

Technische Universität München
Zentrum Mathematik

Variationelle Bildverbesserung mit BV-Funktionalen: Gamma-Konvergenz und Algorithmen

Christian Ludwig

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen
Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Dr. Jürgen Richter-Gebert
Prüfer der Dissertation: 1. Univ.-Prof. Dr. Folkmar Bornemann
2. Univ.-Prof. Dr. Martin Brokate
3. Univ.-Prof. Dr. Otmar Scherzer
Universität Innsbruck / Österreich
(schriftliche Beurteilung)

Die Dissertation wurde am 13.05.2008 bei der Technischen Universität eingereicht und durch die Fakultät für Mathematik am 16.09.2008 angenommen.

Danksagung

Diese Dissertation wäre ohne die große Unterstützung von verschiedenen Seiten nicht zum Abschluss gekommen. An erster Stelle gilt mein Dank meinem Betreuer Herrn Prof. Dr. Folkmar Bornemann für seine Unterstützung, seinen Einsatz und für seine konstruktiven Ratschläge, wobei er mir stets große Freiheiten ließ.

Auch dem von der Deutschen Forschungsgemeinschaft (DFG) geförderten Graduiertenkolleg „Angewandte Algorithmische Mathematik“ der Technischen Universität München, dem ich gut zwei Jahre als Stipendiat angehörte, gilt mein Dank.

Ganz besonderes danke ich Frau Carolyn Kalender für ihre grenzenlose Bereitschaft zum Korrekturlesen und zur Diskussion, sowie für ihre vielen Anregungen und Motivationshilfen. Sie trug damit einen erheblichen Anteil zum Gelingen dieser Arbeit bei.

Nicht zu Letzt geht auch ein herzliches Dankeschön an meine Familie für die Unterstützung in vielerlei Hinsicht.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Kontinuierliche Problemstellung	2
1.2	Diskrete Problemstellung	6
1.3	Überblick	7
2	Konvergenzbetrachtungen	9
2.1	Notation, Grundlagen	9
2.1.1	Maßtheoretische Grundlagen	10
2.1.2	Sobolev Funktionen	12
2.1.3	Funktionen beschränkter Variation	13
2.2	Motivation: Freies Randwertproblem im Eindimensionalen . .	17
2.2.1	Notation	18
2.2.2	Problemstellung und Ziel	18
2.2.3	Gamma-Konvergenz	19
2.2.4	Die \liminf -Ungleichung	21
2.2.5	Die \limsup -Ungleichung	24
2.2.6	Beispiel	24
2.3	Funktionale auf quadratischen Gittern	28
2.3.1	Notation	28
2.3.2	Problemstellung und Ziel	28
2.3.3	Überblick über den Gamma-Konvergenz Beweis	29
2.3.4	Strecken auf uniformen Gittern	29
2.3.5	Funktionen $w \in \mathcal{W}(\Omega)$ auf uniformen Gittern	31
2.3.6	Die \limsup -Ungleichung auf $\mathcal{W}(\Omega)$	32
2.3.7	Die \limsup -Ungleichung auf $SBV^\alpha(\Omega)$	34
2.3.8	Die \liminf -Ungleichung	35
2.3.9	Zusammenfassung	42

3	Existenzaussagen	45
3.1	Existenzresultat in SBV	45
3.2	Existenzresultat in BV^*	47
3.3	Existenzresultat in $W^{1,p}$	52
3.4	Zusammenfassung: Existenzsätze	54
4	Algorithmen	57
4.1	Problemstellung	57
4.2	Wohlgestelltheit und Skalierungsinvarianz	58
4.3	Zusammenhang mit Graphcuts	60
4.3.1	Expansion Moves	61
4.3.2	Swap Moves	65
4.4	MRF und MAP Schätzer	67
4.5	Block-Update-Methode	69
4.5.1	Gauß-Seidel-Verfahren	69
4.5.2	Konstruktion der Grobgitterknoten	70
4.5.3	Konstruktion der Grobgitterkanten	71
4.5.4	Interpolation	72
4.5.5	Update-Mechanismus	72
4.5.6	V-Zyklus	72
4.5.7	Laufzeitverhalten	72
4.5.8	Beispiel	73
4.6	Graduated Non-Convexity	74
4.6.1	Homotopiekonstruktion: Vom konvexen zum gesuchten Funktional	75
4.6.2	Iterationsverfahren	77
4.7	Vergleich mit Mumford-Shah Funktional	79
	Abbildungsverzeichnis	83
	Symbol- und Notationsverzeichnis	85
	Literaturverzeichnis	87

Kapitel 1

Einleitung

Seit mehreren Jahren ist die Bildverbesserung bzw. Bildentrauschung (image denoising) ein aktives Forschungsgebiet der Mathematik. In vielen Anwendungen ist dieses Gebiet eine notwendige und wichtige Vorstufe, ein sogenannter preprocessing step, zu weiteren Algorithmen der Bildverarbeitung oder auch Bilderkennung bzw. -interpretation.

Die Aufgabe der Bildentrauschung besteht darin, aus einem gegebenem gestörten Bild v ein „besseres“ Bild u mit weniger Störungen zu erzeugen, welches „möglichst nahe“ am Originalbild v liegt. Beide Ziele sind (in den meisten Fällen) konkurrierend und müssen natürlich noch mathematisch präzise gefasst werden. Die möglichen Ursachen, warum v verrauscht sein kann, sind vielfältig: Lichtbrechung, unterschiedliche Empfindlichkeit einzelner Aufnahmesensoren, Quantisierungseffekte oder viele andere Einflüsse. Je nach vorhandener (oder auch erwarteter) Störung gibt es unterschiedliche Herangehensweisen, solche Störungen algorithmisch zu entfernen.

In dieser Arbeit werden nur Methoden betrachtet, die sich in der Form

$$F(u; v) = \min_{u \in U} \quad (1.1)$$

mit noch festzulegenden Funktionalen F und Funktionenräumen

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

mit $\Omega \neq \emptyset$ schreiben lassen. Die Werte im Bildraum \mathbb{R} können als Intensität, Helligkeit oder auch als Grauwert interpretiert werden. Natürlich können außer \mathbb{R} auch andere Bildräume betrachtet werden. So wird bei Mehrkanalbildern mit n (Farb-)kanälen oft der \mathbb{R}^n als Bildraum verwendet. Wir werden uns hier aber meist auf Grauwertbilder beschränken.

Durch diese Art der Problemstellung können, bei gegebenem F und U , zwei Kandidaten u_1 und u_2 für ein Eingabebild v objektiv anhand der Funktionalwerte $F(u_1, v)$ bzw. $F(u_2, v)$ verglichen werden. Manchmal heißt F auch

Energiefunktional. Die Probleme der Form (1.1) lassen sich in zwei große Klassen einteilen: diskrete und kontinuierliche Probleme.

- Bei *diskreten Problemen* besteht Ω aus endlich vielen Elementen (meist auch als Pixel bezeichnet). Damit ist (1.1) ein Optimierungsproblem über einem endlichdimensionalen Raum U , also ein diskretes Optimierungsproblem.
- Hingegen ist bei *kontinuierlichen Problemen* Ω nicht mehr endlich. Bei uns ist stets $\Omega \subset \mathbb{R}^2$ einfach zusammenhängend und beschränkt. (Später werden noch weitere Bedingungen an den Rand $\partial\Omega$ von Ω hinzu kommen.) In diesem Fall ist (1.1) ein Optimierungsproblem über einen unendlichdimensionalen Funktionenraum U . Häufig heißt (1.1) dann auch Variationsproblem.

Oft kann F durch

$$F(u; v) = F_D(u, v) + F_R(u)$$

in zwei Teilfunktionale F_D und F_R aufgespaltet werden. Dabei ist F_D ein Datenfitterm, welcher die Abweichung von u und v misst, und F_R ein regularisierender Term, der für die notwendige Glattheit oder auch Regularität von u sorgt.

1.1 Kontinuierliche Problemstellung

Gängige und viel untersuchte Wahlen für F_D bzw. F_R bei der kontinuierlichen Problemstellung sind

$$F_D(u, v) = \lambda \|u - v\|_{L^\gamma(\Omega)}^\gamma \quad \text{und} \quad F_R(u) = \int_{\Omega} g(\nabla u) \, dx \quad (1.2)$$

mit $\gamma \geq 1$ (meist wird $\gamma = 2$ gewählt). Dabei weist g das asymptotische Verhalten $g(z) \sim \|z\|^\alpha$ für $\|z\| \rightarrow \infty$ und $\alpha \geq 0$ auf. $\lambda > 0$ ist ein Parameter zur Gewichtung. α wirkt sich entscheidend auf die theoretischen Untersuchungen (Wahl geeigneter Funktionenräume U sowie Topologien, die Existenz- bzw. Eindeutigkeitsaussagen ermöglichen) sowie auf die Eigenschaften einer Optimallösung aus.

Für $\alpha \geq 2$ liefern große $\|\nabla u\|_2$ einen so großen Beitrag zum Funktional, dass große Gradienten vermieden werden und die Optimallösung sehr „verwaschen“ aussieht. Eventuell vorhandene Kanten im Eingabebild v werden stark ausgeschmiert. Dieses Phänomen ist im Fall $\gamma = \alpha = 2$ für

$$F(u, v) = \lambda \|u - v\|_{L^2(\Omega)}^2 + \int_{\Omega} \|\nabla u\|_2^2 \, dx \quad (1.3)$$

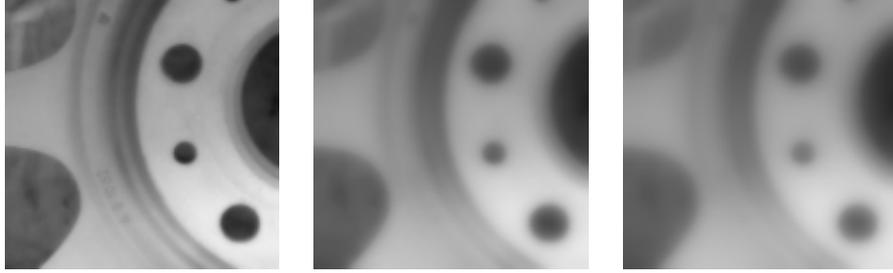


Abbildung 1.1: Optimallösung u für (1.3) mit $\lambda = 1, \frac{1}{10}, \frac{1}{20}$ (Das Eingabebild wurde freundlicherweise von der MVTec Software GmbH zur Verfügung gestellt)

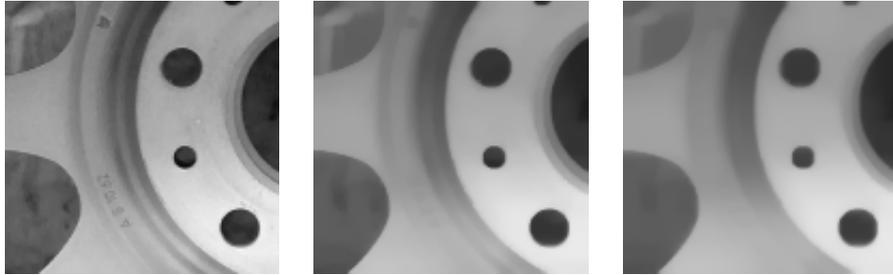


Abbildung 1.2: Approximation an die Optimallösung u für (1.4) mit $\lambda = 1, \frac{1}{10}, \frac{1}{20}$

in Abbildung 1.1 zu sehen. Dieser Fall erfreut(e) sich deshalb großer Beliebtheit, da hier die Euler-Lagrange-Gleichungen linear sind.

Für $\alpha > 1$ gibt es eine Lösungstheorie im Raum der Sobolev-Funktionen $U = W^{1,\alpha}(\Omega)$ (die Notation und einige Eigenschaften der Sobolev-Räume sind im Abschnitt 2.1.2, weitergehende Informationen in [AF03], zu finden). Für eine ausführliche Darstellung dieser Theorie sei auf [Oeh02] verwiesen: Existenz und Eindeutigkeit [Oeh02, Theorem 2.13]; stetige Abhängigkeit von den Anfangsdaten (bzgl. einer geeigneten Topologie) [Oeh02, Theorem 2.16]. Für $1 < \alpha < 2$ sind in einer Optimallösung größere $\|\nabla u\|_2$ möglich als für $\alpha \geq 2$. Deshalb tritt in diesem α -Bereich der oben beschriebene Effekt des Ausschmierens von Kanten nicht bzw. nur schwach auf. Betrachtet man etwa für $\gamma = 1.8$ und $\alpha = 1.2$

$$F(u, v) = \lambda \|u - v\|_{L^{1.8}(\Omega)}^{1.8} + \int_{\Omega} \|\nabla u\|_{1.2}^{1.2} dx, \quad (1.4)$$

so wird dies in Abbildung 1.2 im direkten Vergleich zu Abbildung 1.1 besonders deutlich (der für die Berechnung verwendete Algorithmus wird in

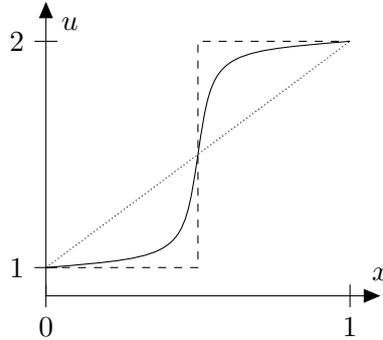


Abbildung 1.3: Drei Funktionen mit derselben Variation

Abschnitt 4.5 beschrieben). Ein $u \in W^{1,\alpha}(\Omega)$ lässt aber keine approximativen Unstetigkeitsmengen mit positivem \mathcal{H}^1 -Hausdorffmaß zu (vgl. (2.1) bzw. Abschnitt 2.1.3 für die Begriffsbildungen). Vielmehr ist jedes $u \in W^{1,\alpha}(\Omega)$ (genauer: es gibt einen Repräsentanten in der Äquivalenzklasse u) auf fast allen achsenparallelen Strecken in Ω absolutstetig [Zie89, Theorem 2.1.4]. Folglich können Sobolev-Funktionen keine Sprünge (in obigem Sinne) aufweisen und man darf deshalb bei $\alpha > 1$ bei der Optimallösung keine Kanten erwarten.

Völlig anders sieht die Lage bei $\alpha = 1$ aus. Rudin, Osher und Fatemi haben in [ROF92] dieses Funktional motiviert. Es wird daher manchmal auch als ROF Funktional bezeichnet. Der passende Rahmen für theoretische Untersuchungen ist hier der Raum der Funktionen beschränkter Variation $BV(\Omega)$ (siehe Abschnitt 2.1.3). Da für hinreichend glatte u

$$\int_{\Omega} \|\nabla u\|_2 \, dx = |Du|(\Omega) = \text{Var}(u, \Omega)$$

gilt, liefert die (totale) Variation $\text{Var}(u, \Omega)$ (vgl. Definition 2.10) die geeignete Verallgemeinerung des Integralterms der linken Seite. Daher wird die Minimierung dieses Funktionals auch als TV-Minimierung bezeichnet. Wie an der Coarea-Formel im Raum $BV(\Omega)$ (Satz 2.14) zu sehen ist, impliziert die Variation $\text{Var}(u, \Omega)$ keine Beschränkungen an die Norm des Gradienten. Dies lässt sich auch im Eindimensionalen veranschaulichen. In Abbildung 1.3 gilt für alle drei Funktionen u_i : $\text{Var}(u_i, [0; 1]) = 1$. Bei den beiden hinreichend glatten, folgt dies direkt aus $\text{Var}(u_i, [0; 1]) = \int_0^1 |u_i'(x)| \, dx = \int_0^1 u_i'(x) \, dx = u_i(1) - u_i(0) = 1$. Die Coarea-Formel zeigt auch, dass sowohl die Sprunghöhen, als auch die Geometrie der Level-Sets $\{u > t\}$ eine Rolle spielen und Einfluss auf den Variationsterm haben. In [CL97] ist die Theorie zur Existenz und Eindeutigkeit von Lösungen zum TV-Funktional zu finden.

Eine weitere Klasse von regularisierenden F_R kann auf mehrere Weisen mo-

tiviert werden. Zum einen kann man vom Fall $g(z) = \|z\|^\alpha$ mit $\alpha > 1$ ausgehen, der weiter oben schon beschrieben wurde. Um Kanten zu ermöglichen, werden nicht nur Sobolev-Funktionen, sondern eine größere Klasse von Funktionen zugelassen, etwa $u \in SBV(\Omega)$. In diesem Fall benötigt man aber einen zusätzlichen Term, der die Sprungmenge und die Sprunghöhen kontrolliert, wie zum Beispiel

$$F_R(u) = \int_{J_u} |u^+ - u^-|^\beta d\mathcal{H}^1 + \int_{\Omega} \|\nabla u\|^\alpha dx \quad (1.5)$$

mit einem $0 \leq \beta < 1$ (dabei bezeichnet J_u die approximative Sprungmenge von u , u^+ bzw. u^- die einseitigen approximativen Grenzwerte entlang J_u und \mathcal{H}^d das d -dimensionale Hausdorff-Maß; die Definitionen befinden sich wieder in Abschnitt 2.1.3). Im Falle der TV-Regularisierung war ein solcher Zusatzterm nicht nötig, da durch die Coarea-Formel Sprungmenge und -höhe implizit kontrolliert wurden. Ein anderer Weg, um ebenfalls bei (1.5) zu landen, besteht darin, für $u \in SBV(\Omega)$ und $0 \leq \beta < 1$ eine sinnvolle Verallgemeinerung von $\int \|\nabla u\|^\beta$ für $\|\nabla u\| \rightarrow \infty$ zu suchen. Es ist aber nicht offensichtlich zu sehen, warum bei diesem Grenzübergang gerade obiges Integral über J_u entstehen soll bzw. kann. Wie dieser Anteil in „natürlicher Weise“ ins Spiel kommt, kann in Abschnitt 2.3.6 betrachtet werden.

Ein sehr bekannter Vertreter dieser Klasse ist das Mumford-Shah Funktional [MS89], welches in der schwachen Formulierung für ein $u \in SBV(\Omega)$ durch

$$F(u; v) = \lambda \|u - v\|_{L^2(\Omega)}^2 + \mu \mathcal{H}^1(S_u) + \int_{\Omega} \|\nabla u\|_2^2.$$

gegeben ist. Dabei sind $\lambda, \mu > 0$ wieder Gewichtsparameter. Bei diesem Funktional hat F_R genau die Form von (1.5) mit $\alpha = 2$ und $\beta = 0$ (da $\int_{J_u} 1 d\mathcal{H}^1 = \mathcal{H}^1(S_u)$). In [AFP00] befasst sich das Kapitel 6 mit der Untersuchung des Mumford-Shah Funktionals und mit [Dav05] gibt es ein umfangreiches Buch zu weiteren Aussagen über die Regularität sowie zur Struktur von Optimallösungen. Dieses Funktional wurde ursprünglich als Funktional zur Bildverbesserung bzw. -segmentierung vorgeschlagen. Nach den Resultaten von [Bon96] hat dieses Funktional viele unerwünschte Eigenschaften: Es werden Optimallösungen favorisiert, die Ecken abrunden, sowie T-Kreuzungen vermeiden, indem drei Kanten symmetrisch mit jeweils einem Winkel von 120° aufeinandertreffen. Gerade die fehlenden T-Kreuzungen (die für die Tiefenwahrnehmung extrem wichtig sind) lassen Lösungen daher flach und cartoonartig wirken. Eine weitere unangenehme, aber für diese Klasse typische Eigenschaft ist die unstetige Abhängigkeit der Lösung von den Anfangsdaten (siehe etwa Abschnitt A.5.d „How many Hong-Kong butterflies?“ in [Dav05]). Trotzdem war und ist das Mumford-Shah Funktional Auslöser und Motor für viele Fragen und Methoden im Bereich der freien Randwertprobleme und der geometrischen Maßtheorie.

1.2 Diskrete Problemstellung

Hier besteht Ω aus endlich vielen Pixeln. Pro Pixel werden endlich viele Freiheitsgrade vergeben und über alle Freiheitsgrade minimiert. Damit ist das gesuchte Bild $u \in \mathbb{R}^d$. Im einfachsten Fall hat jeder Pixel genau einen Freiheitsgrad, nämlich den Grauwert, und es gilt $d = |\Omega|$. In diesem Fall werde der Wert von u am Pixel a mit $u_a := u(a)$ für $a \in \Omega$ abgekürzt. Üblicherweise sind die Pixel in einem uniformen quadratischen Gitter angeordnet und Ω ist ein Rechteck.

$$F_D(u; v) = \lambda \sum_{a \in \Omega} f_a(|u_a - v_a|) \quad \text{und} \quad F_R(u) = \sum_{\substack{a, b \in \Omega \\ a \sim b}} g_{a,b}(|u_a - u_b|)$$

sind typische Datenfit- und Regularisierungsansätze. Dabei bezeichne „ \sim “ die Nachbarschaftsrelation, d. h. $a \sim b$ bedeutet, dass a und b benachbart sind. Es gibt mehrere Festlegungen, wann ein Pixel a zu einem anderen Pixel b benachbart ist. Die zwei häufigsten sind in Abbildung 1.4 zu sehen.

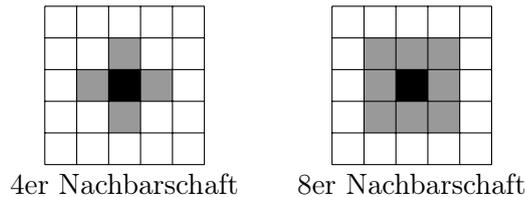


Abbildung 1.4: Gängige Nachbarschaften auf quadratischen Gittern

Es gibt vielfältige Art und Weisen, wie ein diskretes Funktional zustande kommen kann. So kann ein kontinuierliches Funktional als Grundlage dienen, welches dann diskretisiert wird, um die Berechnung einer approximativen Lösung auf Rechnern zu ermöglichen. In diesem Kontext stellt sich die Frage, in welchem Sinne ein diskretes Funktional ein kontinuierliches approximieren kann und soll. Ein bewährtes Framework in diesem Zusammenhang ist die Gamma-Konvergenz (siehe Abschnitt 2.2.3): Man betrachtet für immer kleiner werdende Pixelbreiten h die Folge der diskreten Optimallösungen und fragt sich, ob diese stets bezüglich einer Topologie gegen die kontinuierliche Optimallösung konvergiert.

Etliche diskrete Funktionale wurden aber nicht als Diskretisierung erhalten, sondern direkt angegeben. Ein oft verwendeter stochastischer Ansatz mit Markov Random Fields wird in Abschnitt 4.4 gezeigt. Gerade bei solchen Funktionalen kann es interessant sein zu wissen, ob sie ein kontinuierliches Funktional im Sinne der Gamma-Konvergenz approximieren. So kann auch untersucht werden, ob bestimmte (un-)angenehme Eigenschaften des diskreten Funktionals eine Folge der Diskretisierung sind oder vom kontinuierlichen

vererbt wurden. So sollte man sich bei einer guten diskreten Approximation des oben beschriebenen Mumford-Shah Funktionals nicht wundern, wenn (bei hinreichend kleiner Pixelbreite) keine T-Kreuzungen in der Optimallösung vorkommen.

1.3 Überblick

Diese Arbeit ist im Bereich der variationellen Bildverarbeitung bzw. -verbesserung beheimatet. Ein Ziel ist es, mit Hilfe der Gamma-Konvergenz diskrete und kontinuierliche Problemstellungen in Verbindung zu setzen. Der Grenzübergang vom diskreten zum kontinuierlichen Funktional wird dadurch bewerkstelligt, dass für die Pixelbreite h der Limes $h \rightarrow 0$ betrachtet wird. Für $h > 0$ werden dabei „diskrete Bilder“ u_h als Funktionen aufgefasst, die auf jedem Pixel konstant sind. Diese Betrachtungsweise unterscheidet sich sowohl von [Neg99], wo das Mumford-Shah Funktional im Raum der Finiten Elemente approximiert wurde, als auch von [AT90], wo Approximationen für das Mumford-Shah Funktional mit elliptischen Funktionalen untersucht wurden. Ferner werden in dieser Arbeit allgemeinere kontinuierliche Funktionale als das Mumford-Shah Funktional betrachtet.

Wie oben schon angeklungen, finden alle Untersuchungen im BV -Framework statt, also im Raum der Funktionen beschränkter Variation (oder Teilräume davon). In Kapitel 2 werden, nach einer Motivation im Eindimensionalen (Abschnitt 2.2), diskrete Funktionale der Form

$$F_h(u_h) := \sum_{R \sim T} V_h(u_h|_R, u_h|_T) + \sum_T |T| f(u_h|_T)$$

betrachtet und Funktionen V_h und f so konstruiert, dass bezüglich der L^1 -Topologie ein Gamma-Limes

$$F(u) := \int_{S_u} V(u^+, u^-) \Phi(\nu_u) d\mathcal{H}^1 + \int_{\Omega} f(u) dx + \int_{\Omega} g(\nabla u) dx$$

existiert. Es wurde versucht, die im Laufe der Konstruktion auftretenden Teilergebnisse möglichst allgemein zu halten, auch wenn in der weiteren Darstellung Ergebnisse wegen zusätzlichen Voraussetzungen nur für Spezialfälle benötigt werden.

Im Anschluss werden in Kapitel 3 für unterschiedliche Teilräume von $BV(\Omega)$ Existenzaussagen für das Minimierungsproblem $F(u) = \min!$ vorgestellt. Es werden Lösungen im Raum $SBV(\Omega)$, im der Raum stückweisen konstanten Funktionen und auch in Sobolev-Räumen betrachtet. Je nach Ansatzraum fallen dabei einzelne Terme von $F(u)$ per Definition weg.

Ein weiteres Ziel dieser Arbeit ist es, die Vorteile von Datenfittern F_D und Regularisierern F_R (vgl. (1.2)) mit $1 < \alpha, \gamma < 2$ aufzuzeigen (siehe auch

Abbildungen 1.1 und 1.2). In Kapitel 4 werden Algorithmen für diskrete Probleme

$$\sum_{a \sim b} g_{a,b}(|u_a - u_b|) + \sum_a f_a(|u_a - v_a|) = \min_u!$$

verglichen. Da bei einem konkreten Problem die Pixelbreite h der Eingabebilder fest ist, wurde die h -Abhängigkeit in der Notation im Kapitel 4 unterdrückt. Je nachdem, ob alle $g_{a,b}$ und f_a konvex sind und es sich damit um ein konvexes Optimierungsproblem handelt oder nicht, kommen unterschiedliche Algorithmen zum Einsatz. Die Modellierung von Kanten durch nichtkonvexe $g_{a,b}$ sowie die Schwierigkeiten, die dadurch entstehen, werden näher beleuchtet. Es werden Zusammenhänge zur diskreten Optimierung mit Graphcuts sowie zur Stochastik mit Markov Random Fields hergestellt. In Abschnitt 4.6 wird das konvexe Funktional aus 4.5.8, welches auch zur Berechnung der Bilder in Abbildung 1.2 verwendet wurde, zu einem nichtkonvexen Funktional mit Kantenterm erweitert. Ferner wird aufgezeigt, wie für das entstehende Funktional ein graduated non-convexity Algorithmus konstruiert werden kann. Im letzten Abschnitt 4.7 wird dieser GNC-Algorithmus verwendet, um die Vorteile eines Regularisierers mit $1 < \alpha < 2$ anhand eines Beispiels im direkten Vergleich mit $\alpha = 2$ zu sehen.

Dieser Arbeit liegt eine CD mit allen Quelltexten und Beispielrechnungen bei.

Kapitel 2

Konvergenzbetrachtungen

In diesem Kapitel werden diskrete Funktionale auf uniformen quadratischen Gittern untersucht. Insbesondere wird der Frage nachgegangen, unter welchen Bedingungen sie bzgl. der L^1 -Topologie einen Γ -Limes der Form

$$F(u) = \int_{S_u} V(u^+, u^-) \Phi(\nu_u) d\mathcal{H}^1 + \int_{\Omega} f(u) dx + \int_{\Omega} g(\nabla u) dx$$

aufweisen. Dazu wird in Abschnitt 2.1 zunächst das nötige Rüstzeug bereitgestellt. In 2.2 wird die Problemstellung im Eindimensionalen motiviert. In Abschnitt 2.3 wird schließlich auf die für Bilder relevante Zweidimensionale Situation eingegangen.

2.1 Notation, Grundlagen

Definition 2.1 (Halbstetigkeit von unten)

Ist X ein Banachraum, dann heißt ein Funktional $F: X \rightarrow \mathbb{R} \cup \{+\infty\}$

1. *halbstetig von unten (lower semicontinuous; lsc) oder auch unterhalbstetig, falls*

$$F(x) \leq \liminf_{n \rightarrow \infty} F(x_n) \quad \text{für alle Folgen } (x_n) \text{ mit } x_n \xrightarrow{X} x,$$

2. *schwach halbstetig von unten (weakly lower semicontinuous; wslc), falls*

$$F(x) \leq \liminf_{n \rightarrow \infty} F(x_n) \quad \text{für alle Folgen } (x_n) \text{ mit } x_n \xrightarrow{X} x.$$

Aus schwach halbstetig folgt halbstetig.

2.1.1 Maßtheoretische Grundlagen

Da bei unseren Betrachtungen nicht nur positive Maße, sondern insbesondere auch vektorwertige Maße eine Rolle spielen, werden hier die maßtheoretischen Begriffe zusammengestellt, wie sie etwa in [AFP00] oder teilweise auch in [Els05] und [Bau92] zu finden sind.

Sei Ω eine nichtleere Menge und \mathcal{A} eine σ -Algebra auf Ω .

Definition 2.2 (σ -Additivität)

Eine Abbildung $\mu: \mathcal{A} \rightarrow \mathbb{R} \cup \{\pm\infty\}$ heißt σ -additiv, falls

$$\mu\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \mu(A_n)$$

für eine beliebige Folge $(A_n)_n \subset \mathcal{A}$ mit paarweise disjunkten A_n gilt.

Definition 2.3 (positive und vektorwertige Maße)

Es sei Ω eine nichtleere Menge und \mathcal{A} eine σ -Algebra auf Ω .

1. (positives Maß) Eine Abbildung $\mu: \mathcal{A} \rightarrow [0; \infty]$ heißt positives Maß auf (Ω, \mathcal{A}) , falls $\mu(\emptyset) = 0$ gilt und μ σ -additiv ist.
2. (vektorwertiges Maß) Eine Abbildung $\mu: \mathcal{A} \rightarrow \mathbb{R}^d$ heißt (vektorwertiges) Maß auf (Ω, \mathcal{A}) , falls $\mu(\emptyset) = 0$ gilt und μ σ -additiv ist. Ist $d = 1$, so heißt μ auch reellwertiges Maß.
3. (Variationsmaß) Ist μ ein \mathbb{R}^d -wertiges Maß auf (Ω, \mathcal{A}) so wird für $A \in \mathcal{A}$ durch

$$|\mu|(A) := \sup \left\{ \sum_{n=1}^{\infty} \|\mu(A_n)\|_2 : A_n \in \mathcal{A} \text{ paarweise disjunkt, } A = \bigcup_{n=1}^{\infty} A_n \right\}$$

ein endliches positives Maß, das Variationsmaß $|\mu|$, definiert.

Ist μ reell, so wird durch $\mu^+ := (|\mu| + \mu)/2$ bzw. durch $\mu^- := (|\mu| - \mu)/2$ der Positiv- bzw. Negativteil von μ definiert.

Ab jetzt sei Ω ein lokal kompakter, separabler metrischer Raum. Mit $\mathcal{B}(\Omega)$ werde die Borel- σ -Algebra bezeichnet, die kleinste σ -Algebra, die alle offenen und abgeschlossenen Teilmengen von Ω enthält. Folgende Definition legt speziellere Sprechweisen für die soeben definierten Maße fest, falls wir uns in dem Fall befinden, dass die zugrundeliegende σ -Algebra eine Borel- σ -Algebra ist.

Definition 2.4 (Borel- und Radon-Maße)

Es sei Ω ein lokal kompakter, separabler metrischer Raum und $\mathcal{B}(\Omega)$ die Borel- σ -Algebra.

1. (Borel-Maß) Ein positives Maß μ auf $(\Omega, \mathcal{B}(\Omega))$ heißt Borel-Maß.
2. (positives Radon-Maß) Gilt für ein Borel-Maß μ auf $(\Omega, \mathcal{B}(\Omega))$ stets $\mu(K) < \infty$ für alle kompakten $K \subset \Omega$, so heißt μ positives Radon-Maß.
3. (vektorwertiges Radon-Maß) Eine Abbildung $\mu: \{A \in \mathcal{B}(\Omega) : A \subset\subset \Omega\} \rightarrow \mathbb{R}^d$, die auf $(K, \mathcal{B}(K))$ für jedes kompakte $K \subset \Omega$ ein vektorwertiges Maß ist, heißt vektorwertiges Radon-Maß. Der Raum aller vektorwertigen Radon-Maße werde mit $[\mathcal{M}_{\text{loc}}(\Omega)]^d$ bezeichnet.
4. (endliches Radon-Maß) Ist $\mu: \mathcal{B}(\Omega) \rightarrow \mathbb{R}^d$ ein vektorwertiges Maß, so heißt μ endliches Radon-Maß. Der Raum aller endlichen Radon-Maße werde mit $[\mathcal{M}(\Omega)]^d$ bezeichnet.

Wenn wir zukünftig von Maßen bzw. Radon-Maßen sprechen, dann sind damit stets vektorwertige Maße bzw. vektorwertige Radon-Maße gemeint. Bei allen anderen Maßen wird jeweils das entsprechende Adjektiv angegeben. Mit \mathcal{L}^d wird das Lebesgue-Maß im \mathbb{R}^d und mit \mathcal{H}^k das k -dimensionale Hausdorff-Maß (im \mathbb{R}^d) bezeichnet.

Da auch über vektorwertige Maße integriert wird, wird hier kurz wiederholt, wie die auftretenden Integrale definiert sind. Dabei wird das Lebesgue-Integral über positive Maße für Funktionen mit Werten in $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ vorausgesetzt.

Ist μ ein reelles Maß auf Ω und $u: \Omega \rightarrow \overline{\mathbb{R}}$ ist $|\mu|$ -messbar, dann ist

$$\int_{\Omega} u \, d\mu = \int_{\Omega} u \, d\mu^+ - \int_{\Omega} u \, d\mu^-$$

falls $\int_{\Omega} |u| \, d|\mu| < \infty$. Ist von $u: \Omega \rightarrow \mathbb{R}^d$ jede Komponente u_k $|\mu|$ -messbar mit $\int_{\Omega} |u_k| \, d|\mu| < \infty$, so ist

$$\int_{\Omega} u \, d\mu = \left(\int_{\Omega} u_1 \, d\mu, \dots, \int_{\Omega} u_d \, d\mu \right).$$

Ist $\mu = (\mu_1, \dots, \mu_d)$ ein Maß auf Ω und $u: \Omega \rightarrow \overline{\mathbb{R}}$ wieder $|\mu|$ -messbar mit $\int_{\Omega} |u| \, d|\mu| < \infty$, so ist

$$\int_{\Omega} u \, d\mu = \left(\int_{\Omega} u \, d\mu_1, \dots, \int_{\Omega} u \, d\mu_d \right).$$

Zur Integration Banachraum-wertiger Funktionen auf lokal kompakten Räumen über Banachraum-wertige Maße sei auf den schönen Zugang im Buch [Din74] verwiesen, in welchem Maße als Dualräume zu Räumen stetiger Funktionen eingeführt werden.

Satz 2.5 (Ungleichung von Jensen, [AFP00, Lemma 1.15])

Ist μ ein positives Maß auf (Ω, \mathcal{A}) mit $\mu(\Omega) = 1$ und $f: \mathbb{R}^d \rightarrow \mathbb{R}$ konvex und unterhalbstetig, dann gilt

$$f\left(\int_{\Omega} u \, d\mu\right) \leq \int_{\Omega} f(u) \, d\mu$$

für alle μ -messbaren $u: \Omega \rightarrow \mathbb{R}^d$ mit $\int_{\Omega} |u| \, d\mu < \infty$.

2.1.2 Sobolev Funktionen

In diesem wie im folgenden Abschnitt sollen die benötigten Funktionenräume eingeführt werden. Wir beginnen hier mit den Sobolev-Räumen. Diese Räume werden im nächsten Abschnitt durch die uns interessierenden *SBV* bzw. *BV*-Räume ergänzt.

Sei in diesem Abschnitt $\Omega \subset \mathbb{R}^d$ offen. Bei den Sobolev-Räumen spielt die schwache Ableitung eine entscheidende Rolle.

Definition 2.6 (schwache Ableitung, [AF03, 1.60])

Ist $\alpha \in \mathbb{N}_0^d$ ($\alpha \neq 0$) ein Multiindex und $f, g \in L_{\text{loc}}^1(\Omega)$, dann ist die Funktion g die schwache Ableitung von f von der Ordnung α (Schreibweise: $g = D^\alpha f$), falls

$$\int_{\Omega} f D^\alpha \psi \, dx = (-1)^{|\alpha|} \int_{\Omega} g \psi \, dx$$

für alle $\psi \in C_c^\infty(\Omega)$ gilt.

Dabei bezeichnet $C_c^k(\Omega)$ die Menge der k -mal stetig differenzierbaren Funktionen mit kompaktem Träger in Ω .

Damit ist es möglich, die Räume schwach differenzierbarer Funktionen zu definieren.

Definition 2.7 ($W^{m,p}(\Omega)$, [AF03, 3.1 und 3.2])

Für $m \in \mathbb{N}_0$ und $1 \leq p \leq \infty$ wird durch

$$W^{m,p}(\Omega) := \{ u \in L^p(\Omega) : \exists D^\alpha u \in L^p(\Omega) \text{ für alle } 0 \leq |\alpha| \leq m \}$$

der Raum der m -mal schwach differenzierbaren Sobolev Funktionen mit Ableitungen in $L^p(\Omega)$ bezeichnet. Für $1 \leq p < \infty$ wird durch

$$\|u\|_{W^{m,p}(\Omega)} := \left(\sum_{0 \leq |\alpha| \leq m} \|D^\alpha u\|_{L^p(\Omega)}^p \right)^{1/p}$$

und für $p = \infty$ durch

$$\|u\|_{W^{m,\infty}(\Omega)} := \max_{0 \leq |\alpha| \leq m} \|D^\alpha u\|_{L^\infty(\Omega)}$$

auf $W^{m,p}(\Omega)$ eine Norm definiert. Der Abschluss von $C_c^\infty(\Omega)$ bezüglich der obigen Norm ergibt den Raum

$$W_0^{m,p}(\Omega) := \overline{C_c^\infty(\Omega)}^{W^{m,p}(\Omega)}$$

der Sobolev Funktionen „mit Nullrand“.

Schließlich wollen wir diesen Abschnitt damit beschließen die schwache Konvergenz in $W^{m,p}(\Omega)$ zu charakterisieren.

Satz 2.8 (schwache Konvergenz)

Für $1 \leq p < \infty$ und $u_n, u \in W^{m,p}(\Omega)$ ($n \in \mathbb{N}$) gilt

$$u_n \xrightarrow{W^{m,p}(\Omega)} u \quad :\Leftrightarrow \quad \begin{array}{l} T(u_n) \rightarrow T(u) \\ \forall T \in \left(W^{m,p}(\Omega)\right)' \end{array} \quad \Leftrightarrow \quad \begin{array}{l} D^\alpha u_n \xrightarrow{L^p(\Omega)} D^\alpha u \\ \forall 0 \leq |\alpha| \leq m \end{array} .$$

Ist zusätzlich $\Omega \subset \mathbb{R}^d$ beschränkt mit Lipschitz-Rand und $m \geq 1$, so gilt:

$$u_n \xrightarrow{W^{m,p}(\Omega)} u \quad \Leftrightarrow \quad \begin{array}{l} u_n \xrightarrow{L^p(\Omega)} u, \quad D^\alpha u_n \xrightarrow{L^p(\Omega)} D^\alpha u \\ \forall 1 \leq |\alpha| \leq m \end{array} .$$

Beweis

Die erste Aussage findet sich z. B. in [Alt06], Abschnitt 6.4. Für die zweite Aussage ist die Richtung „ \Leftarrow “ klar. Für „ \Rightarrow “ sei $(u_n)_n \subset W^{m,p}(\Omega)$ ($m \geq 1$) eine gegen u schwach konvergente Folge. Als solche ist (u_n) in $W^{m,p}(\Omega)$ beschränkt. Die Einbettung $W^{m,p}(\Omega) \rightarrow W^{0,p}(\Omega) = L^p(\Omega)$ ist nach dem Rellich-Kondrachov Theorem ([AF03], Theorem 6.3) kompakt. Folglich gibt es eine Teilfolge $(u_{n_k})_k$, welche in $L^p(\Omega)$ gegen ein \bar{u} in der Norm konvergiert. Da $(u_{n_k})_k$ nach Voraussetzung in $L^p(\Omega)$ auch schwach gegen u konvergiert, muss wegen der Eindeutigkeit des schwachen Grenzwertes $\bar{u} = u$ sein und auf die Teilfolgenauswahl kann verzichtet werden. \square

2.1.3 Funktionen beschränkter Variation

Kommen wir nun also zu den für diese Arbeit zentralen Funktionenräumen. Auch in diesem Abschnitt sei wieder $\Omega \subset \mathbb{R}^d$ offen.

Definition 2.9 ($BV(\Omega)$)

Eine Funktion $u \in L^1(\Omega)$ heißt von beschränkter Variation, wenn ihre Ableitung Du (gebildet im Raum der Distributionen) ein endliches \mathbb{R}^d -wertiges Radon-Maß auf $(\Omega, \mathcal{B}(\Omega))$ ist, also $Du \in [\mathcal{M}(\Omega)]^d$ ist. Die Menge aller solchen Funktionen wird mit $BV(\Omega)$ bezeichnet.

Ist $u \in BV(\Omega)$, dann gibt es also Maße $Du = (D_1u, \dots, D_du)$ mit

$$\int_{\Omega} u D_{x_i} \psi \, dx = - \int_{\Omega} \psi \, dD_i u$$

für alle $\psi \in C_c^1(\Omega)$ und $i = 1, \dots, d$ oder auch äquivalent

$$\int_{\Omega} u \operatorname{div} \psi \, dx = - \sum_{i=1}^d \int_{\Omega} \psi_i \, dD_i u$$

für alle $\psi \in [C_c^1(\Omega)]^d$.

Um die Namensgebung der Funktionen beschränkter Variation nachvollziehen zu können, wird der Begriff der Variation benötigt.

Definition 2.10 (Variation)

Für ein $u \in L^1_{\text{loc}}(\Omega)$ ist die Variation $\operatorname{Var}(u, \Omega)$ von u in Ω durch

$$\operatorname{Var}(u, \Omega) := \sup \left\{ \int_{\Omega} u \operatorname{div} \psi \, dx : \psi \in [C_c^1(\Omega)]^d \right. \\ \left. \text{mit } \|\psi(x)\|_2 \leq 1 \text{ für alle } x \in \Omega \right\}$$

definiert.

Folgender Satz liefert die Erklärung für die Namensgebung, da in $BV(\Omega)$ gerade die $L^1(\Omega)$ -Funktionen liegen, deren Variation endlich, also beschränkt ist.

Satz 2.11 ([AFP00, Proposition 3.6])

Für ein $u \in L^1(\Omega)$ sind folgende Aussagen äquivalent:

1. $u \in BV(\Omega)$ und
2. $\operatorname{Var}(u, \Omega) < \infty$.

Für ein $u \in BV(\Omega)$ stimmt die Variation mit dem Variationsmaß der Ableitung überein: $\operatorname{Var}(u, \Omega) = |Du|(\Omega)$.

Für $u \in BV(\Omega)$ wird durch

$$\|u\|_{BV(\Omega)} := \|u\|_{L^1(\Omega)} + |Du|(\Omega)$$

eine Norm auf $BV(\Omega)$ definiert. Jedes $u \in W^{1,1}(\Omega)$ im Sobolev-Raum der einmal (im $L^1(\Omega)$) schwach differenzierbaren Funktionen ist von beschränkter Variation mit

$$\|u\|_{BV(\Omega)} = \|u\|_{W^{1,1}(\Omega)}.$$

Es gilt also $W^{1,1}(\Omega) \subset BV(\Omega)$. Wie wir sehen werden, ist dies eine echte Inklusion.

Für die durch die $\|\cdot\|_{BV(\Omega)}$ festgelegte Topologie sind nur wenige befriedigende Strukturresultate bekannt. Daher wird fast ausschließlich auf $BV(\Omega)$ mit folgender schwach* Topologie gearbeitet.

Definition 2.12 (schwach* Topologie auf $BV(\Omega)$)

Eine Folge $(u_n)_n \subset BV(\Omega)$ heißt schwach* konvergent gegen $u \in BV(\Omega)$

$$u_n \xrightarrow{BV(\Omega)}^* u \quad :\Leftrightarrow \quad u_n \xrightarrow{L^1(\Omega)} u \quad \text{und} \quad Du_n \xrightarrow{[\mathcal{M}(\Omega)]^d}^* Du.$$

Mit Hilfe der Variation lässt sich eine Verallgemeinerung des Umfangs festlegen.

Definition 2.13 (Perimeter)

Für eine \mathcal{L}^d -messbare Menge $E \subset \Omega$ heißt

$$P(E, \Omega) := \text{Var}(\chi_E, \Omega)$$

der Perimeter von E in Ω .

Fleming und Rishel haben in [FR60] folgende Coarea-Formel gezeigt.

Satz 2.14 (Coarea-Formel in BV)

Für $u \in L^1_{\text{loc}}(\Omega)$ gilt

$$\text{Var}(u, \Omega) = \int_{-\infty}^{\infty} P(\{x \in \Omega : u(x) > t\}, \Omega) dt.$$

Ist $u \in BV(\Omega)$, dann gilt $P(\{u > t\}, \Omega) < \infty$ für \mathcal{L}^1 -fast alle $t \in \mathbb{R}$ und für jede Borelmenge A gilt

$$|Du|(A) = \int_{-\infty}^{\infty} |D\chi_{\{u>t\}}|(A) dt \quad \text{und} \quad Du(A) = \int_{-\infty}^{\infty} D\chi_{\{u>t\}}(A) dt.$$

Da schon der Raum $L^1(\Omega)$ keine Funktionen sondern Äquivalenzklassen von Funktionen beinhaltet (für zwei Funktionen u und v gilt $u = v$ in $L^1(\Omega)$, falls $\mathcal{L}^d\{x \in \Omega : u(x) \neq v(x)\} = 0$), erweist sich das klassische Konzept der Stetigkeit als unbrauchbar. Auch das Außerachtlassen von Lebesgue-Nullmengen führt zu keinen brauchbaren Begriffen. Daher werden sowohl Nullmengen, als auch „Mengen ohne Dichte“ ausgenommen, wie es etwa in Definition 2.15 zu sehen ist.

Definition 2.15 (Approximativer Grenzwert)

Ein $u \in L^1_{\text{loc}}(\Omega)$ besitzt bei $x \in \Omega$ den approximativen Grenzwert

$$z = \text{ap-lim}_{y \rightarrow x} u(y) \in \mathbb{R},$$

falls

$$\lim_{0 < r \rightarrow 0} \frac{1}{\mathcal{L}^d(B_r(x))} \int_{B_r(x)} |u(y) - z| dy = 0$$

gilt.

Definition 2.16 (Approximative Unstetigkeitsmenge)

Für ein $u \in L^1_{\text{loc}}(\Omega)$ heißt die Menge

$$S_u := \{x \in \Omega : \text{ap-lim}_{y \rightarrow x} u(y) \text{ existiert nicht}\}$$

die approximative Unstetigkeitsmenge von u .

Für jedes $u \in L^1_{\text{loc}}(\Omega)$ ist S_u messbar und es gilt $\mathcal{L}^d(S_u) = 0$. Außerdem wird durch $\tilde{u}(x) := \text{ap-lim}_{y \rightarrow x} u(y)$ für alle $x \in \Omega \setminus S_u$ eine Funktion $\tilde{u} \in L^1_{\text{loc}}(\Omega)$ festgelegt, die \mathcal{L}^d -fast überall mit u übereinstimmt.

In völlig analoger Weise kann die approximative Differenzierbarkeit definiert werden:

Definition 2.17 (Approximativer Gradient)

Ein $u \in L^1_{\text{loc}}(\Omega)$ besitzt bei $x \in \Omega \setminus S_u$ den approximativen Gradienten $\nabla u(x)$, falls der folgende approximative Grenzwert existiert und

$$\text{ap-lim}_{y \rightarrow x} \frac{u(y) - \tilde{u}(x) - \langle \nabla u(x), y - x \rangle}{\|y - x\|_2} = 0$$

gilt.

Innerhalb der Menge S_u von u werden nun die Punkte ausgezeichnet, bei denen u einen approximativen Sprung zwischen zwei Werten besitzt.

Definition 2.18 (Approximative Sprungmenge)

Für ein $u \in L^1_{\text{loc}}(\Omega)$ findet bei $x \in \Omega$ ein approximativer Sprung zwischen den Werten $a, b \in \mathbb{R}$, $a \neq b$ in Richtung von $\nu \in S^{d-1}$ statt, falls

$$\lim_{0 < r \rightarrow 0} \frac{1}{\mathcal{L}^d(B_r^+(x, \nu))} \int_{B_r^+(x, \nu)} |u(y) - a| dy = 0$$

und

$$\lim_{0 < r \rightarrow 0} \frac{1}{\mathcal{L}^d(B_r^-(x, \nu))} \int_{B_r^-(x, \nu)} |u(y) - b| dy = 0$$

gilt. Dabei ist $B_r^\pm(x, \nu) := \{y \in B_r(x) : \langle y - x, \nu \rangle \gtrless 0\}$. Das Tupel (a, b, ν) ist dabei bis auf Vertauschung von a und b mit gleichzeitigem Vorzeichenwechsel bei ν eindeutig bestimmt. Deshalb werden (a, b, ν) und $(b, a, -\nu)$ als gleich betrachtet. Man schreibt dafür auch: $(u^+(x), u^-(x), \nu_u(x))$. Die Menge J_u der Punkte x , bei denen ein approximativer Sprung vorliegt, heißt *approximative Sprungmenge*.

Mit dem Satz von Radon-Nikodym (siehe z.B. [Els05] Kapitel VII, Abschnitt §2.2) kann für ein $u \in BV(\Omega)$ die Ableitung

$$Du = D^a u + D^s u$$

in einen bzgl. \mathcal{L}^d absolut stetigen Anteil $D^a u$ und einen singulären Anteil $D^s u$ zerlegt werden. Außerdem lässt sich der singuläre Anteil weiter aufspalten in den Sprunganteil $D^j u := D^s u \llcorner J_u$ und den Cantor-Anteil $D^c u := D^s u \llcorner (\Omega \setminus S_u)$. Damit gilt

$$Du = D^a u + D^s u = D^a u + D^j u + D^c u.$$

Für Funktionen beschränkter Variation gilt ferner

$$D^a u = \nabla u \mathcal{L}^d \quad \text{und} \quad D^j u = (u^+ - u^-) \nu_u \mathcal{H}^{d-1} \llcorner J_u$$

und $\mathcal{H}^{d-1}(S_u \setminus J_u) = 0$.

Definition 2.19 (Spezielle Funktionen beschränkter Variation)

Mit

$$SBV(\Omega) := \{u \in BV(\Omega) : D^c u = 0\}$$

wird die Menge der speziellen Funktionen beschränkter Variation bezeichnet. Die Menge $SBV(\Omega)$ ist ein Unterraum von $BV(\Omega)$.

Die Ableitung einer speziellen Funktion u beschränkter Variation ist also allein durch den bzgl. \mathcal{L}^d absolut stetigen Gradientenanteil ∇u und durch den Sprunganteil (u^+, u^-, ν_u) auf J_u eindeutig festgelegt. Für ein $u \in SBV(\Omega)$ gilt

$$u \in W^{1,1}(\Omega) \quad \Leftrightarrow \quad \mathcal{H}^{d-1}(S_u) = 0, \quad (2.1)$$

wenn also u keine $(d-1)$ -dimensionalen (approximativen) Sprünge aufweist.

2.2 Motivation: Freies Randwertproblem im Eindimensionalen

In diesem Abschnitt wird auf dem endlichen Intervall $\Omega =]a; b[$ mit $a < b$

$$F(u) = \int_a^b |u'|^\alpha dx + \int_{S_u} |u^+ - u^-|^\beta d\mathcal{H}^0 + \int_a^b f(u) dx = \min!$$

für $u \in SBV(\Omega)$, $0 \leq \beta < 1 < \alpha \leq 2$ und $f: \mathbb{R} \rightarrow [0; \infty[$ betrachtet. Es wird untersucht, wie F mit Hilfe von finiten Differenzen durch diskrete (nicht-konvexe) Funktionale approximiert werden kann. Da die Approximation des Terms $\int_a^b f(u) dx$ einfach zu behandeln ist (siehe Abschnitte 2.3.6 und 2.3.8) wird in diesem Abschnitt $f \equiv 0$ gesetzt. Natürlich ist in diesem Fall die Minimierungsaufgabe $F(u) = \min!$ trivial durch $u \equiv 0$ lösbar. Aber es geht hier um die Herleitung einer geeigneten Approximation im Diskreten. Diese kann dann zusammen mit der Approximation des Terms für $f \not\equiv 0$ zu einem nicht-trivialen Funktional kombiniert werden.

Für den Fall $\alpha = 2$ und $\beta = 0$, also dem eindimensionalen Mumford-Shah Funktional

$$\int_a^b |u'|^2 dx + \mathcal{H}^0(S_u) = \min!$$

wurde eine entsprechende Approximation in [Bra02, Abschnitt 8.3] angegeben.

Dieser Abschnitt dient als Motivation und Vorbereitung zum Abschnitt 2.3, wo versucht wird, die Überlegungen auf Gebiete $\Omega \subset \mathbb{R}^2$ zu übertragen. Nachdem in 2.2.1 die Notation festgelegt wurde, wird in 2.2.2 das Ziel formuliert. Im Abschnitt 2.2.3 wird der zentrale Konvergenzbegriff, die Gamma-Konvergenz, vorgestellt. In 2.2.4 und 2.2.5 werden die für die Gamma-Konvergenz notwendigen Ungleichungen gezeigt. Zum Abschluss werden in 2.2.6 an einem einfachen Beispiel kontinuierliche und diskrete Minimierer verglichen.

2.2.1 Notation

Zur Diskretisierung wird das Intervall $\Omega =]a; b[$ in n gleichlange Teilintervalle der Länge $h := (b - a)/n$ unterteilt. Für $i \in \{0, \dots, n\}$ sei $x_i^h := a + ih$. Zur übersichtlicheren Notation wird (wenn keine Verwechslungsgefahr besteht) x_i statt x_i^h geschrieben. Mit

$$U_h := \left\{ u_h: \Omega \rightarrow \mathbb{R} : u_h|_{]x_i - \frac{h}{2}; x_i + \frac{h}{2}[\cap \Omega} = \text{const für alle } i \in \{0, \dots, n\} \right\}$$

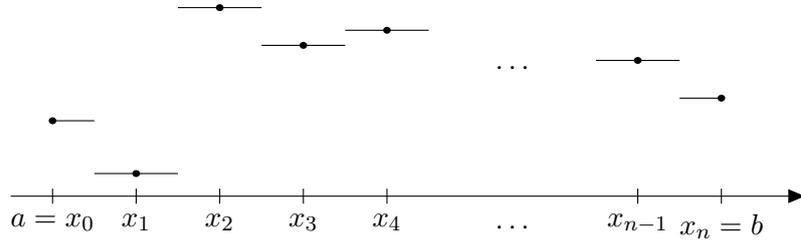
werde der Raum der stückweisen konstanten Funktionen bezeichnet (vgl. Abbildung 2.1).

Fasst man ein $u_h \in U_h$ als „eindimensionales Bild“ auf, so sind die x_i die Pixelmittelpunkte und die $]x_i - \frac{h}{2}; x_i + \frac{h}{2}[$ die Pixel der Breite h .

2.2.2 Problemstellung und Ziel

Ziel ist es, ein diskretes Funktional F_h auf U_h und eine Nachbarschaftsinteraktionsfunktion V_h zu finden, so dass

$$F_h(u_h) = \sum_{i=0}^{n-1} V_h(u_h(x_i), u_h(x_{i+1}))$$

Abbildung 2.1: Skizze für ein $u_h \in U_h$

für $h \rightarrow 0$ das Funktional F im geeigneten Sinne approximiert. Folgender Satz soll gezeigt werden.

Satz 2.20

Sei $K \subset \mathbb{R}$ kompakt und $U_{h,K} := \{u_h \in U_h : u_h(x) \in K \text{ } \mathcal{L}^1\text{-fast überall}\}$ sowie $U := \{u \in SBV(\Omega) : u \in K \text{ } \mathcal{L}^1\text{-fast überall, } \mathcal{H}^0(S_u) < \infty \text{ und } u' \in L^\alpha(\Omega)\}$. Für $0 \leq \beta < 1 < \alpha \leq 2$, $d(z) = \min(|z|^\alpha, |z|^\beta)$, $\ell = \frac{\alpha-1}{\alpha-\beta}$ und $k = \ell\beta$ ist der Γ -Limes von

$$F_h(u_h) = \begin{cases} \sum_{i=0}^{n-1} h^k d\left(\frac{u_h(x_{i+1}^h) - u_h(x_i^h)}{h^\ell}\right) & \text{für } u_h \in U_{h,K} \\ +\infty & \text{sonst} \end{cases}$$

bzgl. der L^1 -Topologie für $h \rightarrow 0$ bzw. $n \rightarrow \infty$ durch

$$F(u) = \begin{cases} \int_{\Omega} |u'|^\alpha dx + \int_{S_u} |u^+ - u^-|^\beta d\mathcal{H}^0 & \text{für } u \in U \\ +\infty & \text{sonst} \end{cases}$$

gegeben.

2.2.3 Gamma-Konvergenz

In diesem Abschnitt werden die Definition sowie die grundlegenden Eigenschaften der Γ -Konvergenz gegeben. Beweise der angegebenen Aussagen sind z. B. auch in [Bra02] zu finden.

Es sei (X, d) ein metrischer Raum.

Definition 2.21 (Γ -lim)

Eine Folge $f_j: X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ ist Γ -konvergent gegen $f_\infty: X \rightarrow \mathbb{R} \cup \{\pm\infty\}$, falls für alle $x \in X$ folgende Bedingungen erfüllt sind:

1. (*lim inf-Ungleichung*) Für jede Folge (x_j) mit $d(x_j, x) \rightarrow 0$ gilt

$$f_\infty(x) \leq \liminf_{j \rightarrow \infty} f_j(x_j). \quad (2.2)$$

2. (*lim sup-Ungleichung*) Es gibt eine Folge (x_j) mit $d(x_j, x) \rightarrow 0$, so dass

$$f_\infty(x) \geq \limsup_{j \rightarrow \infty} f_j(x_j) \quad (2.3)$$

gilt. Für diese Folge gilt dann mit (2.2)

$$f_\infty(x) = \limsup_{j \rightarrow \infty} f_j(x_j).$$

In diesem Fall heißt f_∞ der Γ -Limes der f_j :

$$f_\infty = \Gamma\text{-lim}_{j \rightarrow \infty} f_j.$$

Eine Folge (x_j) , welche (2.3) erfüllt, wird in diesem Kontext auch *recovery Folge* genannt.

Der Γ -Limes hängt von der Metrik d von X ab. Für Betrachtungen des Γ -Limes auf topologischen Räumen X sei auf [DM93] verwiesen.

Direkt aus der Definition 2.21 ist ersichtlich, dass für den Γ -Limes

$$f_\infty(x) = \inf_{x_j \rightarrow x} \liminf_{j \rightarrow \infty} f_j(x_j) = \inf_{x_j \rightarrow x} \limsup_{j \rightarrow \infty} f_j(x_j)$$

für alle $x \in X$ gilt. Dies gibt Anlass zur nächsten Definition.

Definition 2.22 (Γ -lim inf und Γ -lim sup)

Für die Funktionenfolge $f_j: X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ und für ein $x \in X$ heißt

$$\Gamma\text{-lim inf}_{j \rightarrow \infty} f_j(x) := \inf_{x_j \rightarrow x} \liminf_{j \rightarrow \infty} f_j(x_j) \quad \text{bzw.}$$

$$\Gamma\text{-lim sup}_{j \rightarrow \infty} f_j(x) := \inf_{x_j \rightarrow x} \limsup_{j \rightarrow \infty} f_j(x_j)$$

der untere bzw. obere Γ -Limes von f_j am Punkt x .

Γ -lim inf und Γ -lim sup existieren immer. Damit ist es für den Nachweis der Γ -Konvergenz oft vorteilhaft, beide zu berechnen und auf Gleichheit zu prüfen. Außerdem gilt folgender Satz.

Satz 2.23 ([Bra02, Proposition 1.28])

Für jede Funktionenfolge f_j sind Γ -lim inf f_j und Γ -lim sup f_j von unten halbstetig.

Bekanntlich eignet sich die Approximation des Funktionals $f_\infty = \Gamma\text{-lim } f_j$ durch die f_j besonders gut für die Behandlung variationeller Minimierungsprobleme. Das zentrale Argument soll an dieser Stelle nachvollzogen werden: Angenommen, es gibt eine kompakte Menge $K \subset X$ mit

$$\inf_{x \in X} f_j(x) = \inf_{x \in K} f_j(x)$$

für alle $j \in \mathbb{N}$, dann gibt es eine Folge $(x_j)_j \subset K$ mit

$$\lim_{j \rightarrow \infty} f_j(x_j) = \liminf_{j \rightarrow \infty} \inf_{x \in X} f_j(x). \quad (2.4)$$

Ferner gibt es eine Teilfolge, die wieder mit x_j bezeichnet wird, mit $x_j \rightarrow x^* \in K$. Für jedes $\bar{x} \in X$ gibt es eine (von \bar{x} abhängige) recovery Folge \bar{x}_j mit

$$\limsup_{j \rightarrow \infty} \inf_{x \in X} f_j(x) \leq \limsup_{j \rightarrow \infty} f_j(\bar{x}_j) \stackrel{(2.3)}{\leq} f_\infty(\bar{x}) \quad (2.5)$$

Außerdem gilt für das spezielle $\bar{x} = x^*$:

$$\begin{aligned} f_\infty(x^*) &\stackrel{(2.2)}{\leq} \liminf_{j \rightarrow \infty} f_j(x_j) \stackrel{(2.4)}{=} \liminf_{j \rightarrow \infty} \inf_{x \in X} f_j(x) \\ &\leq \limsup_{j \rightarrow \infty} \inf_{x \in X} f_j(x) \stackrel{(2.5)}{\leq} f_\infty(x^*) \end{aligned}$$

Damit sehen wir, dass überall Gleichheit gilt und wegen (2.5)

$$\min_{x \in X} f_\infty(x) = \lim_{j \rightarrow \infty} \inf_{x \in X} f_j(x)$$

gilt. Dies war der Beweis des folgenden Satzes.

Satz 2.24 ([AFP00, Proposition 6.13])

Für $f_j, f_\infty: X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ gelte $f_\infty = \Gamma\text{-lim}_j f_j$. Gibt es eine kompakte Menge $K \subset X$ mit $\inf_X f_j = \inf_K f_j$, so konvergiert $\inf_X f_j$ gegen $\min_X f_\infty$. Ist ferner $(x_j)_j$ eine gegen x^* konvergente Folge mit $f_j(x_j) - \inf_X f_j \rightarrow 0$, so gilt $f_\infty(x^*) = \min_X f_\infty$.

2.2.4 Die lim inf-Ungleichung

Hier muss

$$F(u) \leq \Gamma\text{-lim inf}_{u_h \rightarrow u} F_h(u_h) = \inf \left\{ \liminf_{h \rightarrow 0} F_h(u_h) : u_h \xrightarrow{L^1(\Omega)} u \right\}$$

gezeigt werden. Sei dazu $u_h \in U_h$ eine beliebige Folge mit $u_h \xrightarrow{L^1(\Omega)} u$ mit einem $u \in SBV(\Omega)$. Man kann ohne Einschränkung davon ausgehen, dass $\liminf F_h(u_h) = \lim F_h(u_h) < \infty$ gilt.

Es sei

$$\mathcal{I}_h := \left\{ i \in \{0, \dots, n-1\} : |u_h(x_{i+1}) - u_h(x_i)| > h^\ell \right\}.$$

Zu jedem u_h wird eine Hilfsfunktion $w_h \in SBV(\Omega)$ wie folgt konstruiert: Für $i \notin \mathcal{I}_h$ sei w_h auf dem Intervall $]x_i; x_{i+1}[$ linear mit $w_h(x_i) = u_h(x_i)$ und $w_h(x_{i+1}) = u_h(x_{i+1})$. Für $i \in \mathcal{I}_h$ sei

$$w_h|_{]x_i; x_{i+1}[}(x) = \begin{cases} u_h(x_i) & \text{für } x < x_i + \frac{h}{2} \\ u_h(x_{i+1}) & \text{für } x > x_i + \frac{h}{2} \end{cases}.$$

In Abbildung 2.2 ist für das Beispiel (gestrichelte Linien) aus Abbildung 2.1 das so konstruierte w_h (durchgezogene Linien) zu sehen. Dabei war $\mathcal{I}_h = \{1\}$ und dementsprechend ist im Intervall $]x_1; x_2[$ eine Unstetigkeitsstelle zu sehen.

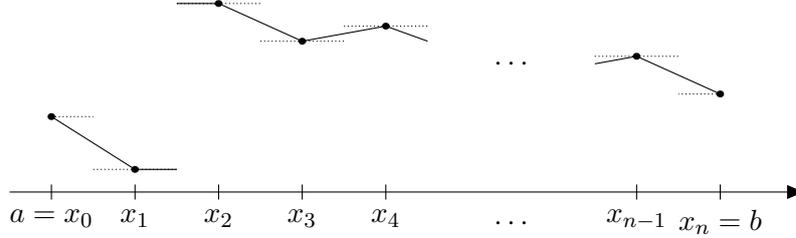


Abbildung 2.2: Skizze zur Konstruktion von w_h mit $\mathcal{I}_h = \{1\}$

Für die so konstruierten w_h gilt

$$\begin{aligned} \int_a^b |w_h'|^\alpha dx &= \sum_{i \notin \mathcal{I}_h} \int_{x_i}^{x_{i+1}} |w_h'|^\alpha dx = \sum_{i \notin \mathcal{I}_h} h \left| \frac{u_h(x_{i+1}) - u_h(x_i)}{h} \right|^\alpha \\ &= \sum_{i \notin \mathcal{I}_h} h^k d \left(\frac{u_h(x_{i+1}) - u_h(x_i)}{h^\ell} \right) \cdot \underbrace{h^{-k} \cdot h \cdot h^{\ell\alpha} \cdot h^{-\alpha}}_{=1} \end{aligned}$$

und

$$\begin{aligned} \int_{S_{w_h}} |w_h^+ - w_h^-|^\beta d\mathcal{H}^0 &= \sum_{i \in \mathcal{I}_h} |w_h^+ - w_h^-|^\beta = \sum_{i \in \mathcal{I}_h} |u_h(x_{i+1}) - u_h(x_i)|^\beta \\ &= \sum_{i \in \mathcal{I}_h} h^k d \left(\frac{u_h(x_{i+1}) - u_h(x_i)}{h^\ell} \right) \cdot \underbrace{h^{-k} h^{\ell\beta}}_{=1}. \end{aligned}$$

Also ist

$$F(w_h) = F_h(u_h) \quad (2.6)$$

für alle h .

Außerdem gilt $\|w_h - u_h\|_{L^1(\Omega)} \rightarrow 0$ für $h \rightarrow 0$. Für $i \in \mathcal{I}_h$ gilt nämlich $w_h = u_h$ auf dem Intervall $]x_i; x_{i+1}[$ und demnach ist

$$\begin{aligned} \|w_h - u_h\|_{L^1(\Omega)} &= \sum_{i \notin \mathcal{I}_h} \|w_h - u_h\|_{L^1(]x_i; x_{i+1}[)} \leq \sum_{i \notin \mathcal{I}_h} h |u_h(x_{i+1}) - u_h(x_i)| \\ &\leq h^{\ell+1} \sum_{i \notin \mathcal{I}_h} 1 \leq h^{\ell+1} \mathcal{O}(h^{-1}) = \mathcal{O}(h^\ell). \end{aligned}$$

Wegen $\|w_h - u\|_{L^1(\Omega)} \leq \|w_h - u_h\|_{L^1(\Omega)} + \|u_h - u\|_{L^1(\Omega)} \rightarrow 0$ für $h \rightarrow 0$ konvergiert auch w_h in $L^1(\Omega)$ gegen u .

Jetzt werden die wichtigen Abgeschlossenheits- und Kompaktheitsresultate in *SBV* benötigt, wie sie in [Amb95] und in [AFP00] zu finden sind.

Satz 2.25 ([AFP00, Theoreme 4.7 und 4.8])

Es seien $\varphi:]0; \infty[\rightarrow]0; \infty[$ und $\theta:]0; \infty[\rightarrow]0; \infty[$ unterhalbstetig, monoton steigend mit

$$\lim_{t \rightarrow \infty} \frac{\varphi(t)}{t} = \infty \quad \text{und} \quad \lim_{t \rightarrow 0} \frac{\theta(t)}{t} = \infty.$$

Ferner sei $\Omega \subset \mathbb{R}^d$ offen und beschränkt und $(u_h) \subset SBV(\Omega)$ eine Folge mit

$$\sup_h \left\{ \int_{\Omega} \varphi(\|\nabla u_h\|_2) dx + \int_{S_{u_h}} \theta(|u_h^+ - u_h^-|) d\mathcal{H}^{d-1} \right\} < \infty. \quad (2.7)$$

Falls (u_h) in $BV(\Omega)$ schwach* gegen u konvergiert, dann gilt:

$$\nabla u_h \xrightarrow{L^1(\Omega)} \nabla u, \quad D^j u_h \xrightarrow{\mathcal{M}(\Omega)^d}^* D^j u,$$

sowie

$$\int_{S_u} \theta(|u^+ - u^-|) d\mathcal{H}^{d-1} \leq \liminf_{h \rightarrow 0} \int_{S_{u_h}} \theta(|u_h^+ - u_h^-|) d\mathcal{H}^{d-1}$$

und falls φ konvex ist auch

$$\int_{\Omega} \varphi(\|\nabla u\|_2) dx \leq \liminf_{h \rightarrow 0} \int_{\Omega} \varphi(\|\nabla u_h\|_2) dx.$$

Jede Folge $(u_h) \subset SBV(\Omega)$, die (2.7) erfüllt und bei der $\sup_h \|u_h\|_{L^\infty(\Omega)} < \infty$ gilt, besitzt eine in $BV(\Omega)$ schwach* konvergente Teilfolge.

Ist $K \subset \mathbb{R}$ kompakt und gilt $u_h(x) \in K$ für \mathcal{L}^1 -fast alle $x \in \Omega$, dann ist $\sup_h \|u_h\|_{L^\infty(\Omega)} < \infty$ und die Voraussetzungen des Satzes 2.25 sind mit $\varphi(z) = z^\alpha$ und $\theta(z) = z^\beta$ erfüllt. Der Satz zeigt, dass F bzgl. der schwach* Topologie von $BV(\Omega)$ unterhalbstetig ist. Ferner gibt es eine Teilfolge von w_h , die wieder mit w_h bezeichnet wird, für die $F(u) \leq \liminf F(w_h)$ gilt, da mit u_h auch w_h in $L^1(\Omega)$ gegen u konvergiert. Zusammen mit (2.6) erhält man

$$F(u) \leq \liminf_{h \rightarrow 0} F(w_h) = \liminf_{h \rightarrow 0} F_h(u_h)$$

und damit die \liminf -Ungleichung, da die Folge (u_h) beliebig war.

2.2.5 Die lim sup-Ungleichung

Für diese Ungleichung muss zu einem $u \in U$ eine recovery Folge (u_h) mit $u_h \in U_h$ konstruiert werden, so dass

$$F(u) \geq \limsup_{h \rightarrow 0} F_h(u_h)$$

erfüllt ist. Ohne Einschränkung ist $F(u) < \infty$. Da u wegen $\mathcal{H}^0(S_u) < \infty$ nur endlich viele approximative Unstetigkeitsstellen hat, ist u eine stückweise $W^{1,1}$ Funktion. Wegen $\|u'\|_{L^\alpha(\Omega)}^\alpha \leq F(u) < \infty$ gilt $u' \in L^\alpha(\Omega)$. Folglich ist u eine stückweise $W^{1,\alpha}$ Funktion und auf den entsprechenden Teilintervallen \mathcal{L}^1 -fast überall differenzierbar [EG92, Theorem 1 in Abschnitt 6.2]. Setzt man

$$u_h(x_i) := u^+(x_i),$$

so gilt für hinreichend kleine h und $]x_i; x_{i+1}[\cap S_u = \emptyset$

$$\begin{aligned} V_h(u_h(x_i), u_h(x_{i+1})) &= h^k d\left(\frac{u_h(x_{i+1}) - u_h(x_i)}{h^\ell}\right) \\ &= h^k d\left(\frac{u_h(x_{i+1}) - u_h(x_i)}{h} h^{1-\ell}\right) = h^k \left|\frac{u_h(x_{i+1}) - u_h(x_i)}{h}\right|^\alpha h^{\alpha(1-\ell)} \\ &= h |u'_h(x_i)|^\alpha + \mathcal{O}(h^2) = \int_{x_i}^{x_{i+1}} |u'_h(x)|^\alpha dx + \mathcal{O}(h). \end{aligned}$$

Für die Teilintervalle mit $]x_i; x_{i+1}[\cap S_u \neq \emptyset$ ist

$$\begin{aligned} V_h(u_h(x_i), u_h(x_{i+1})) &= h^k d\left(\frac{u_h(x_{i+1}) - u_h(x_i)}{h^\ell}\right) \\ &= h^k \frac{|u_h(x_{i+1}) - u_h(x_i)|^\beta}{h^{\ell\beta}} = |u_h(x_{i+1}) - u_h(x_i)|^\beta. \end{aligned}$$

Für die so konstruierten u_h gilt also $F_h(u_h) \rightarrow F(u)$ für $h \rightarrow 0$.

2.2.6 Beispiel

Jetzt wird Satz 2.20 auf das konkrete Problem

$$F(u) = \lambda^2 \int_{-a}^a (u')^2 dx + \int_{-a}^a (u - v)^2 dx + \mu \mathcal{H}^0(S_u) = \min!$$

angewendet, um zu sehen, wie eine Eigenschaft einer Optimallösung u ins Diskrete vererbt wird. Als v sei der Rechtecksimpuls

$$v(x) = \begin{cases} H & \text{falls } |x| \leq 1 \\ 0 & \text{sonst} \end{cases}$$

vorgegeben. Das Beispiel ist so gewählt, dass eine Optimallösung analytisch einfach auszurechnen ist. Wir wollen weiter annehmen, dass $a \gg 1$ ist und so keine Randeffekte auftreten.

Lässt man keine Unstetigkeiten zu und sucht ein optimales $u_1 \in W^{1,2}(\Omega)$, so erhält man (vgl. [BZ87, Appendix A])

$$u_1(x) = \int_{-a}^a G(x,y)v(y) dy.$$

Dabei ist G die sogenannte Green-Funktion. Für $a = \infty$ ist

$$G(x,y) = \frac{1}{2\lambda} \exp\left(-\frac{|x-y|}{\lambda}\right).$$

Damit ergibt sich

$$F(u_1) = \lambda H^2(1 - e^{-2/\lambda}). \quad (2.8)$$

Lässt man Unstetigkeiten zu und sucht ein $u_2 \in SBV(\Omega)$ (zusätzlich mit $u' \in L^2(\Omega)$), so ist ein natürlicher Kandidat für das Optimum $u_2 = v$ mit

$$F(u_2) = 2\mu. \quad (2.9)$$

Mit dem Ansatz $F(u_1) = F(u_2)$ lässt sich der Schwellwert

$$H_s = \sqrt{\frac{2\mu}{\lambda(1 - e^{-2/\lambda})}}$$

berechnen, so dass für $H = H_s$ sowohl u_1 , als auch u_2 Optimallösungen sind. Für $0 \leq H < H_s$ ist die Optimallösung durch u_1 gegeben und für $H_s < H$ ist u_2 die Optimallösung.

Zur Diskretisierung von F gehen wir nach einer Idee von Geman und Geman [GG84] vor und führen einen „line process“ ein (die Idee wurde im Zweidimensionalen in Verbindung mit stochastischen Algorithmen vorgeschlagen, siehe Abschnitt 4.4). Soll im Intervall $]x_i; x_{i+1}[$ eine Unstetigkeit sein, so wird der Index i in die Kantenmenge K aufgenommen. Natürlich ist K unbekannt, so dass

$$\lambda^2 \sum_{i \notin K} h \left(\frac{u_h(x_{i+1}) - u_h(x_i)}{h} \right)^2 + \sum_i h (u_h(x_i) - v_h(x_i))^2 + \mu \mathcal{H}^0(K) = \min_{u_h, K}!$$

eine mögliche Diskretisierung darstellt. Äquivalent dazu ist es

$$F_h(u_h) = \sum_i \min \left(\left(\frac{u_h(x_{i+1}) - u_h(x_i)}{\sqrt{h}} \right)^2, \frac{\mu}{\lambda^2} \right) + \frac{h}{\lambda^2} \sum_i (u_h(x_i) - v_h(x_i))^2 = \min_{u_h}!$$

zu minimieren. F_h hat genau die in Satz 2.20 geforderte Form (durch Umskalierung von u_h und v_h mit $\tilde{u}_h := \lambda u_h / \sqrt{\mu}$ und $\tilde{v}_h := \lambda v_h / \sqrt{\mu}$ kann der

erste Term genau auf das dortige Aussehen gebracht werden) mit $d(z) = \min(z^2, \mu/\lambda^2)$.

Falls die generische Situation vorliegt ($H \neq H_s$) und es für F ein eindeutiges Minimum u gibt, werden folglich die Minimierer u_h von F_h in der L^1 -Norm gegen dieses Minimum konvergieren.

In Abbildung 2.3 ist für folgende Parameter

$$a = 10, \quad \lambda = \sqrt{2}, \quad \mu = \frac{1}{5} \quad \text{und} \quad H = 0.6 < H_s \approx 0.611 \quad (2.10)$$

jeweils das Minimum u von $F(u)$ gepunktet dargestellt und die diskreten Minima u_h von $F_h(u_h)$ (gerechnet wurde auf dem Intervall $[-10; 10]$; gezeichnet wurde aber nur der relevante Bereich um die 0). Benachbarte Intervallmittelpunkte wurden dabei verbunden, sofern zwischen ihnen keine Unstetigkeitsstelle liegt.

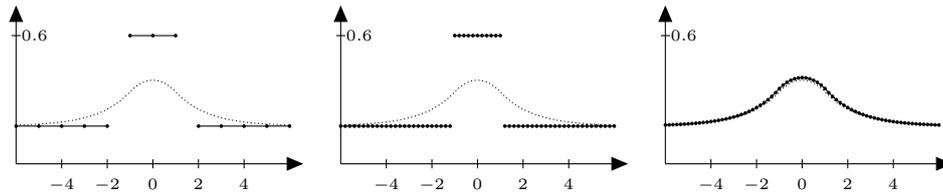


Abbildung 2.3: Vergleich von diskretem Optimum u_h und Optimallösung u für die Parameter aus (2.10) und bei der Einteilung in $n \in \{21, 101, 121\}$ Teilintervalle

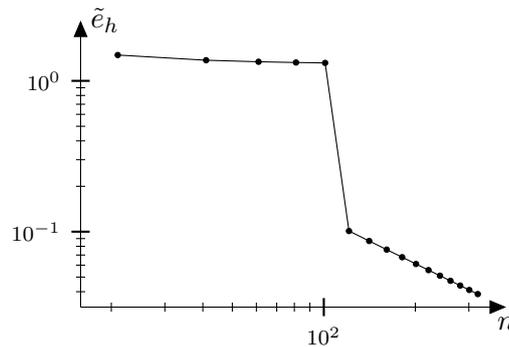


Abbildung 2.4: Konvergenzverlauf

In Abbildung 2.4 ist in Abhängigkeit von h (bzw. n)

$$\tilde{e}_h := h \sum_i |u_h(x_i) - u(x_i)|$$

als Näherung für die L^1 -Abweichung $\|u_h - u\|_{L^1(\Omega)}$ angetragen. Erst bei ca. $n = 100$ ist auch im Diskreten der Lösungskandidat ohne Unstetigkeitsstellen minimal. Erst für $n > 100$ ist deshalb die Konvergenz zu beobachten. Ein weiterer Punkt ist dabei zu beachten: Für eine beliebige monotone Folge $(n_i) \subset \mathbb{N}$ ist der Konvergenzverlauf meist *nicht* monoton. Durch verschiedene diskrete Effekte bei der Einteilung in Teilintervalle ist manchmal die Optimallösung u_{h_1} näher an u als die Optimallösung für ein u_{h_2} obwohl $h_2 < h_1$. Daher wurden für den Konvergenzplot nur n mit $n \bmod 20 = \text{const}$ gewählt. Verändert man bei den Parametern (2.10) lediglich H auf den Wert 0.62, so ist $u = v$ die optimale kontinuierliche Lösung und auch im Diskreten ist hat dann die Lösung zwei Unstetigkeitsstellen (für hinreichend kleine h).

Dynamische Programmierung zur Minimierung von F_h

In diesem Einschub soll erläutert werden, wie in obigem Beispiel die diskreten Optimallösungen u_h von F_h berechnet wurden. Für eindimensionale Probleme gibt es einen Algorithmus, der in endlich vielen Schritten für ein beliebiges v_h die optimale Lösung u_h ermittelt.

Sei

$$E_{\ell,m}^k := \min_{\substack{u_h \\ \mathcal{H}^0(K)=k}} \sum_{\substack{i=\ell \\ i \notin K}}^{m-1} \left(\frac{u_h(x_{i+1}) - u_h(x_i)}{\sqrt{h}} \right)^2 + \sum_{i=\ell}^m h(u_h(x_i) - v_h(x_i)) + \mu k$$

der Wert des Minimums von F_h , wenn nur das Intervall $]x_\ell; x_m[$ betrachtet wird ($\ell \leq m$) und genau k Unstetigkeitsstellen vorliegen müssen. Für alle $\ell \leq m$ kann $E_{\ell,m}^0$ durch Lösen eines tridiagonalen linearen Gleichungssystems berechnet werden.

Es gilt

$$E_{0,i}^k = \min_{j=k, \dots, i-1} (E_{0,j}^{k-1} + \mu + E_{j,i}^0). \quad (2.11)$$

Hinter dieser Formel steckt folgende Überlegung: Wenn man im Intervall $]x_0; x_i[$ genau k Unstetigkeitsstellen setzen darf, dann hat man für die letzte Stelle die möglichen Positionen $j = k, \dots, i-1$; die anderen $k-1$ Unstetigkeiten müssen dann im Intervall $]x_0; x_j[$ optimal gesetzt werden, diese liefern den Beitrag $E_{0,j}^{k-1}$. Das μ ist der Beitrag für die letzte Unstetigkeitsstelle bei Position j und der Term $E_{j,i}^0$ gibt den minimalen Wert des Funktionals für das Intervall $]x_j; x_i[$ ohne weiteren Unstetigkeiten an.

Durch wiederholtes Anwenden von (2.11) lassen sich $E_{0,n}^k$ für alle k berechnen und durch einfachen Vergleich kann eine optimale Anzahl von Kanten k ermittelt werden. Führt man bei (2.11) Buch, bei welchem j das Minimum angenommen wurde, so lässt sich auch das optimale u_h angeben.

2.3 Funktionale auf uniformen quadratischen Gittern und deren Gamma-Limes

Nach der Motivation im Eindimensionalen im vorherigen Abschnitt 2.2 wird in diesem Abschnitt versucht, das Ergebnis ins Zweidimensionale zu übertragen. Im Abschnitt 2.3.1 wird die Notation und in 2.3.2 das Ziel formuliert. Im Anschluss folgt in mehreren Etappen die Untersuchung über die Gamma-Konvergenz, die im Satz 2.28 im Abschnitt 2.3.9 endet. Konvergenzuntersuchungen für das Mumford-Shah Funktional sind in [Neg99] und [Cha95] zu finden.

2.3.1 Notation

Im Folgenden sei $\Omega \subset \mathbb{R}^2$ ein beschränktes Gebiet mit Lipschitz-Rand $\partial\Omega$. Wir betrachten uniforme quadratische Gitter der Gitterweite h , die durch die beiden Richtungen $e_1 = (1, 0)^T$ und $e_2 = (0, 1)^T$ erzeugt werden. Die Menge aller offenen Gitterelemente in Ω werde mit \mathcal{T}_h bezeichnet. Der Vektorraum der Funktionen, die auf jedem $T \in \mathcal{T}_h$ konstant sind, wird mit

$$U_h := \{ u: \Omega \rightarrow \mathbb{R} : u|_T = \text{const für alle } T \in \mathcal{T}_h \}$$

bezeichnet. Zwei Elemente R und T , bei denen $\overline{R} \cap \overline{T}$ eine Strecke ist, heißen benachbart. Die Schreibweise lautet $R \sim T$. Falls die Richtung der Nachbarschaft entscheidend ist, schreiben wir $R \sim_{e_i} T$.

2.3.2 Problemstellung und Ziel

Wir vermuten ein zu Satz 2.20 analoges Ergebnis im Zweidimensionalen. Betrachtet wird also für $h > 0$ und $u_h \in U_h$ ist das Funktional

$$F_h(u_h) := \sum_{R \sim T} V_h(u_h|_R, u_h|_T) + \sum_T |T| f(u_h|_T) \quad (2.12)$$

mit $V_h: \mathbb{R} \times \mathbb{R} \rightarrow [0; \infty[$ und $f: \mathbb{R} \rightarrow [0; \infty[$. Hierbei habe V_h die entsprechende Darstellung

$$V_h(a, b) = h^{k+1} d\left(\frac{|a-b|}{h^\ell}\right) \quad (2.13)$$

mit $d(z) = \min(|z|^\alpha, |z|^\beta)$ und $0 \leq \beta < 1 < \alpha \leq 2$ sowie $\ell = \frac{\alpha-1}{\alpha-\beta}$ und $k = \beta\ell$. Es soll untersucht werden, unter welchen Bedingungen F_h einen Γ -Limes der Form

$$F(u) := \int_{S_u} V(u^+, u^-) \Phi(\nu_u) d\mathcal{H}^1 + \int_\Omega f(u) dx + \int_\Omega g(\nabla u) dx \quad (2.14)$$

bezüglich der L^1 -Topologie besitzt.

Im Laufe der folgenden Abschnitte werden sich nach einigen Annahmen sowohl die Voraussetzungen als auch die entsprechenden Darstellungen für die angegebenen Funktionen ergeben. Diese sind auch im Abschnitt 2.3.9 auf Seite 42 nochmal an zentraler Stelle zusammengefasst. Leider kann in 2.3.8 die notwendige \liminf -Ungleichung nur für den Fall $\beta = 0$ gezeigt werden, so dass Satz 2.28 nur für diesen Fall formuliert wurde.

2.3.3 Überblick über den Gamma-Konvergenz Beweis

In den folgenden Abschnitten wird untersucht, unter welchen Voraussetzungen $F = \Gamma\text{-}\lim F_h$ bzgl. der L^1 -Topologie gilt. Dazu wird in Abschnitt 2.3.4 eine explizite Darstellung für die von einem Gitter erzeugte Anisotropie Φ angegeben. In Abschnitt 2.3.5 wird ein wichtiger Teilraum $\mathcal{W} \subset SBV(\Omega)$ eingeführt, der bei der Herleitung der \limsup -Ungleichung (2.3) eine große Rolle spielen wird. In 2.3.6 wird die \limsup -Ungleichung zunächst in \mathcal{W} bewiesen und im darauffolgenden Abschnitt 2.3.7 wird die Gültigkeit der Ungleichung auf ganz $SBV(\Omega)$ gezeigt. Im Anschluss wird in 2.3.8 die \liminf -Ungleichung (2.2) hergeleitet. In 2.3.9 findet sich dann der ausformulierte Satz mit allen Voraussetzungen und Annahmen, die in den vorherigen Abschnitten eingegangen sind.

2.3.4 Strecken auf uniformen Gittern

Im Γ -Limes F in der Gleichung (2.14) im Abschnitt 2.3.2 beschreibt die Funktion Φ die vom Gitter erzeugte Anisotropie (in Abhängigkeit einer Normalen ν). Um eine explizite Darstellung von Φ herzuleiten, werden in diesem Abschnitt zunächst Strecken betrachtet und ermittelt, welche Länge diese, entlang eines Gitters gemessen, haben. Die Idee (siehe auch [Lud02, Abschnitt 3.5]) orientiert sich an [Neg99]. Dort wurde im Kontext des Mumford-Shah-Funktional die von drei speziellen Dreiecksgittern erzeugte Anisotropie untersucht. Der hier vorgestellte Ansatz lässt sich jedoch auf beliebige uniforme Gitter anwenden und benötigt keine umfangreichen Fallunterscheidungen und Spezialuntersuchungen, wie sie in [Neg99] zu finden sind.

Wir betrachten Strecken S der Länge $\mathcal{H}^1(S) = 1$. Wir wollen S parallel zu den e_i mit Hilfe des uniformen Gitters \mathcal{T}_h messen.

Dazu sei ν ein Normalenvektor von S mit $\|\nu\|_2 = 1$. Mit ν ist natürlich auch $-\nu$ ein Normalenvektor von S . Im folgenden sei ν fest gewählt.

Die Anzahl der Ränder der Gitterelemente, die S für den Grenzfall $h \rightarrow 0$ in Richtung e_i^\perp schneidet, ist durch $\langle e_i, \nu \rangle / h$ gegeben (vgl. Abbildung 2.5). Geben wir jedem Randstück die Länge h , so erhalten wir

$$\Phi(\nu) := h \sum_i \frac{|\langle e_i, \nu \rangle|}{h} = \sum_i |\langle e_i, \nu \rangle| = \|\nu\|_1$$

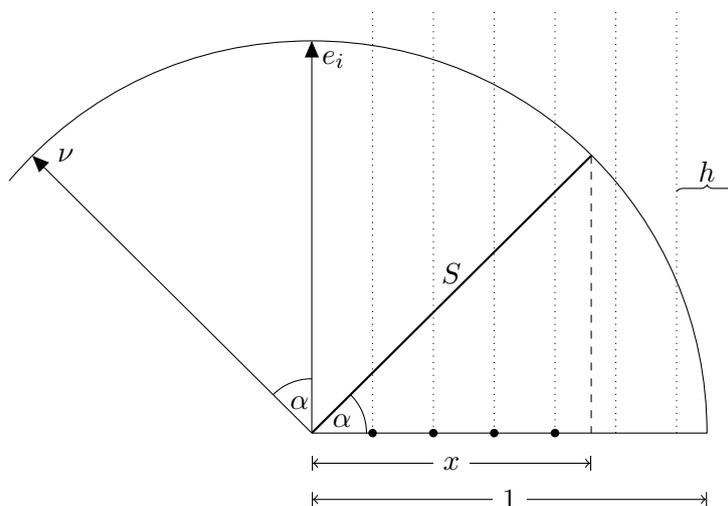


Abbildung 2.5: Asymptotische Anzahl von Gitterelementrändern, die in e_i^\perp Richtung geschnitten werden

als Gesamtlänge von S , gemessen auf dem Gitter. Man beachte, dass die rechte Seite nicht mehr von h abhängt. Um Φ auch für solche ν auswerten zu können, die nicht normiert sind, definieren wir für alle $\nu \neq 0$

$$\Phi(\nu) := \|\nu\|_1. \quad (2.15)$$

Abbildung 2.6 zeigt für eine Beispielstrecke (schwarz) die nach obigem Vorgehen ermittelte Länge auf dem diskreten Gitter (blau).

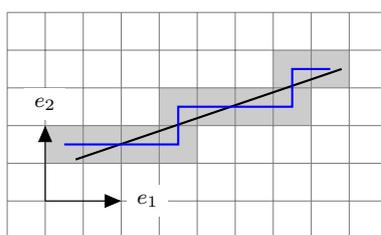


Abbildung 2.6: Strecke mit am Gitter betrachteter Länge

Um Spezialfälle auszuschließen, etwa wenn eine Strecke nicht ein offenes $T \in \mathcal{T}_h$, aber Teile des Randes ∂T schneidet, legen wir folgende Sprechweise fest: Eine Strecke S (mit Normale ν) schneidet T , falls

$$S \cap \bar{T} \neq \emptyset \quad \text{und} \quad \mathcal{L}^2(\{x \in \Omega : x = s + t\nu, s \in S, t \geq 0\} \cap \bar{T}) > 0 \quad (2.16)$$

gilt.

Damit ergibt sich für eine Strecke S mit Einheitsnormale ν und beliebiger Länge

$$\lim_{h \rightarrow 0} \sum_{S \text{ schneidet } T} h = \Phi(\nu) \mathcal{H}^1(S) = \int_S \Phi(\nu) d\mathcal{H}^1.$$

2.3.5 Funktionen $w \in \mathcal{W}(\Omega)$ auf uniformen Gittern

In [CT99] wurde eine spezielle Teilmenge $\mathcal{W} \subset SBV(\Omega)$ beschrieben, die eine hilfreiche Dichteigenschaft besitzt.

Definition 2.26

Der Raum aller Funktionen $w \in SBV(\Omega)$ mit

1. $\mathcal{H}^1(\overline{S_w} \setminus S_w) = 0$,
2. $\overline{S_w}$ ist der Schnitt von Ω mit der Vereinigung endlich vieler paarweise disjunkter Strecken und
3. $w \in W^{k,\infty}(\Omega \setminus \overline{S_w})$ für alle $k \in \mathbb{N}$

wird mit $\mathcal{W}(\Omega)$ bezeichnet.

Der folgende Satz stellt die schon angesprochene, später benötigte Dichteigenschaft parat. Der dabei verwendete Raum $SBV^p(\Omega)$ ist für ein $1 < p \leq 2$ und $\Omega \subset \mathbb{R}^d$ durch

$$SBV^p(\Omega) := \{ u \in SBV(\Omega) : \mathcal{H}^{d-1}(S_u) < \infty \text{ und } \nabla u \in [L^p(\Omega)]^d \}$$

festgelegt. Er ist damit ein Unterraum von $SBV(\Omega)$.

Satz 2.27 ([CT99], Theorem 3.1 zusammen mit Bemerkung 3.5)

Ist $u \in SBV^p(\Omega) \cap L^\infty(\Omega)$ mit $1 < p \leq 2$, dann gibt es eine Folge $(w_k)_k \subset \mathcal{W}(\Omega)$, so dass

$$w_k \xrightarrow{L^1(\Omega)} u, \quad \nabla w_k \xrightarrow{[L^p(\Omega)]^2} \nabla u, \quad \limsup_{k \rightarrow \infty} \|w_k\|_{L^\infty(\Omega)} \leq \|u\|_{L^\infty(\Omega)}$$

und

$$\limsup_{k \rightarrow \infty} \int_{A \cap S_{w_k}} \varphi(x, w_k^+, w_k^-, \nu_{w_k}) d\mathcal{H}^1 \leq \int_{A \cap S_u} \varphi(x, u^+, u^-, \nu_u) d\mathcal{H}^1$$

für jedes $A \subset \subset \Omega$ und für oberhalbstetige Funktionen $\varphi: \Omega \times \mathbb{R} \times \mathbb{R} \times S^1 \rightarrow [0; \infty[$ mit $\varphi(x, a, b, \nu) = \varphi(x, b, a, -\nu)$ für alle $x \in \Omega$, $a, b \in \mathbb{R}$ und $\nu \in S^1$.

2.3.6 Die lim sup-Ungleichung auf $\mathcal{W}(\Omega)$

Nach dieser Vorarbeit beginnt jetzt die Untersuchung des Gamma-Limes von (2.12). In diesem Abschnitt wird die lim sup-Ungleichung (2.3) betrachtet.

Gesucht ist also für ein $w \in \mathcal{W}(\Omega)$ eine recovery Folge $v_h \in U_h$ mit

$$v_h \xrightarrow{L^1(\Omega)} w \quad \text{und} \quad \limsup_{h \rightarrow 0} F_h(v_h) \leq F(w).$$

Aus Regularitätsgründen ist $w \in C^\infty(\Omega \setminus \overline{S_w})$. Betrachten wir zunächst ein Gebiet $D \subset \Omega$, in dem $w \in C^\infty(D)$ gilt. In diesem Fall ist die Festlegung

$$v_h|_T := w(x), \quad x \text{ Mittelpunkt von } T \quad (2.17)$$

für $T \in \mathcal{D}$ wohldefiniert. Da $w|_D$ sogar Riemann integrierbar ist, gilt $\|w - v_h\|_{L^1(D)} \rightarrow 0$ für $h \rightarrow 0$ und damit

$$\lim_{h \rightarrow 0} \sum_{\substack{T \in \mathcal{T}_h \\ T \subset D}} |T| f(v_h|_T) = \int_D f(x) dx.$$

Um den anderen Term des Funktionals (2.12)

$$\sum_{R \sim T} V_h(v_h|_R, v_h|_T)$$

näher zu untersuchen, betrachten wir eine Zeile von Gitterzellen, wobei (x_i, y) die jeweiligen Zellenmittelpunkte seien.

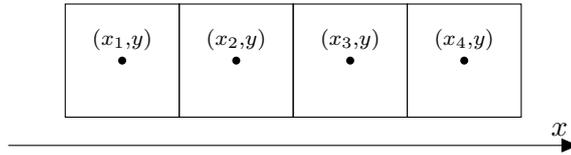


Abbildung 2.7: Zeile von Gitterzellen in e_1 Richtung

Im weiteren verwenden wir die Darstellung (2.13) von V_h zusammen mit dem asymptotischen Verhalten von d :

$$d(t) \sim t^\alpha \quad \text{für } t \rightarrow 0 \quad \text{und} \quad d(t) \sim t^\beta \quad \text{für } t \rightarrow \infty. \quad (2.18)$$

Dann gilt für die Nachbarn $R \sim_{e_1} T$ falls h hinreichend klein ist

$$\begin{aligned} V_h(v_h|_R, v_h|_T) &= V_h(w(x_{i+1}, y), w(x_i, y)) \\ &= h^{k+1} d \left(\frac{|w(x_{i+1}, y) - w(x_i, y)|}{h} h^{1-l} \right) \\ &\sim h^{k+1} \left(\frac{|w(x_{i+1}, y) - w(x_i, y)|}{h} \right)^\alpha h^{\alpha(1-l)} \\ &= h^2 |e_1^T \nabla w(x_i, y)|^\alpha + \mathcal{O}(h^3). \end{aligned}$$

Somit ist für eine einzige Zeile

$$\sum_{\substack{T \sim_{e_1} R \\ \text{eine Zeile}}} V_h(v_h|_R, v_h|_T) = h \int_{x_{\min}}^{x_{\max}} |e_1^T \nabla w(x, y)|^\alpha dx + \mathcal{O}(h^2).$$

Weil es $\frac{1}{h}$ viele Zeilen gibt, ist

$$\sum_{T \sim_{e_1} R} V_h(v_h|_R, v_h|_T) = \int_{\Omega} |e_1^T \nabla w(x, y)|^\alpha dx + \mathcal{O}(h)$$

und für beide Nachbarschaftsrichtungen zusammen gilt

$$\lim_{h \rightarrow 0} \sum_{T \sim R} V_h(v_h|_R, v_h|_T) = \int_{\Omega} \|\nabla w\|_{\alpha}^{\alpha} dx.$$

Für das quadratische Gitter ist also die Funktion g im Funktional (2.14) durch

$$g(v) = \|v\|_{\alpha}^{\alpha}. \quad (2.19)$$

festgelegt.

Kommen wir nun zu den $T \in \mathcal{T}_h$, die von $\overline{S_w}$ im Sinne von (2.16) geschnitten werden. Da es sich bei $\overline{S_w}$ um die Vereinigung endlich vieler Strecken (in Ω) handelt, gibt es nur $\mathcal{O}(\frac{1}{h})$ viele solche T . Speziell ist

$$\lim_{h \rightarrow 0} \sum_{\substack{T \in \mathcal{T}_h \\ \overline{S_w} \text{ schneidet } T}} h^2 f(v_h|_T) = 0$$

und zwar unabhängig davon, wie $v_h|_T$ festgelegt wird. Wir legen für solche T

$$v_h|_T := \min_{x \in T} w(x) \quad (2.20)$$

fest.

Da nach Definition 2.26 alle Strecken in $\overline{S_w}$ paarweise disjunkt sind, können wir annehmen, dass jedes T , welches von einer Strecke geschnitten wird, nur

von einer Strecke geschnitten wird, falls h hinreichend klein ist. Für den zweiten Term des Funktionals gilt

$$\begin{aligned} \sum_{S \sim T} V_h(v_h|_S, v_h|_T) &= \sum_{S \sim T} h^{k+1} d\left(\left|\frac{v_h|_S - v_h|_T}{h^l}\right|\right) \\ &\sim \sum_{S \sim T} h^{k+1} \left(\left|\frac{v_h|_S - v_h|_T}{h^l}\right|\right)^\beta = h \sum_{S \sim T} |v_h|_S - v_h|_T|^\beta \end{aligned}$$

und mit den Vorbereitungen aus 2.3.4

$$\lim_{h \rightarrow 0} \sum_{\substack{S \sim T \\ \overline{S_w} \text{ schneidet } T}} V_h(v_h|_S, v_h|_T) = \int_{S_w} V(w^+, w^-) \Phi(\nu_w) d\mathcal{H}^1.$$

mit

$$V(a, b) := |a - b|^\beta. \quad (2.21)$$

Für den Fall $\beta = 0$ bedeutet dies $V(a, b) \equiv 1$.

Da die Mengen der Unstetigkeitsstellen von w bzw. v_h jeweils L^2 -Nullmengen sind, sind w und v_h sogar Riemann integrierbar und mit den Festlegungen (2.17) und (2.20) gilt demnach $v_h \xrightarrow{L^1(\Omega)} w$ sowie

$$\begin{aligned} F''(w) &:= \Gamma\text{-lim sup}_{u_h \rightarrow w} F_h(u_h) = \inf \left\{ \limsup_{h \rightarrow 0} F_h(u_h) : u_h \xrightarrow{L^1(\Omega)} w \right\} \\ &\leq \limsup_{h \rightarrow 0} F_h(v_h) = \lim_{h \rightarrow 0} F_h(v_h) = F(w), \end{aligned} \quad (2.22)$$

womit die geforderte Ungleichung im Raum \mathcal{W} gezeigt wäre.

2.3.7 Die lim sup-Ungleichung auf $SBV^\alpha(\Omega)$

Sei nun $u \in SBV^\alpha(\Omega) \cap L^\infty(\Omega)$ mit $1 < \alpha \leq 2$. Mit Hilfe des Satzes 2.27 (mit $p = \alpha$) soll nun die Ungleichung (2.22) für den größeren Raum gezeigt werden. Man kann annehmen, dass $F(u) < \infty$ (anderenfalls ist die Ungleichung automatisch erfüllt). Sei also $(w_h) \subset \mathcal{W}(\Omega)$ eine Folge mit den im Satz beschriebenen Eigenschaften.

Der Γ -lim sup ist von unten halbstetig (Satz 2.23)

$$F''(u) \leq \liminf_{h \rightarrow 0} F''(w_h).$$

Für jedes dieser w_h sei $v_k^h \in U_k$ eine nach 2.3.6 konstruierte recovery Folge (für $k \rightarrow 0$). Dann gilt wegen (2.22)

$$\begin{aligned} F''(u) &\leq \liminf_{h \rightarrow 0} F''(w_h) = \liminf_{h \rightarrow 0} \left\{ \inf_{u_k \rightarrow w_h} \left\{ \limsup_{k \rightarrow 0} F_k(u_k) \right\} \right\} \\ &\leq \liminf_{h \rightarrow 0} \left\{ \limsup_{k \rightarrow 0} F_k(v_k^h) \right\} = \liminf_{h \rightarrow 0} F(w_h) \\ &\leq \limsup_{h \rightarrow 0} F(w_h). \end{aligned}$$

Ohne Einschränkung (durch Auswahl einer Teilfolge, die wieder w_h benannt wird) kann man davon ausgehen, dass

$$w_h(x) \rightarrow u(x) \quad \text{und} \quad \nabla w_h(x) \rightarrow \nabla u(x)$$

für \mathcal{L}^2 -fast alle $x \in \Omega$ gilt. Der Satz 2.27 liefert weiter, dass

$$\limsup_{h \rightarrow 0} \|w_h\|_{L^\infty(\Omega)} \leq \|u\|_{L^\infty(\Omega)} < \infty$$

und somit (für eine Teilfolge) alle w_h nur Werte in einem Kompaktum $K \subset \mathbb{R}$ annehmen. Ist f lokal Lipschitz-stetig, so ist f auf K Lipschitz-stetig. Daher sind $f \circ u$ und $f \circ w_h$ aus $L^1(\Omega)$ mit

$$\int_{\Omega} |f(u) - f(w_h)| dx \leq L \int_{\Omega} |u - w_h| dx = L \|u - w_h\|_{L^1(\Omega)} \rightarrow 0.$$

Also gilt

$$\int_{\Omega} f(w_h) dx \rightarrow \int_{\Omega} f(u) dx.$$

Nach Satz 2.27 konvergieren die Gradienten ∇w_h in der $[L^\alpha(\Omega)]^2$ -Norm gegen ∇u . Mit der Gestalt von g aus (2.19) folgt damit

$$\int_{\Omega} g(\nabla w_h) dx \rightarrow \int_{\Omega} g(\nabla u) dx.$$

Mit der weiteren Aussage, dass die Ungleichung

$$\limsup_{h \rightarrow 0} \int_{S_{w_h}} V(w_h^+, w_h^-) \Phi(\nu_{w_h}) d\mathcal{H}^1 \leq \int_{S_u} V(u^+, u^-) \Phi(\nu_u) d\mathcal{H}^1$$

gilt, erhält man insgesamt

$$F''(u) \leq \limsup_{h \rightarrow 0} F(w_h) \leq F(u)$$

und damit die gewünschte lim sup-Ungleichung auf $SBV^\alpha(\Omega)$.

2.3.8 Die lim inf-Ungleichung

Es muss

$$F(u) \leq F'(u) := \Gamma\text{-lim inf}_{u_h \rightarrow u} F_h(u_h) = \inf \left\{ \liminf_{h \rightarrow 0} F_h(u_h) : u_h \xrightarrow{L^1(\Omega)} u \right\}$$

gezeigt werden. Sei dazu $u_h \in U_h$ eine beliebige Folge mit $u_h \xrightarrow{L^1(\Omega)} u$ mit einem $u \in SBV^\alpha(\Omega)$.

Wir beginnen mit dem Integral-Term über Ω . Ist f Lipschitz-stetig, dann sind auch $f \circ u, f \circ u_h \in L^1(\Omega)$. Aus

$$\begin{aligned} \int_{\Omega} f(u) dx &= \sum_{T \in \mathcal{T}_h} \int_T f(u) dx \\ &= \sum_{T \in \mathcal{T}_h} \left(\int_T (f(u) - f(u_h|_T)) dx + \int_T f(u_h|_T) dx \right) \\ &\leq \sum_{T \in \mathcal{T}_h} \int_T |f(u) - f(u_h|_T)| dx + \sum_{T \in \mathcal{T}_h} |T| f(u_h|_T) \\ &= \|u - u_h\|_{L^1(\Omega)} + \sum_{T \in \mathcal{T}_h} |T| f(u_h|_T) \end{aligned}$$

folgt für $u_h \xrightarrow{L^1(\Omega)} u$

$$\int_{\Omega} f(u) dx \leq \liminf_{h \rightarrow 0} \sum_{T \in \mathcal{T}_h} |T| f(u_h|_T).$$

Überblick

Die Herleitung der Γ -lim-inf-Ungleichung für den Sprung- bzw. Gradiententerm wird in mehreren Etappen stattfinden. Dazu seien

$$\tilde{F}(u) := \int_{S_u} V(u^+, u^-) \Phi(\nu_u) d\mathcal{H}^1 + \int_{\Omega} g(\nabla u) dx$$

und

$$\tilde{F}_h(u_h) := \sum_{R \sim T} V_h(u_h|_R, u_h|_T).$$

Hier wird zunächst ein grober Überblick über die Schritte gegeben. In den folgenden Abschnitten sind dann die Einzelheiten zu finden.

Ohne Einschränkung kann $\liminf \tilde{F}_h(u_h) < \infty$ angenommen werden, da ansonsten nichts zu zeigen ist. Außerdem kann zu einer Teilfolge, die wieder mit u_h bezeichnet wird, übergegangen werden, so dass $\liminf \tilde{F}_h(u_h) = \lim \tilde{F}_h(u_h)$ gilt. Außerdem gelte $u(x), u_h(x) \in K$ für \mathcal{L}^2 -fast alle $x \in \Omega$ und alle h für ein Kompaktum $K \subset \mathbb{R}$.

Zunächst wird zu jedem u_h ein $w_h \in SBV^\alpha(\Omega)$ konstruiert, so dass einerseits $\|u_h - w_h\|_{L^1(\Omega)} \rightarrow 0$ für $h \rightarrow 0$ gilt und andererseits

$$\tilde{F}(w_h) \leq \tilde{F}_h(u_h) + \mathcal{O}(h). \quad (2.23)$$

Dann wird die Unterhalbstetigkeit von \tilde{F} verwendet, um

$$\tilde{F}(u) \leq \liminf_{h \rightarrow 0} \tilde{F}(w_h) \quad (2.24)$$

zu erhalten. Für die Unterhalbstetigkeit sind Kompaktheits- und Abgeschlossenheitsresultate in $SBV^\alpha(\Omega)$ bezüglich der schwach* Topologie notwendig.

Und schließlich liefert der Übergang zu $\liminf_{h \rightarrow 0}$

$$\tilde{F}(u) \leq \liminf_{h \rightarrow 0} \tilde{F}(w_h) \leq \liminf_{h \rightarrow 0} \tilde{F}_h(u_h).$$

Da die Folge $u_h \in U_h$ mit $\|u_h - u\|_{L^1(\Omega)} \rightarrow 0$ beliebig war, folgt hieraus die gewünschte Γ -lim inf-Ungleichung.

Konstruktion von w_h

Zur Beschreibung der Konstruktion sei h beliebig aber fest. Mit $x_{i,j}$ werden die Mittelpunkte der Gitterzellen $T_h \in \mathcal{T}_h$ bezeichnet. Dabei sei $x_{i,j+1}$ der Mittelpunkt der rechten Nachbarzelle und $x_{i+1,j}$ der Mittelpunkt der oberen Nachbarzelle (siehe Abbildung 2.8).

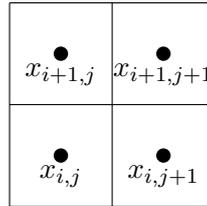


Abbildung 2.8: Zellenmittelpunkte $x_{i,j}$

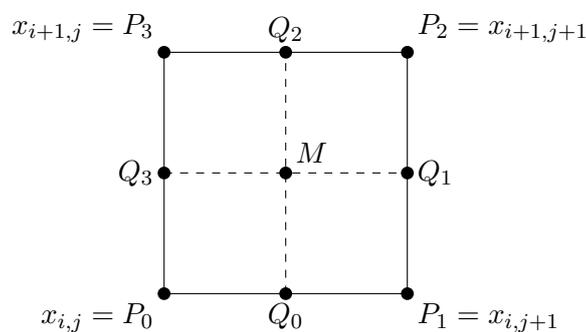
Für vier so benachbarte Zellen wird nun angegeben, wie w_h im Inneren von $\text{conv}(x_{i,j}, x_{i,j+1}, x_{i+1,j}, x_{i+1,j+1})$ definiert wird. Dadurch ist dann w_h (bis auf eine $\mathcal{L}^2(\Omega)$ -Nullmenge) eindeutig festgelegt. Zur Vereinfachung der Notation werden die Bezeichnungen aus Abbildung 2.9 eingeführt.

Alle Indizes für P und Q sind im Rahmen dieser Konstruktion „modulo 4“ zu verstehen, d. h. $P_i := P_{i \bmod 4}$ und analog für Q . Ein Q_i liegt in der Mitte von P_i und P_{i+1} und M ist der Mittelpunkt des Quadrats $Z = \text{conv}(P_0, P_1, P_2, P_3)$. Abkürzend sei $u_i := u_h(P_i)$.

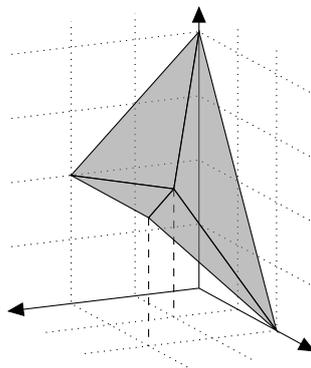
Ferner sei $\lambda: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ durch $\lambda(a, b) := |a - b|/h^\ell - 1$ festgelegt. Mit Hilfe von $\lambda_i := \lambda(u_i, u_{i+1})$ wird entschieden, ob w_h entlang der Strecke $Q_i M$ eine Sprungstelle hat.

$$\lambda_i \geq 0 \quad \Rightarrow \quad \text{Strecke } Q_i M \subset \overline{S(w_h)}.$$

Für die Konstruktion sind jetzt mehrere Fälle zu unterscheiden.

Abbildung 2.9: Bezeichnungen im Quadrat Z **Konstruktion von w_h für $\lambda_i < 0$ für alle i**

Seien zunächst alle $\lambda_i < 0$. In diesem Fall sei w_h auf den Dreiecken $P_i P_{i+1} M$ eine affin-lineare Funktion mit $w_h(P_i) = u_i$ und $w_h(M) = \frac{1}{4} \sum_{i=0}^3 u_i$ (siehe Abbildung 2.10).

Abbildung 2.10: Konstruktion w_h für $\lambda_i < 0$ für alle i

Dann gilt

$$\begin{aligned} \int_Z \|\nabla w_h\|_\alpha^\alpha dx &= \frac{1}{4} h^{2-\alpha} \sum_{i=0}^3 |u_i - u_{i+1}|^\alpha + \frac{1}{2} h^{2-\alpha} \left| \frac{u_2 - u_1}{2} + \frac{u_3 - u_0}{2} \right|^\alpha + \\ &+ \frac{1}{2} h^{2-\alpha} \left| \frac{u_0 - u_1}{2} + \frac{u_3 - u_2}{2} \right|^\alpha. \end{aligned}$$

Da $t \mapsto |t|^\alpha$ konvex ist, folgt (auch aus der Ungleichung von Jensen, Satz 2.5)

$$\left| \frac{a}{2} + \frac{b}{2} \right|^\alpha \leq \frac{|a|^\alpha}{2} + \frac{|b|^\alpha}{2}.$$

Mit $a = u_2 - u_1$ und $b = u_3 - u_0$ bzw. $a = u_0 - u_1$ und $b = u_3 - u_2$ erhält man die Abschätzung

$$\int_Z \|\nabla w_h\|_\alpha^\alpha dx \leq \frac{1}{2} h^{2-\alpha} \sum_{i=0}^3 |u_i - u_{i+1}|^\alpha. \quad (2.25)$$

Konstruktion von w_h für $\lambda_i, \lambda_{i+2} < 0$ und $\lambda_{i-1}, \lambda_{i+1} \geq 0$

Für den Fall, dass genau λ_i und λ_{i+2} negativ sind, also etwa $\lambda_0, \lambda_2 \geq 0$ und $\lambda_1, \lambda_3 < 0$, wird für $x \in \text{conv}(P_0, Q_0, Q_2, P_3)$

$$w_h(x) := \frac{1}{h}(u_3 - u_0)e_2^T(x - (P_3 + P_0)/2) + \frac{1}{2}(u_3 + u_0)$$

und für $x \in \text{conv}(Q_0, P_1, P_2, Q_2)$

$$w_h(x) := \frac{1}{h}(u_2 - u_1)e_2^T(x - (P_2 + P_1)/2) + \frac{1}{2}(u_2 + u_1)$$

gesetzt (siehe Abbildung 2.11).

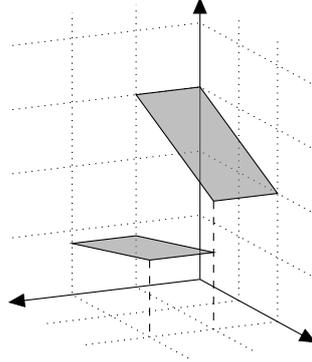


Abbildung 2.11: Konstruktion w_h für $\lambda_1, \lambda_3 < 0$, $\lambda_0, \lambda_2 \geq 0$

Für dieses w_h gilt

$$\int_Z \|\nabla w_h\|_\alpha^\alpha dx = \frac{1}{2} h^{2-\alpha} (|u_3 - u_0|^\alpha + |u_2 - u_1|^\alpha) \quad (2.26)$$

und

$$\begin{aligned} & \int_{S_{w_h} \cap Z} V(w_h^+, w_h^-) \Phi(\nu_{w_h}) d\mathcal{H}^1 \\ &= \int_{-h/2}^{h/2} \left| \frac{u_2 - u_1}{h} y + \frac{u_2 + u_1}{2} - \frac{u_3 - u_0}{h} y - \frac{u_3 + u_0}{2} \right|^\beta dy \\ &\leq h \left(\frac{1}{h} \int_{-h/2}^{h/2} \left| \frac{u_2 - u_1}{h} y + \frac{u_2 + u_1}{2} - \frac{u_3 - u_0}{h} y - \frac{u_3 + u_0}{2} \right| dy \right)^\beta. \end{aligned}$$

Dabei folgt die letzte Ungleichung aus der Ungleichung von Jensen (Satz 2.5) für konkave Funktionen.

Innerhalb des Betrages findet kein Vorzeichenwechsel statt, denn aus $\lambda_0, \lambda_2 \geq 0$ und $\lambda_1, \lambda_3 < 0$ folgt

$$\begin{aligned} |u_0 - u_1| &\geq h^\ell, & |u_0 - u_3| &< h^\ell, \\ |u_2 - u_3| &\geq h^\ell, & |u_1 - u_2| &< h^\ell, \end{aligned}$$

und daher $\text{conv}(u_0, u_3) \cap \text{conv}(u_1, u_2) = \emptyset$. Somit gilt weiter

$$\begin{aligned} h \left(\frac{1}{h} \cdot h \left| \frac{u_1 + u_2 - (u_0 + u_3)}{2} \right| \right)^\beta &\leq h \left(\frac{|u_1 - u_0| + |u_3 - u_2|}{2} \right)^\beta \\ &\leq \frac{h}{2^\beta} (|u_1 - u_0|^\beta + |u_3 - u_2|^\beta). \end{aligned}$$

Für die gewünschte Ungleichung (2.23) ist eine Abschätzung der Form

$$\int_{S_{w_h} \cap Z} V(w_h^+, w_h^-) \Phi(\nu_{w_h}) d\mathcal{H}^1 \leq \frac{h}{2} (|u_1 - u_0|^\beta + |u_3 - u_2|^\beta)$$

erforderlich. Dies ist mit dem angegebenen w_h leider nur für $\beta = 0$ möglich:

$$\int_{S_{w_h} \cap Z} V(w_h^+, w_h^-) \Phi(\nu_{w_h}) d\mathcal{H}^1 = h \leq \frac{h}{2} (|u_1 - u_0|^\beta + |u_3 - u_2|^\beta). \quad (2.27)$$

Obwohl die hier vorgestellte Konstruktion nur den Fall $\beta = 0$ erlaubt, vermutet der Autor, dass sich eine geschicktere Approximation finden lässt, die die obige Abschätzung auch für $0 \leq \beta < 1$ erlaubt. Deshalb werden die folgenden Ungleichungen auch für ein allgemeines $0 \leq \beta < 1$ angegeben.

Konstruktion von w_h für die restlichen Fälle

Für alle anderen Fälle von Vorzeichenkombinationen der λ_i wird w_h lokal konstant durch $w_h(x) = u_i$ für $x \in \text{conv}(P_i, Q_i, M, Q_{i+3})$ festgelegt (siehe Abbildung 2.12).

Hier ist

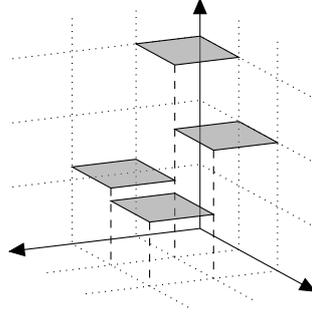
$$\int_{S_{w_h} \cap Z} V(w_h^+, w_h^-) \Phi(\nu_{w_h}) d\mathcal{H}^1 \leq \frac{h}{2} \sum_{i=0}^3 |u_i - u_{i+1}|^\beta.$$

und $\int_Z \|\nabla w_h\|_\alpha^\alpha dx = 0$.

Zur Ungleichung (2.23)

Die Ungleichung (2.23) wird nur für den Fall

$$d(t) = \begin{cases} t^\alpha & \text{für } 0 \leq t < 1 \\ t^\beta & \text{für } 1 \leq t \end{cases}$$

Abbildung 2.12: Konstruktion w_h für restliche Fälle

gezeigt. In diesem Fall ist (vgl. (2.13))

$$V_h(a, b) = h^{k+1} d\left(\frac{|a-b|}{h^\ell}\right) = \begin{cases} h^{2-\alpha} |a-b|^\alpha & \text{falls } \lambda(a, b) < 0 \\ h |a-b|^\beta & \text{falls } \lambda(a, b) \geq 0. \end{cases}$$

Alle Quadrate, innerhalb deren w_h keine (approximative) Sprungstellen besitzt, werden mit \square bezeichnet, die Quadrate, bei denen w_h eine vertikale bzw. horizontale Sprungstelle aufweist, werden mit \blacksquare bzw. \blacksquare bezeichnet.

Mit Hilfe von (2.25), (2.26) und (2.27) kann nun abgeschätzt werden

$$\begin{aligned} \tilde{F}(w_h) &= \sum_Z \int_Z \|\nabla w_h\|_\alpha^\alpha dx + \sum_Z \int_{S_{w_h} \cap Z} V(w_h^+, w_h^-) \Phi(\nu_{w_h}) d\mathcal{H}^1 \\ &\leq \sum_{Z:\square} \frac{1}{2} h^{2-\alpha} (|u_0 - u_1|^\alpha + |u_1 - u_2|^\alpha + |u_2 - u_3|^\alpha + |u_3 - u_0|^\alpha) + \\ &\quad \sum_{Z:\blacksquare} \frac{1}{2} h^{2-\alpha} (|u_3 - u_0|^\alpha + |u_2 - u_1|^\alpha) + \frac{h}{2} (|u_1 - u_0|^\beta + |u_3 - u_2|^\beta) + \\ &\quad \sum_{Z:\blacksquare} \frac{1}{2} h^{2-\alpha} (|u_1 - u_0|^\alpha + |u_2 - u_3|^\alpha) + \frac{h}{2} (|u_3 - u_0|^\beta + |u_2 - u_2|^\beta) + R \\ &= \tilde{F}_h(u_h) + R, \end{aligned}$$

dabei sei R die Summen über alle restlichen Quadrate, die nicht durch die obigen Fälle abgedeckt sind.

Für $u \in SBV^\alpha(\Omega)$ ist die Menge $S_u \setminus J_u$ der approximativen Sprungstellen, bei denen u keine approximative Sprungrichtung besitzt, eine \mathcal{H}^1 -Nullmenge. Daher gibt es für die Folge (u_h) höchstens $\mathcal{O}(1)$ Quadrate, die einen Beitrag zu R liefern. Jeder Summand von R ist ein $\mathcal{O}(h)$, so dass also $R = \mathcal{O}(h)$ gilt.

Unterhalbstetigkeit von \tilde{F}

Für die Unterhalbstetigkeit verwenden wir wie im eindimensionalen Fall den Satz 2.25 mit den Kompaktheits- und Abgeschlossenheitsresultate in *SBV*

Da mit $u_h(x) \in K$ für \mathcal{L}^2 -fast alle $x \in \Omega$ gilt, ist auch w_h gleichmäßig in h bzgl. L^∞ -Norm beschränkt. Außerdem ist für ein $C > 0$

$$\begin{aligned} \sup_h \left\{ \int_{\Omega} \|\nabla w_h\|_{\alpha}^{\alpha} dx + \int_{S_{w_h}} |w_h^+ - w_h^-|^{\beta} d\mathcal{H}^1 \right\} &\leq C \sup_h \tilde{F}(w_h) \stackrel{(2.23)}{\leq} \\ &\leq C \sup_h (\tilde{F}_h(u_h) + \mathcal{O}(h)) < \infty. \end{aligned} \quad (2.28)$$

Nach dem oben zitierten Kompaktheitsresultat gibt es eine Teilfolge, die wieder mit w_h bezeichnet wird, die schwach* in $BV(\Omega)$ gegen u konvergiert. Außerdem gilt dann

$$\int_{\Omega} \|\nabla u\|_{\alpha}^{\alpha} dx \leq \liminf_{h \rightarrow 0} \int_{\Omega} \|\nabla w_h\|_{\alpha}^{\alpha} dx.$$

Damit fehlt noch die Unterhalbstetigkeit des Integralterms über S_u . Hierzu wird [AFP00, Theorem 5.22] verwendet. Die Voraussetzungen sind gerade die Punkte 5 und 6 aus dem noch folgenden Satz 3.1, deren Nachweis dort zu finden sind. Wegen (2.28) ist $\|\nabla w_h\|_2$ auch equiintegabel und ist gilt

$$\int_{S_u} V(u^+, u^-) \Phi(\nu_u) d\mathcal{H}^1 \leq \liminf_{h \rightarrow 0} \int_{S_{w_h}} V(w_h^+, w_h^-) \Phi(\nu_{w_h}) d\mathcal{H}^1.$$

Zusammen ist also (2.24) gezeigt.

2.3.9 Zusammenfassung

Bevor die vorangegangenen Abschnitte in einem Satz zusammengefasst werden, werden an dieser Stelle die eingeführten Hilfsgrößen und -funktionen notiert.

Seien $\beta = 0$ und $1 < \alpha \leq 2$ und

$$d(t) = \begin{cases} t^{\alpha} & \text{falls } 0 \leq t < 1 \\ t^{\beta} = 1 & \text{falls } 1 \leq t. \end{cases}$$

Ferner seien $\ell = \frac{\alpha-1}{\alpha-\beta} = \frac{\alpha-1}{\alpha}$, $k = \beta\ell = 0$,

$$V(a, b) = |a - b|^{\beta} = 1 \quad \text{und} \quad V_h(a, b) = h^{k+1} d\left(\frac{|a - b|}{h^{\ell}}\right).$$

Außerdem sei $\Phi(\nu_u) = \|\nu_u\|_1$ und $g(v) = \|v\|_{\alpha}^{\alpha}$.

Mit diesen Bezeichnungen wurde in den vorangegangenen Abschnitten also folgender Satz bewiesen.

Satz 2.28

Sei $K \subset \mathbb{R}$ kompakt, $U_{h,K} := \{u_h \in U_h : u_h(x) \in K \text{ } \mathcal{L}^2\text{-fast überall}\}$ sowie $U := \{u \in SBV^\alpha(\Omega) : u \in K \text{ } \mathcal{L}^2\text{-fast überall}\}$ und ist $f: K \rightarrow \mathbb{R}$ Lipschitzstetig, dann ist der Γ -Limes von

$$F_h(u_h) = \begin{cases} \sum_{R \sim T} h d\left(\frac{u_h|_R - u_h|_T}{h^\ell}\right) + \sum_T |T| f(u_h|_T) & \text{für } u_h \in U_{h,K} \\ +\infty & \text{sonst} \end{cases}$$

bzgl. der L^1 -Topologie für $h \rightarrow 0$ durch

$$F(u) = \begin{cases} \int_{S_u} \Phi(\nu_u) d\mathcal{H}^1 + \int_\Omega f(u) dx + \int_\Omega \|\nabla u\|_\alpha^\alpha dx & \text{für } u \in U \\ +\infty & \text{sonst} \end{cases}$$

gegeben.

Kapitel 3

Existenzaussagen

In diesem Kapitel werden Funktionale der Form

$$F(u) = \int_{S_u} V(u^+, u^-) \Phi(\nu_u) d\mathcal{H}^1 + \int_{\Omega} f(u) dx + \int_{\Omega} g(\nabla u) dx$$

untersucht. In den Abschnitten 3.1, 3.2 bzw. 3.3 wird für das Problem $F(u) = \min!$ die Existenztheorie von Lösungen im Raum spezieller Funktionen beschränkter Variation, im Raum stückweiser konstanter Funktionen bzw. in Sobolev-Räumen dargestellt. Die Existenzaussagen, die in 3.4 zusammengefasst sind, werden jeweils auf das Funktional aus Satz 2.28 angewendet.

3.1 Existenzresultat in SBV

Zur Untersuchung, wann bzw. ob das Problem (2.14) Lösungen besitzt, können wir auf folgenden Satz von Ambrosio, Fusco und Pallara zurückgreifen.

Satz 3.1 ([AFP00], Theorem 5.24 mit Theoremen 5.8 und 5.22)

Sei $\Omega \subset \mathbb{R}^d$ offen und beschränkt. Ferner sei $W: \Omega \times \mathbb{R}^m \times \mathbb{R}^{md} \rightarrow [0; \infty]$, $K \subset \mathbb{R}^m$ kompakt und $\varphi: K \times K \times \mathbb{R}^d \rightarrow [0; \infty]$. Sind die folgenden Voraussetzungen erfüllt

1. W ist $\mathcal{L}^d \times \mathcal{B}(\mathbb{R}^{m+md})$ messbar,
2. $(s, z) \mapsto W(x, s, z)$ ist auf \mathbb{R}^{m+md} von unten halbstetig für \mathcal{L}^d -fast alle $x \in \Omega$,
3. $z \mapsto W(x, s, z)$ ist in \mathbb{R}^{md} konvex für alle $x \in \Omega$ und alle $s \in \mathbb{R}^m$,
4. es gibt eine von unten halbstetige Funktion $\rho: [0; \infty[\rightarrow [0; \infty]$ mit $\lim_{t \rightarrow \infty} \rho(t)/t = \infty$, so dass

$$W(x, s, z) \geq \rho(\|z\|_2) \quad \text{für alle } x \in \Omega, s \in \mathbb{R}^m \text{ und } z \in \mathbb{R}^{md}$$

gilt,

5. φ ist verbunden konvex (jointly convex), d. h. es gibt eine Folge $(g_n) \subset [C(K)]^d$ mit

$$\varphi(a, b, p) = \sup_{n \in \mathbb{N}} \langle g_n(a) - g_n(b), p \rangle \quad \text{für alle } (a, b, p) \in K \times K \times \mathbb{R}^d,$$

6. es gibt ein $c > 0$ und ein $0 < \gamma < 1$, so dass

$$\varphi(a, b, p) \geq \inf\{c, \|a - b\|_2^\gamma\} \|p\|_2 \quad \text{für alle } (a, b, p) \in K \times K \times \mathbb{R}^d$$

gilt,

so hat mit $X := \{u \in [SBV(\Omega)]^m : u(x) \in K \text{ für } \mathcal{L}^d\text{-fast alle } x \in \Omega\}$

$$\min \left\{ \int_{\Omega} W(x, u, \nabla u) dx + \int_{J_u} \varphi(u^+, u^-, \nu_u) d\mathcal{H}^{d-1} : u \in X \right\}$$

mindestens eine Lösung.

Damit liegt für unser Problem (2.14) folgende Struktur zu Grunde:

$$W(x, s, z) = W(s, z) = f(s) + g(z) \quad \text{und} \quad \varphi(a, b, p) = V(a, b)\Phi(p)$$

mit g aus (2.19). Da unser f und g jeweils messbar ist, ist die Voraussetzung 1 erfüllt. f ist unterhalbstetig und g ist stetig, damit gilt auch die Voraussetzung 2.

Da g konvex ist, ist auch die Voraussetzung 3 erfüllt. Aus

$$W(x, s, z) = f(s) + g(z) = f(s) + \|z\|_\alpha^\alpha \geq \|z\|_\alpha^\alpha \geq \|z\|_2^\alpha = \rho(\|z\|_2)$$

mit $\rho(t) = t^\alpha$ folgt wegen $\alpha > 1$, dass die Voraussetzung 4 zutrifft.

Zum Nachweis, dass φ verbunden konvex ist, reicht es nach [AFP00, 5.23] zu zeigen, dass Φ von unten halbstetig, gerade, positiv homogen vom Grad 1 und konvex ist sowie dass V positiv und symmetrisch ist und auch die Dreiecksungleichung erfüllt. Für $\beta = 0$ erfüllt $V \equiv 1$ diese Bedingungen. Für $0 < \beta < 1$ ist V (siehe (2.21)) stetig, positiv und per Konstruktion symmetrisch mit $V(i, i) = 0$. Die Dreiecksungleichung, die V zu erfüllen hat, lässt sich wie folgt umschreiben.

Satz 3.2

Für $V: \mathbb{R} \times \mathbb{R} \rightarrow [0; \infty[$, $V(a, b) := d(|a - b|)$ mit $d: [0; \infty[\rightarrow [0; \infty[$ und $d(x) = 0$ genau dann wenn $x = 0$ gilt:

$$V \text{ ist Metrik} \quad \Leftrightarrow \quad \begin{cases} i) & d(a + b) \leq d(a) + d(b) \quad \forall a, b \geq 0 \\ ii) & d(a - b) \geq d(b) - d(a) \quad \forall a \geq b \geq 0 \end{cases} \quad (3.1)$$

Beweis

V sei Metrik. Dann gilt für $a, b \geq 0$

$$V(a, -b) \leq V(a, 0) + V(0, -b) = V(a, 0) + V(0, b),$$

woraus i) folgt. ii) folgt aus der Dreiecksungleichung

$$V(x, y) \leq V(x, z) + V(z, y)$$

mit $x = b$, $y = 0$ und $z = a$.

Ist umgekehrt i) und ii) erfüllt, dann muss für beliebige $x, y, z \in \mathbb{R}$ die Dreiecksungleichung gezeigt werden. Aus i) folgt für $x \leq z \leq y$ mit $a = z - x \geq 0$ und $b = y - z \geq 0$

$$\begin{aligned} V(x, y) &= d(y - x) = d(a + b) \leq d(a) + d(b) = \\ &= d(z - x) + d(y - z) = V(x, z) + V(z, y). \end{aligned}$$

Aus ii) folgt für $x \leq y \leq z$ mit $a = z - x \geq 0$ und $b = y - x \geq 0$

$$\begin{aligned} V(y, z) &= d(z - y) = d(a - b) \geq d(b) - d(a) = \\ &= d(y - x) - d(z - x) = V(x, y) - V(z, x), \end{aligned}$$

also auch $V(x, y) \leq V(x, z) + V(z, y)$. Ebenfalls aus ii) folgt für $z \leq x \leq y$ mit $a = y - z \geq 0$ und $b = y - x \geq 0$

$$\begin{aligned} V(x, z) &= d(x - z) = d(a - b) \geq d(b) - d(a) = \\ &= d(y - x) - d(y - z) = V(x, y) - V(z, y), \end{aligned}$$

die Dreiecksungleichung. Wegen der Symmetrie von V ist daher die Dreiecksungleichung für alle $x, y, z \in \mathbb{R}$ gezeigt. \square

Da die Funktion $t \mapsto t^\beta$ den beiden Ungleichungen (3.1) genügt, ist die Voraussetzung 5 des Satzes 3.1 auch für ein $0 < \beta < 1$ gegeben.

Da Voraussetzung 6 ist wegen (2.21) erfüllt.

3.2 Existenzresultat in BV^*

In diesem Abschnitt wird das Funktional (2.14) untersucht, wenn nur Funktionen u zulässig sind, die stückweise konstant und nur endlich viele Werte annehmen können. Dazu muss zunächst festgelegt werden, was unter einer stückweisen konstanten Funktion verstanden wird, also auf welchen Mengen die zulässigen Funktionen konstant sind.

Hierfür benötigen wir zunächst den Begriff der Caccioppoli Partition.

Definition 3.3 (Caccioppoli Partition)

Ist $L \subset \mathbb{R}^m$ und L endlich oder abzählbar, dann heißt eine Partition $\{E_\alpha\}_{\alpha \in L}$ von Ω Caccioppoli Partition, falls für die Perimeter $P(E_\alpha, \Omega)$

$$\sum_{\alpha \in L} P(E_\alpha, \Omega) < \infty$$

gilt.

Damit können wir nun $BV^*(\Omega, L)$ festlegen.

Definition 3.4 (stückweise konstante Funktionen)

Für ein $L \subset \mathbb{R}^m$ und L endlich oder abzählbar wird mit

$$BV^*(\Omega, L) := \left\{ u: \Omega \rightarrow L : \text{es gibt eine Caccioppoli Partition } \{E_\alpha\}_{\alpha \in L} \text{ von } \Omega \text{ mit } u = \sum_{\alpha \in L} \alpha \chi_{E_\alpha} \right\}$$

der Raum der stückweise konstanten Funktionen bezeichnet. Als Norm wird auf $BV^*(\Omega, L)$ die L^1 -Norm verwendet.

Die Bezeichnung BV^* erklärt folgender Satz.

Satz 3.5

Für $u: \Omega \rightarrow L$ sind folgende Aussagen für ein endliches L äquivalent.

1. $u \in BV^*(\Omega, L)$
2. $u \in BV(\Omega)$ und das Maß Du ist auf J_u konzentriert und ferner gilt $\mathcal{H}^1(J_u) < \infty$
3. $u \in BV(\Omega)$
4. $u \in SBV(\Omega)$

Beweis

Die Äquivalenz von 1 und 2 wird in [AFP00] Proposition 5.9 gezeigt. Aus 2 folgt 4 und daraus 3. Bleibt damit zu zeigen, dass aus 3 etwa 1 folgt.

Für $\alpha \in L$ sei mit $E_\alpha := u^{-1}(\{\alpha\})$ die messbare Urbildmenge von α bezeichnet. Die Coarea-Formel (Satz 2.14) liefert, dass die Mengen

$$\{u > t\} = \bigcup_{\alpha > t} E_\alpha$$

für \mathcal{L}^1 -fast alle $t \in \mathbb{R}$ endlichen Perimeter in Ω haben. Wegen

$$P(A, \Omega) \leq P(A \cup B, \Omega) + P(B, \Omega)$$

folgt damit, dass $P(E_\alpha, \Omega) < \infty$ für alle $\alpha \in L$ gilt. Damit lässt sich u in der Form $u = \sum_{\alpha \in L} \alpha \chi_{E_\alpha}$ schreiben. Da L endlich ist, ist $\{E_\alpha\}_{\alpha \in L}$ auch eine Caccioppoli Partition und es gilt $u \in BV^*(\Omega, L)$. \square

Damit gilt für eine stückweise konstante Funktion $u \in BV^*(\Omega, L)$ nun $\nabla u(x) = 0$ für \mathcal{L}^d -fast alle $x \in \Omega$. Daher fällt im Funktional (2.14) der letzte Summand von $F(u)$ weg.

Der folgende Existenzsatz verwendet die direkte Methode der Variationsrechnung (siehe z. B. [Dac89]). Bei dieser Methode wird in folgenden Schritten vorgegangen.

1. Man geht von einer minimierenden Folge $(u_n)_n$ aus.
2. Man zeigt, dass (u_n) in einer (bzgl. einer geeigneten Topologie) abgeschlossenen und beschränkten Menge liegt.
3. Man wählt eine gegen \bar{u} konvergente Teilfolge. Bei den hier vorkommenden unendlichdimensionalen Räumen muss die Existenz einer solchen Teilfolge erst sichergestellt werden.
4. Man zeigt, dass das Funktional (bzgl. der Topologie aus 2) von unten halbstetig ist.

Zusammen erhält man dann

$$F(\bar{u}) \leq \liminf_{n \rightarrow \infty} F(u_n)$$

und damit, dass F bei \bar{u} ein globales Minimum besitzt. Die Wahl der geeigneten Topologie ist dabei oft schwierig. Eine Topologie mit vielen offenen Mengen erleichtert zwar den Nachweis der Halbstetigkeit, erschwert bzw. verhindert aber die Existenz einer konvergenten Teilfolge.

Zur Formulierung des Existenzsatzes wird noch folgende Eigenschaft benötigt.

Definition 3.6 (BV-Elliptizität)

$\varphi: L \times L \times S^{d-1} \rightarrow [0; \infty[$ heißt *BV-elliptisch*, falls für jede beschränkte stückweise konstante Funktion $v: Q(\nu) \rightarrow L$ mit $\{v \neq u_{\alpha, \beta, \nu}\} \subset\subset Q(\nu)$

$$\int_{J_v} \varphi(v^+, v^-, \nu_v) d\mathcal{H}^{d-1} \geq \varphi(\alpha, \beta, \nu)$$

für jedes $(\alpha, \beta, \nu) \in L \times L \times S^{d-1}$ gilt. Dabei ist $Q(\nu)$ jeder d -dimensionale offene Würfel mit Zentrum 0 , Kantenlänge 1 und Seitenflächen, die entweder parallel oder orthogonal zu ν sind. Ferner werde mit $u_{\alpha, \beta, \nu}$ die Funktion

$$u_{\alpha, \beta, \nu}(x) := \begin{cases} \alpha & \text{falls } \langle x, \nu \rangle > 0 \\ \beta & \text{falls } \langle x, \nu \rangle < 0 \end{cases}$$

bezeichnet.

Die BV -Elliptizität ist eine charakterisierende Bedingung für die benötigte Halbstetigkeit des Integralterms über die Sprungmenge J_u im betrachteten Funktional:

Satz 3.7 ([AFP00], Theorem 5.14)

Ist $\varphi: L \times L \times S^{d-1} \rightarrow [0; \infty[$ beschränkt und stetig, dann ist

$$u \mapsto \int_{\Omega} \varphi(u^+, u^-, \nu_u) d\mathcal{H}^{d-1}$$

genau dann in $BV^*(\Omega, L)$ bzgl. der L^1 -Topologie von unten halbstetig, wenn φ BV -elliptisch ist.

Satz 3.8 (Existenzsatz)

Sei $\Omega \subset \mathbb{R}^d$ offen und beschränkt mit Lipschitz-Rand $\partial\Omega$. Ferner sei $W: \Omega \times \mathbb{R}^m \rightarrow [0; \infty]$, $L \subset \mathbb{R}^m$ endlich und $\varphi: L \times L \times \mathbb{R}^d \rightarrow [0; \infty]$. Sind die folgenden Voraussetzungen erfüllt

1. W ist $\mathcal{L}^d \times \mathcal{B}(\mathbb{R}^m)$ messbar,
2. $s \mapsto W(x, s)$ ist auf \mathbb{R}^m von unten halbstetig für \mathcal{L}^d -fast alle $x \in \Omega$,
3. φ ist stetig und BV -elliptisch,
4. es gibt ein $c > 0$, so dass

$$\varphi(a, b, p) \geq c\|p\|_2 \quad \text{für alle } (a, b, p) \in L \times L \times \mathbb{R}^d$$

gilt,

so hat

$$\min \left\{ F^*(u) := \int_{\Omega} W(x, u) dx + \int_{J_u} \varphi(u^+, u^-, \nu_u) d\mathcal{H}^{d-1} : u \in BV^*(\Omega, L) \right\} \quad (3.2)$$

mindestens eine Lösung.

Beweis

Sei $(u_n)_n \subset BV^*(\Omega, L)$ eine minimierende Folge von F^* mit

$$0 \leq m := \lim_{n \rightarrow \infty} F^*(u_n) \leq M < \infty.$$

Da L endlich ist, ist $\|u_n\|_{L^\infty(\Omega)} \leq \max_{\alpha \in L} \|\alpha\|_2 < \infty$ für jedes n . Außerdem gilt wegen der Voraussetzung 4

$$\begin{aligned} c\mathcal{H}^{d-1}(J_{u_n}) &= \int_{J_{u_n}} c\|\nu_{u_n}\|_2 d\mathcal{H}^{d-1} \\ &\leq \int_{J_{u_n}} \varphi(u_n^+, u_n^-, \nu_{u_n}) d\mathcal{H}^{d-1} \leq F^*(u_n) \leq M \end{aligned}$$

Nach dem Federer-Vol’pert Theorem [AFP00, Theorem 3.78] ist

$$\mathcal{H}^{d-1}(S_{u_n}) = \mathcal{H}^{d-1}(J_{u_n}).$$

Somit ist $\|u_n\|_{L^\infty(\Omega)} + \mathcal{H}^{d-1}(S_{u_n})$ beschränkt. Mit der Voraussetzung, dass Ω einen Lipschitz-Rand besitzt, existiert (siehe [AFP00], Theorem 4.25) eine Teilfolge, die wieder mit u_n bezeichnet werde, die nach Maß gegen eine stückweise konstante Funktion $\bar{u} \in BV^*(\Omega, L)$ konvergiert, d. h. für alle $\epsilon > 0$ gilt

$$\mathcal{L}^d(\{x \in \Omega : \|u_n(x) - \bar{u}(x)\|_2 > \epsilon\}) \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Wegen

$$\begin{aligned} \mathcal{L}^d\left(\bigcup_{\alpha \in L} \{u_n = \alpha\} \Delta \{\bar{u} = \alpha\}\right) &= \mathcal{L}^d(\{x \in \Omega : u_n(x) \neq \bar{u}(x)\}) = \\ &= \mathcal{L}^d(\{x \in \Omega : \|u_n(x) - \bar{u}(x)\|_2 > d_{min}/2\}) \rightarrow 0 \end{aligned}$$

mit $0 < d_{min} = \min_{\alpha \neq \beta \in L} \|\alpha - \beta\|_2$, konvergieren die Levelsets der u_n nach Maß gegen die Levelsets \bar{u} . Damit gilt $u_n \xrightarrow{L^1(\Omega)} \bar{u}$.

Wegen Voraussetzungen 1 und 2 gilt zusammen mit dem Lemma von Fatou

$$\int_{\Omega} W(x, \bar{u}) dx \leq \int_{\Omega} \liminf_{n \rightarrow \infty} W(x, u_n) dx \leq \liminf_{n \rightarrow \infty} \int_{\Omega} W(x, u_n) dx.$$

Mit Voraussetzung 3 liefert Satz 3.7 die Halbstetigkeit (von unten) des zweiten Terms.

Damit ist F^* auf $BV^*(\Omega, L)$ von unten halbstetig und es gilt

$$m \leq F^*(\bar{u}) \leq \liminf_{n \rightarrow \infty} F^*(u_n) = m$$

und \bar{u} ist ein Minimum von F^* . □

Es werden nun die Voraussetzungen des Satzes für das Funktional (2.14) überprüft. Die Bedingungen 1 und 2 wurden schon im Anschluss an Satz 3.1 nachgewiesen.

Wegen

$$\varphi(a, b, p) = \|a - b\|_2^\beta \|p\|_1 \geq m^\beta \|p\|_1 \geq m^\beta \|p\|_2$$

ist mit $0 < m := \min_{a \neq b \in L} \|a - b\|_2$ auch die Koerzivitätsbedingung 4 erfüllt.

Die Stetigkeit von φ ist klar. Zum Nachweis der BV -Elliptizität reicht der Nachweis, dass φ verbunden konvex ist ([AFP00], Theorem 5.20). Das war aber gerade die Voraussetzung 5 von Satz 3.1 und sie wurde ebenfalls schon in Abschnitt 3.1 nachgewiesen.

3.3 Existenzresultat in $W^{1,p}$

In diesem Abschnitt wird das Funktional (2.14) untersucht, wenn für u nur Sobolev-Funktionen zulässig sind.

Für jedes $u \in W^{1,p}(\Omega)$ mit $p \geq 1$ ist $\mathcal{H}^{d-1}(S_u) = 0$. Daher fällt im Funktional (2.14) das Integral über S_u weg.

Auch hier ist wieder entscheidend, unter welchen Bedingungen das Funktional unterhalbstetig ist. Der dafür benötigte Begriff wurde von Morrey in [Mor52] eingeführt.

Definition 3.9 (Quasikonvexität)

Eine $\mathcal{B}(\mathbb{R}^{md})$ messbare Funktion $W: \mathbb{R}^{md} \rightarrow \mathbb{R}$ mit $W \in L^1_{\text{loc}}(\mathbb{R}^{md})$ heißt quasikonvex, falls

$$\mathcal{L}^d(\Omega)W(A) \leq \int_{\Omega} W(A + D\psi(x)) dx \quad (3.3)$$

für ein beschränktes Gebiet $\Omega \subset \mathbb{R}^d$, für jedes $A \in \mathbb{R}^{md}$ und für jedes $\psi \in C_c^\infty(\Omega, \mathbb{R}^m)$ gilt.

In [Mor52] wird gezeigt, dass die Definition in folgendem Sinne unabhängig vom Gebiet Ω ist: Ist W quasikonvex für irgendein beschränktes Gebiet Ω , dann gilt (3.3) auch in jedem anderen beschränkten Gebiet.

Die Quasikonvexität (zusammen mit Wachstumsbedingungen an W) ist notwendig und hinreichend für die Unterhalbstetigkeit von

$$u \mapsto \int_{\Omega} W(x, u, \nabla u) dx \quad (3.4)$$

bezüglich der schwach* Topologie in $[W^{1,\infty}(\Omega)]^m$. Dies wird etwa in [Dac89] in den Theoremen 2.1 und 2.3 aus Abschnitt 4.2.1 gezeigt.

Marcellini zeigte mit folgendem Satz, dass aus Quasikonvexität die Unterhalbstetigkeit von (3.4) in $[W^{1,p}(\Omega)]^m$ folgt. Zur Formulierung der Voraussetzungen wird der Begriff der Carathéodory Funktion benötigt.

Definition 3.10 (Carathéodory Funktion)

Eine Funktion $W: \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}$ heißt Carathéodory Funktion, falls sie $\mathcal{L}^d \times \mathcal{B}(\mathbb{R}^n)$ messbar ist und $s \mapsto W(x, s)$ auf \mathbb{R}^n für \mathcal{L}^d -fast alle $x \in \Omega$ stetig ist.

Satz 3.11 ([Mar85], Theorem 1.1)

Ist $\Omega \subset \mathbb{R}^d$ offen, beschränkt mit Lipschitz-Rand und $W: \Omega \times \mathbb{R}^m \times \mathbb{R}^{md} \rightarrow \mathbb{R}$, $(x, s, z) \mapsto W(x, s, z)$ eine Carathéodory Funktion (auf $\Omega \times \mathbb{R}^{m+md}$), quasikonvex bezüglich z und erfüllt W die Wachstumsbedingung

$$-C_1 \|z\|_2^r - C_2 \|s\|_2^t - C_3(x) \leq W(x, s, z) \leq \overline{W}(x, s)(1 + \|z\|_2^p)$$

mit $C_1, C_2 \geq 0$, $C_3 \in L^1(\Omega)$, $p \geq 1$, $1 \leq r < p$ ($r = 1$, falls $p = 1$), $1 \leq t < dp/(d-p)$ ($t \geq 1$ falls $p \geq d$) und mit einer Carathéodory Funktion $\overline{W} \geq 0$, so ist (3.4) unterhalbstetig bezüglich der schwachen Topologie von $[W^{1,p}(\Omega)]^m$.

Es gilt folgender Satz.

Satz 3.12 (Existenzsatz)

Sei $\Omega \subset \mathbb{R}^d$ offen und beschränkt mit Lipschitz-Rand $\partial\Omega$. Ferner sei $W: \Omega \times \mathbb{R}^m \times \mathbb{R}^{md} \rightarrow [0; \infty[$ und $p > 1$. Sind die folgenden Voraussetzungen erfüllt

1. W ist (auf $\Omega \times \mathbb{R}^{md+m}$) eine Carathéodory Funktion,
2. $(x, s, z) \mapsto W(x, s, z)$ ist bezüglich z quasikonvex,
3. W erfüllt die Wachstumsbedingung

$$W(x, s, z) \leq \overline{W}(x, s)(1 + \|z\|_2^p) \quad \text{für alle } (x, s, z) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{md}$$

mit einer Carathéodory Funktion $\overline{W} \geq 0$,

4. W erfüllt die Koerzivitätsbedingung

$$W(x, s, z) \geq C\|z\|_2^p \quad \text{für alle } (x, s, z) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{md}$$

mit einem $C > 0$,

so hat

$$\min \left\{ F^W(u) := \int_{\Omega} W(x, u, \nabla u) dx : u \in u_0 + [W_0^{1,p}(\Omega)]^m \right\}$$

für jedes $u_0 \in [W^{1,p}(\Omega)]^m$ mit $F^W(u_0) < \infty$ mindestens eine Lösung.

Beweis

Sei $(u_n)_n \subset [W^{1,p}(\Omega)]^m$ eine minimierende Folge von F^W mit

$$0 \leq m := \lim_{n \rightarrow \infty} F^W(u_n) \leq M < \infty,$$

da $F^W(u_0) < \infty$. Aus der Bedingung 4 folgt

$$\|\nabla u_n\|_{L^p} \leq \frac{1}{C} \int_{\Omega} W(x, u_n, \nabla u_n) dx \leq \frac{M}{C} < \infty,$$

also die (in n gleichmäßige) Beschränktheit von $\|\nabla u_n\|_{L^p(\Omega)}$. Die Poincaré-Ungleichung in $W^{1,p}$ [Dac89, Theorem 1.7] liefert, dass auch $\|u_n\|_{W^{1,p}}$ gleichmäßig beschränkt ist. Wegen $p > 1$ ist $[W^{1,p}(\Omega)]^m$ reflexiv und daher schwach

folgenkompakt. Es gibt also eine Teilfolge, die wieder mit (u_n) bezeichnet wird und ein $\bar{u} \in [W^{1,p}(\Omega)]^m$, so dass

$$u_n \xrightarrow{[W^{1,p}(\Omega)]^m} \bar{u}.$$

Die Voraussetzungen 1, 2 und 3 ermöglichen die Anwendung von Satz 3.11 (mit $C_1, C_2 = 0$, $C_3 \equiv 0$ und etwa $r = 1$). Damit ist F^W bezüglich der schwachen Topologie in $[W^{1,p}(\Omega)]^m$ unterhalbstetig und bei \bar{u} nimmt F^W sein Minimum an. \square

Dieser Satz wird nun auf das Funktional (2.14) angewendet. Wieder ist

$$W(x, s, z) = f(s) + g(z) \quad \text{und} \quad g(z) = \|z\|_\alpha^\alpha.$$

Ist f stetig, so ist W eine Carathéodory Funktion, da g ebenfalls stetig ist. Da g zudem konvex ist, ist g auch quasikonvex. Damit sind die Voraussetzungen 1 und 2 erfüllt.

Wegen

$$\begin{aligned} W(x, s, z) &= f(s) + \|z\|_\alpha^\alpha \leq \max(1, f(s)) \cdot (1 + \|z\|_\alpha^\alpha) \\ &\leq \bar{W}(x, s) \cdot (1 + \|z\|_2^\alpha) \end{aligned}$$

ist mit $\bar{W}(x, s) = C_1 \max(1, f(s))$ die Wachstumsbedingung 3 erfüllt.

Die Koerzivitätsbedingung 4 folgt aus

$$W(x, s, z) = f(s) + g(z) \geq g(z) = \|z\|_\alpha^\alpha \geq C_2 \|z\|_2^\alpha.$$

Die Konstanten C_1 und C_2 ergeben sich aus der Normäquivalenzgleichung

$$\|z\|_\ell \leq \|z\|_k \leq \|z\|_\ell \cdot d^{\frac{1}{k} - \frac{1}{\ell}}$$

für $1 \leq k \leq \ell \leq \infty$ und $z \in \mathbb{R}^d$.

3.4 Zusammenfassung: Existenzsätze

In diesem Abschnitt werden die Ergebnisse aus den Abschnitten 3.1, 3.2 und 3.3 im Überblick zusammengefasst. Außerdem wird die konkrete Form des Funktional (2.14) angegeben, die es in den jeweiligen Lösungsräumen annimmt.

Die V_h haben die in (2.13) angegebene Gestalt. Ferner beschreiben $0 < \beta < 1 < \alpha$ aus (2.18) das asymptotische Verhalten von d in (2.13).

Es gilt:

1. Im Raum $\{u : u \in [SBV(\Omega)]^m \text{ mit } u(x) \in K \subset \mathbb{R}^m \text{ kompakt}\}$ hat

$$F(u) = \int_{S_u} |u^+ - u^-|^{\beta} \|\nu_u\|_1 d\mathcal{H}^1 + \int_{\Omega} f(u) dx + \int_{\Omega} \|\nabla u\|_{\alpha}^{\alpha} dx = \min!$$

für ein unterhalbstetiges f mindestens eine Lösung.

2. Im Raum $\{u : u \in BV^*(\Omega, L) \text{ mit } L \subset \mathbb{R}^m \text{ endlich}\}$ hat

$$F^*(u) = \int_{S_u} |u^+ - u^-|^{\beta} \|\nu_u\|_1 d\mathcal{H}^1 + \int_{\Omega} f(u) dx = \min!$$

für ein unterhalbstetiges f mindestens eine Lösung.

3. Im Raum $\{u : u = u_0 + [W_0^{1,p}(\Omega)]^2\}$ mit $1 < p < \infty$, wobei $u_0 \in [W^{1,p}(\Omega)]^2$ sowie $F^W(u_0) < \infty$ hat

$$F^W(u) = \int_{\Omega} f(u) dx + \int_{\Omega} \|\nabla u\|_{\alpha}^{\alpha} dx = \min!$$

für ein stetiges f mindestens eine Lösung.

Kapitel 4

Algorithmen

In diesem Kapitel werden einige der Algorithmen untersucht, die zur Minimierung von Funktionalen der Form (2.12) verwendet werden können. Da hierbei der Diskretisierungsparameter h stets fest ist, wird in Abschnitt 4.1 die Notation vereinfacht und die Problemstellung mit Hilfe von Graphen formuliert. In Abschnitt 4.2 wird die Wohlgestellttheit untersucht. Insbesondere werden die Probleme bei nichtkonvexen Funktionalen beschrieben. In 4.3 wird die Verbindung zu Max-Flow- bzw. Min-Cut-Algorithmen aus der Graphentheorie hergestellt. In 4.4 wird die Problemstellung aus stochastischer Sicht beleuchtet. Abschnitt 4.5 stellt die Block-Update-Methode vor, die für einige Beispiele zur Minimierung der konvexen (4.1) verwendet wurde. In 4.6 wird gezeigt, wie sich für nichtkonvexe Funktionale (4.1) ein GNC (graduated non-convexity) Algorithmus konstruieren lässt. Zum Abschluss finden sich in 4.7 einige Ergebnisse, die mit diesem GNC-Algorithmus berechnet wurden.

4.1 Problemstellung

Gegeben ist ein ungerichteter Graph mit einer endlichen Knotenmenge V ($d := |V|$). Die Nachbarschaftsrelation werde mit \sim bezeichnet: Zwei Knoten $a, b \in V$ sind also genau dann benachbart, wenn $a \sim b$ gilt. Mit $N_a := \{b \in V : a \sim b\}$ wird die Menge der Nachbarknoten von a bezeichnet. Ist $u: V \rightarrow \mathbb{R}$ eine Funktion auf V , so wird mit $u_a := u(a)$ der Funktionswert am Knoten $a \in V$ bezeichnet. Bringt man die Knoten in V in eine feste Reihenfolge, so lässt sich $u = (u(v_i))_{v_i \in V} \in \mathbb{R}^d$ auch als Vektor schreiben.

Betrachtet wird das Optimierungsproblem

$$F(u) = F(u; v) = \sum_{\substack{a, b \in V \\ a \sim b}} g_{a,b}(|u_a - u_b|) + \sum_{a \in V} f_a(|u_a - v_a|) = \min_u! \quad (4.1)$$

Dabei sind die $g_{a,b}: [0; \infty[\rightarrow [0; \infty[$ für $a \sim b$ und die $f_a: [0; \infty[\rightarrow [0; \infty[$ stetige Funktionen. $v: V \rightarrow \mathbb{R}$ ist eine vorgegebene Funktion.

4.2 Wohlgestelltheit und Skalierungsinvarianz

Ein Problem ist im Sinne von Hadamard wohlgestellt, falls eine eindeutige Lösung existiert, die stetig von den Eingabedaten abhängt. Dem Autor ist keine notwendige und zugleich auch hinreichende Bedingung an die $g_{a,b}$ und f_a bekannt, die diese Wohlgestelltheit charakterisieren. Indes finden einige hinreichende Voraussetzungen in der Praxis Verwendung. So ist folgender Satz eine Verallgemeinerung der Homotopie-Variante [Lud02, Satz 3.2].

Satz 4.1 (Bornemann, 2000)

Ist $U \subset \mathbb{R}^d$ kompakt, $(u, v) \mapsto F(u; v)$ stetig und gibt es zu jedem v ein eindeutiges globales Minimum $u(v) \in U$ mit $F(u(v); v) = \min_{w \in U} F(w, v)$, dann ist die Abbildung $v \mapsto u(v)$ stetig.

Beweis

Angenommen es gibt eine Folge $v_n \rightarrow v^*$, bei der $u(v_n)$ nicht gegen $u(v^*)$ konvergiert. Da U kompakt ist, hat $u(v_n)$ eine konvergente Teilfolge, die wieder mit $u(v_n)$ bezeichnet wird, mit $u(v_n) \rightarrow u^* \neq u(v^*)$. Stets gilt

$$F(u(v_n); v_n) \leq F(u(v^*), v_n). \quad (4.2)$$

Wegen der Stetigkeit von F ist

$$\begin{aligned} & |F(u(v_n); v_n) - F(u^*; v^*)| \\ & \leq |F(u(v_n); v_n) - F(u(v_n); v^*)| + |F(u(v_n); v^*) - F(u^*; v^*)| \rightarrow 0 \end{aligned}$$

und der Grenzübergang $n \rightarrow \infty$ in (4.2) ergibt

$$F(u^*; v^*) \leq F(u(v^*); v^*),$$

was ein Widerspruch zu $F(u^*; v^*) > F(u(v^*); v^*)$ ist. \square

Die häufig anzutreffende Bedingung, die $g_{a,b}$ bzw. die f_a müssen (strikt) konvex sein, ermöglicht es, die Voraussetzungen von Satz 4.1 einfach nachzuprüfen.

Mit zusätzlichen Monotonieforderungen gilt folgendes Einschlussprinzip.

Satz 4.2 (Einschlussprinzip)

Sind die $g_{a,b}$ und f_a streng monoton steigend, dann gilt für jede Lösung u^* von (4.1) folgendes Einschlussprinzip

$$v_{\min} = \min_a v(a) \leq u^* \leq \max_a v(a) = v_{\max}.$$

Sind die $g_{a,b}$ und f_a nur monoton steigend, so ist zu einer Lösung u^* von (4.1) durch $\bar{u} = \max(v_{\min}, \min(v_{\max}, u^*))$ eine Lösung \bar{u} festgelegt, die das Einschlussprinzip erfüllt.

Beweis

Angenommen, u^* ist Lösung und erfüllt nicht die Behauptung. Setzt man $\bar{u} = \max(v_{\min}, \min(v_{\max}, u^*))$, so gilt $|\bar{u}_a - \bar{u}_b| \leq |u_a^* - u_b^*|$ und $|\bar{u}_a - v_a| \leq |u_a^* - v_a|$ für alle $a, b \in V$. Ferner gibt es mindestens ein $c \in V$, so dass $|\bar{u}_c - v_c| < |u_c^* - v_c|$. Im Falle der strengen Monotonie der $g_{a,b}$ und f_a folgt $F(\bar{u}) < F(u^*)$, was ein Widerspruch zur Annahme ist, dass u^* Minimum ist. Im Falle der Monotonie folgt $F(\bar{u}) \leq F(u^*)$ und \bar{u} ist damit auch eine Lösung. \square

Direkt an der Problemstellung erkennbar ist, dass folgende Verschiebungsinvarianz gilt: Ist u_* eine Lösung von (4.1), so ist für ein $t \in \mathbb{R}$ dann $u_* + t$ eine Lösung zu $F(u; v + t) = \min!$.

Zusätzlich zu dieser Verschiebungsinvarianz kann man sich auch fragen, wann eine Skalierungsinvarianz vorliegt, also wann für ein $s \in \mathbb{R} \setminus \{0\}$ und eine Lösung u^* von (4.1) su^* eine Lösung von $F(u; sv) = \min!$ ist. Hinreichend dafür ist die von Bouman und Sauer in [BS93, Definition 3] eingeführte Eigenschaft der Skalierbarkeit.

Definition 4.3 (Skalierbare Funktionen)

F heißt skalierbar, falls es für jedes $s \neq 0$ ein $m = m(s) > 0$ und ein $c = c(s) \in \mathbb{R}$ gibt, so dass $F(su) = m(s)F(u) + c(s)$ gilt.

Bouman und Sauer haben folgende Charakterisierung gezeigt.

Satz 4.4 ([BS93, Theorem 2])

F ist genau dann konvex und skalierbar mit $\int_{\mathbb{R}^d} \exp(-F(x)) dx < \infty$, falls für irgendeine Norm $\|\cdot\|$, ein $p \geq 1$ und ein $C \in \mathbb{R}$ gilt: $F(x) = \|x\|^p + C$.

Verzichtet man auf die Konvexität, so muss man damit rechnen, dass die Eindeutigkeit der Lösung verloren geht. Dies ist etwa auch bei $g_{a,b}(x) = g_{BZ}(x) := \min(x^2, 1)$ (vgl. Abbildung 4.9 rechts) der Fall. Das mit diesem g_{BZ} entstehende Funktional hat zudem noch sehr viele lokale Minima, so dass zur Suche eines globalen Minimums speziell angepasste Algorithmen zum Einsatz kommen müssen. So haben Blake und Zisserman in [BZ87] einen graduated non-convexity (GNC) Algorithmus vorgeschlagen. Hierbei handelt es sich um ein Verfahren, welches eine Homotopie verwendet, die bei einem konvexen Funktional startet und bei F endet. Im Laufe der Homotopie wird versucht das globale Minimum zu verfolgen (vgl. Abbildung 4.1). Die Details sowie eine Implementierung dieses Algorithmus war Teil der Diplomarbeit [Lud02].

Ein anderer Ansatz ist in [GG84] beschrieben: Simulated Annealing. Dabei wird ein Algorithmus, der in Abstiegsrichtung zum nächsten lokalen Minimum läuft, um eine stochastische Komponente erweitert: mittels eines Zufallsprozesses werden „globale Sprünge“ eingebaut, um zu vermeiden, dass

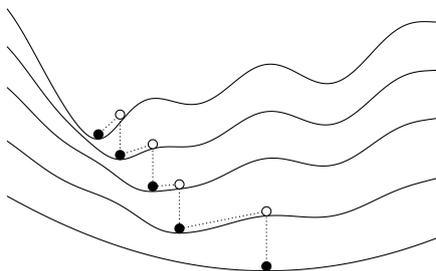


Abbildung 4.1: Minimum Tracking beim GNC-Algorithmus

der Algorithmus sich in einem lokalen Minimum festfrisst. Ein Parameter, den man als Temperatur deuten kann, steuert, wie häufig solche Sprünge stattfinden.

Wie im Abschnitt 2.3 zu sehen, scheint die Modellierung von Kanten im Diskreten (so dass im Gamma-Limes Integralterme der Form $\int_{S_u} \dots d\mathcal{H}^1$ auftauchen) und die Nichtkonvexität der $g_{a,b}$ eng miteinander verknüpft zu sein. Bei der TV-Minimierung (vgl. Kapitel 1), bei der kein Kantenterm in obiger expliziter Form auftaucht, sind jedoch konvexe Ansätze möglich, z. B. [Cha04].

4.3 Zusammenhang mit Graphcuts

Falls zur Minimierung des Funktionals (4.1) nur Funktionen zulässig sind, die nur endliche viele Werte annehmen dürfen, also ein $u \in U^L$ mit $U^L := \{u: \Omega \rightarrow L\}$ gesucht wird und $L \subset \mathbb{R}$ nur endlich viele Elemente enthält, so können Methoden aus der Graphentheorie zum Einsatz kommen. Insbesondere werden Graphcuts zur Anwendung kommen.

Definition 4.5 (Graphcut und Kosten eines Graphcuts)

Ist $G = (N, E)$ ein gerichteter und gewichteter Graph mit nichtnegativen Kantengewichten auf der Knotenmenge N und der Kantenmenge E . In N seien zwei Knoten $s, t \in N$ ($s \neq t$) ausgezeichnet. Eine disjunkte Zerlegung von N in zwei Mengen S und T ($V = S \cup T$, $S \cap T = \emptyset$) mit $s \in S$ und $t \in T$ heißt Graphcut von G .

Die Kosten $\text{cut}(S, T)$ eines Graphcuts mit den Mengen S und T sind durch die Summe aller Kantengewichte $w_{a,b}$, deren Kante $(a, b) \in E$ von S nach T laufen,

$$\text{cut}(S, T) := \sum_{\substack{(a,b) \in E \\ a \in S, b \in T}} w_{a,b}$$

festgelegt (siehe auch Abbildung 4.2).

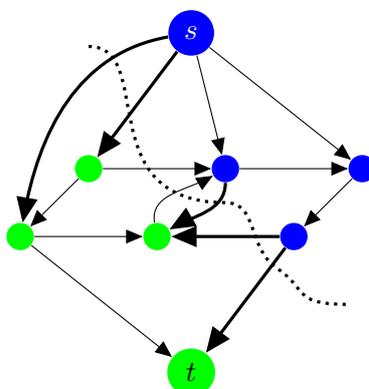


Abbildung 4.2: Graphcut; Menge S blau, Menge T grün, Kanten mit Beitrag zu $\text{cut}(S, T)$ sind dicker

In folgenden Experimenten diente zum Auffinden eines Graphcuts mit minimalen Kosten das C++ Programm „graphcut“ von Walter Bell [Bel01] in der Version 1.03 als Grundlage. Es wurde umgeschrieben, so dass es dem ISO C90 Standard erfüllt und ein Matlab Mex-Interface wurde programmiert. So steht Bell's Implementierung direkt unter Matlab zur Verfügung.

4.3.1 Expansion Moves

Von Boykov, Veksler und Zabih wurde in [BVZ99] Expansion Moves zur Lösung von (4.1) vorgeschlagen.

Definition 4.6 (Expansion Move)

Sei $u^0 \in U^L$ und $\alpha \in L$ beliebig, dann heißt ein $u \in U^L$ ein Expansion Move von u^0 bezüglich α , falls

$$\begin{aligned} u_a &= \alpha & \text{für } u_a^0 &= \alpha \\ u_a &\in \{u_a^0, \alpha\} & \text{für } u_a^0 &\neq \alpha \end{aligned}$$

für alle $a \in V$ erfüllt ist.

Ein Expansion Move von u^0 bezüglich α ist also eine Funktion, bei der an beliebig vielen Stellen der ursprüngliche Wert von u^0 durch α ersetzt wurde. Zur Minimierung von (4.1) wurde vorgeschlagen, ein $u \in U^L$ zu suchen, welches

$$F(u) \leq F(v) \quad \text{für alle Expansion Moves } v \text{ von } u \text{ bzgl. } \alpha \quad \forall \alpha \in L \quad (4.3)$$

erfüllt. Ein solches u ist also ein lokales Minimum bezüglich allen möglichen Expansion Moves.

Um ein solches u zu finden, wird iterativ vorgegangen und zu einem gegebenem $\alpha \in L$ und gegebenem $u^0 \in U^L$ ein $u^1 \in U^L$ gesucht mit

$$F(u^1) \leq F(v) \quad \text{für alle Expansion Moves } v \text{ von } u^0 \text{ bzgl. } \alpha. \quad (4.4)$$

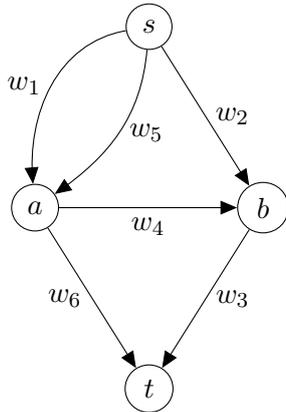
Dies wird mit verschiedenen α so lange wiederholt, bis für kein $\alpha \in L$ mehr eine Veränderung eintritt. Oft wird L in eine bestimmte Reihenfolge gebracht und die Expansion Moves in dieser Reihenfolge durchgeführt. $|L|$ solche Iterationen heißen auch ein Durchlauf (oder ein Cycle). Natürlich wird durch dieses Vorgehen nicht sichergestellt, dass das globale Minimum erreicht wird, aber in Satz 4.7 wird eine Schranke angegeben, wie weit die so gefundenen lokalen Minima höchstens vom Optimum entfernt sind.

In [KZ04] haben Kolmogorov und Zabih notwendige und hinreichende Bedingungen angegeben, wann (4.4) als Graphcut Problem dargestellt und gelöst werden kann.

Theorem 4.1 in [KZ04] liefert: Genau dann, wenn die $g_{a,b}$ die Bedingung

$$g_{a,b}(0) + g_{a,b}(|\beta - \gamma|) \leq g_{a,b}(|\alpha - \gamma|) + g_{a,b}(|\beta - \alpha|) \quad \forall a, b \in V, \forall \alpha, \beta, \gamma \in L \quad (4.5)$$

erfüllen, lässt sich für jedes α ein Graph $G = (N, E)$ mit $N = V \cup \{s, t\}$ konstruieren, so dass das Problem (4.4) zu diesem α äquivalent zur Suche eines Graphcuts von G mit minimalen Kosten ist. Das Konstruktionsprinzip ist in Abbildung 4.3 für den Knoten a zusammen mit den Kanten zu einem Nachbarn b dargestellt.



$$w_1 = \max(0, C - A)$$

$$w_2 = \max(0, D - C)$$

$$w_3 = \max(0, C - D)$$

$$w_4 = B + C - A - D$$

$$w_5 = \max(0, F - E)$$

$$w_6 = \max(0, E - F)$$

$$A = g_{a,b}(0), \quad B = g_{a,b}(|\alpha - u_b|),$$

$$C = g_{a,b}(|u_a - \alpha|), \quad D = g_{a,b}(|u_a - u_b|),$$

$$E = f_a(|\alpha - v_a|), \quad F = f_a(|u_a - v_a|)$$

Abbildung 4.3: Kanten des zu (4.4) gehörenden Graphen nach [KZ04] für $a \sim b$

Erfüllen die $g_{a,b}$ stets $g_{a,b}(0) = 0$ für alle $a, b \in V$, dann besagt die Bedingung (4.5) gerade, dass jedes $g_{a,b}$ eine Metrik sein muss und damit die beiden Ungleichungen aus (3.1) aus Satz 3.2 erfüllt sein müssen.

In [Vek99] wurde folgende Optimalitätseigenschaft gezeigt.

Satz 4.7 ([Vek99, Abschnitt 4.3.4], Quasioptimalität)

Für alle $a, b \in V$ seien die $g_{a,b}$ Metriken. Ferner sei $u^* \in U^L$ eine Lösung von (4.1). Für jedes $u \in U^L$, welches (4.3) erfüllt, also minimal bezüglich Expansion Moves ist, gilt

$$F(u) \leq 2cF(u^*),$$

wobei c durch

$$c = \max_{a,b \in V} \left(\frac{\max_{\alpha \neq \beta} g_{a,b}(\alpha, \beta)}{\min_{\alpha \neq \beta} g_{a,b}(\alpha, \beta)} \right)$$

gegeben ist.

In Abbildung 4.4 ist ein typischer Energieverlauf für dieses Verfahren angetragen. In diesem Beispiel wurde

$$F(u) = \sum_{\substack{a,b \in V \\ a \sim b}} |u_a - u_b|^{0.1} + \sum_{a \in V} |u_a - v_a|^{1.1} \quad (4.6)$$

minimiert. Das vorgegebene Eingangsbild v ist in Abbildung 4.5 links zu sehen. Es handelt sich um ein 60×60 Bild mit 31 verschiedenen Graustufen. Dieses Bild wurde zugleich als Startbild u der Iteration verwendet. Als Reihenfolge für die Expansion Moves wurde $\alpha = 1, 2, \dots, 31$ gewählt und zwei Durchläufe gerechnet (damit wurden 62 Expansion Moves ausgeführt).

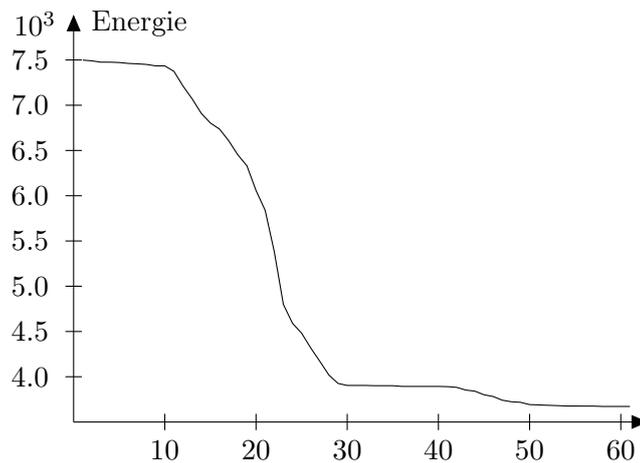


Abbildung 4.4: Energieverlauf im Laufe der iterativen Minimierung von (4.6) mit Expansion Moves

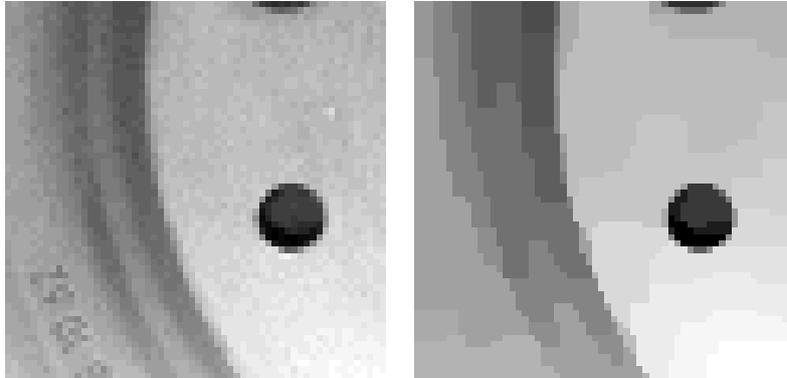
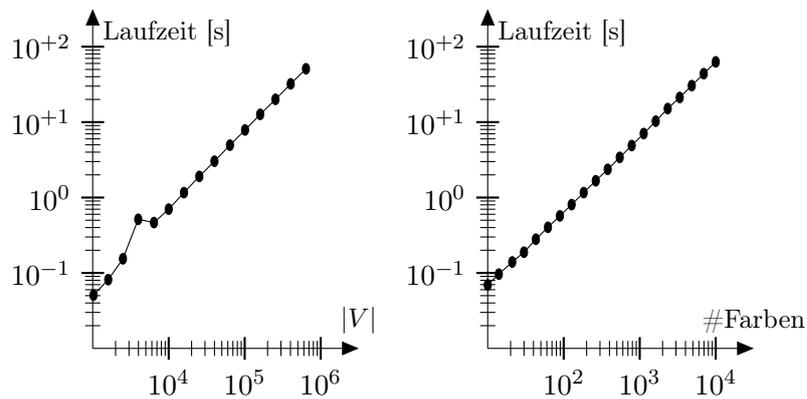


Abbildung 4.5: links: Eingabe- und gleichzeitig Startbild; rechts: Ausgabebild

Abbildung 4.6: Experimentelles Laufzeitverhalten eines Expansion Move Cycles: links in Abhängigkeit der Bildgröße; rechts in Abhängigkeit der Farb-
raumgröße

Wegen der Bedingung (4.5) können Nachbarschaftsfunktionen $g_{a,b}$ der Form $g_{a,b}(u_a, u_b) = |u_a - u_b|^\gamma$ nur für $0 < \gamma \leq 1$ mittels Graphcuts behandelt werden.

In Abbildung 4.6 ist die Dauer eines Durchlaufs zu sehen. Auf der linken Seite wurde die Anzahl der Farben (31) festgehalten und die Größe des Bildes verändert. Auf der rechten Seite wurden die Bildabmessungen (60×60) nicht verändert, aber die Anzahl der Farben variiert.

Die Laufzeit zum Auffinden eines Graphcuts mit minimalen Kosten für einen Expansion Move ist demnach ein $\mathcal{O}(|V|)$. Für einen Cycle sind $|L|$ Expansion Moves notwendig. Damit beträgt die Komplexität für einen Cycle $\mathcal{O}(|L| \cdot |V|)$.

Das lokale Minimum wird in endlich vielen Schritten erreicht: Da der Energieverlauf monoton fallend ist und L nur endlich viele Element besitzt, gibt es ein $\Delta E > 0$, so dass bei jedem Cycle die Energie mindestens um ΔE fällt, sofern ein lokales Minimum noch nicht erreicht ist. Als obere Abschätzung, wie viele Cycles notwendig sind, bis ein lokales Minimum erreicht wird, ist in Abschnitt 4.3.3 in [Vek99] $\mathcal{O}(|V|)$ zu finden. Experimentell beobachten lässt sich ein $\mathcal{O}(1)$.

4.3.2 Swap Moves

Um größere Problemklassen behandeln zu können, deren Nachbarschaftsfunktion $g_{a,b}$ nicht alle für die Theorie der Expansion Moves notwendigen Voraussetzungen erfüllen, wurden in Abschnitt 4.3.1 in [Vek99] Swap Moves vorgeschlagen.

Definition 4.8 (Swap Move)

Sei $u^0 \in U^L$ und $\alpha, \beta \in L$ beliebig mit $\alpha \neq \beta$, dann heißt ein $u \in U^L$ ein Swap Move von u^0 bezüglich α und β , falls

$$\begin{aligned} u_a \in \{\alpha, \beta\} & \quad \text{für } u_a^0 \in \{\alpha, \beta\} \\ u_a = u_a^0 & \quad \text{für } u_a^0 \notin \{\alpha, \beta\} \end{aligned}$$

für alle $a \in V$ erfüllt ist.

Ein Swap Move von u^0 bezüglich α und β ist also eine Funktion, bei der beliebig viele α -Stellen durch β und beliebige viele β -Stellen durch α ersetzt wurden.

Ganz analog zu den Expansion Moves kann man nun $u \in U^L$ suchen, die

$$F(u) \leq F(v) \quad \text{für alle Swap Moves } v \text{ von } u \text{ bzgl. } \alpha \text{ und } \beta \quad \forall \alpha, \beta \in L$$

erfüllen. Ein solches u ist dementsprechend ein lokales Minimum bezüglich allen möglichen Swap Moves.

Auch hier wird iterativ vorgegangen, um ein solches u zu finden. Zu einem gegebenem Paar $\alpha, \beta \in L$ und einem gegebenem $u^0 \in U^L$ wird ein $u^1 \in U^L$ gesucht mit

$$F(u^1) \leq F(v) \quad \text{für alle Swap Moves } v \text{ von } u^0 \text{ bzgl. } \alpha \text{ und } \beta. \quad (4.7)$$

Dies wird mit verschiedenen Paaren α, β wieder so lange wiederholt, bis für kein Paar mehr eine Veränderung eintritt.

Der Graph (N, E) wird dabei wie folgt konstruiert: $N = \{s, t\} \cup \{a \in V \mid u_a \in \{\alpha, \beta\}\}$. In Abbildung 4.7 sind für ein $a \in V \cap N$ die Kanten angegeben. Dabei ist $b \in V \cap N$ ein Nachbar von a .

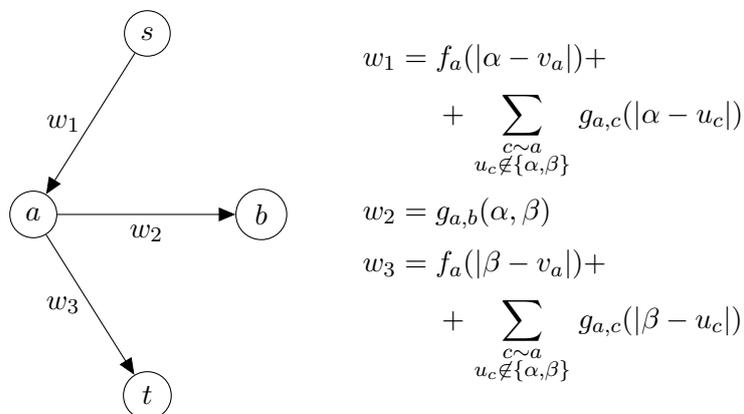


Abbildung 4.7: Kanten des zu (4.7) gehörenden Graphen nach [Vek99, Abschnitt 4.3.1] für $a \sim b \in N$

In Abbildung 4.8 sind wieder die Laufzeiten für das gleiche Beispiel wie bei den Expansion Moves angetragen. Es wurde wieder die Bildgröße und die Farbraumgröße variiert. Auch hier ist die Laufzeit für das Auffinden eines Graphcuts mit minimalen Kosten für einen Swap Move ein $\mathcal{O}(|V|)$. Pro Cycle sind diesmal aber $|L|^2$ viele Iterationen notwendig, da alle Paare $(\alpha, \beta) \in L \times L$ mit $\alpha \neq \beta$ behandelt werden müssen. Damit liegt der Aufwand pro Cycle bei $\mathcal{O}(|L|^2|V|)$. Auch hier beobachtet man, dass nach $\mathcal{O}(1)$ Cycles ein lokales Minimum erreicht wird.

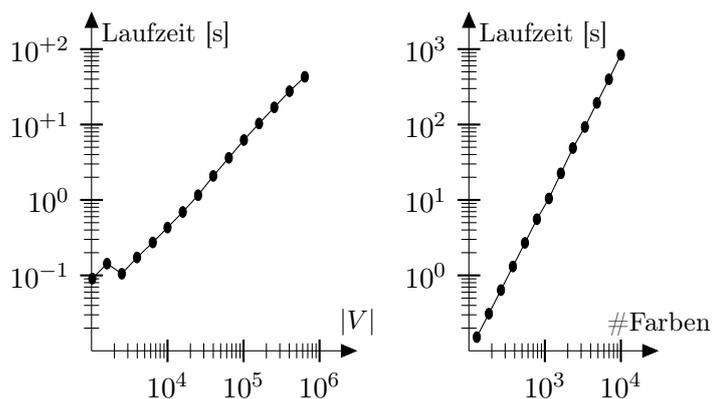


Abbildung 4.8: Experimentelles Laufzeitverhalten eines Swap Move Cycles: links in Abhängigkeit der Bildgröße; rechts in Abhängigkeit der Farbraumgröße

4.4 Markov Random Fields (MRF) und Maximum A-posteriori (MAP) Schätzer

In diesem Abschnitt soll kurz darauf eingegangen werden, wo im Bereich der stochastischen Bildverarbeitung Funktionale der Form (4.1) auftauchen, da insbesondere von diesem Bereich etliche Impulse ausgegangen sind. Eine ausführliche Einführung in diesen Bereich ist im Buch [Win03] zu finden.

Zur Umsetzung dieser stochastischen Idee benötigt man zwei Wahrscheinlichkeitsverteilungen: Auf dem Raum U der möglichen Ergebnisbilder muss eine Prior-Verteilung bekannt sein. Sie gibt an mit welcher Wahrscheinlichkeit $P(u)$ ein Bild u vorliegt. Außerdem wird ein Modell für die Transformation/Störungen benötigt, welches für ein fest vorgegebenes u eine Wahrscheinlichkeitsverteilung für die bedingte Wahrscheinlichkeit $P(\cdot|u)$ auf dem Raum der möglichen Eingabebilder liefert. So ergibt sich eine Gesamtverteilung von

$$P(u, v) = P(u) \cdot P(v|u).$$

Um von einem gegebenen Eingabebild v auf ein Ausgabebild u zu kommen, betrachtet man die Posterior-Verteilung

$$P(u|v) = \frac{P(u, v)}{P(\cdot, v)} = \frac{P(u) \cdot P(v|u)}{P(\cdot, v)}.$$

Hier kann man nun etwa die Frage nach dem Bild u stellen, welches $P(u|v)$ maximiert, also zum gegebenen v am maximal wahrscheinlichsten erscheint. Deshalb heißt dieses Vorgehen auch „Maximum A-posteriori“ Schätzung. Für den Fall, dass stets $P(u) > 0$ und $P(v|u) > 0$ gilt, ist

$$P(u|v) = \max! \quad \Leftrightarrow \quad \underbrace{-\log P(u)}_{=: E_D(u)} - \underbrace{\log P(v|u)}_{=: E_R(u, v)} = \min!$$

Nimmt man an, dass die Pixel des beobachtbaren Eingabebildes v einfach nur dadurch entstehen, dass die Pixel von u unabhängig voneinander mit additivem Gauß'schem Rauschen versehen werden, also für $a \in V$

$$v_a = u_a + n_a$$

gilt, dann sind die n_a jeweils Normalverteilt mit Erwartungswert 0 und Standardabweichung σ . Wegen der Unabhängigkeit ist

$$\begin{aligned} P(n) &= c \prod_{a \in V} \exp\left(-\frac{n_a^2}{2\sigma^2}\right) = c \exp\left(-\frac{1}{2\sigma^2} \sum_{a \in V} n_a^2\right) \\ &= c \exp\left(-\frac{1}{2\sigma^2} \sum_{a \in V} (v_a - u_a)^2\right) \end{aligned}$$

und demnach (der von u und v unabhängige Summand $-\log c$ wurde weggelassen)

$$E_D(u; v) = \frac{1}{2\sigma^2} \sum_{a \in V} (v_a - u_a)^2$$

ein bekannter Datenfitterm, wie er schon in der Einleitung erwähnt wurde. Wahrscheinlichkeitsverteilungen der Form $P(n) = c \exp(-G(n))$ heißen Gibbs Verteilungen. Jede Wahrscheinlichkeitsverteilung mit strikt positiver Dichtefunktion lässt sich in diese Form überführen.

Für die Modellierung des Priors wird oft ein Markov Random Field verwendet.

Definition 4.9 (Markov Random Field)

Das Wahrscheinlichkeitsmaß $P(u)$ heißt Markov Random Field (MRF), falls für alle u stets $P(u) > 0$ und

$$P(u_a | \{u_b : b \neq a\}) = P(u_a | \{u_b : b \sim a\})$$

gilt.

Bei einem MRF hängt also die Wahrscheinlichkeit eines Pixels nur von seinen Nachbarn ab. In der Gibbs'schen Form

$$P(u_a | \{u_b : b \neq a\}) = \exp\left(-\sum_{b \sim a} g_{a,b}(|u_a - u_b|)\right) \quad (4.8)$$

erhält man für E_R die Darstellung aus (4.1). Die Interaktionsfunktionen g bestimmen, wie stark sich Werte der Nachbarn eines Pixels auf diesen auswirken. Ist $g_{a,b}(t) = g_G(t) := t^2$, so spricht man von einem Gauß'schen MRF. Etliche weitere Priors sind untersucht worden. So wurde in [Hub81]

$$g_H(x) = \begin{cases} t^2 & \text{für } |t| \leq T \\ T(2|t| - T) & \text{für } |t| > T \end{cases}$$

vorgeschlagen (siehe Abbildung 4.9), um den Prior robuster gegenüber „Ausreißern“ werden zu lassen.

Geman und Geman haben in [GG84] einen sogenannten „line process“ vorgeschlagen. Diese Idee wurde schon beim eindimensionalen Beispiel in Abschnitt 2.2.6 verwendet. Dabei wird pro Nachbarschaft $a \sim b$ eine zusätzliche Variable $\ell_{a,b} \in \{0, 1\}$ eingeführt, welche für $\ell_{a,b} = 0$ eine Entkopplung von a und b repräsentiert. Regularisierungsfunktionale der Form

$$E_R(u) = \sum_{a \sim b} \ell_{a,b} \tilde{g}_{a,b}(|u_a - u_b|) + \sum_{a \sim b} (1 - \ell_{a,b}) C$$

mit einem $C > 0$ können durch Elimination dieses line process durch

$$\min_{\ell} E_R(u) = \sum_{a \sim b} \min(\tilde{g}_{a,b}(|u_a - u_b|), C)$$

mit $g_{a,b}(x) := \min(\tilde{g}_{a,b}(x), C)$ ebenfalls auf die Form (4.8) gebracht werden. So ist auch die Interaktionsfunktion von Blake und Zisserman entstanden [BZ87, Abschnitt 3.3 und 6.1]: $g_{BZ}(x) = \min(x^2, C)$. In Abbildung 4.9 sind die drei hier erwähnten Funktionen g_G , g_H und g_{BZ} zu sehen.

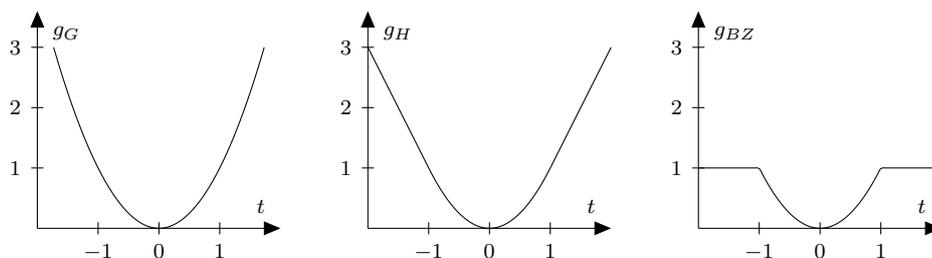


Abbildung 4.9: Typische Prior-Interaktionsfunktionen; g_H mit $T = 1$ und g_{BZ} mit $C = 1$

4.5 Block-Update-Methode

In diesem Abschnitt wird ein mehrgitterartiges Verfahren zur Minimierung von 4.1 für konvexe und stetig differenzierbare $g_{a,b}$ und f_a vorgestellt. Die Konstruktion dieser „Block-Update-Methode“ wurde für

$$g_{a,b}(x) := \alpha_{a,b}|x|^p \quad \text{und} \quad f_a(x) := \beta_a|x|^q \quad (4.9)$$

für $1 < p, q \leq 2$ und $\alpha_{a,b}, \beta_a \in [0; \infty[$ implementiert. Als Basis- und Vergleichsverfahren wird in Abschnitt 4.5.1 das Gauß-Seidel-Verfahren beschrieben. Im Anschluss wird die Block-Update-Methode mit zwei Gittern (einem Fein- und einem Grobgitter) beschrieben. Durch rekursives Vorgehen lässt sich das Verfahren auf eine Gitterhierarchie mit beliebiger Tiefe erweitern.

4.5.1 Gauß-Seidel-Verfahren

Für dieses Iterationsverfahren wird ein Startvektor $u^{(0)} \in \mathbb{R}^d$ benötigt. Der Schritt von einer Iterierten $u^{(k)}$ zu $u^{(k+1)}$ wird dadurch bewerkstelligt, dass bezüglich jeder Komponente genau einmal minimiert wird:

$$\begin{aligned} u^{(k+1)} &\leftarrow u^{(k)}: \\ \text{for } i &= 1, \dots, d \\ u_i^{(k+1)} &\leftarrow \operatorname{argmin}_{x \in \mathbb{R}} F\left(u_1^{(k+1)}, \dots, u_{i-1}^{(k+1)}, x, u_{i+1}^{(k+1)}, \dots, u_d^{(k+1)}\right) \end{aligned}$$

Da für $k \rightarrow \infty$ die Folge $F(u^{(k)})$ monoton fällt und F nach unten beschränkt ist, existiert $F^* := \lim F(u^{(k)})$. Für ein strikt konvexes F liegen alle $u^{(k)}$ in

einem Kompaktum und es gibt eine konvergente Teilfolge mit $u^{(k)} \rightarrow u^*$. Wegen der Stetigkeit von F ist $F^* = F(u^*) = \lim F(u^{(k)})$. Per Konstruktion (in keiner Achsenrichtung wird $F(u^*)$ kleiner) ist u^* ein lokales Minimum und damit auch ein globales Minimum.

4.5.2 Konstruktion der Grobgitterknoten

Ausgangspunkt sind die Knotenmenge V des Feingitters mit $d := |V|$ Knoten und die $\alpha_{a,b}$ bzw. die β_a . Ziel dieses Abschnitts ist die Konstruktion einer Knotenmenge $\tilde{V} \subset V$, welche als Grobgitter dient. Die $\alpha_{a,b}$ geben an, wie stark der Knoten b mit dem Knoten a verkoppelt bzw. verbunden ist. Mit folgendem Auswahlverfahren für die Grobgitterknotenmenge wird versucht, mehrere stark miteinander verbundene Knoten zu einer „Gruppe“ zusammenzufassen und durch einen Grobgitterknoten zu ersetzen. Das hier vorgestellte Einteilungsverfahren orientiert sich daran, wie typischerweise bei einem linearen algebraischen Mehrgitterverfahren die Grobgitter konstruiert werden, vgl. [Stü99].

Dazu definieren wir zunächst zu einem $a \in V$ die Menge

$$S_a := \{ b \in N_a : \alpha_{a,b} \geq \epsilon_{\text{str}} \}$$

der stark mit a verbundenen Knoten. Dabei ist ϵ_{str} ein vorzugebender Parameter. Zur Beschreibung des Entscheidungsprozesses, welche Knoten in \tilde{V} kommen, bezeichne im Folgenden $U \subset V$ die Menge der Knoten, bei denen diese Entscheidung noch nicht gefallen ist. Mit $C := \tilde{V}$ wird die Menge der Grobgitterknoten und mit $F := V \setminus \tilde{V}$ die Menge der Feingitterknoten bezeichnet. Für jedes $a \in U$ wird ein Indikator $\lambda_a \in \mathbb{N}_0$ durch

$$\lambda_a := |S_a \cap U| + 2|S_a \cap F|$$

festgelegt, welcher die Dringlichkeit angibt, mit der der Knoten a in C aufgenommen werden soll. Die Idee ist nun, einen Knoten in U mit höchster Dringlichkeit in C aufzunehmen und alle Knoten, die stark mit diesem verbunden sind, in F aufzunehmen. Da sich dabei die Mengen U , C und F verändern, verändern sich auch die Indikatoren λ_a . Diese Veränderungen können jedoch inkrementell und effizient berechnet werden. Wegen $\lambda_a \in \mathbb{N}_0$ kann mit Hilfe von Bucket-Sort die Bestimmung des Maximums in linearer Laufzeit erfolgen.

1. $C \leftarrow \emptyset; F \leftarrow \emptyset; U \leftarrow V; \lambda_a := |S_a| \quad \forall a \in U$
2. wähle $a = \operatorname{argmax}_{b \in U} \lambda_b; C \leftarrow C \cup \{a\}; U \leftarrow U \setminus \{a\}$
3. $\forall b \in S_a \cap U:$

- (a) $F \leftarrow F \cup \{b\}$; $U \leftarrow U \setminus \{b\}$
 (b) $\forall c \in S_b \cap U : \lambda_c \leftarrow \lambda_c + 1$

4. falls $U \neq \emptyset$, gehe zu 2; sonst setze $\tilde{V} := C$

Damit sind die Mengen C und F eine disjunkte Zerlegung von V . Deshalb wird der soeben beschriebene Einteilungsalgorithmus auch als C/F-Splitting bezeichnet.

4.5.3 Konstruktion der Grobgitterkanten

Nachdem im vorherigen Abschnitt der Algorithmus zur Bestimmung der Grobgitterknotenmenge \tilde{V} vorgestellt wurde, wird nun angegeben, wie die neuen Kanten auf dem Grobgitter berechnet werden. Es muss also für jedes Paar $\tilde{a}, \tilde{b} \in \tilde{V}$ angegeben werden, ob \tilde{a} und \tilde{b} verbunden sind und falls dies der Fall ist, wie $\tilde{\alpha}_{\tilde{a}, \tilde{b}}$ lautet. Auch hier bezeichne $F := V \setminus \tilde{V}$ die Knoten, welche ausschließlich im Feingitter vorhanden sind. Wir setzen

$$\tilde{\alpha}_{\tilde{a}, \tilde{b}} := 2\alpha_{\tilde{a}, \tilde{b}} + \sum_{\substack{\tilde{a} \sim f \sim \tilde{b} \\ f \in F}} \frac{1}{2}(\alpha_{\tilde{a}, f} + \alpha_{f, \tilde{b}}).$$

Durch obige Grobgitterkonstruktion von \tilde{V} und $\tilde{\alpha}_{\tilde{a}, \tilde{b}}$ wurde versucht, die Verbindungsstruktur mit weniger Knoten abzubilden: \tilde{a} und \tilde{b} sind verbunden, falls die beiden schon im Feingitter verbunden waren oder falls es mindestens einen Feingitterknoten f gibt, den beide zum Nachbarn haben.

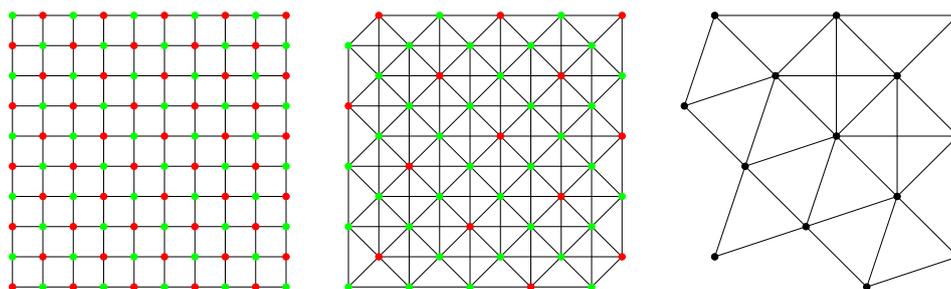


Abbildung 4.10: Beispiel einer Gitterhierarchie (rot: Menge C , grün: Menge F)

In Abbildung 4.10 sind für 10×10 Pixel das C/F-Splitting sowie die nächsten beiden Grobgitter zu sehen. Bei diesem Beispiel haben im Ausgangsgitter alle $\alpha_{a,b}$ den gleichen Wert, so dass jeder Pixel mit seinen Nachbarn (4er Nachbarschaft) gleich stark gekoppelt ist.

4.5.4 Interpolation

Der Interpolationsoperator \tilde{I} beschreibt, wie aus einer Grobgitterfunktion $\tilde{u}: \tilde{V} \rightarrow \mathbb{R}$ eine Funktion $u: V \rightarrow \mathbb{R}$ berechnet wird. In diesem Kontext wird eine lineare Interpolation $\tilde{I}: \mathbb{R}^{\tilde{d}} \rightarrow \mathbb{R}^d$ vorgeschlagen, welche u_a bei einem Knoten $a \in V \setminus \tilde{V}$ als gewichtete Summe der Nachbarwerte u_b mit $b \in \tilde{V}$ berechnet:

$$u_a = u(a) := \begin{cases} \tilde{u}_a & \text{falls } a \in \tilde{V}, \\ \left(\sum_{\tilde{V} \ni b \sim a} \alpha_{a,b} \tilde{u}_b \right) / \left(\sum_{\tilde{V} \ni b \sim a} \alpha_{a,b} \right) & \text{falls } a \in V \setminus \tilde{V}. \end{cases}$$

4.5.5 Update-Mechanismus

Seien nun $V_L \subset V_{L-1} \subset \dots \subset V_1 \subset V_0$ eine nach 4.5.2 und 4.5.3 konstruierte Graphenhierarchie. Ferner bezeichne I_k^{k-1} (für $k \in \{1, \dots, L\}$) die Interpolation vom Gitter k zum feineren Gitter $k-1$, so wie in 4.5.4 beschrieben. Sei abkürzend

$$I_k^0 := I_1^0 I_2^1 \dots I_k^{k-1}.$$

Bei einer Update-Iteration mit Hilfe des Levels k wird nun zu einem gegebenem Vektor $u_0^{(0)}$ auf Level 0 zur Berechnung der nächsten Iterierten $u_0^{(1)}$

$$\bar{u}_k = \underset{u_k}{\operatorname{argmin}} F(u_0^{(0)} + I_k^0 u_k) \quad (4.10)$$

betrachtet und

$$u_0^{(1)} = u_0^{(0)} + I_k^0 \bar{u}_k \quad (4.11)$$

gesetzt.

Zur approximativen Lösung von (4.10) wird das Gauß-Seidel-Verfahren auf jede Komponente von u_k angewendet. Als Startvektor wird $u_k = 0$ verwendet.

4.5.6 V-Zyklus

Der im vorherigen Abschnitt beschriebene Update-Mechanismus auf einem Level ℓ kann nun wie bei einer Mehrgittermethode in einem V-Zyklus verwendet werden. Abbildung 4.11 zeigt für drei Levels den Verlauf. Jeder „•“ entspricht einem Update nach (4.10) und (4.11).

4.5.7 Laufzeitverhalten

Da für ein Block-Update nach (4.10) und (4.11) auf Level ℓ auch der Level 0 „besucht werden muss“, ist nicht damit zu rechnen, dass diese Methode optimale Mehrgitterkomplexität aufweist.

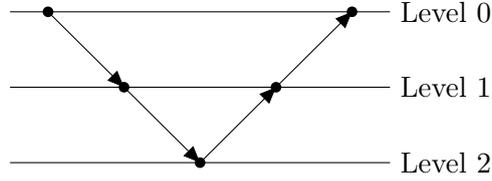


Abbildung 4.11: V-Zyklus der Block-Update-Methode

Sei d die Anzahl der Pixel auf Level 0 und N die maximale Anzahl von Nachbarn, die irgendein Graphknoten hat. Geht man davon aus, dass pro Level die Anzahl der Knoten halbiert wird, so gibt es $\mathcal{O}(d2^{-\ell})$ Knoten pro Level und $L = \mathcal{O}(\log_2 d)$ Levels. Für die Minimierung von (4.10) auf Level ℓ müssen pro zu minimierende Komponente $\mathcal{O}(N^\ell)$ Pixel auf Level 0 verändert werden. Die Kosten U_ℓ für das Update (4.10) und (4.11) belaufen sich daher auf

$$U_\ell = \mathcal{O}((N/2)^\ell d).$$

Der Aufwand eines V-Zyklus ist

$$\sum_{\ell=0}^L U_\ell = \mathcal{O}\left(d \sum_{\ell=0}^L (N/2)^\ell\right) = \mathcal{O}(d(N/2)^L) = \mathcal{O}(d^{\log_2 N}).$$

Damit weist die Block-Update-Methode eine Laufzeitkomplexität auf, die weit von einem effizienten Algorithmus entfernt ist. So ergibt sich für (durchschnittlich) $N = 4$ Nachbarn eine Komplexität von $\mathcal{O}(d^2)$.

In Abbildung 4.12 ist für $p = q = 1.4$, $\alpha_{a,b} \equiv \alpha$, $\beta_a \equiv \beta$ und $\alpha : \beta = 2 : 1$ für das Autofelgentestbild (siehe Abbildung 1.1) der Laufzeitvergleich der Block-Update-Methode mit dem Gauß-Seidel-Verfahren zu sehen.

4.5.8 Beispiel

Abbildung 4.13 zeigt für eine konkrete Wahl von p und q sowie verschiedene α und β die mit der Block-Update-Methode approximierten Minima des Funktionals

$$F(u; v) = \alpha \sum_{a \sim b} |u_a - u_b|^p + \beta \sum_a |u_a - v_a|^q. \quad (4.12)$$

Man sieht über weite Parameterbereiche von α und β brauchbare Ergebnisse. In den Fällen bei denen kleine Störungen durch Glättung beseitigt wurden, bleiben trotzdem kontrastreiche Kanten optisch erkennbar und werden kaum verschmiert. Dies ist ein erhebliche Fortschritt im Vergleich zur oft

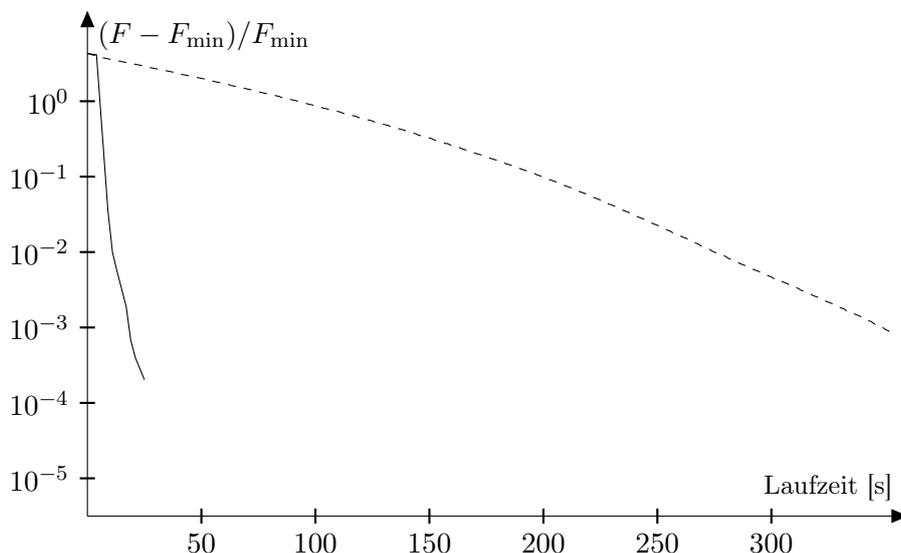


Abbildung 4.12: Vergleich: Block-Update-Methode (durchgezogene Linie) und Gauß-Seidel-Verfahren (gestrichelte Linie)

verwendeten Wahl $p = q = 2$. So führt eine Wahl von $1 < p, q < 2$ zu großen Verbesserungen, ohne dass die Konvexitätseigenschaften verloren gehen.

Für $\alpha/\beta \rightarrow \infty$ ist das Optimum ein konstantes Bild mit dem mittleren Grauwert. Aus diesem Grund empfiehlt es sich, die Bilder im Anschluss auf den ursprünglichen Grauwertbereich zu reskalieren. Dies wurde zum Zwecke der Vergleichbarkeit in keiner der Abbildungen vorgenommen.

4.6 Graduated Non-Convexity

Es wird das Problem

$$F(u, \ell; v) := \alpha \sum_{a \sim b} \ell_{a,b} |u_a - u_b|^p + \beta \sum_{a \in V} |u_a - v_a|^q + \sum_{a \sim b} (1 - \ell_{a,b}) = \min! \quad (4.13)$$

mit $1 < p, q \leq 2$, $\alpha, \beta > 0$ und $\ell_{a,b} \in \{0, 1\}$ für $a \sim b$ betrachtet. Wie in den Abschnitten 2.2.6 und 4.4 handelt es sich bei ℓ um die Idee des „line process“. Auch hier kann dieser mittels

$$g_0: [0; \infty[\rightarrow \mathbb{R}, \quad t \mapsto g_0(t) := \min(1, t)$$

eliminiert werden und das Problem (4.13) lässt sich in der Form

$$F_0(u; v) = F_0(u) = \sum_{a \sim b} g_0(\alpha |u_a - u_b|^p) + \beta \sum_a |u_a - v_a|^q \quad (4.14)$$

schreiben. Dieses Funktional ist nicht konvex.

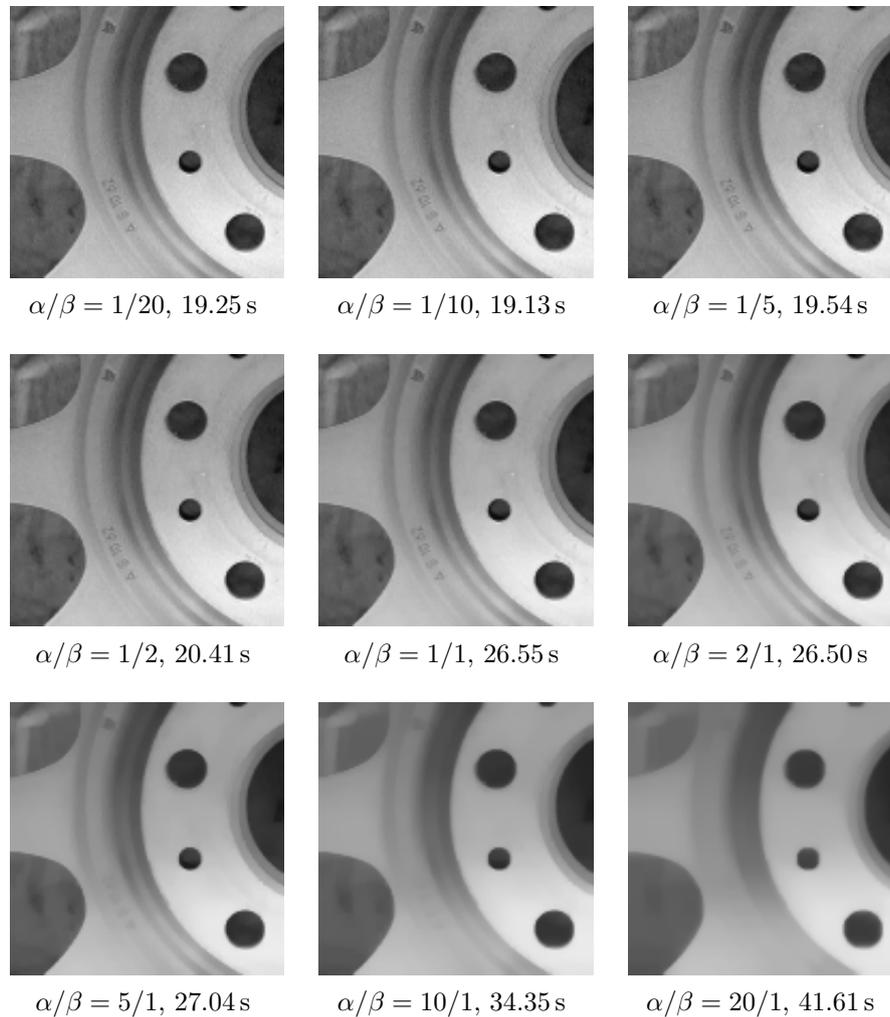


Abbildung 4.13: Mit Block-Update-Methode approximierte Minima von (4.12) mit $p = 1.2$, $q = 1.8$ und verschiedene α, β -Kombinationen

4.6.1 Homotopiekonstruktion: Vom konvexen zum gesuchten Funktional

In diesem Abschnitt wird mit Hilfe einer Funktionalfamilie F_τ eine Homotopie von einem konvexen Funktional zu F_0 konstruiert. Für $p = 2$ ist eine solche Konstruktion in [Lud02, Abschnitt 5.3] zu finden. Eine solche Homotopie kann als Grundlage für einen GNC-Algorithmus dienen, um ein globales Minimum von F_0 zu approximieren. Hat jedes Funktional F_τ nur ein globales Minimum, so hängt nach Satz 4.1 die Position des Minimums stetig von τ ab und die Verfolgung des globalen Minimums ist möglich, so wie es

in Abbildung 4.1 zu sehen ist.

Dazu betrachtet man die Funktionen $P: t \mapsto P(t) := |t|^p$, $Q: t \mapsto Q(t) := |t|^q$ und

$$F_\tau(u) := \underbrace{\sum_{a \sim b} g_\tau[\alpha P(u_a - u_b)]}_{=: G_\tau(u)} + \beta \underbrace{\sum_a Q(u_a - f_a)}_{=: H(u, f)} \quad (4.15)$$

mit noch näher zu bestimmenden g_τ . Für $t \neq 0$ gilt $P'(t) = pt|t|^{p-2}$ und $P''(t) = p(p-1)|t|^{p-2}$ bzw. $Q'(t) = qt|t|^{q-2}$ und $Q''(t) = q(q-1)|t|^{q-2}$.

Satz 4.10

Erfüllen die $g_\tau \in C^2([0; \infty[, \mathbb{R})$ die Differentialungleichung

$$\alpha^2 g_\tau''[\alpha P(x)][P'(x)]^2 + \alpha g_\tau'[\alpha P(x)]P''(x) \geq -\frac{1}{\tau}, \quad (4.16)$$

und ist $\tau \geq 4|V|/Q''(f_{\max} - f_{\min})$, dann ist F_τ auf der Menge $\{u : f_{\min} \leq u \leq f_{\max}\}$ konvex.

Beweis

Sei $f_{\min} \leq u \leq f_{\max}$. Es wird nun die 2. Ableitung von $G_\tau(u + t\Delta u) + H(u + t\Delta u)$ in Richtung von Δu untersucht. Dazu sei $\max_a |\Delta u_a| \leq 1$ und dieses Maximum werde angenommen. Dann gilt mit $g(t) := G(u + t\Delta u)$ und $h(t) := H(u + t\Delta u)$

$$g'(t) = \sum_{a \sim b} \alpha g_\tau'[u_a - u_b + t(\Delta u_a - \Delta u_b)] \cdot P'[u_a - u_b + t(\Delta u_a - \Delta u_b)](\Delta u_a - \Delta u_b),$$

außerdem

$$\begin{aligned} g''(0) &= \sum_{a \sim b} \left\{ \alpha^2 g_\tau''[\alpha P(u_a - u_b)][P'(u_a - u_b)]^2 + \right. \\ &\quad \left. \alpha g_\tau'[\alpha P(u_a - u_b)]P''(u_a - u_b) \right\} (\Delta u_a - \Delta u_b)^2 \\ &\stackrel{(4.16)}{\geq} -\frac{1}{\tau} \sum_{a \sim b} (\Delta u_a - \Delta u_b)^2 \geq -\frac{4}{\tau}|V| \end{aligned}$$

und

$$h''(0) = \sum_a Q''(u_a - f_a) \Delta u_a^2 \geq Q''(f_{\max} - f_{\min}).$$

Somit gilt

$$g''(0) + h''(0) \geq -\frac{4}{\tau}|V| + Q''(f_{\max} - f_{\min}) \geq 0.$$

Damit ist die 2. Ableitung von $G_\tau + H$ in allen Richtungen niemals negativ. Folglich ist F_τ konvex. \square

Satz 4.11

Für $C_1 \in \mathbb{R}$, $C_2 > 0$ und $\tau > 0$ ist $g:]0; \infty[\rightarrow \mathbb{R}$ mit

$$g(x) = \frac{C_1}{\alpha} + \frac{C_2}{\alpha} x^{1/p} - \frac{1}{2\tau} \alpha^{(-2/p)} x^{2/p}$$

eine Lösung der zur Differentialungleichung (4.16) gehörenden Differentialgleichung

$$\alpha^2 g''[\alpha P(x)][P'(x)]^2 + \alpha g'[\alpha P(x)]P''(x) = -\frac{1}{\tau}.$$

Diese Behauptung lässt sich durch zweimaliges Differenzieren von g und Einsetzen in die linke Seite verifizieren.

Für die Konstruktion der Homotopie werden nun $x_0 = x_0(\tau)$, $x_1 = x_1(\tau)$, $C_1 = C_1(\tau)$ und $C_2 = C_2(\tau)$ in Abhängigkeit von τ so bestimmt, dass

$$\begin{aligned} g(x_0) &= x_0, & g'(x_0) &= 1, & \text{und} \\ g(x_1) &= 1, & g'(x_1) &= 0 \end{aligned}$$

erfüllt sind.

Für ein $\tau > 0$ wird dann g_τ durch

$$g_\tau(x) := \begin{cases} x & \text{falls } 0 \leq x \leq x_0(\tau) \\ \frac{C_1}{\alpha} + \frac{C_2}{\alpha} x^{1/p} - \frac{1}{2\tau} \alpha^{(-2/p)} x^{2/p} & \text{falls } x_0(\tau) < x < x_1(\tau) \\ 1 & \text{falls } x_1(\tau) \leq x \end{cases} \quad (4.17)$$

festgelegt. Per Konstruktion ist $g_\tau \in C^1$ für alle $\tau > 0$. Alle g_τ sind, außer bei $x \in \{x_0, x_1\}$, zweimal stetig differenzierbar.

Somit gilt

$$\begin{aligned} & \alpha^2 g_\tau''[\alpha P(x)][P'(x)]^2 + \alpha g_\tau'[\alpha P(x)]P''(x) \\ &= \left\{ \begin{array}{ll} \alpha P''(x) & \text{für } 0 < x < x_0 \\ -\frac{1}{\tau} & \text{für } x_0 < x < x_1 \\ 0 & \text{für } x_1 < x \end{array} \right\} \geq -\frac{1}{\tau}. \end{aligned}$$

Damit erfüllen die g_τ die Differentialungleichung (4.16). Außerdem gilt für $\tau \rightarrow 0$: $\|g_\tau - g_0\|_\infty \rightarrow 0$. In Abbildung 4.14 sind einige g_τ zu sehen.

4.6.2 Iterationsverfahren

Somit bleibt noch zu klären, wie man das Minimierungsproblem (4.15) algorithmisch angeht. Hier hilft [Lud02, Satz 4.3] weiter. Er besagt, dass für $x \geq 0$

$$g_\tau(x) = \min_{w \in [0;1]} (xw + \Psi_\tau(w))$$

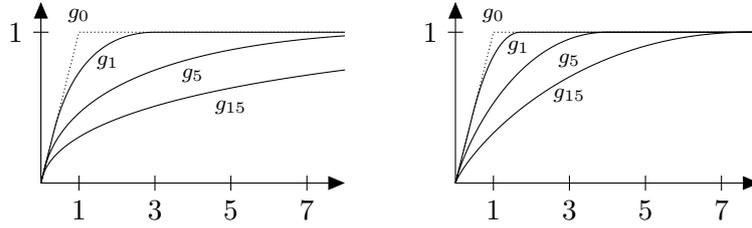


Abbildung 4.14: Funktionen g_τ ($\tau \in \{1, 5, 15\}$) für $\alpha = 1$ und $p = 2$ (links) und $p = 1.2$ (rechts)

gilt. Dabei ist $\Psi_\tau(w) := (-g_\tau)^*(-w)$ und $(-g_\tau)^*$ bezeichnet die Legendre-Fenchel Transformierte von $-g_\tau$. Das Minimum wird bei $w = g'_\tau(x)$ angenommen.

Damit ist

$$F_\tau(u) : \sum_{a \sim b} g_\tau[\alpha P(u_a - u_b)] + \beta \sum_a Q(u_a - f_a) = \min!$$

$$\Downarrow$$

$$\bar{F}_\tau(u, w) := \sum_{a \sim b} \alpha w_{a,b} |u_a - u_b|^p + \beta \sum_a |u_a - f_a|^q + \sum_{a \sim b} \Psi_\tau(w_{a,b}) = \min!_{u,v}$$

wobei $w_{a,b} \in [0; 1]$ für $a \sim b$ gilt. Damit kann man für festes τ und der Startiterierten u^0 wie folgt vorgehen: Zunächst wird \bar{F}_τ für festgehaltenes u^n minimiert:

$$w^{n+1} := \operatorname{argmin}_w \bar{F}_\tau(u^n, w). \quad (4.18)$$

Wie oben zitiert ist $w_{a,b}^{n+1} = g'_\tau(\alpha P(u_a^n - u_b^n))$. Danach wird \bar{F}_τ für festgehaltenes w^{n+1} minimiert:

$$u^{n+1} := \operatorname{argmin}_u \bar{F}_\tau(u, w^{n+1}). \quad (4.19)$$

Für festgehaltenes w^{n+1} handelt es sich hierbei um ein konvexes Minimierungsproblem, welches in den vorherigen Abschnitten behandelt wurde. Diese Iteration wird solange wiederholt bis das Minimum von \bar{F}_τ hinreichend genau ermittelt wurde. Anschließend kann τ verkleinert werden und das letzte Minimum als Startwert für die neue Iteration verwendet werden.

In Abbildung 4.15 sind die geschachtelten Iterationen des GNC-Algorithmus zu sehen.

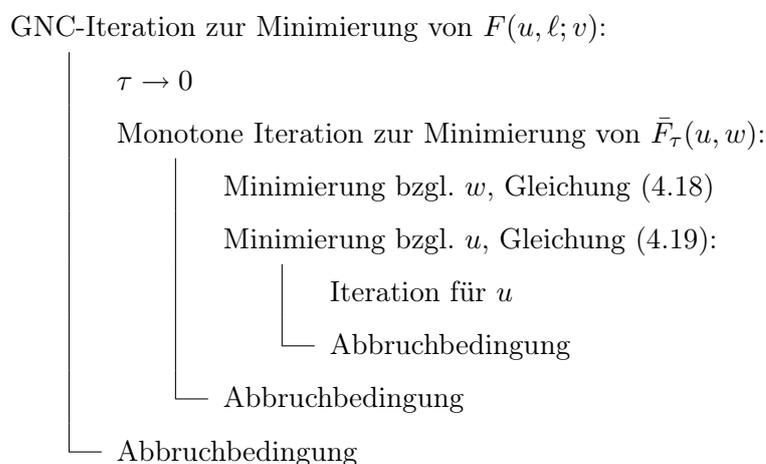


Abbildung 4.15: Überblick: Iterationsschleifen bei GNC-Algorithmus

4.7 Vergleich mit Mumford-Shah Funktional

In diesem abschließenden Abschnitt tritt der GNC-Algorithmus in Aktion. Es werden die beiden kontinuierlichen Funktionale

$$F^1(u) = \lambda^2 \int_{\Omega} \|\nabla u\|_2^2 dx + \mu \int_{J_u} \|\nu_u\|_1 d\mathcal{H}^1 + \|u - v\|_{L^2(\Omega)}^2$$

und

$$F^2(u) = \lambda^2 \int_{\Omega} \|\nabla u\|_{1,2}^{1,2} dx + \mu \int_{J_u} \|\nu_u\|_1 d\mathcal{H}^1 + \|u - v\|_{L^2(\Omega)}^2$$

betrachtet. Bei folgenden Beispielbildern wurde mit den nach Satz 2.28 zugehörigen diskreten Funktionalen der Form (4.13) gerechnet.

Betrachtet man für F^1 die Ergebnisse in Abbildung 4.16, so erkennt man, dass in großen Parameterbereichen zu viele Kanten entstehen. Dies ist insbesondere auch an Stellen der Fall, bei denen man keine Kante im Ergebnisbild erwartet. So entstehen Sprünge, die fast konstante Bereiche voneinander abgrenzen.

Eine Ursache dieser Probleme kann mit Hilfe der Darstellung von F_0 (siehe Gleichung (4.14)) erklärt werden. Betrachtet man die Interaktionsfunktion $s \mapsto g_0(s^2) = \min(1, s^2)$ und versucht man Kanten stärker zu bestrafen (siehe Abbildung 4.17), also z. B. mit $s \mapsto \min(m^2, s^2)$ (mit $m > 1$), so bemerkt man, dass man für $m \gg 1$ viel zu diffusive Ergebnisse erhält (der Fall $m = \infty$ ist in Abbildung 1.1 zu sehen), da s^2 für $m \gg 1$ zu schnell ansteigt.

Völlig anders verhält es sich im Fall $1 < p < 2$, etwa für $p = 1.2$. Wie schon in der Abbildung 1.2 zu sehen, hat man hier sogar für $m = \infty$ annehmbare

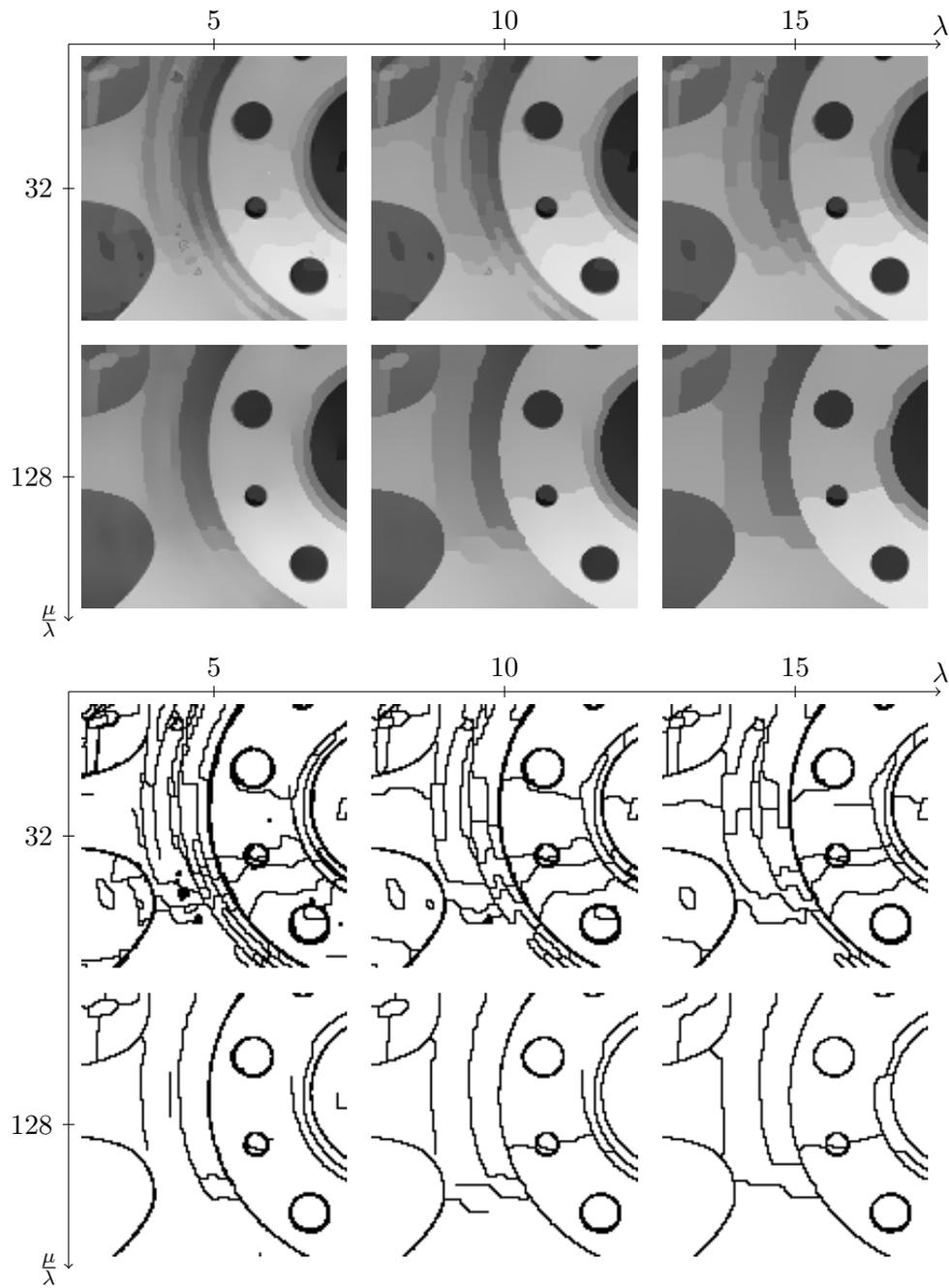
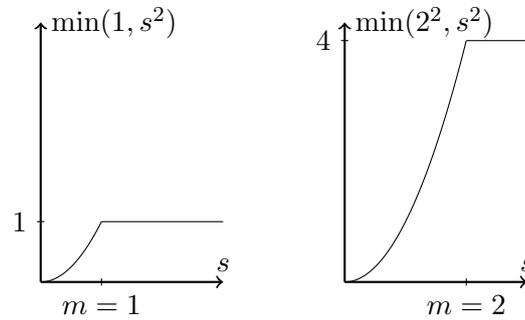
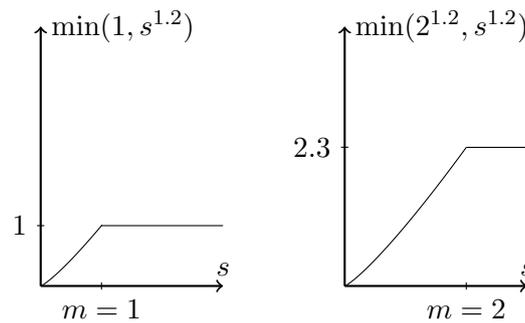


Abbildung 4.16: Approximierte Minima von $F^1(u)$ für verschiedene λ, μ -Kombinationen

Abbildung 4.17: Interaktionsfunktion bei Mumford-Shah ($p = 2$)

Ergebnisse. In dieser Situation kann man also $m \gg 1$ wählen und damit Kanten stärker bestrafen, ohne zu diffusive Anteile zu bekommen (siehe Abbildung 4.18).

Abbildung 4.18: Interaktionsfunktion für $p = 1.2$

In Abbildung 4.19 sind die Ergebnisbilder für F^2 , also für $p = 1.2$, zu sehen. Ein Vergleich mit den Ergebnissen für F^1 zeigt, dass für einen großen Parameterbereich die ungewollten „Zusatzkanten“ verschwunden sind.

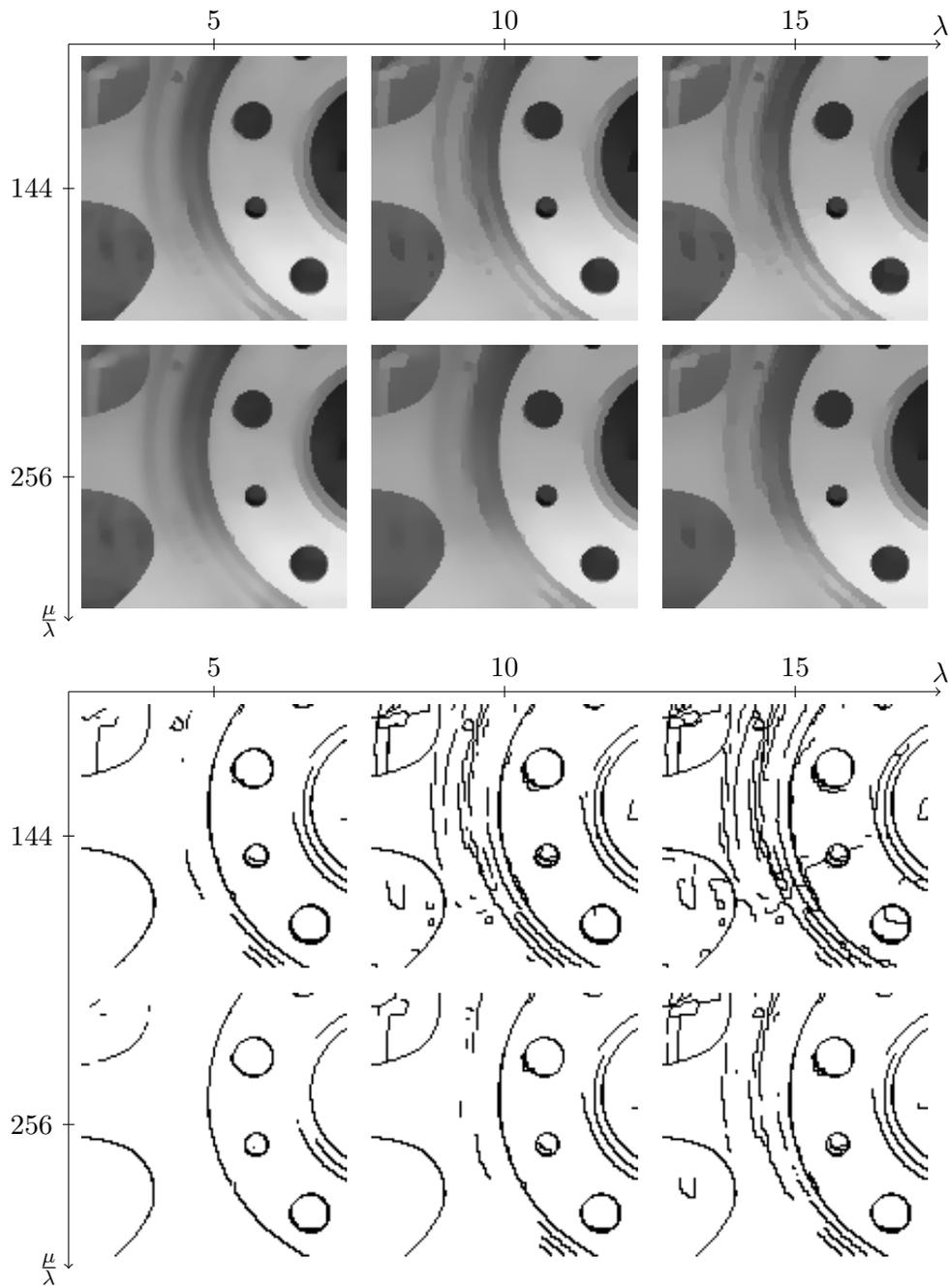


Abbildung 4.19: Approximierte Minima von $F^2(u)$ für verschiedene λ, μ -Kombinationen

Abbildungsverzeichnis

1.1	Optimallösung u für (1.3) mit $\lambda = 1, \frac{1}{10}, \frac{1}{20}$	3
1.2	Approximation an die Optimallösung u für (1.4) mit $\lambda = 1, \frac{1}{10}, \frac{1}{20}$	3
1.3	Drei Funktionen mit derselben Variation	4
1.4	Gängige Nachbarschaften auf quadratischen Gittern	6
2.1	Skizze für ein $u_h \in U_h$	19
2.2	Skizze zur Konstruktion von w_h mit $\mathcal{I}_h = \{1\}$	22
2.3	Vergleich von diskretem Optimum u_h und Optimallösung u für die Parameter aus (2.10) und bei der Einteilung in $n \in \{21, 101, 121\}$ Teilintervalle	26
2.4	Konvergenzverlauf	26
2.5	Asymptotische Anzahl von Gitterelementrändern, die in e_i^\perp Richtung geschnitten werden	30
2.6	Strecke mit am Gitter betrachteter Länge	30
2.7	Zeile von Gitterzellen in e_1 Richtung	32
2.8	Zellenmittelpunkte $x_{i,j}$	37
2.9	Bezeichnungen im Quadrat Z	38
2.10	Konstruktion w_h für $\lambda_i < 0$ für alle i	38
2.11	Konstruktion w_h für $\lambda_1, \lambda_3 < 0, \lambda_0, \lambda_2 \geq 0$	39
2.12	Konstruktion w_h für restliche Fälle	41
4.1	Minimum Tracking beim GNC-Algorithmus	60
4.2	Graphcut; Menge S blau, Menge T grün, Kanten mit Beitrag zu $\text{cut}(S, T)$ sind dicker	61
4.3	Kanten des zu (4.4) gehörenden Graphen nach [KZ04] für $a \sim b$	62
4.4	Energieverlauf im Laufe der iterativen Minimierung von (4.6) mit Expansion Moves	63
4.5	Eingabe- und Ausgabebild für Expansion Move Beispiel	64

4.6	Laufzeitverhalten für Expansion Moves	64
4.7	Kanten des zu (4.7) gehörenden Graphen nach [Vek99, Abschnitt 4.3.1] für $a \sim b \in N$	66
4.8	Laufzeitverhalten für Swap Moves	66
4.9	Typische Prior-Interaktionsfunktionen	69
4.10	Beispiel einer Gitterhierarchie (rot: Menge C , grün: Menge F)	71
4.11	V-Zyklus der Block-Update-Methode	73
4.12	Vergleich: Block-Update-Methode und Gauß-Seidel-Verfahren	74
4.13	Mit Block-Update-Methode approximierte Minima von (4.12) mit $p = 1.2$, $q = 1.8$ und verschiedene α, β -Kombinationen . .	75
4.14	Funktionen g_τ ($\tau \in \{1, 5, 15\}$) für $\alpha = 1$ und $p = 2$ (links) und $p = 1.2$ (rechts)	78
4.15	Überblick: Iterationsschleifen bei GNC-Algorithmus	79
4.16	Approximierte Minima von $F^1(u)$ für verschiedene λ, μ -Kombinationen	80
4.17	Interaktionsfunktion bei Mumford-Shah ($p = 2$)	81
4.18	Interaktionsfunktion für $p = 1.2$	81
4.19	Approximierte Minima von $F^2(u)$ für verschiedene λ, μ -Kombinationen	82

Symbol- und Notationsverzeichnis

Mengen

\mathbb{N}	natürliche Zahlen	
\mathbb{N}_0	natürliche Zahlen mit Null: $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$	
\mathbb{R}	reelle Zahlen	
$\overline{\mathbb{R}}$	reelle Zahlen zusammen mit $\pm\infty$	
$\mathcal{B}(\Omega)$	Borel- σ -Algebra auf Ω	S. 10
S_u	approximative Unstetigkeitsmenge von u	S. 16
J_u	approximative Sprungmenge von u	S. 16

Maße und Maßräume

\mathcal{L}^d	d -dimensionales Lebesgue-Maß im \mathbb{R}^d	S. 11
\mathcal{H}^k	k -dimensionales Hausdorff-Maß (im \mathbb{R}^d)	S. 11
$[\mathcal{M}_{\text{loc}}(\Omega)]^d$	Radon-Maße mit Werten in \mathbb{R}^d	S. 11
$[\mathcal{M}(\Omega)]^d$	endliche Radon-Maße mit Werten in \mathbb{R}^d	S. 11
$\mu \llcorner A$	Maß μ eingeschränkt auf A : $(\mu \llcorner A)(B) := \mu(A \cap B)$	

Konvergenzbegriffe

$x_n \xrightarrow{X} x$	starke Konv. in X : $\ x - x_n\ _X \rightarrow 0$	
$x_n \rightharpoonup_X x$	schwache Konv. in X : $x'(x_n) \rightarrow x'(x) \forall x' \in X'$	
$x'_n \xrightarrow{X'}^* x'$	schwach* Konv. in X' : $x'_n(x) \rightarrow x'(x) \forall x \in X$	
$\text{ap-lim}_{y \rightarrow x} u(y)$	approximativer Grenzwert	S. 16

$\Gamma\text{-}\liminf_{j \rightarrow \infty} f_j$	untere Γ -Limes der f_j	S. 20
$\Gamma\text{-}\limsup_{j \rightarrow \infty} f_j$	obere Γ -Limes der f_j	S. 20
$\Gamma\text{-}\lim_{j \rightarrow \infty} f_j$	Γ -Limes der f_j	S. 20

Funktionsräume

X'	stetige, lineare Funktionale auf X ; Dualraum von X	
C	stetige Funktionen	
C_c	stetige Funktionen mit kompaktem Träger	
BV	Funktionen mit beschränkter Variation	S. 13
SBV	spezielle Funktionen beschränkter Variation	S. 17
BV^*	stückweise konstante Funktionen	S. 48
$W^{m,p}$	m -mal schwach differenzierbare Sobolev Funktionen mit Ableitungen in L^p	S. 12
SBV^p	SBV -Funktionen u mit $\nabla u \in L^p$	S. 31
\mathcal{W}	Teilraum von SBV nach [CT99]	S. 31

Literaturverzeichnis

- [AF03] ADAMS, R. und FOURNIER, J.: *Sobolev Spaces*. Elsevier Science, 2003. (zitiert auf Seiten: 3, 12, 13)
- [AFP00] AMBROSIO, L., FUSCO, N. und PALLARA, D.: *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford University Press, 2000. (zitiert auf Seiten: 5, 10, 12, 14, 21, 23, 42, 45, 46, 48, 50, 51)
- [Alt06] ALT, H.W.: *Lineare Funktionalanalysis*. Springer Verlag, 2006. (zitiert auf Seite: 13)
- [Amb95] AMBROSIO, L.: *A new proof of the SBV compactness theorem*. Calc. Var., 3(1):127–137, 1995. (zitiert auf Seite: 23)
- [AT90] AMBROSIO, L. und TORTORELLI, V.M.: *Approximation of functionals depending on jumps by elliptic functionals via Γ -convergence*. Commun. Pure Appl. Math, 43(8):999–1036, 1990. (zitiert auf Seite: 7)
- [Bau92] BAUER, H.: *Maß- und Integrationstheorie*. de Gruyter-Verlag, 1992. (zitiert auf Seite: 10)
- [Bel01] BELL, W.: *Graphcut*. <https://www.cs.cornell.edu/vision/wbell/implementation.html#graphcut>, 2001. (zitiert auf Seite: 61)
- [Bon96] BONNET, A.: *On the regularity of edges in image segmentation*. Ann. Inst. H. Poincaré Anal. Non Linéaire, 13(4):485–528, 1996. (zitiert auf Seite: 5)
- [Bra02] BRAIDES, A.: *Γ -convergence for beginners*. Oxford University Press, 2002. (zitiert auf Seiten: 18, 19, 20)
- [BS93] BOUMAN, C. und SAUER, K.: *A Generalized Gaussian Image Model for Edge-Preserving MAP Estimation*. IEEE Trans. on Image Processing, 2(3):296–310, 1993. (zitiert auf Seite: 59)

- [BVZ99] BOYKOV, Y., VEKSLER, O. und ZABIH, R.: *Fast Approximate Energy Minimization via Graph Cuts*. ICCV, 1:377–384, 1999. (zitiert auf Seite: 61)
- [BZ87] BLAKE, A. und ZISSERMAN, A.: *Visual Reconstruction*. MIT Press, 1987. (zitiert auf Seiten: 25, 59, 69)
- [Cha95] CHAMBOLLE, A.: *Image Segmentation by Variational Methods: Mumford and Shah Functional and the Discrete Approximations*. SIAM Journal on Applied Mathematics, 55(3):827–863, 1995. (zitiert auf Seite: 28)
- [Cha04] CHAMBOLLE, A.: *An Algorithm for Total Variation Minimization and Applications*. Journal of Mathematical Imaging and Vision, 20(1-2):89–97, 2004. (zitiert auf Seite: 60)
- [CL97] CHAMBOLLE, A. und LIONS, P.-L.: *Image recovery via total variation minimization and related problems*. Numer. Math., 76(2):167–188, 1997. (zitiert auf Seite: 4)
- [CT99] CORTESANI, G. und TOADER, R.: *A density result in SBV with respect to non-isotropic energies*. Nonlinear Anal., 38:585–604, 1999. (zitiert auf Seiten: 31, 86)
- [Dac89] DACOROGNA, B.: *Direct Methods in the Calculus of Variations*. Applied Mathematical Sciences 78. Springer Verlag, 1989. (zitiert auf Seiten: 49, 52, 53)
- [Dav05] DAVID, G.: *Singular Sets of Minimizers for the Mumford-Shah Functional*. Birkhäuser Verlag, 2005. (zitiert auf Seite: 5)
- [Din74] DINCULEANU, N.: *Integration on Locally Compact Spaces*. Springer Verlag, 1974. (zitiert auf Seite: 11)
- [DM93] DAL MASO, G.: *An Introduction to Γ -Convergence*. Birkhäuser, 1993. (zitiert auf Seite: 20)
- [EG92] EVANS, L. und GARIEPY, R.: *Measure Theory and Fine Properties of Functions*. CRC Press, 1992. (zitiert auf Seite: 24)
- [Els05] ELSTRODT, J.: *Maß- und Integrationstheorie*. Springer-Verlag, 2005. (zitiert auf Seiten: 10, 17)
- [FR60] FLEMING, W. H. und RISHEL, R.: *An integral formula for total gradient variation*. Arch. Math., 11:218–222, 1960. (zitiert auf Seite: 15)

- [GG84] GEMAN, S. und GEMAN, D.: *Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images*. IEEE Trans. Patt. Anal. Machine Intell., Vol. PAMI-6:721–741, 1984. (zitiert auf Seiten: 25, 59, 68)
- [Hub81] HUBER, P.: *Robust statistics*. John Wiley & Sons, 1981. (zitiert auf Seite: 68)
- [KZ04] KOLMOGOROV, V. und ZABIH, R.: *What Energy Functions Can Be Minimized via Graph Cuts?* IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(2):147–159, 2004. (zitiert auf Seiten: 62, 83)
- [Lud02] LUDWIG, C.: *Mehrgitterbasierte Beschleunigung von GNC-Verfahren in der Bildsegmentierung*. Diplomarbeit, 2002. (zitiert auf Seiten: 29, 58, 59, 75, 77)
- [Mar85] MARCELLINI, P.: *Approximation of quasiconvex functions, and lower semicontinuity of multiple integrals*. manuscripta math., 51:1–28, 1985. (zitiert auf Seite: 52)
- [Mor52] MORREY, C. B.: *Quasi-convexity and the lower semicontinuity of multiple integrals*. Pac. J. Math., 2:25–53, 1952. (zitiert auf Seite: 52)
- [MS89] MUMFORD, D. und SHAH, J.: *Optimal Approximation by Piecewise Smooth Functions and Associated Variational Problems*. Comm. Pure Appl. Math., 42:577–685, 1989. (zitiert auf Seite: 5)
- [Neg99] NEGRI, M.: *The Anisotropy Introduced by the Mesh in the Finite Element Approximation of the Mumford-Shah Functional*. Numer. Funct. Anal. Optim., 20:957–982, 1999. (zitiert auf Seiten: 7, 28, 29)
- [Oeh02] OEHSEN, M. VON: *Multiscale Methods for Variational Image Denoising*. Logos Verlag, 2002. (zitiert auf Seite: 3)
- [ROF92] RUDIN, L., OSHER, S. und FATEMI, E.: *Nonlinear total variation based noise removal algorithms*. Physica D, 60(1-4):259–268, 1992. (zitiert auf Seite: 4)
- [Stü99] STÜBEN, K.: *Algebraic Multigrid (AMG): An Introduction with Applications*. Technischer Bericht 53, GMD, 1999. (zitiert auf Seite: 70)
- [Vek99] VEKSLER, O.: *Efficient Graph-Based Energy Minimization Methods in Computer Vision*. Doktorarbeit, 1999. (zitiert auf Seiten: 63, 65, 66, 84)

- [Win03] WINKLER, G.: *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods*. Springer Verlag, 2003. (zitiert auf Seite: 67)
- [Zie89] ZIEMER, W.: *Weakly Differentiable Functions*. Springer Verlag, 1989. (zitiert auf Seite: 4)