partitionen

$$\begin{aligned} D_i^{(1)} &:= \{\, j \in D_i \,|\, \text{es tritt Fall I oder Fall II auf} \,\} \\ D_i^{(2)} &:= \{\, j \in D_i \,|\, \text{es tritt Fall IV oder Fall VII auf} \,\} \\ D_i^{(3)} &:= \{\, j \in D_i \,|\, \text{es tritt Fall III oder Fall V auf} \,\} \\ D_i^{(4)} &:= \{\, j \in D_i \,|\, \text{es tritt Fall VI auf} \,\} \end{aligned}$$

ein und setzt man (6.25) und (6.26) in (6.23) ein, so erhält man für alle $i \in F$ die Interpolationsformel

$$e_i = \sum_{k\in C\cap N_i} w_{ik}e_k,$$

wobei

$$\begin{aligned} w_{ik} &= \frac{-\bar{a}_{ik}}{\bar{a}_{ii}} \\ \bar{a}_{ii} &= a_{ii} - \sum_{j\in D_i^{(1)}} |a_{ij}| - \sum_{j\in D_i^{(3)}} a_{ij} + \frac{1}{2}\sum_{j\in D_i^{(4)}} a_{ij} \\ \bar{a}_{ik} &= a_{ik} + \sum_{j\in D_i^{(2)}} a_{ij}g_{jk} + 2\sum_{j\in D_i^{(3)}} a_{ij}g_{jk} + \frac{1}{2}\sum_{j\in D_i^{(4)}} a_{ij}g_{jk} \end{aligned} \tag{6.27}$$

Konvergenzaussagen, die auf den Sätzen 6.7 und 6.8 beruhen, findet man in [CWF96].

7 Numerische Beispiele

Zum Abschluss wird jetzt der in den vorigen Kapiteln dargestellte Algorithmus auf einige Beispiele angewendet.

Der Algorithmus benötigt die vier Parameter $A_{\min}$, $A_{\max}$, η und Δu (vgl. Abschnitt 5.6) zur Steuerung der verschachtelten Iterationen. Die Parameter $A_{\min}$ und $A_{\max}$ als Entscheidungskriterien für die Güte eines GNC-Schrittes und die Parameter η und Δu für das Abbruchkriterium der HQR-Iteration. Als Faustformel für $A_{\max}$ wurde in allen Beispielen $A_{\max} \approx \alpha mn \cdot 1\%$ verwendet. Sie basiert auf folgender Überlegung: Wenn die Energie des Differenzbildes A nur von unterschiedlichen Kantenmengen der beiden Vergleichsbilder rührt, dann ist der GNC-Schritt akzeptabel, wenn in den beiden zu vergleichenden Bildern mindestens 99% der Kanten übereinstimmen.

Alle Zeitangaben in diesem Kapitel wurden mit bei einem Intel Pentium IV mit 1,8 GHz Taktfrequenz gemessen.

7.1 Autofelge

Im Abschnitt 2.2 wurde zur Erläuterung der Parameter λ und h_0 auf ein 128×128 Pixel großes Bild einer Autofelge der Algorithmus mit verschiedenen Parameterpaaren (λ, h_0) angewendet. Die Parameter zur Steuerung, sowie der Aufwand, sind in folgender Tabelle aufgelistet. In der Spalte mit der Überschrift „#GNC“ stehen Einträge der Form $a|b$. Das bedeutet, es waren a erfolgreiche GNC-Schritte nötig und b GNC-Schritte wurden verworfen. In der Spalte mit dem Titel „#HQR“ steht die Anzahl aller insgesamt benötigten HQR-Iterationen.

$\lambda,\ h_0$	Δu	$A_{\min}$	$A_{\max}$	η	#GNC	#HQR	Zeit (s)
$\lambda = 5,\ h_0 = 8$	0,50	$15 \cdot 10^3$	$30 \cdot 10^3$	0,10	12|3	314	67
$\lambda = 5,\ h_0 = 16$	0,50	$20 \cdot 10^3$	$10 \cdot 10^4$	0,15	16|2	349	73
$\lambda = 10,\ h_0 = 8$	1,00	$10 \cdot 10^3$	$60 \cdot 10^3$	0,15	17|2	274	61
$\lambda = 10,\ h_0 = 16$	1,00	$30 \cdot 10^3$	$20 \cdot 10^4$	0,20	24|3	420	92
$\lambda = 15,\ h_0 = 8$	1,00	$30 \cdot 10^3$	$10 \cdot 10^4$	0,20	21|7	437	98
$\lambda = 15,\ h_0 = 16$	1,25	$50 \cdot 10^3$	$30 \cdot 10^4$	0,20	23|5	432	97

7.2 Verrauschte Buchstaben

Als nächstes werden verrauschte Buchstaben betrachtet. In Abbildung 7.1 sind das 100×60 Pixel große Originalbild und einige Zwischenergebnisse zu sehen. Die verwendeten Parameter und der Aufwand sind aus der folgenden Tabelle ersichtlich:

$\lambda,\ \alpha$	Δu	$A_{\min}$	$A_{\max}$	η	#GNC	#HQR	Zeit (s)
$\lambda = 8,\ \alpha = 20 \cdot 10^3$	2,0	$1 \cdot 10^5$	$12 \cdot 10^5$	0,20	20|2	456	34

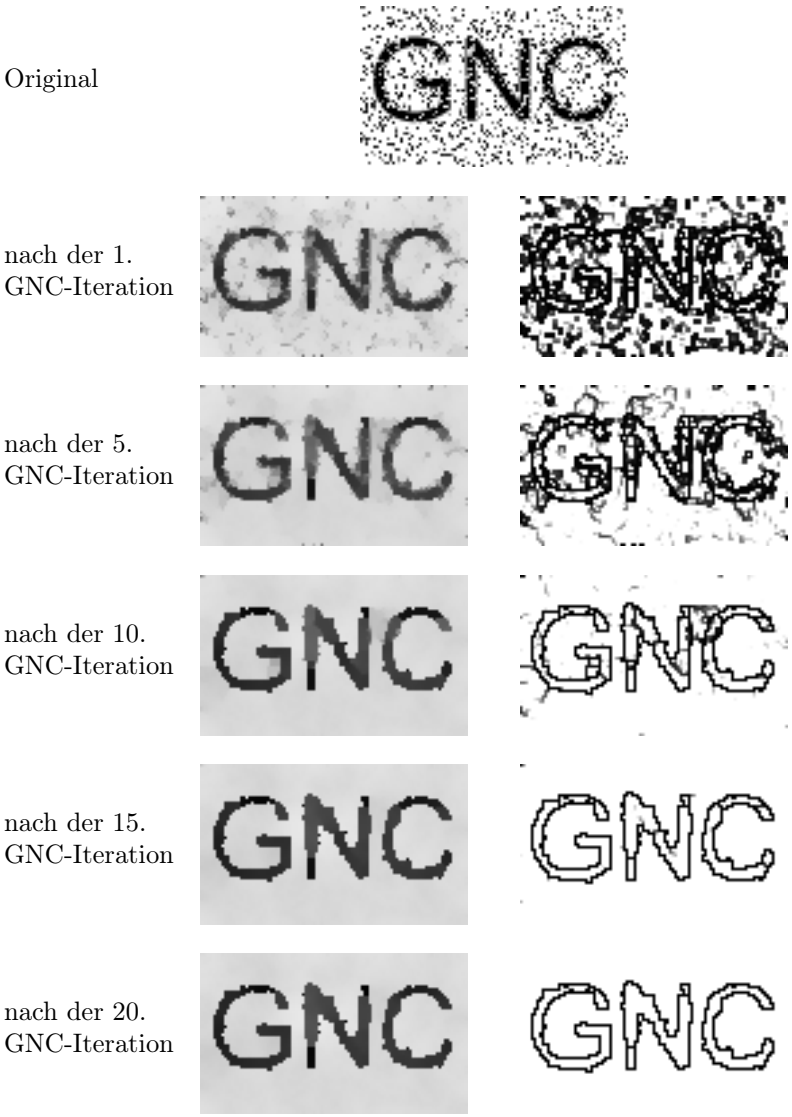

Abbildung 7.1: Verrauschte Buchstaben (Zwischenergebnisse)

7.3 Luftaufnahme

In Abbildung 7.2 sieht man ein 256×256 Pixel großes Luftbild einer Straßenkreuzung und das vom Algorithmus abgelieferte Bild. Die Kantenmenge ist in Abbildung 7.3 dargestellt. Die verwendeten Parameter und der Aufwand waren:

$\lambda,\ \alpha$	Δu	$A_{\min}$	$A_{\max}$	η	#GNC	#HQR	Zeit (s)
$\lambda = 10,\ \alpha = 1 \cdot 10^3$	1,0	$1 \cdot 10^5$	$8 \cdot 10^5$	0,50	25|2	334	309

Abbildung 7.2: Luftbildaufnahme

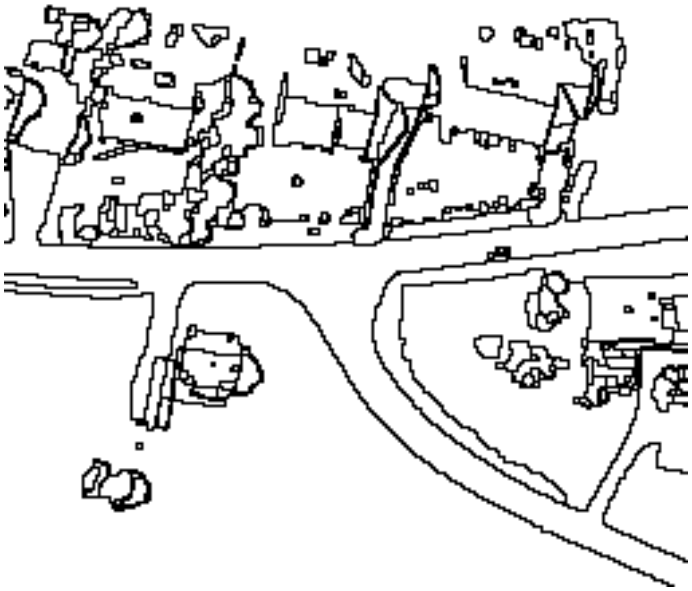

Abbildung 7.3: Luftbildaufnahme (gefundene Kanten)

Abbildungsverzeichnis

Literaturverzeichnis

[Amb89] L. Ambrosio, *Variational Problems in SBV and Image Segmentation*, Acta Appl. Math. 17 (1989), no. 1, 1–40.

[AT90] L. Ambrosio and V.M. Tortorelli, *Approximation of Functionals Depending on Jumps by Elliptic Functionals via* Γ*-Convergence*, Comm. Pure Appl. Math. 43 (1990), 999–1036.

[BC94] G. Bellettini and A. Coscia, *Discrete Approximation of a Free Discontinuity Problem*, Numer. Funct. Anal. Optim. 15 (1994), 201–224.

[BC00] B. Bourdin and A. Chambolle, *Implementation of an adaptive finite-element approximation of the Mumford-Shah functional*, Numer. Math. 85 (2000), 609–646.

[BDM97] A. Braides and G. Dal Maso, *Non-Local Approximation of the Mumford-Shah Functional*, Calc. Var. Partial Differential Equations 5 (1997), 293–322.

[Bon96] A. Bonnet, *On the regularity of edges in image segmentation*, Ann. Inst. H. Poincaré Anal. Non Linéaire 13 (1996), no. 4, 485–528.

[Bor00] F. Bornemann, *Towards a Multigrid Graduated Non-Convexity Algorithm for Free-Discontinuity Problems*, International Conference on Mathematical Modeling and Scientific Computing, April 2000.

[BZ87] A. Blake and A. Zisserman, *Visual Reconstruction*, MIT Press, 1987.

[CBFAB97] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Berlaud, *Deterministic Edge-Preserving Regularization in Computed Imaging*, IEEE Trans. Image Processing 6 (1997), no. 2, 298–311.

[CDM99] A. Chambolle and G. Dal Maso, *Discrete Approximation of the Mumford-Shah Functional in Dimension Two*, M2AN Math. Model. Numer. Anal. 33 (1999), no. 4, 651–672.

[CWF96] Q. Chang, Y.S. Wong, and H. Fu, *On the Algebraic Multigrid Method*, J. Comput. Phys. 125 (1996), 279–292.

[DG88] E. De Giorgi, *Free discontinuity problems in calculus of variations. Analyse Mathématique et Applications*, Gauthier-Villars, 1988.

[DGCL89] E. De Giorgi, M. Carriero, and A. Leaci, *Existence Theorem for a Minimum Problem with Free Discontinuity Set*, Arch. Rational Mech. Anal. 108 (1989), 195–218.

[DM93] G. Dal Maso, *An Introduction to* Γ*-convergence*, Birkhäuser, Boston, 1993.

[DMMS92] G. Dal Maso, J.-M. Morel, and S. Solimini, *A variational method in image segmentation: Existence and approximation results*, Acta Math. 168 (1992), 89–151.

[ET76] I. Ekeland and R. Temam, *Convex analysis and variational problems*, North-Holland, Amsterdam, 1976.

[GG84] S. Geman and D. Geman, *Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images*, IEEE Trans. Patt. Anal. Machine Intell., Vol. PAMI-6 (1984), 721–741.

[GY95] D. Geman and C. Yang, *Nonlinear Image Recovery with Half-Quadratic Regularization*, IEEE Trans. Image Processing 4 (1995), no. 7, 932–946.

[Hac85] W. Hackbusch, *Multi-Grid Methods and Applications*, Springer-Verlag, 1985.

[MS89] D. Mumford and J. Shah, *Optimal Approximation by Piecewise Smooth Functions and Associated Variational Problems*, Comm. Pure Appl. Math. 42 (1989), 577–685.

[MS95] J.-M. Morel and S. Solimini, *Variational Methods in Image Segmentation*, Birkhäuser, 1995.

[Neg99] M. Negri, *The Anisotropy Introduced by the Mesh in the Finite Element Approximation of the Mumford-Shah Functional*, Numer. Funct. Anal. Optim. 20 (1999), 957–982.

[Rit94] K. Ritter, *Nichtlineare Optimierung*, Technische Universität München, Institut für Angewandte Mathematik und Statistik, 1994.

[Roc97] R.T. Rockafellar, *Convex Analysis*, NJ: Princeton University Press, 1997.

[RS86] J.W. Ruge and K. Stüben, *Algebraic Multigrid (AMG)*, In "Multigrids Methods" (Mc-Cormick, S.F., ed.), SIAM, Frontiers in Applied Mathematics, Vol 5, 1986.

[Stü99] K. Stüben, *Algebraic Multigrid (AMG): An Introduction with Applications*, Tech. Report 53, GMD, 1999.