

HUMBOLDT-UNIVERSITÄT ZU BERLIN  
MATHEMATISCH-NATURWISSENSCHAFTLICHE FAKULTÄT  
INSTITUT FÜR MATHEMATIK



# Die Crouzeix-Raviart-Finite-Elemente-Methode für eine nichtkonforme Formulierung des Rudin-Osher-Fatemi-Modellproblems

Bachelorarbeit

zur Erlangung des akademischen Grades  
Bachelor of Science (B. Sc.)

eingereicht von: Enrico Bergmann  
geboren am: 13.10.1995  
geboren in: Berlin  
Gutachter/innen: Prof. Dr. Carsten Carstensen  
Dr. Philipp Bringmann

Eingereicht am Institut für Mathematik der Humboldt-Universität zu Berlin am:

# Inhaltsverzeichnis

<b>1 Einleitung</b>	<b>3</b>
<b>2 Theoretische Grundlagen</b>	<b>6</b>
2.1 Notation . . . . .	6
2.2 Crouzeix-Raviart-Finite-Elemente-Funktionen . . . . .	7
2.3 Variationsrechnung auf Banachräumen . . . . .	8
2.4 Subdifferentiale . . . . .	9
2.5 Funktionen beschränkter Variation . . . . .	10
<b>3 Das kontinuierliche Problem</b>	<b>12</b>
3.1 Existenz eines eindeutigen Minimierers . . . . .	12
3.2 Konstruktion eines Eingangssignals zu einer gegebenen Lösung . . . . .	16
<b>4 Das diskrete Problem</b>	<b>18</b>
4.1 Formulierung . . . . .	18
4.2 Charakterisierung und Existenz eines eindeutigen Minimierers . . . . .	18
4.3 Verfeinerungsindikator und garantierter Energieschranken . . . . .	23
<b>5 Iterative Lösung</b>	<b>24</b>
5.1 Primale-duale Iteration . . . . .	24
5.2 Konvergenz der Iteration . . . . .	25
<b>6 Implementierung</b>	<b>29</b>
6.1 Hinweise zur Benutzung des Programms . . . . .	29
6.2 Programmablauf . . . . .	30
6.3 Realisierung der primalen-dualen Iteration . . . . .	36
6.4 Mathematische Grundlagen ausgewählter Methoden . . . . .	37
6.4.1 Berechnung lokaler Crouzeix-Raviart-Basisfunktionen . . . . .	37
6.4.2 Berechnung lokaler Knotenwerte einer Crouzeix-Raviart-Funktion . . . . .	37
6.4.3 Berechnung von Sprungtermen . . . . .	38
<b>7 Numerische Beispiele</b>	<b>39</b>
7.1 Wahl der Parameter für die primale-duale Iteration . . . . .	44
7.2 Eigenschaften der primalen-dualen Iteration . . . . .	47
7.3 Experimente mit bekannter exakter Lösung . . . . .	49
7.4 Graufarbenbilder als Eingangssignale . . . . .	57
7.5 Stetige Approximation eines unstetigen Eingangssignals . . . . .	57
7.6 Fazit und Ausblick . . . . .	62
<b>Literatur</b>	<b>67</b>

# 1 Einleitung

In der Bildverarbeitung kann ein gegebenes Signal häufig nur durch eine unstetige Funktion dargestellt werden. Deshalb stellt sich zunächst die Frage, welcher Funktionenraum zum Beschreiben dieser Signale geeignet ist.

Sei  $\Omega \subset \mathbb{R}^2$  ein polygonal berandetes Lipschitz-Gebiet und  $g : \Omega \rightarrow \mathbb{R}$  stelle ein gegebenes Signal auf  $\Omega$  dar. Das Signal  $g$  könnte im Sobolev-Raum  $W^{1,1}(\Omega)$  vermutet werden, da Elemente dieses Raums im Allgemeinen nicht stetig sein müssen. Allerdings lassen Sobolev-Funktionen die oftmals benötigten Sprünge über Teilmengen niedrigerer Dimension von  $\Omega$  nicht zu. Dieses Problem kann gelöst werden, indem der Raum der Funktionen von beschränkter Variation  $BV(\Omega)$  betrachtet wird. Dieser ist eine echte Obermenge von  $W^{1,1}(\Omega)$  und hat sich als geeignet für die Modellierung von Signalen in der Bildverarbeitung und weitere Anwendungen erwiesen (cf. [ABM14, S. 393; AK06, S. 42; Bar15b, S. 297; Bra98, S. 1 f.]).

Eine mögliche Problemstellung in der Bildverarbeitung ist die Rauschunterdrückung, das heißt, der Versuch unerwünschtes Rauschen in einem Signal zu verringern. In [ROF92] beschrieben Rudin, Osher und Fatemi 1992 das heute als ROF-Modell bekannte Minimierungsproblem dafür (cf. [Bar15a, S. 1217; CP10, S. 132; Get12, S. 74 f.]). Dabei ist für das gegebene Signal  $g \in L^2(\Omega)$  und eine Funktion  $v \in BV(\Omega) \cap L^2(\Omega)$  die Minimierung der Summe der zwei folgenden Terme relevant. Der erste Term ist die Seminorm

$$|v|_{BV(\Omega)} := \sup_{\substack{\phi \in C_0^\infty(\Omega; \mathbb{R}^2) \\ |\phi| \leq 1}} \int_\Omega v \operatorname{div}(\phi) dx < \infty.$$

von  $v$  auf  $BV(\Omega)$  [Bar12, S. 1162]. Diese entspricht der totalen Variation der distributiven Ableitung  $Dv$  von  $v$  und ihre Minimierung vermindert Oszillationen in der Lösung, lässt aber Unstetigkeiten zu [Get12, S. 75]. Außerdem stimmt diese, falls  $v \in W^{1,1}(\Omega)$ , mit der Seminorm auf  $W^{1,1}(\Omega)$  überein. Der zweite Term ist die Normdifferenz von  $v$  und  $g$  in  $L^2(\Omega)$ . Die Minimierung dieses Terms bewirkt, dass die Lösung dem Eingangssignal ähnelt. Mit diesen Termen und mit einem positiven Parameter  $\alpha \in \mathbb{R}_+$ , der das Verhältnis zwischen Rauschverminderung und Ähnlichkeit der Lösung zum Eingangssignal gewichtet, sucht das ROF-Modell eine Funktion  $u \in BV(\Omega) \cap L^2(\Omega)$ , die das Funktional

$$I(v) := |v|_{BV(\Omega)} + \frac{\alpha}{2} \|v - g\|_{L^2(\Omega)}^2 \quad (1.1)$$

unter allen  $v \in BV(\Omega) \cap L^2(\Omega)$  minimiert. Wird hierbei  $\alpha$  zu klein gewählt, führt das zu einer zu stark geglätteten, verwaschen aussehenden Lösung, zu sehen zum Beispiel in den Abbildungen 1.1c und 1.1d. Wird andererseits  $\alpha$  zu groß gewählt, ist die Verminderung des Rauschens im Vergleich zum Eingangssignal  $g$  nur gering, zu sehen zum Beispiel in den Abbildungen 1.1g und 1.1h. Für weitere Details und Referenzen zur Rauschunterdrückung und zur Wahl von  $\alpha$  siehe [Get12].

Zur numerischen Behandlung dieses Problems gibt es bereits einige Ansätze in der Literatur. Dazu gehören die Regularisierung der Seminorm  $|\cdot|_{BV(\Omega)}$ , indem die Betragsfunktion  $|\cdot|$  durch eine stetig differenzierbare Approximation  $|\cdot|_\varepsilon$  ersetzt wird, und die Nutzung von höheren Ableitungen in der Definition von  $|\cdot|_{BV(\Omega)}$ . Vor- und Nachteile dieser Ansätze und entsprechende Referenzen werden in [Bar12, S. 1165] zusammengefasst. Außerdem wird ebenda auf Arbeiten verwiesen, in denen verschiedene iterative Lösungsmethoden

## 1 Einleitung

für das ROF-Modellproblem diskutiert werden. Professor Bartels untersucht in [Bar15b, Section 10.2] eine  $W^{1,1}$ -konforme Diskretisierung des ROF-Modells mit Courant-Finite-Elemente-Funktionen. Zur numerischen Lösung dieser diskreten Formulierung nutzt er eine primale-duale Iteration, welche durch Betrachtung der primalen und der dualen Formulierung des Minimierungsproblems motiviert ist. Eine Regularisierung oder die Nutzung höherer Ableitungen für die BV-Seminorm werden dabei nicht benötigt.

In dieser Arbeit möchten wir die Anwendung dieser primalen-dualen Iteration auf eine nichtkonforme, mit Crouzeix-Raviart-Finite-Elemente-Funktionen diskretisierten Formulierung des ROF-Modells untersuchen. Dabei nutzen wir einen, von Professor Carstensen zur Verfügung gestellten, Verfeinerungsindikator, um die Iteration im Solve-Schritt der AFEM-Routine aus Abbildung 6.1 nutzen zu können. Außerdem erlaubt uns die nichtkonforme Formulierung die Betrachtung einer garantierten unteren Energieschranke, welche ebenfalls von Professor Carstensen zur Verfügung gestellt wurde. Die Implementierung des adaptiven Algorithmus basiert auf dem Matlab-Softwarepaket AFEM [Car09a].

Abschließend sei angemerkt, dass wir folgende, leicht andere Formulierung des ROF-Modells betrachten. Wir minimieren das Funktional

$$E(v) := \frac{\alpha}{2} \|v\|_{L^2(\Omega)}^2 + |v|_{BV(\Omega)} + \|v\|_{L^1(\partial\Omega)} - \int_{\Omega} fv \, dx$$

unter allen  $v \in BV(\Omega) \cap L^2(\Omega)$ . Dabei ist der Term  $\|v\|_{L^1(\partial\Omega)}$  durch den Spursatz für BV-Funktionen [ABM14, S. 400, Theorem 10.2.1] wohldefiniert und seine Minimierung modelliert homogene Randdaten. Für  $f = \alpha g$  gilt dann

$$I(v) = E(v) - \|v\|_{L^1(\partial\Omega)} + \frac{\alpha}{2} \|g\|_{L^2(\Omega)}^2 \quad \text{für alle } v \in BV(\Omega) \cap L^2(\Omega).$$

Aufgrund der Konstanz von  $\|g\|_{L^2(\Omega)}$  folgt damit, dass die Funktionale  $E$  und  $I$  die gleichen Minimierer in  $\{v \in BV(\Omega) \cap L^2(\Omega) \mid \|v\|_{L^1(\partial\Omega)} = 0\}$  besitzen.

Die Struktur dieser Arbeit ist wie folgt. Nachdem in Kapitel 2 zunächst die Notation eingeführt und die theoretischen Grundlagen aus der Optimierung und zu den Funktionen beschränkter Variation zusammengetragen wurden, wird in Kapitel 3 bewiesen, dass für unsere Formulierung des ROF-Modells ein eindeutiger Minimierer existiert. Anschließend folgt in Kapitel 4 die nichtkonforme Formulierung und Diskretisierung des Minimierungsproblems. Mithilfe der Sattelpunktsformulierung des diskreten Problems werden äquivalente Charakterisierungen für den eindeutigen diskreten Minimierer bewiesen. Außerdem werden der Verfeinerungsindikator und garantierte Schranken für  $E$  aufgeführt. In Kapitel 5 wird die primale-duale Iteration formuliert und bewiesen, dass diese gegen den diskreten Minimierer konvergiert. Es folgen in Kapitel 6 Hinweise zur Benutzung des Programms und Details zur Implementierung des Algorithmus und schließlich in Kapitel 7 die Darstellung der Experimente und deren Auswertung.



Abbildung 1.1: Zu sehen sind das Originalbild<sup>1</sup> (a) und das Originalbild mit additivem weißen gaußschen Rauschen (englisch „additive white Gaussian noise“, AWGN) mit einem Signal-Rausch-Verhältnis (englisch „signal-to-noise ratio“, SNR) von 15 (b), jeweils mit einem nachträglich hinzugefügten graduellen Übergang zu schwarzem Rand, der bei Graufarbenbildern als Eingangssignal Nullranddaten entspricht. Außerdem sind sechs Ergebnisse (c)–(h) des in Kapitel 6 beschriebenen adaptiven Algorithmus mit Eingangssignal (b) und verschiedenen Werten von  $\alpha$  abgebildet.

---

<sup>1</sup>  <https://homepages.cae.wisc.edu/~ece533/images/cameraman.tif>

## 2 Theoretische Grundlagen

In dieser Arbeit werden Grundbegriffe in topologischen Räumen und Kenntnisse zu Banach- und Hilberträumen sowie zu Lebesgue- und Sobolev-Räumen vorausgesetzt. Dazu gehören insbesondere wichtige Ungleichungen (beispielsweise Cauchy-Schwarz, Hölder, Young), grundlegende Einbettungssätze, Dualraumtheorie, Aussagen zur schwachen Konvergenz sowie der Rieszsche Darstellungssatz und seine Implikationen. Benötigte topologische Begriffe und grundlegende Aussagen zu Banach- und Hilberträumen können beispielsweise in [Zei86] nachgeschlagen werden. Grundlagen der Optimierung sind in [Zei85] nachlesbar und alles Weitere eben genannte in [Zei90a] und [Zei90b]. Dabei eignen sich die Register von [Zei90b] und [Zei85] hervorragend zum schnellen Auffinden von Begriffen in allen eben aufgeführten Referenzen.

Schließlich sollte ein Verständnis des AFEM-Softwarepaketes [Car09a], von den darin verwendeten Datenstrukturen und von den mathematischen Konzepten hinter dessen Funktionen vorhanden sein. Dazu sei an dieser Stelle auf die Dokumentation [Car+10] dieses Softwarepaketes verwiesen. Die in dieser Arbeit benötigten Begriffe und Notationen führen wir im nächsten Abschnitt nochmals ein, wobei wir den Definitionen und Aussagen zum Crouzeix-Raviart-Finite-Elemente-Raum einen separaten Abschnitt 2.2 widmen.

### 2.1 Notation

Wir wählen für die natürlichen Zahlen die Konvention  $\mathbb{N} = \{1, 2, 3, \dots\}$  und  $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ . Die Menge der positiven reellen Zahlen notieren wir mit  $\mathbb{R}_+$ . Um auszudrücken, dass eine Menge  $A$  Teilmenge einer Menge  $B$  ist, schreiben wir  $A \subseteq B$ . Falls wir hervorheben wollen, dass  $A$  sogar eine echte Teilmenge von  $B$  ist, so schreiben wir  $A \subset B$ . Ist  $A$  Teilmenge eines topologischen Raumes, so notieren wir den Rand von  $A$  mit  $\partial A$ , das Innere von  $A$  mit  $\text{int}(A)$  und den Abschluss von  $A$  mit  $\overline{A}$ . Wir nennen weiterhin eine Teilmenge  $B \subseteq A$  Umgebung eines Punktes  $x \in A$ , wenn es eine offene Teilmenge von  $B$  gibt, die  $x$  enthält. Ist eine Funktion  $F : X \rightarrow Y$  zwischen nichtleeren Mengen  $X$  und  $Y$  konstant, das heißt, nimmt sie nur genau einen Wert  $y \in Y$  an, schreiben wir  $F \equiv y$ .

Für den Rest dieses Abschnitts seien  $d, m \in \mathbb{N}$ ,  $k \in \mathbb{N}_0$ ,  $p \in [1, \infty]$  und  $U$  eine nichtleere, offene Teilmenge von  $\mathbb{R}^d$ . Die Einheitsmatrix in  $\mathbb{R}^{m \times m}$  bezeichnen wir mit  $I_m$ . Für das euklidische Skalarprodukt zweier Vektoren  $v, w \in \mathbb{R}^m$  schreiben wir  $v \cdot w$ . Betrachten wir einen Funktionenraum mit Werten in  $\mathbb{R}$ , so verzichten wir auf  $\mathbb{R}$  beim Notieren des Funktionenraums. Zum Beispiel schreiben wir  $C(U) := C(U; \mathbb{R})$  für den Raum der stetigen Funktionen von  $U$  nach  $\mathbb{R}$ .

Mit  $\Omega \subset \mathbb{R}^2$  bezeichnen wir stets ein polygonal berandetes Lipschitz-Gebiet. Dazu sei  $\mathcal{T}$  eine reguläre Triangulierung von  $\Omega$  im Sinne von Ciarlet (cf. [Car09b, S. 34; CGR12, S. 345; Car+10, S. 8 f.]), das heißt,  $\mathcal{T}$  sei eine endliche Menge von abgeschlossenen Dreiecken  $T$  mit positivem Flächeninhalt  $|T|$ , sodass

$$\bigcup_{T \in \mathcal{T}} T = \overline{\Omega}$$

und zwei Dreiecke  $T_1, T_2 \in \mathcal{T}$  mit  $T_1 \neq T_2$  und  $T_1 \cap T_2 \neq \emptyset$  genau eine gemeinsame Ecke oder genau eine gemeinsame Kante haben. Die Menge der Knoten der Triangulierung sei  $\mathcal{N}$ , wobei die Menge der inneren Knoten mit  $\mathcal{N}(\Omega)$ , die Menge der Randknoten mit

$\mathcal{N}(\partial\Omega)$  und die Menge der Knoten eines Dreiecks  $T \in \mathcal{T}$  mit  $\mathcal{N}(T)$  bezeichnet werde. Für die Kanten der Triangulierung seien die Mengen  $\mathcal{E}$ ,  $\mathcal{E}(\Omega)$ ,  $\mathcal{E}(\partial\Omega)$  und  $\mathcal{E}(T)$  analog definiert. Außerdem definieren wir für  $z \in \mathcal{N}$  die Menge  $\mathcal{T}(z) := \{T \in \mathcal{T} \mid z \in \mathcal{N}(T)\}$  und für  $E \in \mathcal{E}$  die Menge  $\mathcal{T}(E) := \{T \in \mathcal{T} \mid E \in \mathcal{E}(T)\}$ . Den Mittelpunkt einer Kante  $E \in \mathcal{E}$  bezeichnen wir mit  $\text{mid}(E)$ . Der Normaleneinheitsvektor auf dem Rand eines Dreiecks  $T \in \mathcal{T}$  sei  $\nu_T$  und der Normaleneinheitsvektor auf einer Kante  $E \in \mathcal{E}$  sei  $\nu_E$ . Für eine Innenkante  $E \in \mathcal{E}(\Omega)$  bezeichnen wir dann die beiden Dreiecke in  $\mathcal{T}(E)$  so mit  $T_+$  und  $T_-$ , dass  $\nu_{T_+}$  und  $\nu_E$  gleich orientiert sind, also  $\nu_{T_+} \cdot \nu_E = 1$ . Damit können wir den Sprung entlang einer Innenkante  $E \in \mathcal{E}(\Omega)$  definieren als  $[\cdot]_E := \cdot|_{T_+} - \cdot|_{T_-}$ . Für eine Randkante  $E \in \mathcal{E}(\partial\Omega)$  definieren wir  $[\cdot]_E := \cdot|_E$ . Die Menge der stückweise konstanten Funktionen auf der Triangulierung  $\mathcal{T}$  mit Werten in  $\mathbb{R}^m$  notieren wir dann mit  $P_0(\mathcal{T}, \mathbb{R}^m)$  und die Menge der stückweise affinen Funktionen auf  $\mathcal{T}$  mit Werten in  $\mathbb{R}^m$  mit  $P_1(\mathcal{T}, \mathbb{R}^m)$ . Weiterhin ist der Courant-Finite-Elemente-Raum (cf. [Car+10, S. 12]) definiert als

$$S^1(\mathcal{T}) := P_1(\mathcal{T}) \cap C(\overline{\Omega}).$$

Für ein Dreieck  $T \in \mathcal{T}$  sei die Länge der längsten Seite  $h_T$ . Damit können wir die stückweise konstante Funktion  $h_T \in P_0(\mathcal{T})$  für alle  $T \in \mathcal{T}$  durch  $h_T|_T := h_T$  und die Länge der längsten Seite der Triangulierung durch  $h := \max_{T \in \mathcal{T}} h_T$  definieren.

Mit  $|\cdot|$  bezeichnen wir, je nach Argument, die euklidischen Norm eines Vektors  $v \in \mathbb{R}^m$ , das Lebesgue-Maß einer Menge  $M \subset \mathbb{R}^2$ , die Länge einer Kante  $E \in \mathcal{E}$  oder die Kardinalität einer endlichen Menge  $A$ .

Ist  $V$  ein Vektorraum, so notieren wir die konvexe Hülle einer Teilmenge  $X \subseteq V$  mit  $\text{conv } X$ . Falls  $V$  ein normierter Vektorraum ist, dann bezeichnen wir die entsprechende Norm auf  $V$  mit  $\|\cdot\|_V$ . Die Einheitskugel auf  $V$  ist damit  $B_V := \{v \in V \mid \|v\|_V < 1\}$  und die abgeschlossene Einheitskugel auf  $V$  ist  $\overline{B_V} = \{v \in V \mid \|v\|_V \leq 1\}$ . Ist  $V$  sogar ein Prähilbertraum, so bezeichnen wir das Skalarprodukt auf  $V$ , welches  $\|\cdot\|_V$  induziert, mit  $(\cdot, \cdot)_V$ . Betrachten wir eine Eigenschaft, für die gegeben sein muss, bezüglich welcher Norm Folgen auf  $V$  konvergieren, so ist stets die Konvergenz in der Norm  $\|\cdot\|_V$  gewählt, sofern nicht anders angegeben. Beispielsweise meinen wir mit der Folgenstetigkeit eines Funktionalen  $F : V \rightarrow \mathbb{R}$  im Detail die Folgenstetigkeit mit der Normkonvergenz bezüglich  $\|\cdot\|_V$  in  $V$  und der Konvergenz bezüglich  $|\cdot|$  in  $\mathbb{R}$ .

Die Signumfunktion auf dem  $\mathbb{R}^m$  definieren wir für einen Vektor  $v \in \mathbb{R}^m$  durch

$$\text{sign}(v) := \begin{cases} \left\{ \frac{v}{\|v\|} \right\}, & \text{falls } v \in \mathbb{R}^m \setminus \{0\}, \\ \overline{B_{\mathbb{R}^m}}, & \text{sonst.} \end{cases} \quad (2.1)$$

Für den Dualraum eines Banachraums  $X$  über  $\mathbb{R}$  schreiben wir  $X^*$ . Die Auswertung eines Funktionalen  $F \in X^*$  an der Stelle  $u \in X$  notieren wir, vor eventueller Anwendung des Rieszschen Darstellungssatzes, mit  $\langle F, u \rangle$ . Identifizieren wir einen Raum  $Y$  mit dem Dualraum  $X^*$ , so schreiben wir  $Y \cong X^*$ .

Weiterhin benutzen wir die übliche Notation für Lebesgue-Räume  $L^p(U; \mathbb{R}^m)$  und die Sobolev-Räume  $W^{k,p}(U)$  sowie  $H^k(U) := W^{k,2}(U)$  und  $H_0^k(U)$ . Die Normen auf diesen Räumen definieren wir ebenfalls wie üblich. Außerdem schreiben wir kurz  $\|\cdot\| := \|\cdot\|_{L^2(\Omega)}$  und  $(\cdot, \cdot) := (\cdot, \cdot)_{L^2(\Omega)}$ . Mit  $H^1(\mathcal{T})$  bezeichnen wir den Raum der stückweisen  $H^1$ -Funktionen auf  $\mathcal{T}$ .

## 2.2 Crouzeix-Raviart-Finite-Elemente-Funktionen

In Abschnitt 4.1 wollen wir für eine nichtkonforme Diskretisierung des Minimierungsproblems 3.1 den Crouzeix-Raviart-Finite-Elemente-Raum nutzen. Die dafür benötigten

## 2 Theoretische Grundlagen

Definitionen und Aussagen tragen wir in diesem Abschnitt aus [Car09b; CGR12; Car+10] zusammen.

Die Räume der Crouzeix-Raviart-Finite-Elemente-Funktionen sind definiert durch

$$\begin{aligned}\text{CR}^1(\mathcal{T}) &:= \{v_{\text{CR}} \in P_1(\mathcal{T}) \mid v_{\text{CR}} \text{ ist stetig im } \text{mid}(E) \text{ für alle } E \in \mathcal{E}\} \quad \text{und} \\ \text{CR}_0^1(\mathcal{T}) &:= \{v_{\text{CR}} \in \text{CR}^1(\mathcal{T}) \mid \forall E \in \mathcal{E}(\partial\Omega) \quad v_{\text{CR}}(\text{mid}(E)) = 0\}.\end{aligned}$$

Wir wollen  $\text{CR}_0^1(\mathcal{T})$  mit einem diskreten Skalarprodukt ausstatten. Dafür führen wir den stückweisen Gradienten  $\nabla_{\text{NC}} : H^1(\mathcal{T}) \rightarrow L^2(\Omega; \mathbb{R}^2)$  ein. Dieser ist für alle  $v \in H^1(\mathcal{T})$  und alle  $T \in \mathcal{T}$  definiert durch  $(\nabla_{\text{NC}}v)|_T := \nabla v|_T$ . Das ermöglicht uns die Definition der diskreten Bilinearform  $a_{\text{NC}}(\cdot, \cdot) : \text{CR}_0^1(\mathcal{T}) \times \text{CR}_0^1(\mathcal{T}) \rightarrow \mathbb{R}$  durch

$$a_{\text{NC}}(u_{\text{CR}}, v_{\text{CR}}) := \int_{\Omega} \nabla_{\text{NC}} u_{\text{CR}} \cdot \nabla_{\text{NC}} v_{\text{CR}} \, dx \quad \text{für alle } u_{\text{CR}}, v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T}).$$

Diese ist ein Skalarprodukt auf  $\text{CR}_0^1(\mathcal{T})$ . Die von  $a_{\text{NC}}$  induzierte Norm bezeichnen wir mit  $\|\cdot\|_{\text{NC}}$ .

Nun definieren wir für jede Kante  $E \in \mathcal{E}$  eine Funktion  $\psi_E \in \text{CR}^1(\mathcal{T})$  durch ihre Werte

$$\psi_E(\text{mid}(E)) := 1 \quad \text{sowie} \quad \psi_E(\text{mid}(F)) := 0 \quad \text{für alle } F \in \mathcal{E}.$$

Die Menge  $\{\psi_E \mid E \in \mathcal{E}\}$  bildet dann eine Basis von  $\text{CR}^1(\mathcal{T})$  und die Menge  $\{\psi_E \mid E \in \mathcal{E}(\Omega)\}$  bildet eine Basis von  $\text{CR}_0^1(\mathcal{T})$ . Für Kapitel 6 werden wir außerdem die folgenden Aussagen benötigen. Sei  $T \in \mathcal{T}$  mit  $T = \text{conv}\{P_1, P_2, P_3\}$ . Dann sind die baryzentrischen Koordinaten  $\lambda_1, \lambda_2, \lambda_3 \in P_1(T)$  für  $j, k \in \{1, 2, 3\}$  charakterisiert durch

$$\lambda_j(P_k) = \begin{cases} 1, & \text{falls } j = k, \\ 0, & \text{sonst.} \end{cases}$$

Seien die Kanten von  $T$  nun so mit  $E_1, E_2$  und  $E_3$  bezeichnet, dass für  $j \in \{1, 2, 3\}$  die Kante  $E_j$  gegenüber von  $P_j$  liegt. Dann können wir die lokalen Crouzeix-Raviart-Basisfunktionen  $\psi_{E_j}|_T$ ,  $j \in \{1, 2, 3\}$ , für alle  $x \in T$  darstellen durch

$$\psi_{E_j}|_T(x) = 1 - 2\lambda_j(x). \tag{2.2}$$

Dadurch ist für alle  $k_1, k_2, k_3 \in \mathbb{N}$  die Integrationsformel

$$\int_T \lambda_1^{k_1} \lambda_2^{k_2} \lambda_3^{k_3} \, dx = 2|T| \frac{k_1! k_2! k_3!}{(2 + k_1 + k_2 + k_3)!} \tag{2.3}$$

auch für die Berechnung von Integralen über Produkte der lokalen Crouzeix-Raviart-Basisfunktionen auf  $T$  nutzbar. Ein Beweis der Formel (2.3) ist in [EM73] zu finden.

### 2.3 Variationsrechnung auf Banachräumen

In dieser Arbeit setzen wir Kenntnisse über die Variationsrechnung voraus. Da wir aber auch Funktionale betrachten, die auf unendlichdimensionalen, reellen Banachräumen definiert sind, führen wir in diesem Abschnitt die grundlegenden Notationen dafür ein und formulieren schließlich die notwendige Optimalitätsbedingung erster Ordnung.

Dabei folgen wir [Zei85, S. 189-194]. Dort werden einige der Aussagen auf einem reellen, lokal konvexen Raum formuliert. Da nach [Zei86, S. 781, (43)] alle Banachräume lokal konvex sind und wir die Aussagen in dieser Arbeit nur auf Banachräumen benötigen, formulieren wir sie hier auf einem reellen Banachraum  $X$ . Außerdem betrachten wir eine Teilmenge  $V \subseteq X$ , einen inneren Punkt  $u$  von  $V$  und ein Funktional  $F : V \rightarrow \mathbb{R}$ . Schließlich definieren wir noch für alle  $h \in X$  eine Funktion  $\varphi_h : \mathbb{R} \rightarrow \mathbb{R}$ , die für alle  $t \in \mathbb{R}$  gegeben ist durch  $\varphi_h(t) := F(u + th)$ . Damit können wir die  $n$ -te Variation, das Gâteaux- und das Fréchet-Differential von  $F$  definieren.

**Definition 2.1** (*n*-te Variation). Die *n*-te Variation von  $F$  an der Stelle  $u$  in Richtung  $h \in X$  ist, falls die *n*-te Ableitung  $\varphi_h^{(n)}(0)$  von  $\varphi_h$  in 0 existiert, definiert durch

$$\delta^n F(u; h) := \varphi_h^{(n)}(0) = \left. \frac{d^n F(u + th)}{dt^n} \right|_{t=0}.$$

Wir schreiben  $\delta$  für  $\delta^1$ .

**Definition 2.2** (Gâteaux- und Fréchet-Differential).  $F$  heißt Gâteaux-differenzierbar an der Stelle  $u$ , falls ein Funktional  $F'(u) \in X^*$  existiert, sodass

$$\lim_{t \rightarrow 0} \frac{F(u + th) - F(u)}{t} = \langle F'(u), h \rangle \quad \text{für alle } h \in X.$$

$F'(u)$  heißt dann Gâteaux-Differential von  $F$  an der Stelle  $u$ .  $F$  heißt Fréchet-differenzierbar an der Stelle  $u$ , falls ein Funktional  $F'(u) \in X^*$  existiert, sodass

$$\lim_{\|h\|_X \rightarrow 0} \frac{|F(u + th) - F(u) - \langle F'(u), h \rangle|}{\|h\|_X} = 0.$$

$F'(u)$  heißt dann Fréchet-Differential von  $F$  an der Stelle  $u$ . Das Fréchet-Differential von  $F$  an der Stelle  $u$  in Richtung  $h \in X$  ist definiert durch  $dF(u; h) := \langle F'(u), h \rangle$ .

*Bemerkung 2.3.* Existiert das Fréchet-Differential  $F'(u)$  von  $F$  an der Stelle  $u$ , so ist  $F'(u)$  auch das Gâteaux-Differential von  $F$  an der Stelle  $u$  und es gilt

$$\delta F(u; h) = dF(u; h) = \langle F'(u), h \rangle \quad \text{für alle } h \in X.$$

Nachdem nun die relevante Notation eingeführt ist, können wir zum Abschluss die notwendige Bedingung erster Ordnung für einen lokalen Minimierer von  $F$  formulieren.

**Theorem 2.4** (Notwendige Optimalitätsbedingung erster Ordnung). Sei  $u \in \text{int}(V)$  lokaler Minimierer von  $F$ , das heißt, es existiere eine Umgebung  $U \subseteq V$  von  $u$ , sodass  $F(v) \geq F(u)$  für alle  $v \in U$ . Dann gilt für alle  $h \in X$ , dass  $\delta F(u; h) = 0$ , falls diese Variation für alle  $h \in X$  existiert, beziehungsweise  $F'(u) = 0$ , falls  $F'(u)$  als Gâteaux- oder Fréchet-Differential existiert.

*Beweis.* Die Aussage folgt direkt aus [Zei85, S. 193 f., Theorem 40.A, Theorem 40.B].  $\square$

## 2.4 Subdifferentiale

Für diesen Abschnitt betrachten wir stets einen reellen Banachraum  $X$  und, falls nicht anders spezifiziert, ein Funktional  $F : X \rightarrow [-\infty, \infty]$ . Wir wollen die in dieser Arbeit benötigten Notationen und Eigenschaften des Subdifferentials von  $F$  zusammentragen. Zuvor starten wir mit einer grundlegenden Definition.

**Definition 2.5** ([Zei85, S. 245, Definition 42.1]). Sei  $V$  ein Vektorraum,  $M \subseteq V$  und  $F : M \rightarrow \mathbb{R}$ . Dann heißt die Menge  $M$  konvex, wenn für alle  $u, v \in M$  und alle  $t \in [0, 1]$  gilt

$$(1 - t)u + tv \in M.$$

Ist  $M$  konvex, so heißt  $F$  konvex, falls für alle  $u, v \in M$  und alle  $t \in [0, 1]$  gilt

$$F((1 - t)u + tv) \leq (1 - t)F(u) + tF(v). \quad (2.4)$$

Gilt Ungleichung (2.4) für alle  $t \in (0, 1)$  mit „ $<$ “, so heißt  $F$  strikt konvex. Falls  $-F$  konvex ist, so heißt  $F$  konkav.

## 2 Theoretische Grundlagen

Für den Rest dieses Abschnitts folgen wir [Zei85, S. 385-397]. Analog zur Begründung zum Beginn von Abschnitt 2.3, schränken wir auch hier die Definitionen und Aussagen, die in [Zei85] auf reellen, lokal konvexen Räumen formuliert sind, auf den reellen Banachraum  $X$  ein. Zunächst definieren wir das Subdifferential von  $F$  an einer Stelle  $u \in X$ .

**Definition 2.6** (Subdifferential). Für  $u \in X$  mit  $F(u) \neq \pm\infty$  heißt

$$\partial F(u) := \{u^* \in X^* \mid \forall v \in X \quad F(v) \geq F(u) + \langle u^*, v - u \rangle\} \quad (2.5)$$

Subdifferential von  $F$  an der Stelle  $u$ . Für  $F(u) = \pm\infty$  ist  $\partial F(u) := \emptyset$ . Ein Element  $u^* \in \partial F(u)$  heißt Subgradient von  $F$  an der Stelle  $u$ .

Es folgen für Optimierungsprobleme wichtige Aussagen über das Subdifferential von  $F$ .

**Theorem 2.7** ([Zei85, S. 387, Proposition 47.12]). *Falls  $F : X \rightarrow (-\infty, \infty]$  mit  $F \not\equiv \infty$ , gilt  $F(u) = \inf_{v \in X} F(v)$  genau dann, wenn  $0 \in \partial F(u)$ .*

**Theorem 2.8** ([Zei85, S. 387, Proposition 47.13 (i)]). *Falls  $F$  konvex und Gâteaux-differenzierbar an der Stelle  $u \in X$  mit Gâteaux-Differential  $F'(u)$  ist, gilt  $\partial F(u) = \{F'(u)\}$ .*

Das folgende Theorem folgt aus [Zei85, S. 389, Theorem 47.B] unter Beachtung der Tatsache, dass die Addition von Funktionen  $F_1, F_2, \dots, F_n : X \rightarrow (-\infty, \infty]$  und die Addition von Mengen in  $X^*$  kommutieren.

**Theorem 2.9.** *Seien für  $n \geq 2$  die Funktionale  $F_1, F_2, \dots, F_n : X \rightarrow (-\infty, \infty]$  konvex und es existiere ein  $u_0 \in X$ , sodass  $F_k(u_0) < \infty$  für alle  $k \in \{1, 2, \dots, n\}$ . Außerdem seien mindestens  $n-1$  der  $n$  Funktionale  $F_1, F_2, \dots, F_n$  stetig an der Stelle  $u_0$ . Dann gilt*

$$\partial(F_1 + F_2 + \dots + F_n)(u) = \partial F_1(u) + \partial F_2(u) + \dots + \partial F_n(u) \quad \text{für alle } u \in X.$$

Zum Abschluss formulieren wir noch die Monotonie des Subdifferentials. Diese folgt aus [Zei85, S. 396 f., Definition 47.15, Theorem 47.F (1)].

**Theorem 2.10.** *Sei  $F : X \rightarrow (-\infty, \infty]$  konvex und unterhalbstetig mit  $F \not\equiv \infty$ . Dann ist  $\partial F(\cdot)$  monoton, das heißt,*

$$\langle u^* - v^*, u - v \rangle \geq 0 \quad \text{für alle } u, v \in X, u^* \in \partial F(u), v^* \in \partial F(v).$$

## 2.5 Funktionen beschränkter Variation

In diesem Abschnitt führen wir den Raum der Funktionen beschränkter Variation ein. Wir vermeiden dabei für den weiteren Verlauf dieser Arbeit nicht benötigte Notation und Theorie, indem wir die Definitionen und Aussagen entsprechend aus- und umformulieren. Für detaillierte Ausführungen und die maßtheoretischen Hintergründe siehe zum Beispiel [ABM14; EG92; Bra98]. Soweit nicht anders angegeben, folgen die Definitionen und Aussagen dieses Abschnitts aus [ABM14, S. 393-395]. Sei im Weiteren  $U$  eine offene Teilmenge des  $\mathbb{R}^d$ . Zunächst definieren wir den Raum der Funktionen beschränkter Variation.

**Definition 2.11** (Funktionen beschränkter Variation). Eine Funktion  $u \in L^1(U)$  ist von beschränkter Variation, wenn

$$|u|_{BV(U)} := \sup_{\substack{\phi \in C_C^1(U; \mathbb{R}^d) \\ \|\phi\|_{L^\infty(U)} \leq 1}} \int_U u \operatorname{div}(\phi) dx < \infty. \quad (2.6)$$

Die Menge aller Funktionen beschränkter Variation ist  $BV(U)$ .

*Bemerkung 2.12.* Durch  $|\cdot|_{BV(U)}$  ist eine Seminorm auf  $BV(U)$  gegeben. Ausgestattet mit der Norm

$$\|\cdot\|_{BV(U)} := \|\cdot\|_{L^1(U)} + |\cdot|_{BV(U)}$$

ist  $BV(U)$  ein Banachraum. Außerdem gilt  $W^{1,1}(U) \subset BV(U)$  und  $\|u\|_{BV(U)} = \|u\|_{W^{1,1}(U)}$  für alle  $u \in W^{1,1}(U)$ .

In der Anwendung ist Konvergenz in  $BV(U)$  bezüglich der Norm  $\|\cdot\|_{BV(U)}$  zu restriktiv (cf. [Bar15b, S. 300]). Deshalb führen wir einen schwächeren Konvergenzbegriff ein.

**Definition 2.13.** Sei  $(u_n)_{n \in \mathbb{N}} \subset BV(U)$  und sei  $u \in BV(U)$  mit  $u_n \rightarrow u$  in  $L^1(U)$  für  $n \rightarrow \infty$ . Dann konvergiert die Folge  $(u_n)_{n \in \mathbb{N}}$  schwach gegen  $u$  in  $BV(U)$ , wenn für alle  $\phi \in C_0(U; \mathbb{R}^d)$  gilt, dass

$$\int_U u_n \operatorname{div}(\phi) dx \rightarrow \int_U u \operatorname{div}(\phi) dx \quad \text{für } n \rightarrow \infty.$$

Wir schreiben dann  $u_n \rightharpoonup u$  für  $n \rightarrow \infty$ .

Damit können wir das folgende Theorem formulieren, welches unmittelbar die schwache Unterhalbstetigkeit der Seminorm  $|\cdot|_{BV(U)}$  auf  $BV(U)$  impliziert.

**Theorem 2.14** ([ABM14, S. 394, Proposition 10.1.1]). *Sei  $u \in L^1(U)$  und sei  $(u_n)_{n \in \mathbb{N}} \subset BV(U)$  mit  $\sup_{n \in \mathbb{N}} |u_n|_{BV(U)} < \infty$  und  $u_n \rightarrow u$  in  $L^1(U)$  für  $n \rightarrow \infty$ . Dann gilt  $u \in BV(U)$  und  $|u|_{BV(U)} \leq \liminf_{n \rightarrow \infty} |u_n|_{BV(U)}$ . Außerdem gilt dann  $u_n \rightharpoonup u$  in  $BV(U)$ .*

Für den Rest dieses Abschnitts sei  $U$  ein beschränktes Lipschitz-Gebiet. Unter dieser Voraussetzung können wir zeigen, dass jede in  $BV(U)$  beschränkte Folge eine in  $BV(U)$  schwach konvergente Teilfolge besitzt mit schwachem Grenzwert in  $BV(U)$ . Für den Beweis dieser Aussage benötigen wir zunächst das folgende Theorem.

**Theorem 2.15** ([EG92, S. 176, Theorem 4]). *Sei  $(u_n)_{n \in \mathbb{N}} \subset BV(U)$  eine beschränkte Folge. Dann existiert eine Teilfolge  $(u_{n_k})_{k \in \mathbb{N}}$  von  $(u_n)_{n \in \mathbb{N}}$  und ein  $u \in BV(U)$ , sodass  $u_{n_k} \rightarrow u$  in  $L^1(U)$  für  $k \rightarrow \infty$ .*

Damit können wir nun das folgende Theorem beweisen.

**Theorem 2.16.** *Sei  $(u_n)_{n \in \mathbb{N}} \subset BV(U)$  eine beschränkte Folge. Dann existiert eine Teilfolge  $(u_{n_k})_{k \in \mathbb{N}}$  und ein  $u \in BV(U)$ , sodass  $u_{n_k} \rightharpoonup u$  in  $BV(U)$  für  $k \rightarrow \infty$ .*

*Beweis.* Nach Theorem 2.15 besitzt  $(u_n)_{n \in \mathbb{N}}$  eine Teilfolge  $(u_{n_k})_{k \in \mathbb{N}}$ , die in  $L^1(U)$  gegen ein  $u \in BV(U)$  konvergiert. Diese Teilfolge ist nach Voraussetzung beschränkt in  $BV(U)$ , woraus nach Definition der Norm  $\|\cdot\|_{BV(U)}$  insbesondere folgt, dass  $\sup_{k \in \mathbb{N}} |u_{n_k}|_{BV(U)} < \infty$ . Somit ist Theorem 2.14 anwendbar und impliziert die schwache Konvergenz von  $(u_{n_k})_{k \in \mathbb{N}}$  in  $BV(U)$  gegen  $u \in BV(U)$ .  $\square$

### 3 Das kontinuierliche Problem

In diesem Kapitel wollen wir für einen Parameter  $\alpha \in \mathbb{R}_+$  und eine Funktion  $f \in L^2(\Omega)$  folgendes Minimierungsproblem untersuchen.

**Problem 3.1.** Finde  $u \in \text{BV}(\Omega) \cap L^2(\Omega)$ , sodass  $u$  das Funktional

$$E(v) := \frac{\alpha}{2} \|v\|^2 + |v|_{\text{BV}(\Omega)} + \|v\|_{L^1(\partial\Omega)} - \int_{\Omega} fv \, dx \quad (3.1)$$

unter allen  $v \in \text{BV}(\Omega) \cap L^2(\Omega)$  minimiert.

Dabei ist der Term  $\|v\|_{L^1(\partial\Omega)}$  wohldefiniert, da nach [ABM14, S. 400, Theorem 10.2.1] eine lineare, stetige Abbildung  $T : \text{BV}(\Omega) \rightarrow L^1(\partial\Omega)$  existiert mit  $T(v) = v|_{\partial\Omega}$  für alle  $v \in \text{BV}(\Omega) \cap C(\overline{\Omega})$ . Wie in Kapitel 1 beschrieben, hat Problem 3.1 für homogene Randdaten die gleichen Minimierer wie das ROF-Modellproblem mit Eingangssignal  $g \in L^2(\Omega)$ , falls  $f = \alpha g$ .

*Bemerkung 3.2.* Für  $d \in \{2, 3\}$  und ein beschränktes Lipschitz-Gebiet  $U \subset \mathbb{R}^d$  ist nach [Bar15b, S. 302, Remark 10.5 (i)] die Einbettung  $\text{BV}(U) \hookrightarrow L^p(U)$  stetig, wenn  $1 \leq p \leq d/(d-1)$ . Damit ist  $\text{BV}(\Omega)$  für das polygonal berandete Lipschitz-Gebiet  $\Omega \subset \mathbb{R}^2$  Teilmenge von  $L^2(\Omega)$  und die Lösung von Problem 3.1 kann in  $\text{BV}(\Omega)$  gesucht werden. Wir vernachlässigen diese Vereinfachung und erreichen dadurch, dass alle Aussagen im nachfolgenden Abschnitt auch gelten, falls  $\Omega \subset \mathbb{R}^d$  mit  $d \in \mathbb{N}$  ein beschränktes Lipschitz-Gebiet ist.

In den folgenden Abschnitten zeigen wir, dass Problem 3.1 eine eindeutige Lösung  $u \in \text{BV}(\Omega) \cap L^2(\Omega)$  besitzt. Außerdem beschreiben wir, wie  $f$  für eine gegebene Lösung  $u$  konstruiert werden kann und welche Eigenschaften  $u$  dafür erfüllen muss.

#### 3.1 Existenz eines eindeutigen Minimierers

Nach Bemerkung 3.2 kann in diesem Abschnitt  $\Omega \subset \mathbb{R}^d$  für  $d \in \mathbb{N}$  ein beliebiges beschränktes Lipschitz-Gebiet sein. Zunächst zeigen wir, dass Problem 3.1 eine Lösung besitzt. Dafür benötigen wir die folgende Formulierung der Youngschen Ungleichung.

**Lemma 3.3** (Youngsche Ungleichung). *Seien  $a, b \in \mathbb{R}$  und  $\varepsilon \in \mathbb{R}_+$  beliebig. Dann gilt*

$$ab \leq \frac{1}{\varepsilon} a^2 + \frac{\varepsilon}{4} b^2.$$

Außerdem wird die folgende Aussage benötigt, die direkt aus [EG92, S. 183, Theorem 1] folgt, da  $0 \in \text{BV}(\mathbb{R}^d \setminus \overline{\Omega})$  mit  $|0|_{\text{BV}(\mathbb{R}^d \setminus \overline{\Omega})} = 0$  und  $0|_{\partial\Omega} = 0$ .

**Lemma 3.4.** *Sei  $v \in \text{BV}(\Omega)$ . Definiere die Fortsetzung  $\tilde{v}$  von  $v$  für alle  $x \in \mathbb{R}^d$  durch*

$$\tilde{v}(x) := \begin{cases} v(x), & \text{falls } x \in \Omega, \\ 0, & \text{falls } x \in \mathbb{R}^d \setminus \overline{\Omega}. \end{cases}$$

*Dann gilt  $\tilde{v} \in \text{BV}(\mathbb{R}^d)$  und  $|\tilde{v}|_{\text{BV}(\mathbb{R}^d)} = |v|_{\text{BV}(\Omega)} + \|v\|_{L^1(\partial\Omega)}$ .*

**Theorem 3.5** (Existenz einer Lösung). *Problem 3.1 besitzt eine Lösung  $u \in \text{BV}(\Omega) \cap L^2(\Omega)$ .*

*Beweis.* Die Beweisidee ist die Anwendung der direkten Methode der Variationsrechnung (cf. [Dac89]) unter Nutzung der in Abschnitt 2.5 aufgeführten Eigenschaften der schwachen Konvergenz in  $\text{BV}(\Omega)$ .

Für alle  $v \in L^2(\Omega)$  gilt mit der Hölderschen Ungleichung und der Youngschen Ungleichung aus Lemma 3.3, dass

$$\int_{\Omega} fv \, dx \leq \|f\| \|v\| \leq \frac{1}{\alpha} \|f\|^2 + \frac{\alpha}{4} \|v\|^2. \quad (3.2)$$

Die Höldersche Ungleichung impliziert außerdem für alle  $v \in L^2(\Omega) \subseteq L^1(\Omega)$ , dass

$$\|v\|_{L^1(\Omega)} = \|1 \cdot v\|_{L^1(\Omega)} \leq \|1\| \|v\| = \sqrt{|\Omega|} \|v\|. \quad (3.3)$$

Für das Funktional  $E$  aus (3.1) folgt dann mit den Ungleichungen (3.2) und (3.3) für alle  $v \in \text{BV}(\Omega) \cap L^2(\Omega)$ , dass

$$\begin{aligned} E(v) &\geq \frac{\alpha}{4} \|v\|^2 + |v|_{\text{BV}(\Omega)} + \|v\|_{L^1(\partial\Omega)} - \frac{1}{\alpha} \|f\|^2 \\ &\geq \frac{\alpha}{4|\Omega|} \|v\|_{L^1(\Omega)}^2 + |v|_{\text{BV}(\Omega)} + \|v\|_{L^1(\partial\Omega)} - \frac{1}{\alpha} \|f\|^2 \geq -\frac{1}{\alpha} \|f\|^2. \end{aligned} \quad (3.4)$$

Somit ist  $E$  nach unten beschränkt, was die Existenz einer infimierenden Folge  $(u_n)_{n \in \mathbb{N}} \subset \text{BV}(\Omega) \cap L^2(\Omega)$  von  $E$  impliziert, das heißt,  $(u_n)_{n \in \mathbb{N}}$  erfüllt

$$\lim_{n \rightarrow \infty} E(u_n) = \inf_{v \in \text{BV}(\Omega) \cap L^2(\Omega)} E(v).$$

Ungleichung (3.4) impliziert außerdem, dass  $E(u_n) \rightarrow \infty$  für  $n \rightarrow \infty$ , falls  $|u_n|_{\text{BV}(\Omega)} \rightarrow \infty$  oder  $\|u_n\|_{L^1(\Omega)} \rightarrow \infty$  für  $n \rightarrow \infty$ . Daraus folgt insbesondere, dass  $E(u_n) \rightarrow \infty$  für  $n \rightarrow \infty$ , falls  $\|u_n\|_{\text{BV}(\Omega)} \rightarrow \infty$  für  $n \rightarrow \infty$ . Deshalb muss die Folge  $(u_n)_{n \in \mathbb{N}}$  beschränkt in  $\text{BV}(\Omega)$  sein. Nun garantiert Theorem 2.16 die Existenz einer in  $\text{BV}(\Omega)$  schwach konvergenten Teilfolge  $(u_{n_k})_{k \in \mathbb{N}}$  von  $(u_n)_{n \in \mathbb{N}}$  mit schwachem Grenzwert  $u \in \text{BV}(\Omega)$ . Ohne Beschränkung der Allgemeinheit sei  $(u_{n_k})_{k \in \mathbb{N}} = (u_n)_{n \in \mathbb{N}}$ . Aus der schwachen Konvergenz von  $(u_n)_{n \in \mathbb{N}}$  in  $\text{BV}(\Omega)$  gegen  $u$  folgt nach Definition, dass  $(u_n)_{n \in \mathbb{N}}$  stark, und damit insbesondere auch schwach, in  $L^1(\Omega)$  gegen  $u$  konvergiert.

Weiterhin folgt aus (3.4), dass  $E(v) \rightarrow \infty$  für  $\|v\| \rightarrow \infty$ . Somit muss  $(u_n)_{n \in \mathbb{N}}$  auch beschränkt sein bezüglich der Norm  $\|\cdot\|$  und besitzt deshalb, wegen der Reflexivität von  $L^2(\Omega)$ , eine Teilfolge (ohne Beschränkung der Allgemeinheit weiterhin bezeichnet mit  $(u_n)_{n \in \mathbb{N}}$ ), die in  $L^2(\Omega)$  schwach gegen einen Grenzwert  $\bar{u} \in L^2(\Omega)$  konvergiert. Damit gilt für alle  $w \in L^2(\Omega) \cong L^2(\Omega)^*$  und, da  $L^\infty(\Omega) \subseteq L^2(\Omega)$ , insbesondere auch für alle  $w \in L^\infty(\Omega) \cong L^1(\Omega)^*$ , dass

$$\lim_{n \rightarrow \infty} \int_{\Omega} u_n w \, dx = \int_{\Omega} \bar{u} w \, dx.$$

Das bedeutet, dass  $(u_n)_{n \in \mathbb{N}}$  auch schwach in  $L^1(\Omega)$  gegen  $\bar{u} \in L^2(\Omega) \subseteq L^1(\Omega)$  konvergiert. Da schwache Grenzwerte eindeutig bestimmt sind, gilt insgesamt  $u = \bar{u} \in L^2(\Omega)$ , das heißt,  $u \in \text{BV}(\Omega) \cap L^2(\Omega)$ . Nun definieren wir für alle  $n \in \mathbb{N}$  und für alle  $x \in \mathbb{R}^d$  die Fortsetzungen

$$\tilde{u}_n(x) := \begin{cases} u_n(x), & \text{falls } x \in \Omega, \\ 0, & \text{falls } x \in \mathbb{R}^d \setminus \overline{\Omega} \end{cases} \quad \text{und} \quad \tilde{u}(x) := \begin{cases} u(x), & \text{falls } x \in \Omega, \\ 0, & \text{falls } x \in \mathbb{R}^d \setminus \overline{\Omega}. \end{cases}$$

### 3 Das kontinuierliche Problem

Dann gilt nach Lemma 3.4 sowohl

$$\begin{aligned}\tilde{u}_n &\in \text{BV}(\mathbb{R}^d) \quad \text{und} \quad |\tilde{u}_n|_{\text{BV}(\mathbb{R}^d)} = |u_n|_{\text{BV}(\Omega)} + \|u_n\|_{L^1(\partial\Omega)} \quad \text{für alle } n \in \mathbb{N} \text{ als auch} \\ \tilde{u} &\in \text{BV}(\mathbb{R}^d) \quad \text{und} \quad |\tilde{u}|_{\text{BV}(\mathbb{R}^d)} = |u|_{\text{BV}(\Omega)} + \|u\|_{L^1(\partial\Omega)}.\end{aligned}$$

Da  $(u_n)_{n \in \mathbb{N}}$  infimierende Folge von  $E$  ist, muss die Folge

$$\left(|\tilde{u}_n|_{\text{BV}(\mathbb{R}^d)}\right)_{n \in \mathbb{N}} = \left(|u_n|_{\text{BV}(\Omega)} + \|u_n\|_{L^1(\partial\Omega)}\right)_{n \in \mathbb{N}}$$

beschränkt sein. Außerdem folgt aus den Definitionen von  $\tilde{u}$  und  $\tilde{u}_n$  für alle  $n \in \mathbb{N}$  und der bereits bekannten Eigenschaft  $u_n \rightarrow u$  in  $L^1(\Omega)$  für  $n \rightarrow \infty$ , dass

$$\|\tilde{u}_n - \tilde{u}\|_{L^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |\tilde{u}_n - \tilde{u}| \, dx = \int_{\Omega} |u_n - u| \, dx = \|u_n - u\|_{L^1(\Omega)} \rightarrow 0 \quad \text{für } n \rightarrow \infty,$$

das heißt,  $\tilde{u}_n \rightarrow \tilde{u}$  in  $L^1(\mathbb{R}^d)$  für  $n \rightarrow \infty$ . Insgesamt ist also  $(\tilde{u}_n)_{n \in \mathbb{N}}$  eine Folge in  $\text{BV}(\mathbb{R}^d)$ , die in  $L^1(\mathbb{R}^d)$  gegen  $\tilde{u} \in \text{BV}(\mathbb{R}^d) \subseteq L^1(\mathbb{R}^d)$  konvergiert und  $\sup_{n \in \mathbb{N}} |\tilde{u}_n|_{\text{BV}(\mathbb{R}^d)} < \infty$  erfüllt. Somit folgt mit Theorem 2.14

$$\begin{aligned}|u|_{\text{BV}(\Omega)} + \|u\|_{L^1(\partial\Omega)} &= |\tilde{u}|_{\text{BV}(\mathbb{R}^d)} \leq \liminf_{n \rightarrow \infty} |\tilde{u}_n|_{\text{BV}(\mathbb{R}^d)} \\ &= \liminf_{n \rightarrow \infty} (|u_n|_{\text{BV}(\Omega)} + \|u_n\|_{L^1(\partial\Omega)}).\end{aligned}\tag{3.5}$$

Die Funktionen  $\|\cdot\|^2$  und  $-\int_{\Omega} f \cdot \, dx$  sind auf  $L^2(\Omega)$  stetig und konvex, was impliziert, dass sie schwach unterhalbstetig auf  $L^2(\Omega)$  sind. Da wir bereits wissen, dass  $u_n \rightarrow u$  in  $L^2(\Omega)$  für  $n \rightarrow \infty$ , folgt

$$\frac{\alpha}{2} \|u\|^2 - \int_{\Omega} f u \, dx \leq \liminf_{n \rightarrow \infty} \left( \frac{\alpha}{2} \|u_n\|^2 - \int_{\Omega} f u_n \, dx \right).$$

Damit und mit Ungleichung (3.5) gilt insgesamt

$$\inf_{v \in \text{BV}(\Omega) \cap L^2(\Omega)} E(v) \leq E(u) \leq \liminf_{n \rightarrow \infty} E(u_n) = \lim_{n \rightarrow \infty} E(u_n) = \inf_{v \in \text{BV}(\Omega) \cap L^2(\Omega)} E(v),$$

das heißt,  $\min_{v \in \text{BV}(\Omega) \cap L^2(\Omega)} E(v) = E(u)$ . □

Nachdem wir nun gezeigt haben, dass eine Lösung von Problem 3.1 existiert, können wir aus dem folgenden Theorem unmittelbar die Eindeutigkeit dieser Lösung folgern.

**Theorem 3.6** (Stabilität). *Seien  $u_1, u_2 \in \text{BV}(\Omega) \cap L^2(\Omega)$  die Lösungen von Problem 3.1 mit  $f_1, f_2 \in L^2(\Omega)$  anstelle von  $f$ , das heißt, für  $\ell \in \{1, 2\}$  minimiere  $u_\ell$  das Funktional*

$$E_\ell(v) := \frac{\alpha}{2} \|v\|^2 + |v|_{\text{BV}(\Omega)} + \|v\|_{L^1(\partial\Omega)} - \int_{\Omega} f_\ell v \, dx$$

*unter allen  $v \in \text{BV}(\Omega) \cap L^2(\Omega)$ . Dann gilt*

$$\|u_1 - u_2\| \leq \frac{1}{\alpha} \|f_1 - f_2\|.$$

*Beweis.* Wir folgen der Argumentation im Beweis von [Bar15b, S. 304, Theorem 10.6].

Zunächst definieren wir die Funktionale  $F : L^2(\Omega) \rightarrow [0, \infty]$  und  $G_\ell : L^2(\Omega) \rightarrow \mathbb{R}$ ,  $\ell \in \{1, 2\}$ , für alle  $v \in L^2(\Omega)$  durch

$$F(v) := \begin{cases} |v|_{\text{BV}(\Omega)} + \|v\|_{L^1(\partial\Omega)}, & \text{falls } v \in \text{BV}(\Omega) \cap L^2(\Omega), \\ \infty, & \text{falls } v \in L^2(\Omega) \setminus \text{BV}(\Omega) \end{cases} \quad \text{und}$$

$$G_\ell(v) := \frac{\alpha}{2} \|v\|^2 - \int_{\Omega} f_\ell v \, dx.$$

### 3.1 Existenz eines eindeutigen Minimierers

Damit gilt für  $\ell \in \{1, 2\}$  und alle  $v \in \text{BV}(\Omega) \cap L^2(\Omega)$ , dass  $E_\ell(v) = F(v) + G_\ell(v)$ . Für  $\ell \in \{1, 2\}$  ist  $G_\ell$  Fréchet-differenzierbar und die Fréchet-Ableitung  $G'_\ell(v) : L^2(\Omega) \rightarrow \mathbb{R}$  von  $G_\ell$  an der Stelle  $v \in L^2(\Omega)$  ist für alle  $w \in L^2(\Omega)$  gegeben durch

$$dG_\ell(v; w) = \alpha(v, w) - \int_{\Omega} f_\ell w \, dx = (\alpha v - f_\ell, w).$$

Das Funktional  $F$  ist konvex, unterhalbstetig und es gilt  $F \not\equiv \infty$ . Deshalb ist nach Theorem 2.10 das Subdifferential  $\partial F(\bullet)$  von  $F$  monoton. Damit gilt für alle  $\mu_\ell \in \partial F(u_\ell)$ ,  $\ell \in \{1, 2\}$ , dass

$$(\mu_1 - \mu_2, u_1 - u_2) \geq 0. \quad (3.6)$$

Für  $\ell \in \{1, 2\}$  gilt, dass  $E_\ell$  konvex ist und von  $u_\ell$  in  $\text{BV}(\Omega) \cap L^2(\Omega)$  minimiert wird. Außerdem gilt  $E_\ell \not\equiv \infty$  und  $G_\ell$  ist stetig. Somit gilt nach den Theoremen 2.7 – 2.9, dass

$$0 \in \partial E_\ell(u_\ell) = \partial F(u_\ell) + \partial G_\ell(u_\ell) = \partial F(u_\ell) + \{G'_\ell(u_\ell)\}.$$

Daraus folgt  $-G'_\ell(u_\ell) \in \partial F(u_\ell)$ . Zusammen mit Ungleichung (3.6) impliziert das

$$(-(\alpha u_1 - f_1) - (-(\alpha u_2 - f_2)), u_1 - u_2) \geq 0.$$

Durch Umformen und Anwenden der Cauchy-Schwarzschen Ungleichung erhalten wir

$$\alpha \|u_1 - u_2\|^2 \leq (f_1 - f_2, u_1 - u_2) \leq \|f_1 - f_2\| \|u_1 - u_2\|.$$

Falls  $\|u_1 - u_2\| = 0$ , gilt die zu zeigende Aussage. Ansonsten führt die Division durch  $\alpha \|u_1 - u_2\| \neq 0$  den Beweis zum Abschluss.  $\square$

Zum Ende dieses Abschnitts beweisen wir, dass der Abstand einer Funktion zu einem Minimierer des Funktionals  $E$  durch die Differenz der Werte des Funktionals kontrolliert wird. Auch aus diesem Theorem folgt die Eindeutigkeit der nach Theorem 3.5 existierenden Lösung von Problem 3.1.

**Theorem 3.7.** *Sei  $u \in \text{BV}(\Omega) \cap L^2(\Omega)$  Lösung von Problem 3.1. Dann gilt*

$$\frac{\alpha}{2} \|u - v\|^2 \leq E(v) - E(u) \quad \text{für alle } v \in \text{BV}(\Omega) \cap L^2(\Omega).$$

*Beweis.* Wir folgen der Argumentation im Beweis von [Bar15b, S. 309, Lemma 10.2].

Wir betrachten die konvexen Funktionale  $F : L^2(\Omega) \rightarrow [0, \infty]$  und  $G : L^2(\Omega) \rightarrow \mathbb{R}$ , wobei  $F$  wie im Beweis von Theorem 3.6 definiert ist und  $G$  für alle  $v \in L^2(\Omega)$  gegeben ist durch

$$G(v) := \frac{\alpha}{2} \|v\|^2 - \int_{\Omega} fv \, dx.$$

Es gilt  $E(v) = F(v) + G(v)$  für alle  $v \in \text{BV}(\Omega) \cap L^2(\Omega)$ . Die Fréchet-Ableitung  $G'(u) : L^2(\Omega) \rightarrow \mathbb{R}$  von  $G$  an der Stelle  $u \in \text{BV}(\Omega) \cap L^2(\Omega)$  ist für alle  $v \in L^2(\Omega)$  gegeben durch

$$dG(u; v) = \alpha(u, v) - \int_{\Omega} fv \, dx = (\alpha u - f, v).$$

Das impliziert mit wenigen Rechenschritten

$$dG(u; v - u) + \frac{\alpha}{2} \|u - v\|^2 + G(u) = G(v) \quad \text{für alle } v \in L^2(\Omega). \quad (3.7)$$

### 3 Das kontinuierliche Problem

Da  $u$  der Minimierer von  $E$  ist, erhalten wir mit den Theoremen 2.7 – 2.9 die Aussage

$$0 \in \partial E(u) = \partial F(u) + \{G'(u)\},$$

woraus folgt  $-G'(u) \in \partial F(u)$ . Das ist nach Definition 2.6 äquivalent zu

$$-dG(u; v - u) \leq F(v) - F(u) \quad \text{für alle } v \in \mathrm{BV}(\Omega) \cap L^2(\Omega).$$

Für alle  $v \in \mathrm{BV}(\Omega) \cap L^2(\Omega)$  folgt daraus, zusammen mit Gleichung (3.7), dass

$$\frac{\alpha}{2} \|u - v\|^2 + G(u) - G(v) + F(u) = -dG(u; v - u) + F(u) \leq F(v).$$

Da  $E = F + G$  auf  $\mathrm{BV}(\Omega) \cap L^2(\Omega)$ , impliziert das die zu zeigende Aussage.  $\square$

## 3.2 Konstruktion eines Eingangssignals zu einer gegebenen Lösung

Für die numerische Untersuchung der primalen-dualen Iteration aus Kapitel 5 ist es sinnvoll, Eingangssignale  $f$  für Problem 3.1 gegeben zu haben, für die der entsprechende gesuchte Minimierer bekannt ist. Die folgende Konstruktion solcher Signale basiert auf einer Aussage von Professor Carstensen.

Sei  $u : \Omega \rightarrow \mathbb{R}$  gegeben als Funktion in Polarkoordinaten. Dabei beschränken wir uns auf vom Polarwinkel unabhängige Funktionen, das heißt, für alle  $x \in \Omega$  gelte  $u(x) := u_P(|x|)$  für  $u_P : [0, \infty) \rightarrow \mathbb{R}$ . Weiterhin fordern wir  $u_P(r) = 0$  für  $r \geq 1$  und die Existenz der partiellen Ableitung  $\partial_r u_P$  fast überall in  $[0, \infty)$ . Außerdem existiere fast überall in  $[0, \infty)$  die partielle Ableitung des für  $r \in [0, \infty)$  mithilfe einer Funktion  $q : \mathbb{R} \rightarrow [0, 1]$  definierten Ausdrucks

$$\operatorname{sgn}(\partial_r u_P(r)) := \begin{cases} -1 & \text{für } \partial_r u_P(r) < 0, \\ q(r) & \text{für } \partial_r u_P(r) = 0, \\ 1 & \text{für } \partial_r u_P(r) > 0. \end{cases}$$

Des Weiteren fordern wir  $\operatorname{sgn}(\partial_r u_P(r))/r \rightarrow 0$  für  $r \rightarrow 0$ , damit  $f_P$  in der folgenden Definition stetig in 0 fortgesetzt werden kann. Sei  $f_P : [0, \infty) \rightarrow \mathbb{R}$  mit  $\alpha$  aus Problem 3.1 gegeben durch

$$f_P(r) := \alpha u_P(r) - \partial_r(\operatorname{sgn}(\partial_r u_P(r))) - \frac{\operatorname{sgn}(\partial_r u_P(r))}{r} \quad \text{für alle } r \in [0, \infty). \quad (3.8)$$

Dann ist  $u$  Lösung von Problem 3.1, wenn das Eingangssignal auf  $\Omega \supseteq \{w \in \mathbb{R}^2 \mid |w| \leq 1\}$  für fast alle  $x \in \Omega$  gegeben ist durch  $f(x) := f_P(|x|)$ .

Für die Experimente in Kapitel 7 wollen wir die garantierte untere Energieschranke aus Theorem 4.9 berechnen können. Da dieses Theorem für das Eingangssignal voraussetzt, dass  $f \in H_0^1(\Omega)$ , müssen wir noch die folgenden Bedingungen an  $u_P$  formulieren. Hinreichend für  $f \in H_0^1(\Omega)$  ist nach Gleichung (3.8), dass  $u_P$ ,  $\partial_r(\operatorname{sgn}(\partial_r u_P))$  und  $\operatorname{sgn}(\partial_r u_P)$  stetig sind und

$$u_P(1) = \partial_r(\operatorname{sgn}(\partial_r u_P(1))) = \operatorname{sgn}(\partial_r u_P(1)) = 0.$$

Mit diesen Einschränkung gilt insbesondere  $u \in H_0^1(\Omega)$ , weshalb die exakte Energie  $E(u)$  nach Bemerkung 2.12 berechnet werden kann durch

$$E(u) = \frac{\alpha}{2} \|u\|^2 + \|u\|_{W^{1,1}(\Omega)} - \int f u \, dx.$$

### 3.2 Konstruktion eines Eingangssignals zu einer gegebenen Lösung

Um also  $E(u)$  berechnen zu können, wird der schwache Gradient  $\nabla u$  von  $u$  benötigt und um die gesuchte untere Energieschranke  $E_{\text{GLEB}, \mathcal{T}}$  aus Gleichung (4.12) zu berechnen, wird der schwache Gradient  $\nabla f$  von  $f$  benötigt. Deshalb betrachten wir an dieser Stelle noch kurz die benötigten Zusammenhänge zwischen den partiellen Ableitungen in kartesischen Koordinaten und in Polarkoordinaten für eine hinreichend glatte Funktion  $g_P : \mathbb{R}^2 \rightarrow \mathbb{R}$ . Für  $x_1, x_2 \in \mathbb{R}$  sei

$$\text{atan2}(x_2, x_1) := \begin{cases} \arctan\left(\frac{x_2}{x_1}\right), & \text{wenn } x_1 > 0, \\ \arctan\left(\frac{x_2}{x_1}\right) + \pi, & \text{wenn } x_1 < 0, x_2 \geq 0, \\ \arctan\left(\frac{x_2}{x_1}\right) - \pi, & \text{wenn } x_1 < 0, x_2 < 0, \\ \frac{\pi}{2}, & \text{wenn } x_1 = 0, x_2 > 0, \\ -\frac{\pi}{2}, & \text{wenn } x_1 = 0, x_2 < 0, \\ \text{undefiniert,} & \text{wenn } x_1 = x_2 = 0. \end{cases}$$

Ein Argument  $x = (x_1, x_2) \in \mathbb{R}^2$  von  $g_P$  kann dann in Polarkoordinaten charakterisiert werden durch die Länge  $r = \sqrt{x_1^2 + x_2^2}$  und den Winkel  $\varphi = \text{atan2}(x_2, x_1)$ . Mit dieser Notation gelten für die partiellen Ableitungen die Zusammenhänge

$$\partial_{x_1} = \cos(\varphi)\partial_r - \frac{1}{r}\sin(\varphi)\partial_\varphi \quad \text{und} \quad \partial_{x_2} = \sin(\varphi)\partial_r - \frac{1}{r}\cos(\varphi)\partial_\varphi.$$

Ist nun  $g_P$  vom Winkel  $\varphi$  unabhängig, so erhalten wir

$$\nabla g_P = \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix} \partial_r g_P. \quad (3.9)$$

Unter Beachtung der trigonometrischen Zusammenhänge

$$\sin(\arctan(y)) = \frac{y}{\sqrt{1+y^2}} \quad \text{und} \quad \cos(\arctan(y)) = \frac{1}{\sqrt{1+y^2}}$$

für alle  $y \in \mathbb{R}$  gilt

$$\begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix} = \frac{1}{r} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Aus Gleichung (3.9) folgt damit

$$\nabla g_P = \frac{\partial_r g_P}{r} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Zum Bestimmen des Gradienten in kartesischen Koordinaten einer in Polarkoordinaten gegebenen Funktion  $g_P$ , die vom Polarwinkel unabhängig ist, muss also lediglich die partielle Ableitung  $\partial_r g_P$  berechnet werden.

Konkrete Beispiele formulieren wir in Kapitel 7, wobei wir dort zugunsten einer übersichtlichen Schreibweise nicht zwischen  $f$  und  $f_P$  beziehungsweise zwischen  $u$  und  $u_P$  unterscheiden werden.

# 4 Das diskrete Problem

## 4.1 Formulierung

Bevor wir Problem 3.1 diskretisieren, merken wir an, dass  $\text{CR}^1(\mathcal{T}) \subset \text{BV}(\Omega)$ , da

$$|v_{\text{CR}}|_{\text{BV}(\Omega)} = \|\nabla_{\text{NC}} v_{\text{CR}}\|_{L^1(\Omega)} + \sum_{F \in \mathcal{E}(\Omega)} \|[v_{\text{CR}}]_F\|_{L^1(F)} \quad \text{für alle } v_{\text{CR}} \in \text{CR}^1(\mathcal{T}).$$

Dies wird für  $|\mathcal{T}| = 2$  zum Beispiel von [ABM14, S. 404, Example 10.2.1; Bar15b, S. 301, Proposition 10.1] impliziert und kann analog für beliebige reguläre Triangulierungen von  $\Omega$  bewiesen werden. Damit gilt für alle  $v_{\text{CR}} \in \text{CR}^1(\mathcal{T})$  insbesondere

$$|v_{\text{CR}}|_{\text{BV}(\Omega)} + \|v_{\text{CR}}\|_{L^1(\partial\Omega)} = \|\nabla_{\text{NC}} v_{\text{CR}}\|_{L^1(\Omega)} + \sum_{F \in \mathcal{E}} \|[v_{\text{CR}}]_F\|_{L^1(F)}.$$

Um eine nichtkonforme Formulierung von Problem 3.1 zu erhalten, ersetzen wir die Terme  $|\cdot|_{\text{BV}(\Omega)} + \|\cdot\|_{L^1(\partial\Omega)}$  des Funktionals  $E$  durch  $\|\nabla_{\text{NC}} \cdot\|_{L^1(\Omega)}$ , das heißt, wir vernachlässigen bei der nichtkonformen Formulierung die Terme  $\sum_{F \in \mathcal{E}} \|[ \cdot ]_F\|_{L^1(F)}$ . Somit erhalten wir das folgende Minimierungsproblem für den Parameter  $\alpha \in \mathbb{R}_+$  und das Eingangssignal  $f \in L^2(\Omega)$ .

**Problem 4.1.** Finde  $u_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$ , sodass  $u_{\text{CR}}$  das Funktional

$$E_{\text{NC}}(v_{\text{CR}}) := \frac{\alpha}{2} \|v_{\text{CR}}\|^2 + \|\nabla_{\text{NC}} v_{\text{CR}}\|_{L^1(\Omega)} - \int_{\Omega} f v_{\text{CR}} \, dx \quad (4.1)$$

unter allen  $v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$  minimiert.

## 4.2 Charakterisierung und Existenz eines eindeutigen Minimierers

In diesem Abschnitt führen wir die Argumente in [Bar15b, S. 313], angepasst für unsere Formulierung in Problem 4.1, detailliert aus. Zunächst zeigen wir, dass Problem 4.1 eine eindeutige Lösung besitzt. Dafür benötigen wir folgendes Lemma.

**Lemma 4.2.** Das Funktional  $E_{\text{NC}}$  aus Gleichung (4.1) ist stetig bezüglich der Konvergenz in  $L^2(\Omega)$ .

*Beweis.* Die Folge  $(v_n)_{n \in \mathbb{N}} \subset \text{CR}_0^1(\mathcal{T})$  konvergiere gegen  $v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$  bezüglich der Norm  $\|\cdot\|$ . Damit ist  $(v_n)_{n \in \mathbb{N}}$  insbesondere beschränkt in  $L^2(\Omega)$  und es gilt mit einer binomischen Formel und der umgekehrten Dreiecksungleichung, dass

$$\begin{aligned} |\|v_{\text{CR}}\|^2 - \|v_n\|^2| &= \||v_{\text{CR}}| - \|v_n\|\| \||v_{\text{CR}}| + \|v_n\| \\ &\leq \|v_{\text{CR}} - v_n\| (\|v_{\text{CR}}\| + \|v_n\|) \rightarrow 0 \quad \text{für } n \rightarrow \infty. \end{aligned}$$

Außerdem gilt mit der Hölderschen Ungleichung

$$\left| \int_{\Omega} f(v_{\text{CR}} - v_n) \, dx \right| \leq \|f\| \|v_{\text{CR}} - v_n\| \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

## 4.2 Charakterisierung und Existenz eines eindeutigen Minimierers

Schließlich gilt für alle  $n \in \mathbb{N}$  und alle  $T \in \mathcal{T}$  mit der inversen Ungleichung (cf. [Bar15b, S. 53, Lemma 3.5]) mit Konstante  $c_T \in \mathbb{R}_+$  und der Hölderschen Ungleichung, dass

$$\|\nabla_{NC}(v_{CR} - v_n)\|_{L^1(T)} \leq c_T h_T^{-1} \|v_{CR} - v_n\|_{L^1(T)} \leq c_T h_T^{-1} \sqrt{|T|} \|v_{CR} - v_n\|_{L^2(T)}.$$

Damit folgt zusammen mit der umgekehrten Dreiecksungleichung

$$\begin{aligned} |\|\nabla_{NC}v_{CR}\|_{L^1(\Omega)} - \|\nabla_{NC}v_n\|_{L^1(\Omega)}| &\leq \|\nabla_{NC}(v_{CR} - v_n)\|_{L^1(\Omega)} \\ &= \sum_{T \in \mathcal{T}} \|\nabla_{NC}(v_{CR} - v_n)\|_{L^1(T)} \\ &\leq \max_{T \in \mathcal{T}} (c_T h_T^{-1} \sqrt{|T|}) \sum_{T \in \mathcal{T}} \|v_{CR} - v_n\|_{L^2(T)} \\ &= \max_{T \in \mathcal{T}} (c_T h_T^{-1} \sqrt{|T|}) \|v_{CR} - v_n\| \rightarrow 0 \quad \text{für } n \rightarrow \infty. \end{aligned}$$

Somit ist  $E_{NC}$  Summe von drei Termen, die bezüglich der Norm  $\|\cdot\|$  folgenstetig sind, und deshalb stetig bezüglich der Konvergenz in  $L^2(\Omega)$ .  $\square$

**Theorem 4.3.** *Es existiert eine eindeutige Lösung  $u_{CR} \in CR_0^1(\mathcal{T})$  von Problem 4.1.*

*Beweis.* Mit analogen Abschätzungen wie in (3.4) erhalten wir für das Funktional  $E_{NC}$  aus Problem 4.1 für alle  $v_{CR} \in CR_0^1(\mathcal{T}) \subset L^2(\Omega)$  die Ungleichung

$$E_{NC}(v_{CR}) \geq \frac{\alpha}{4} \|v_{CR}\|^2 + \|\nabla_{NC}v_{CR}\|_{L^1(\Omega)} - \frac{1}{\alpha} \|f\|^2 \geq -\frac{1}{\alpha} \|f\|^2. \quad (4.2)$$

Somit ist  $E_{NC}$  nach unten beschränkt und es existiert eine infimierende Folge  $(v_n)_{n \in \mathbb{N}} \subset CR_0^1(\mathcal{T})$  von  $E_{NC}$ . Ungleichung (4.2) impliziert weiterhin, dass diese Folge beschränkt bezüglich der Norm  $\|\cdot\|$  sein muss. Der endlichdimensionale Raum  $CR_0^1(\mathcal{T})$  ist, ausgestattet mit der Norm  $\|\cdot\|$ , ein Banachraum und damit reflexiv. Demnach existiert eine in  $CR_0^1(\mathcal{T})$  schwach konvergente Teilfolge von  $(v_n)_{n \in \mathbb{N}}$ . Da  $CR_0^1(\mathcal{T})$  endlichdimensional ist, konvergiert diese sogar stark in  $L^2(\Omega)$ . Weil  $CR_0^1(\mathcal{T})$  ein Banachraum und damit abgeschlossen bezüglich der Konvergenz in  $\|\cdot\|$  ist, gilt für den Grenzwert  $u_{CR}$  dieser Teilfolge, dass  $u_{CR} \in CR_0^1(\mathcal{T})$ . Nach Lemma 4.2 ist  $E_{NC}$  stetig bezüglich der Konvergenz in  $L^2(\Omega)$ , was impliziert, dass  $u_{CR}$  Minimierer von  $E_{NC}$  in  $CR_0^1(\mathcal{T})$  sein muss. Dieser Minimierer  $u_{CR}$  ist eindeutig, da  $E_{NC}$  strikt konvex ist.  $\square$

Als Nächstes wollen wir äquivalente Charakterisierungen der eindeutigen Lösung von Problem 4.1 beweisen, die von Professor Carstensen formuliert wurden. Dazu leiten wir zunächst ein zu Problem 4.1 äquivalentes Minimaxproblem nach [Roc70, Section 36] her. Wir betrachten die konvexe Menge

$$K := \{\Lambda \in L^\infty(\Omega; \mathbb{R}^2) \mid |\Lambda(\bullet)| \leq 1 \text{ fast überall in } \Omega\}$$

und das dazugehörige Indikatorfunktional  $I_K : L^\infty(\Omega; \mathbb{R}^2) \rightarrow \mathbb{R} \cup \{\infty\}$ , das für  $\Lambda \in L^\infty(\Omega; \mathbb{R}^2)$  gegeben ist durch

$$I_K(\Lambda) := \begin{cases} \infty, & \text{falls } \Lambda \notin K, \\ 0, & \text{falls } \Lambda \in K. \end{cases}$$

Aufgrund der Konvexität von  $K$  ist  $I_K$  konvex. Für  $v_{CR} \in CR_0^1(\mathcal{T})$  und  $\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2) \subset L^\infty(\Omega; \mathbb{R}^2)$  können wir damit die Sattelfunktion  $L : CR_0^1(\mathcal{T}) \times P_0(\mathcal{T}; \mathbb{R}^2) \rightarrow [-\infty, \infty)$  nach [Roc70, Section 33] definieren durch

$$L(v_{CR}, \Lambda_0) := \int_{\Omega} \Lambda_0 \cdot \nabla_{NC}v_{CR} \, dx + \frac{\alpha}{2} \|v_{CR}\|^2 - \int_{\Omega} f v_{CR} \, dx - I_K(\Lambda_0). \quad (4.3)$$

#### 4 Das diskrete Problem

Nun wählen wir  $v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$  beliebig. Mit der Cauchy-Schwarzschen Ungleichung gilt für alle  $\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2) \cap K$ , dass

$$\int_{\Omega} \Lambda_0 \cdot \nabla_{\text{NC}} v_{\text{CR}} \, dx \leq \int_{\Omega} |\Lambda_0| |\nabla_{\text{NC}} v_{\text{CR}}| \, dx \leq \|\nabla_{\text{NC}} v_{\text{CR}}\|_{L^1(\Omega)}.$$

Daraus folgt

$$\sup_{\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2) \cap K} L(v_{\text{CR}}, \Lambda_0) \leq E_{\text{NC}}(v_{\text{CR}}). \quad (4.4)$$

Weiterhin gilt, wenn wir  $\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2) \cap K$  mit der Signumfunktion aus Gleichung (2.1) elementweise auf allen  $T \in \mathcal{T}$  definieren durch  $\Lambda_0(x) \in \text{sign}(\nabla_{\text{NC}} v_{\text{CR}}(x))$  für alle  $x \in \text{int}(T)$ , dass  $L(v_{\text{CR}}, \Lambda_0) = E_{\text{NC}}(v_{\text{CR}})$  und deshalb auch

$$E_{\text{NC}}(v_{\text{CR}}) \leq \sup_{\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2) \cap K} L(v_{\text{CR}}, \Lambda_0). \quad (4.5)$$

Außerdem ist  $L(v_{\text{CR}}, \Lambda_0) > -\infty$  genau dann, wenn  $\Lambda_0 \in K$ . Damit folgt aus den Ungleichungen (4.4) und (4.5) insgesamt

$$E_{\text{NC}}(v_{\text{CR}}) = \sup_{\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2) \cap K} L(v_{\text{CR}}, \Lambda_0) = \sup_{\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2)} L(v_{\text{CR}}, \Lambda_0).$$

Wenn also das folgende Minimaxproblem 4.4 eine Lösung  $(\tilde{u}_{\text{CR}}, \bar{\Lambda}_0) \in \text{CR}_0^1(\mathcal{T}) \times P_0(\mathcal{T}; \mathbb{R}^2)$  hat, dann löst die Funktion  $\tilde{u}_{\text{CR}}$  Problem 4.1.

**Problem 4.4.** Finde  $(\tilde{u}_{\text{CR}}, \bar{\Lambda}_0) \in \text{CR}_0^1(\mathcal{T}) \times P_0(\mathcal{T}; \mathbb{R}^2)$ , sodass

$$L(\tilde{u}_{\text{CR}}, \bar{\Lambda}_0) = \inf_{v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})} \sup_{\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2)} L(v_{\text{CR}}, \Lambda_0).$$

**Lemma 4.5.** Es existiert eine Lösung  $(\tilde{u}_{\text{CR}}, \bar{\Lambda}_0) \in \text{CR}_0^1(\mathcal{T}) \times (P_0(\mathcal{T}; \mathbb{R}^2) \cap K)$  von Problem 4.4.

*Beweis.* Die Sattelfunktion  $L$  aus Gleichung (4.3) ist, wenn ihre zweite Komponente in  $P_0(\mathcal{T}; \mathbb{R}^2) \cap K$  fixiert ist, in ihrer ersten Komponente eine konvexe, unterhalbstetige, auf  $\text{CR}_0^1(\mathcal{T})$  reellwertige Funktion und in ihrer zweiten Komponente eine konkave, oberhalbstetige, auf  $P_0(\mathcal{T}; \mathbb{R}^2) \cap K$  reellwertige Funktion. Somit ist  $L$  in beiden Komponenten abgeschlossen nach [Roc70, S. 52, 308]. Insgesamt ist  $L$  damit eine konvex-konkave, propere und abgeschlossene Funktion nach [Roc70, S. 349, 362 f.], deren effektiver Definitionsbereich nach [Roc70, S. 362] die Menge  $\text{CR}_0^1(\mathcal{T}) \times (P_0(\mathcal{T}; \mathbb{R}^2) \cap K)$  ist. Unter Beachtung der Isomorphie von  $\text{CR}_0^1(\mathcal{T})$  zu  $\mathbb{R}^{|\mathcal{E}(\Omega)|}$  und der Isomorphie von  $P_0(\mathcal{T}; \mathbb{R}^2)$  zu  $\mathbb{R}^{2|\mathcal{T}|}$ , folgt aus [Roc70, S. 397, Theorem 37.6] die Existenz eines Sattelpunkts  $(\tilde{u}_{\text{CR}}, \bar{\Lambda}_0) \in \text{CR}_0^1(\mathcal{T}) \times (P_0(\mathcal{T}; \mathbb{R}^2) \cap K)$  von  $L$  nach [Roc70, S. 380]. Für diesen impliziert [Roc70, S. 380, Lemma 36.2], dass

$$\sup_{\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2)} \inf_{v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})} L(v_{\text{CR}}, \Lambda_0) = L(\tilde{u}_{\text{CR}}, \bar{\Lambda}_0) = \inf_{v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})} \sup_{\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2)} L(v_{\text{CR}}, \Lambda_0).$$

Somit löst  $(\tilde{u}_{\text{CR}}, \bar{\Lambda}_0) \in \text{CR}_0^1(\mathcal{T}) \times (P_0(\mathcal{T}; \mathbb{R}^2) \cap K)$  Problem 4.4.  $\square$

Nachdem diese Vorbereitungen abgeschlossen sind, können wir nun folgendes Theorem beweisen.

**Theorem 4.6.** Für eine Funktion  $\tilde{u}_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$  sind die folgenden drei Aussagen äquivalent.

## 4.2 Charakterisierung und Existenz eines eindeutigen Minimierers

(i) Problem 4.1 wird von  $\tilde{u}_{\text{CR}}$  gelöst.

(ii) Es existiert ein  $\bar{\Lambda}_0 \in P_0(\mathcal{T}; \mathbb{R}^2)$  mit  $|\bar{\Lambda}_0(\bullet)| \leq 1$  fast überall in  $\Omega$ , sodass

$$\bar{\Lambda}_0(\bullet) \cdot \nabla_{\text{NC}} \tilde{u}_{\text{CR}}(\bullet) = |\nabla_{\text{NC}} \tilde{u}_{\text{CR}}(\bullet)| \quad \text{fast überall in } \Omega \quad (4.6)$$

und

$$(\bar{\Lambda}_0, \nabla_{\text{NC}} v_{\text{CR}}) = (f - \alpha \tilde{u}_{\text{CR}}, v_{\text{CR}}) \quad \text{für alle } v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T}). \quad (4.7)$$

(iii) Für alle  $v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$  gilt

$$(f - \alpha \tilde{u}_{\text{CR}}, v_{\text{CR}} - \tilde{u}_{\text{CR}}) \leq \|\nabla_{\text{NC}} v_{\text{CR}}\|_{L^1(\Omega)} - \|\nabla_{\text{NC}} \tilde{u}_{\text{CR}}\|_{L^1(\Omega)}. \quad (4.8)$$

*Beweis.* Sei  $\tilde{u}_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$ .

(i)  $\Rightarrow$  (ii). Sei  $\tilde{u}_{\text{CR}}$  Lösung von Problem 4.1. Nach Lemma 4.5 existiert eine Lösung  $(\hat{u}_{\text{CR}}, \bar{\Lambda}_0) \in \text{CR}_0^1(\mathcal{T}) \times (P_0(\mathcal{T}; \mathbb{R}^2) \cap K)$  von Problem 4.4. Außerdem wissen wir, dass damit  $\hat{u}_{\text{CR}}$  Lösung von Problem 4.1 ist. Daraus folgt, da nach Theorem 4.3 die Lösung von Problem 4.1 eindeutig ist, dass  $\hat{u}_{\text{CR}} = \tilde{u}_{\text{CR}}$  in  $\text{CR}_0^1(\mathcal{T})$ . Weiterhin wissen wir aus dem Beweis von Lemma 4.5, dass  $(\tilde{u}_{\text{CR}}, \bar{\Lambda}_0)$  Sattelpunkt der Funktion  $L$  aus Gleichung (4.3) ist. Das bedeutet nach [Roc70, S. 380] insbesondere, dass  $\tilde{u}_{\text{CR}}$  Minimierer von  $L(\bullet, \bar{\Lambda}_0)$  in  $\text{CR}_0^1(\mathcal{T})$  und  $\bar{\Lambda}_0$  Maximierer von  $L(\tilde{u}_{\text{CR}}, \bullet)$  in  $P_0(\mathcal{T}; \mathbb{R}^2)$  ist. Mit dieser Erkenntnis können wir nun die entsprechenden Optimalitätsbedingungen diskutieren. Zunächst gilt, da  $L(\tilde{u}_{\text{CR}}, \bullet) : P_0(\mathcal{T}; \mathbb{R}^2) \rightarrow [-\infty, \infty)$  konkav und  $\bar{\Lambda}_0$  Maximierer von  $L(\tilde{u}_{\text{CR}}, \bullet)$  in  $P_0(\mathcal{T}; \mathbb{R}^2)$  ist, dass das konvexe Funktional  $-L(\tilde{u}_{\text{CR}}, \bullet) : P_0(\mathcal{T}; \mathbb{R}^2) \rightarrow (-\infty, \infty]$  von  $\bar{\Lambda}_0$  in  $P_0(\mathcal{T}; \mathbb{R}^2)$  minimiert wird. Nach den Theoremen 2.7 – 2.9 gilt somit

$$0 \in \partial(-L(\tilde{u}_{\text{CR}}, \bullet))(\bar{\Lambda}_0) = \{-(\nabla_{\text{NC}} \tilde{u}_{\text{CR}}, \bullet)\} + \partial I_K(\bar{\Lambda}_0).$$

Äquivalent zu dieser Aussage ist, dass  $(\nabla_{\text{NC}} \tilde{u}_{\text{CR}}, \bullet) \in \partial I_K(\bar{\Lambda}_0)$ . Da  $\bar{\Lambda}_0 \in K$ , folgt mit Definition 2.6, dass für alle  $\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2)$  gilt

$$(\nabla_{\text{NC}} \tilde{u}_{\text{CR}}, \Lambda_0 - \bar{\Lambda}_0) \leq I_K(\Lambda_0) - I_K(\bar{\Lambda}_0) = I_K(\Lambda_0).$$

Falls  $\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2) \cap K$ , folgt insbesondere

$$\begin{aligned} (\nabla_{\text{NC}} \tilde{u}_{\text{CR}}, \Lambda_0 - \bar{\Lambda}_0) &\leq 0, \quad \text{also} \\ (\nabla_{\text{NC}} \tilde{u}_{\text{CR}}, \Lambda_0) &\leq (\nabla_{\text{NC}} \tilde{u}_{\text{CR}}, \bar{\Lambda}_0). \end{aligned} \quad (4.9)$$

Sei nun  $\Lambda_0 \in P_0(\mathcal{T}; \mathbb{R}^2) \cap K$  elementweise auf allen  $T \in \mathcal{T}$  durch  $\Lambda_0(x) \in \text{sign}(\nabla_{\text{NC}} \tilde{u}_{\text{CR}}(x))$  definiert für alle  $x \in \text{int}(T)$ . Mit dieser Wahl von  $\Lambda_0$ , Ungleichung (4.9), der Cauchy-Schwarzschen Ungleichung und  $\bar{\Lambda}_0 \in K$  erhalten wir die Abschätzung

$$\begin{aligned} \int_{\Omega} |\nabla_{\text{NC}} \tilde{u}_{\text{CR}}| dx &= \int_{\Omega} \nabla_{\text{NC}} \tilde{u}_{\text{CR}} \cdot \Lambda_0 dx \leq \int_{\Omega} \nabla_{\text{NC}} \tilde{u}_{\text{CR}} \cdot \bar{\Lambda}_0 dx \\ &\leq \int_{\Omega} |\nabla_{\text{NC}} \tilde{u}_{\text{CR}}| |\bar{\Lambda}_0| dx \leq \int_{\Omega} |\nabla_{\text{NC}} \tilde{u}_{\text{CR}}| dx, \quad \text{das heißt,} \\ \int_{\Omega} |\nabla_{\text{NC}} \tilde{u}_{\text{CR}}| dx &= \int_{\Omega} \nabla_{\text{NC}} \tilde{u}_{\text{CR}} \cdot \bar{\Lambda}_0 dx \quad \text{beziehungsweise} \\ \sum_{T \in \mathcal{T}} |T| |(\nabla_{\text{NC}} \tilde{u}_{\text{CR}})|_T &= \sum_{T \in \mathcal{T}} |T| |(\nabla_{\text{NC}} \tilde{u}_{\text{CR}} \cdot \bar{\Lambda}_0)|_T. \end{aligned} \quad (4.10)$$

Außerdem gilt für alle  $T \in \mathcal{T}$  mit der Cauchy-Schwarzschen Ungleichung und  $\bar{\Lambda}_0 \in K$ , dass

$$(\nabla_{\text{NC}} \tilde{u}_{\text{CR}} \cdot \bar{\Lambda}_0)|_T \leq |(\nabla_{\text{NC}} \tilde{u}_{\text{CR}})|_T |\bar{\Lambda}_0|_T \leq |(\nabla_{\text{NC}} \tilde{u}_{\text{CR}})|_T.$$

#### 4 Das diskrete Problem

Mit Gleichung (4.10) folgt daraus für alle  $T \in \mathcal{T}$ , dass  $(\nabla_{NC}\tilde{u}_{CR} \cdot \bar{\Lambda}_0)|_T = |(\nabla_{NC}\tilde{u}_{CR})|_T$ , das heißt, fast überall in  $\Omega$  gilt  $\bar{\Lambda}_0(\bullet) \cdot \nabla_{NC}\tilde{u}_{CR}(\bullet) = |\nabla_{NC}\tilde{u}_{CR}(\bullet)|$ . Damit ist Gleichung (4.6) gezeigt. Als Nächstes betrachten wir das reellwertige Funktional  $L(\bullet, \bar{\Lambda}_0) : CR_0^1(\mathcal{T}) \rightarrow \mathbb{R}$ . Es ist Fréchet-differenzierbar mit

$$dL(\bullet, \bar{\Lambda}_0)(\tilde{u}_{CR}; v_{CR}) = \int_{\Omega} \bar{\Lambda}_0 \cdot \nabla_{NC}v_{CR} dx + \alpha(\tilde{u}_{CR}, v_{CR}) - \int_{\Omega} fv_{CR} dx$$

für alle  $v_{CR} \in CR_0^1(\mathcal{T})$ . Da  $\tilde{u}_{CR}$  Minimierer von  $L(\bullet, \bar{\Lambda}_0)$  in  $CR_0^1(\mathcal{T})$  ist, gilt nach Theorem 2.4, dass  $0 = dL(\bullet, \bar{\Lambda}_0)(\tilde{u}_{CR}; v_{CR})$  für alle  $v_{CR} \in CR_0^1(\mathcal{T})$ . Diese Bedingung ist für alle  $v_{CR} \in CR_0^1(\mathcal{T})$  äquivalent zu  $(\bar{\Lambda}_0, \nabla_{NC}v_{CR}) = (f - \alpha\tilde{u}_{CR}, v_{CR})$ . Somit ist auch Gleichung (4.7) gezeigt.

(ii)  $\Rightarrow$  (iii). Die Funktion  $\bar{\Lambda}_0 \in P_0(\mathcal{T}; \mathbb{R}^2)$  erfülle  $|\bar{\Lambda}_0(\bullet)| \leq 1$  fast überall in  $\Omega$  sowie die Gleichungen (4.6) und (4.7). Sei  $v_{CR} \in CR_0^1(\mathcal{T})$ . Mit den Gleichungen (4.7) und (4.6) gilt

$$\begin{aligned} (f - \alpha\tilde{u}_{CR}, v_{CR} - \tilde{u}_{CR}) &= (\bar{\Lambda}_0, \nabla_{NC}v_{CR}) - (\bar{\Lambda}_0, \nabla_{NC}\tilde{u}_{CR}) \\ &= \int_{\Omega} \bar{\Lambda}_0 \cdot \nabla_{NC}v_{CR} dx - \int_{\Omega} |\nabla_{NC}\tilde{u}_{CR}| dx. \end{aligned} \quad (4.11)$$

Weiterhin gilt mit der Cauchy-Schwarzschen Ungleichung und  $|\bar{\Lambda}_0(\bullet)| \leq 1$  fast überall in  $\Omega$ , dass

$$\int_{\Omega} \bar{\Lambda}_0 \cdot \nabla_{NC}v_{CR} dx \leq \int_{\Omega} |\bar{\Lambda}_0| |\nabla_{NC}v_{CR}| dx \leq \int_{\Omega} |\nabla_{NC}v_{CR}| dx.$$

Zusammen mit Gleichung (4.11) folgt daraus Ungleichung (4.8).

(iii)  $\Rightarrow$  (i). Es gelte Ungleichung (4.8) für alle  $v_{CR} \in CR_0^1(\mathcal{T})$ , also

$$(f - \alpha\tilde{u}_{CR}, v_{CR} - \tilde{u}_{CR}) \leq \|\nabla_{NC}v_{CR}\|_{L^1(\Omega)} - \|\nabla_{NC}\tilde{u}_{CR}\|_{L^1(\Omega)}.$$

Nach Theorem 4.3 existiert eine eindeutige Lösung  $u_{CR} \in CR_0^1(\mathcal{T})$  von Problem 4.1. Wir haben bereits gezeigt, dass somit für alle  $v_{CR} \in CR_0^1(\mathcal{T})$  gilt

$$(f - \alpha u_{CR}, v_{CR} - u_{CR}) \leq \|\nabla_{NC}v_{CR}\|_{L^1(\Omega)} - \|\nabla_{NC}u_{CR}\|_{L^1(\Omega)}.$$

Um nun zu beweisen, dass  $\tilde{u}_{CR}$  Problem 4.1 löst, genügt es  $\tilde{u}_{CR} = u_{CR}$  in  $CR_0^1(\mathcal{T})$  zu zeigen. Es gilt

$$\begin{aligned} (f - \alpha u_{CR}, \tilde{u}_{CR} - u_{CR}) &\leq \|\nabla_{NC}\tilde{u}_{CR}\|_{L^1(\Omega)} - \|\nabla_{NC}u_{CR}\|_{L^1(\Omega)} \quad \text{und} \\ (f - \alpha\tilde{u}_{CR}, u_{CR} - \tilde{u}_{CR}) &\leq \|\nabla_{NC}u_{CR}\|_{L^1(\Omega)} - \|\nabla_{NC}\tilde{u}_{CR}\|_{L^1(\Omega)}. \end{aligned}$$

Die Addition dieser Ungleichungen impliziert

$$\alpha \|\tilde{u}_{CR} - u_{CR}\|^2 = (-\alpha u_{CR}, \tilde{u}_{CR} - u_{CR}) + (-\alpha\tilde{u}_{CR}, u_{CR} - \tilde{u}_{CR}) \leq 0.$$

Da  $\alpha > 0$ , folgt daraus  $\|\tilde{u}_{CR} - u_{CR}\|^2 = 0$ , also  $\tilde{u}_{CR} = u_{CR}$  in  $CR_0^1(\mathcal{T})$ .  $\square$

Zum Abschluss dieses Abschnitts wollen wir noch zwei Bemerkungen von Professor Carstensen erwähnen und kurz deren Gültigkeit begründen. Die erste Bemerkung ist eine äquivalente Charakterisierung der dualen Variable  $\bar{\Lambda}_0 \in P_0(\mathcal{T}; \mathbb{R}^2)$  aus Theorem 4.6 zur diskreten Lösung  $u_{CR} \in CR_0^1(\mathcal{T})$  von Problem 4.1.

*Bemerkung 4.7.* Dass  $\bar{\Lambda}_0 \in P_0(\mathcal{T}; \mathbb{R}^2)$  fast überall in  $\Omega$  Gleichung (4.6) und  $|\bar{\Lambda}_0(\bullet)| \leq 1$  erfüllt, ist äquivalent zu der Bedingung  $\bar{\Lambda}_0(x) \in \text{sign}(\nabla_{NC}u_{CR}(x))$  für alle  $x \in \text{int}(T)$  für alle  $T \in \mathcal{T}$ .

### 4.3 Verfeinerungsindikator und garantierter Energieschranken

*Beweis.* Dass die genannte Bedingung hinreichend ist, folgt direkt aus der Definition der Signumfunktion. Ihre Notwendigkeit folgt aus der folgenden Beobachtung. Da  $|\bar{\Lambda}_0(\cdot)| \leq 1$  fast überall in  $\Omega$ , ist Gleichung (4.6) eine Cauchy-Schwarzsche Ungleichung, bei der sogar Gleichheit gilt. Dies ist genau dann der Fall, wenn  $\bar{\Lambda}_0(\cdot)$  und  $\nabla_{NC} u_{CR}(\cdot)$  fast überall in  $\Omega$  linear abhängig sind.  $\square$

Daraus können wir folgern, unter welchen Umständen die duale Variable  $\bar{\Lambda}_0$  auf einem Dreieck  $T \in \mathcal{T}$  eindeutig bestimmt ist.

*Bemerkung 4.8.* Falls  $\nabla_{NC} u_{CR} \neq 0$  auf  $T \in \mathcal{T}$ , gilt nach Definition der Signumfunktion, dass  $\bar{\Lambda}_0 = \nabla_{NC} u_{CR} / |\nabla_{NC} u_{CR}|$  eindeutig bestimmt ist auf  $T$ . Im Allgemeinen ist  $\bar{\Lambda}_0$  allerdings nicht eindeutig bestimmbar. Betrachten wir zum Beispiel  $f \equiv 0$  in Problem 4.1 mit eindeutiger Lösung  $u_{CR} \equiv 0$  fast überall in  $\Omega$ . Dann erfüllt nach der diskreten Helmholtz-Zerlegung [Car09b, S. 193, Theorem 3.32] die Wahl  $\bar{\Lambda}_0 := \operatorname{Curl} v_C$  für ein beliebiges  $v_C \in S^1(\mathcal{T})$  mit  $|\operatorname{Curl} v_C| \leq 1$  die Charakterisierung (ii) aus Theorem 4.6.

## 4.3 Verfeinerungsindikator und garantierter Energieschranken

Professor Carstensen stellte für die numerischen Untersuchungen eine Aussage über eine garantierter untere Energieschranke und einen Verfeinerungsindikator zur adaptiven Netzverfeinerung zur Verfügung, die wir in diesem Abschnitt aufführen wollen.

**Theorem 4.9** (Garantierte untere Energieschranke). *Sei  $\Omega$  konvex,  $f \in H_0^1(\Omega)$  das Eingangssignal für Problem 3.1 mit Lösung  $u \in H_0^1(\Omega)$  sowie für Problem 4.1 mit Lösung  $u_{CR} \in CR_0^1(\mathcal{T})$ . Dann gilt*

$$E_{NC}(u_{CR}) + \frac{\alpha}{2} \|u - u_{CR}\|^2 - \frac{\kappa_{CR}}{\alpha} \|h_{\mathcal{T}}(f - \alpha u_{CR})\| \|\nabla f\| \leq E(u).$$

Dabei ist die Konstante  $\kappa_{CR} := \sqrt{1/48 + 1/j_{1,1}^2}$  mit der kleinsten positiven Nullstelle  $j_{1,1}$  der Bessel-Funktion erster Art. Insbesondere gilt dann für

$$E_{GLEB,\mathcal{T}} := E_{NC}(u_{CR}) - \frac{\kappa_{CR}}{\alpha} \|h_{\mathcal{T}}(f - \alpha u_{CR})\| \|\nabla f\|, \quad (4.12)$$

dass  $E_{NC}(u_{CR}) \geq E_{GLEB,\mathcal{T}}$  und  $E(u) \geq E_{GLEB,\mathcal{T}}$ .

**Definition 4.10** (Verfeinerungsindikator). Für  $d \in \mathbb{N}$  (in dieser Arbeit stets  $d = 2$ ) und  $0 < \gamma \leq 1$  definieren wir für alle  $T \in \mathcal{T}$  und  $u_{CR} \in CR_0^1(\mathcal{T})$  die Funktionen

$$\eta_{V,\mathcal{T}}(T) := |T|^{2/d} \|f - \alpha u_{CR}\|_{L^2(T)}^2 \quad \text{und}$$

$$\eta_{J,\mathcal{T}}(T) := |T|^{\gamma/d} \sum_{F \in \mathcal{E}(T)} \| [u_{CR}]_F \|_{L^1(F)}.$$

Damit definieren wir den Verfeinerungsindikator  $\eta_{\mathcal{T}} := \sum_{T \in \mathcal{T}} \eta_{\mathcal{T}}(T)$ , wobei

$$\eta_{\mathcal{T}}(T) := \eta_{V,\mathcal{T}}(T) + \eta_{J,\mathcal{T}}(T) \quad \text{für alle } T \in \mathcal{T}. \quad (4.13)$$

Darüber hinaus können wir eine garantierter obere Energieschranke formulieren. Dabei nutzen wir den Operator  $J_{1,\mathcal{T}} : CR^1(\mathcal{T}) \rightarrow P_1(\mathcal{T}) \cap C_0(\Omega)$  (cf. [CH18, Section 4]), wobei  $J_{1,\mathcal{T}} v_{CR}$  für eine Funktion  $v_{CR} \in CR^1(\mathcal{T})$  in allen Innenknoten  $z \in \mathcal{N}(\Omega)$  definiert ist durch

$$J_{1,\mathcal{T}} v_{CR}(z) := |\mathcal{T}(z)|^{-1} \sum_{T \in \mathcal{T}(z)} v_{CR}|_T(z). \quad (4.14)$$

Da für die Lösung  $u$  von Problem 3.1 und die Lösung  $u_{CR}$  von Problem 4.1 gilt

$$E(u) \leq E(J_{1,\mathcal{T}} u_{CR}) = E_{NC}(J_{1,\mathcal{T}} u_{CR}) =: E_{GUEB,\mathcal{T}}, \quad (4.15)$$

wählen wir  $E_{GUEB,\mathcal{T}}$  als garantierter obere Energieschranke.

# 5 Iterative Lösung

## 5.1 Primale-duale Iteration

In diesem Abschnitt präsentieren wir ein iteratives Verfahren mit dem wir Problem 4.1 numerisch lösen möchten. Dieses basiert auf der primalen-dualen Iteration [Bar15b, S. 314, Algorithm 10.1] unter Beachtung von [Bar15b, S. 314, Remark 10.11]. Details dazu und weitere Referenzen finden sich in [Bar12; Bar15b, S. 118-121]. Angepasst an unser Problem und die Notation dieser Arbeit lautet der Algorithmus wie folgt.

**Algorithmus 5.1** (Primale-duale Iteration).

**Input:**  $(u_0, \Lambda_0) \in \text{CR}_0^1(\mathcal{T}) \times P_0(\mathcal{T}; \overline{B_{\mathbb{R}^2}}), \tau > 0$

Initialisiere  $v_0 := 0$  in  $\text{CR}_0^1(\mathcal{T})$ .

**for**  $j = 1, 2, \dots$

$$\begin{aligned}\tilde{u}_j &:= u_{j-1} + \tau v_{j-1}, \\ \Lambda_j &:= \frac{\Lambda_{j-1} + \tau \nabla_{\text{NC}} \tilde{u}_j}{\max \{1, |\Lambda_{j-1} + \tau \nabla_{\text{NC}} \tilde{u}_j|\}},\end{aligned}\tag{5.1}$$

bestimme  $u_j \in \text{CR}_0^1(\mathcal{T})$  als Lösung des linearen Gleichungssystems

$$\begin{aligned}\frac{1}{\tau} a_{\text{NC}}(u_j, \cdot) + \alpha(u_j, \cdot) &= \frac{1}{\tau} a_{\text{NC}}(u_{j-1}, \cdot) + (f, \cdot) - (\Lambda_j, \nabla_{\text{NC}} \cdot) \quad \text{in } \text{CR}_0^1(\mathcal{T}), \\ v_j &:= \frac{u_j - u_{j-1}}{\tau}.\end{aligned}\tag{5.2}$$

**Output:** Folge  $(u_j, \Lambda_j)_{j \in \mathbb{N}}$  in  $\text{CR}_0^1(\mathcal{T}) \times P_0(\mathcal{T}; \overline{B_{\mathbb{R}^2}})$

In [Bar15b] wird in Gleichung (5.2) anstelle des diskreten Skalarprodukts  $a_{\text{NC}}(\cdot, \cdot)$  ein diskretes Skalarprodukt  $(\cdot, \cdot)_{h,s}$ , welches ungleich dem  $L^2$ -Skalarprodukt sein kann, genutzt. Die Iteration wird ebenda abgebrochen, wenn die von  $(\cdot, \cdot)_{h,s}$  induzierte Norm des Terms  $(u_j - u_{j-1})/\tau$  kleiner einer gegebenen Toleranz ist. Dementsprechend nutzen wir mit  $\varepsilon_{\text{stop}} > 0$  für Algorithmus 5.1 das Abbruchkriterium

$$\left\| \frac{u_j - u_{j-1}}{\tau} \right\|_{\text{NC}} < \varepsilon_{\text{stop}}.\tag{5.3}$$

*Bemerkung 5.2.* Mit den kantenorientierten Crouzeix-Raviart-Basisfunktionen  $\{\psi_E \mid E \in \mathcal{E}\}$  aus Abschnitt 2.2 können wir die Steifigkeitsmatrix  $A \in \mathbb{R}^{|\mathcal{E}| \times |\mathcal{E}|}$  und die Massenmatrix  $M \in \mathbb{R}^{|\mathcal{E}| \times |\mathcal{E}|}$  für alle  $k, \ell \in \{1, 2, \dots, |\mathcal{E}|\}$  definieren durch

$$A_{k\ell} := a_{\text{NC}}(\psi_{E_k}, \psi_{E_\ell}) \quad \text{und} \quad M_{k\ell} := (\psi_{E_k}, \psi_{E_\ell}).$$

Außerdem definieren wir mit  $u_{j-1}$  und  $\Lambda_j$  aus Algorithmus 5.1 den Vektor  $b \in \mathbb{R}^{|\mathcal{E}|}$  für alle  $k \in \mathbb{N}$  durch

$$b_k := \left( \frac{1}{\tau} \nabla_{\text{NC}} u_{j-1} - \Lambda_j, \nabla_{\text{NC}} \psi_{E_k} \right) + (f, \psi_{E_k}).$$

Sei nun  $\mathcal{E} = \{E_1, E_2, \dots, E_{|\mathcal{E}|}\}$  und sei ohne Beschränkung der Allgemeinheit  $\mathcal{E}(\Omega) := \{E_1, E_2, \dots, E_{|\mathcal{E}(\Omega)|}\}$ . Dann ist  $J := \{|\mathcal{E}(\Omega)| + 1, |\mathcal{E}(\Omega)| + 2, \dots, |\mathcal{E}|\}$  die Menge der Indizes der Randkanten in  $\mathcal{E}$ . Damit definieren wir die Matrix  $\bar{A} \in \mathbb{R}^{|\mathcal{E}(\Omega)| \times |\mathcal{E}(\Omega)|}$ , die durch

Streichen der Zeilen und Spalten von  $A$  mit den Indizes aus  $J$  entsteht, die Matrix  $\bar{M} \in \mathbb{R}^{|\mathcal{E}(\Omega)| \times |\mathcal{E}(\Omega)|}$ , die ebenso aus  $M$  hervorgeht und den Vektor  $\bar{b} \in \mathbb{R}^{|\mathcal{E}(\Omega)|}$ , der durch Streichen der Komponenten von  $b$  mit Indizes in  $J$  entsteht. Weiterhin sei  $x \in \mathbb{R}^{|\mathcal{E}(\Omega)|}$ , wobei  $x_k$  für alle  $k \in \{1, 2, \dots, |\mathcal{E}(\Omega)|\}$  der Koeffizient der Lösung  $u_j \in \text{CR}_0^1(\mathcal{T})$  des Gleichungssystems (5.2) zur  $k$ -ten Basisfunktion von  $\text{CR}_0^1(\mathcal{T})$  sei, das heißt, es gelte

$$u_j = \sum_{k=1}^{|\mathcal{E}(\Omega)|} x_k \psi_{E_k}.$$

Da wir das Gleichungssystem (5.2) in  $\text{CR}_0^1(\mathcal{T})$  lösen, lässt sich somit  $u_j$  durch Lösen einer Matrixgleichung nach  $x$  bestimmen. Diese lautet

$$\left( \frac{1}{\tau} \bar{A} + \alpha \bar{M} \right) x = \bar{b}. \quad (5.4)$$

## 5.2 Konvergenz der Iteration

In diesem Abschnitt beweisen wir die Konvergenz der Iterate von Algorithmus 5.1 gegen die Lösung von Problem 4.1. Dabei bedienen wir uns unter anderem der äquivalenten Charakterisierungen aus Theorem 4.6.

**Theorem 5.3.** *Sei  $u_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$  Lösung von Problem 4.1 und  $\bar{\Lambda}_0 \in P_0(\mathcal{T}; \mathbb{R}^2)$  erfülle  $|\bar{\Lambda}_0(\cdot)| \leq 1$  fast überall in  $\Omega$  sowie Gleichung (4.6) und Gleichung (4.7) aus Theorem 4.6 mit  $\tilde{u}_{\text{CR}} = u_{\text{CR}}$ . Falls  $\tau \in (0, 1]$ , dann konvergieren die Iterate  $(u_j)_{j \in \mathbb{N}}$  von Algorithmus 5.1 in  $L^2(\Omega)$  gegen  $u_{\text{CR}}$ .*

*Beweis.* Der Beweis folgt einer Skizze von Professor Carstensen.

Sei  $j \in \mathbb{N}$ . Seien weiterhin  $u_0, \Lambda_0$  und  $v_0$  sowie  $\tilde{u}_j, \Lambda_j, u_j$  und  $v_j$  definiert wie in Algorithmus 5.1. Außerdem definieren wir  $\mu_j := \max\{1, |\Lambda_{j-1} + \tau \nabla_{\text{NC}} \tilde{u}_j|\}$  und für alle  $k \in \mathbb{N}_0$  die Abkürzungen  $e_k := u_{\text{CR}} - u_k, E_k := \bar{\Lambda}_0 - \Lambda_k$ . Dabei nutzen wir die Konvention  $e_{-1} := e_0$ . Wir testen zunächst (5.2) mit  $e_j$  und formen das Resultat um. Damit erhalten wir

$$a_{\text{NC}}(v_j, e_j) + \alpha(u_j, e_j) + (\Lambda_j, \nabla_{\text{NC}} e_j) = (f, e_j).$$

Zusammen mit Gleichung (4.7) folgt daraus

$$a_{\text{NC}}(v_j, e_j) = \alpha(u_{\text{CR}} - u_j, e_j) + (\bar{\Lambda}_0 - \Lambda_j, \nabla_{\text{NC}} e_j) = \alpha \|e_j\|^2 + (E_j, \nabla_{\text{NC}} e_j). \quad (5.5)$$

Als Nächstes betrachten wir Gleichung (5.1). Es gilt

$$\Lambda_{j-1} - \Lambda_j + \tau \nabla_{\text{NC}} \tilde{u}_j = (\mu_j - 1) \Lambda_j \quad \text{fast überall in } \Omega. \quad (5.6)$$

Außerdem folgt aus Gleichung (5.1) und einer Fallunterscheidung zwischen  $1 \geq |\Lambda_{j-1} + \tau \nabla_{\text{NC}} \tilde{u}_j|$  und  $1 < |\Lambda_{j-1} + \tau \nabla_{\text{NC}} \tilde{u}_j|$ , dass

$$(1 - |\Lambda_j|)(\mu_j - 1) = 0 \quad \text{fast überall in } \Omega. \quad (5.7)$$

Testen wir nun Gleichung (5.6) in  $L^2(\Omega)$  mit  $E_j$ , erhalten wir unter Nutzung von  $\mu_j \geq 1$ , der Cauchy-Schwarzschen Ungleichung,  $|\bar{\Lambda}_0(\cdot)| \leq 1$  fast überall in  $\Omega$  und Gleichung (5.7), dass

$$\begin{aligned} (\Lambda_{j-1} - \Lambda_j + \tau \nabla_{\text{NC}} \tilde{u}_j, E_j) &= ((\mu_j - 1) \Lambda_j, \bar{\Lambda}_0 - \Lambda_j) \\ &\leq \int_{\Omega} (\mu_j - 1) (|\Lambda_j| - |\Lambda_j|^2) \, dx \\ &= \int_{\Omega} |\Lambda_j| (1 - |\Lambda_j|)(\mu_j - 1) \, dx = 0. \end{aligned}$$

## 5 Iterative Lösung

Daraus folgt mit  $\Lambda_{j-1} - \Lambda_j = E_j - E_{j-1}$  und  $\tilde{u}_j = u_{j-1} - (e_{j-1} - e_{j-2})$ , dass nach Division durch  $\tau$  gilt

$$\left( \frac{E_j - E_{j-1}}{\tau} + \nabla_{\text{NC}} u_{j-1} - \nabla_{\text{NC}}(e_{j-1} - e_{j-2}), E_j \right) \leq 0. \quad (5.8)$$

Aus der Cauchy-Schwarzschen Ungleichung, Gleichung (4.6) und  $|\Lambda_j(\bullet)| \leq 1$  fast überall in  $\Omega$  folgt, dass

$$\begin{aligned} \nabla_{\text{NC}} u_{\text{CR}} \cdot E_j &= \nabla_{\text{NC}} u_{\text{CR}} \cdot \bar{\Lambda}_0 - \nabla_{\text{NC}} u_{\text{CR}} \cdot \Lambda_j \\ &\geq \nabla_{\text{NC}} u_{\text{CR}} \cdot \bar{\Lambda}_0 - |\nabla_{\text{NC}} u_{\text{CR}}| |\Lambda_j| \\ &= |\nabla_{\text{NC}} u_{\text{CR}}| (1 - |\Lambda_j|) \geq 0 \quad \text{fast überall in } \Omega. \end{aligned}$$

Daraus folgt

$$(\nabla_{\text{NC}} u_{\text{CR}}, E_j) = \int_{\Omega} \nabla_{\text{NC}} u_{\text{CR}} \cdot E_j \, dx \geq 0. \quad (5.9)$$

Aus den Ungleichungen (5.8) und (5.9) folgt insgesamt

$$\left( \frac{E_j - E_{j-1}}{\tau} + \nabla_{\text{NC}} u_{j-1} - \nabla_{\text{NC}}(e_{j-1} - e_{j-2}), E_j \right) \leq (\nabla_{\text{NC}} u_{\text{CR}}, E_j).$$

Das ist äquivalent zu

$$\left( \frac{E_j - E_{j-1}}{\tau} - \nabla_{\text{NC}}(2e_{j-1} - e_{j-2}), E_j \right) \leq 0. \quad (5.10)$$

Weiterhin gilt

$$\begin{aligned} \|e_j\|_{\text{NC}}^2 - \|e_{j-1}\|_{\text{NC}}^2 + \|E_j\|^2 - \|E_{j-1}\|^2 + \|e_j - e_{j-1}\|_{\text{NC}}^2 + \|E_j - E_{j-1}\|^2 \\ = 2a_{\text{NC}}(e_j, e_j - e_{j-1}) + 2(E_j, E_j - E_{j-1}). \end{aligned} \quad (5.11)$$

Unter Nutzung von  $e_j - e_{j-1} = -\tau v_j$  und Gleichung (5.5) gilt außerdem

$$\begin{aligned} 2a_{\text{NC}}(e_j, e_j - e_{j-1}) + 2(E_j, E_j - E_{j-1}) &= -2\tau a_{\text{NC}}(e_j, v_j) + 2(E_j, E_j - E_{j-1}) \\ &= -2\tau \alpha \|e_j\|^2 + 2\tau \left( E_j, -\nabla_{\text{NC}} e_j + \frac{E_j - E_{j-1}}{\tau} \right). \end{aligned}$$

Daraus folgt durch Ungleichung (5.10) zusammen mit  $\tau > 0$ , dass

$$\begin{aligned} 2a_{\text{NC}}(e_j, e_j - e_{j-1}) + 2(E_j, E_j - E_{j-1}) \\ \leq -2\tau \alpha \|e_j\|^2 + 2\tau \left( E_j, -\nabla_{\text{NC}} e_j + \frac{E_j - E_{j-1}}{\tau} \right) \\ - 2\tau \left( \frac{E_j - E_{j-1}}{\tau} - \nabla_{\text{NC}}(2e_{j-1} - e_{j-2}), E_j \right) \\ = -2\tau \alpha \|e_j\|^2 - 2\tau (E_j, \nabla_{\text{NC}}(e_j - 2e_{j-1} + e_{j-2})). \end{aligned}$$

Damit und mit Gleichung (5.11) erhalten wir insgesamt

$$\begin{aligned} \|e_j\|_{\text{NC}}^2 - \|e_{j-1}\|_{\text{NC}}^2 + \|E_j\|^2 - \|E_{j-1}\|^2 + \|e_j - e_{j-1}\|_{\text{NC}}^2 + \|E_j - E_{j-1}\|^2 \\ \leq -2\tau \alpha \|e_j\|^2 - 2\tau (E_j, \nabla_{\text{NC}}(e_j - 2e_{j-1} + e_{j-2})). \end{aligned}$$

Für jedes  $J \in \mathbb{N}$  führt die Summation dieser Ungleichung über  $j = 1, \dots, J$  und eine Äquivalenzumformung zu

$$\begin{aligned} & \|e_J\|_{\text{NC}}^2 + \|E_J\|^2 + \sum_{j=1}^J (\|e_j - e_{j-1}\|_{\text{NC}}^2 + \|E_j - E_{j-1}\|^2) \\ & \leq \|e_0\|_{\text{NC}}^2 + \|E_0\|^2 - 2\tau\alpha \sum_{j=1}^J \|e_j\|^2 - 2\tau \sum_{j=1}^J (E_j, \nabla_{\text{NC}}(e_j - 2e_{j-1} + e_{j-2})). \end{aligned} \quad (5.12)$$

Für die letzte Summe auf der rechten Seite dieser Ungleichung gilt, unter Beachtung von  $e_{-1} = e_0$ , dass

$$\begin{aligned} & \sum_{j=1}^J (E_j, \nabla_{\text{NC}}(e_j - 2e_{j-1} + e_{j-2})) \\ &= \sum_{j=1}^J (E_j, \nabla_{\text{NC}}(e_j - e_{j-1})) - \sum_{j=0}^{J-1} (E_{j+1}, \nabla_{\text{NC}}(e_j - e_{j-1})) \\ &= \sum_{j=1}^{J-1} (E_j - E_{j+1}, \nabla_{\text{NC}}(e_j - e_{j-1})) + (E_J, \nabla_{\text{NC}}(e_J - e_{J-1})) - (E_1, \nabla_{\text{NC}}(e_0 - e_{-1})) \\ &= \sum_{j=1}^{J-1} (E_j - E_{j+1}, \nabla_{\text{NC}}(e_j - e_{j-1})) + (E_J, \nabla_{\text{NC}}(e_J - e_{J-1})). \end{aligned}$$

Mit dieser Umformung erhalten wir aus Ungleichung (5.12) für jedes  $\tau \in (0, 1]$ , das heißt,  $\tau^{-1} \geq 1$ , dass

$$\begin{aligned} & \|e_J\|_{\text{NC}}^2 + \|E_J\|^2 + \sum_{j=1}^J (\|e_j - e_{j-1}\|_{\text{NC}}^2 + \|E_j - E_{j-1}\|^2) \\ & \leq \tau^{-1} (\|e_0\|_{\text{NC}}^2 + \|E_0\|^2) - 2\alpha \sum_{j=1}^J \|e_j\|^2 \\ & \quad - 2 \sum_{j=1}^{J-1} (E_j - E_{j+1}, \nabla_{\text{NC}}(e_j - e_{j-1})) - 2(E_J, \nabla_{\text{NC}}(e_J - e_{J-1})). \end{aligned} \quad (5.13)$$

Außerdem gilt

$$\begin{aligned} 2\alpha \sum_{j=1}^J \|e_j\|^2 &\leq 2\alpha \sum_{j=1}^J \|e_j\|^2 + \|E_J + \nabla_{\text{NC}}(e_J - e_{J-1})\|^2 + \|e_J\|_{\text{NC}}^2 + \|E_1 - E_0\|^2 \\ &\quad + \sum_{j=1}^{J-1} \|\nabla_{\text{NC}}(e_j - e_{j-1}) - (E_{j+1} - E_j)\|^2 \\ &= 2\alpha \sum_{j=1}^J \|e_j\|^2 + \|e_J\|_{\text{NC}}^2 + \|E_J\|^2 + \sum_{j=1}^J (\|e_j - e_{j-1}\|_{\text{NC}}^2 + \|E_j - E_{j-1}\|^2) \\ &\quad + 2 \sum_{j=1}^{J-1} (E_j - E_{j+1}, \nabla_{\text{NC}}(e_j - e_{j-1})) + 2(E_J, \nabla_{\text{NC}}(e_J - e_{J-1})). \end{aligned}$$

Zusammen mit Ungleichung (5.13) folgt daraus

$$2\alpha \sum_{j=1}^J \|e_j\|^2 \leq \tau^{-1} (\|e_0\|_{\text{NC}}^2 + \|E_0\|^2). \quad (5.14)$$

## 5 Iterative Lösung

Somit gilt, dass  $\sum_{j=1}^{\infty} \|e_j\|^2$  beschränkt ist, was impliziert  $\|u_{\text{CR}} - u_j\| = \|e_j\| \rightarrow 0$  für  $j \rightarrow \infty$ .  $\square$

# 6 Implementierung

## 6.1 Hinweise zur Benutzung des Programms

Ziel der für diese Arbeit implementierten Methoden ist die Realisierung von Algorithmus 5.1 im Solve-Schritt des AFEM-Algorithmus aus Abbildung 6.1. Wir gehen davon aus, dass dieser und die im AFEM-Softwarepaket [Car09a] realisierten Methoden sowie deren Datenstrukturen bekannt sind und verweisen für weitere Details auf [Car+10]. Im Estimate-Schritt nutzen wir anstelle eines Fehlerschätzers den Verfeinerungsindikator aus Definition 4.10. Alle in diesem Kapitel aufgeführten Pfadnamen sind relative Pfade in Relation zum Ordner `code` im Anhang. Die zur korrekten Funktionsweise dieses Programms nötigen Methoden und Dateien des AFEM-Softwarepakets sind enthalten in den Ordnern

```
./utils/afemPackage/ sowie ./utils/geometries/.
```

Dort ist eine veränderte Version der Methode `plotTriangulation` gespeichert, die nun den Plot einer Triangulierung als Schwarz-Weiß-Bild ohne Titel erzeugt. Alle Ein- und Ausgabeparameter der im Rahmen dieser Arbeit implementierten Methoden sind in den entsprechenden Dateien im Anhang detailliert dokumentiert. Ausgenommen davon sind die bekannten Datenstrukturen aus dem AFEM-Softwarepaket, welche die Triangulierung beschreiben. Diese werden lediglich aufgelistet. Entwickelt und getestet wurde das Programm in Matlab R2017b. Die mathematischen Grundlagen für die Realisierung einiger Methoden diskutieren wir in Abschnitt 6.4.

Das Hauptprogramm, das heißt die ausführbare Methode, welche den AFEM-Algorithmus realisiert, ist

```
./nonconforming/startAlgorithmCR.m.
```

Beim Ausführen dieser Methode sollte das aktuelle Verzeichnis `./nonconforming/` sein. Als optionaler Eingabeparameter ist dabei ein nichtleeres `char array` mit genau einer Zeile möglich, das den Namen einer Matlab-Funktion enthält. Diese muss die Parameter und Einstellungen für ein Experiment als Felder eines Structure Arrays zurückgeben und in

```
./nonconforming/benchmarks/
```

gespeichert sein. Als Muster für eine solche Benchmark-Datei dient

```
./nonconforming/benchmarks/editable.m.
```

Wird das Hauptprogramm ohne Übergabe eines Parameters ausgeführt, nutzt es den Standardwert `benchmark = 'editable'`. Für jedes in dieser Arbeit dokumentierte Experiment verweisen wir in Kapitel 7 an entsprechender Stelle auf das dafür benutzte Benchmark, welches in `./nonconforming/benchmarks/` hinterlegt ist und somit die Reproduzierbarkeit



Abbildung 6.1: AFEM-Schleife

## 6 Implementierung

des Experiments garantiert. Eine Übersicht über die in einer Benchmark-Datei wählbaren Parameter ist zu finden in den Tabellen 6.1 – 6.5 sowie über die definierbaren Funktionen in Tabelle 6.6. Details zu den Datentypen sind in `editable.m` aufgeführt. Dass die zahlreichen Parameter, die während des Programmablaufs über- oder ausgegeben werden müssen, als Felder von Structure Arrays gespeichert werden, dient der Modifizierbarkeit des Programms. So haben Änderungen am Programm häufig nur zur Folge, dass einige `structs` um Felder ergänzt werden müssen, während Methodenköpfe unverändert bleiben können. Das Eingangssignal  $f$  und eventuell weitere Funktionen, wie die exakte Lösung  $u$  von Problem 3.1 oder die schwache Ableitung  $\nabla f$  eines schwach differenzierbaren Eingangssignals, müssen in der Benchmark-Datei definiert werden, um sie dem Programm zu übergeben. Die für die Experimente in Kapitel 7 genutzten Funktionen sind zu finden in

```
./utils/functions/.
```

Ist eine Lösung  $u \in H_0^1(\Omega)$  von Problem 3.1 bekannt, so kann die exakte Energie  $E(u)$  approximiert werden mit der Methode

```
./nonconforming/computeExactEnergyBV.m
```

Die so berechneten Energien werden gespeichert im Ordner

```
./nonconforming/knownExactEnergies/
```

und können anschließend manuell in ein Benchmark aufgenommen werden. Soll als Eingangssignal ein Graufarbenbild gegeben sein, so muss es in einem mit der Matlab-Funktion `imread` kompatiblen Format gespeichert sein in

```
./utils/functions/images/.
```

Um Dirichlet-Nullranddaten des Bildes zu garantieren, was einem schwarzen Rand entspricht, kann die Methode

```
./utils/functions/images/addBoundary2image.m
```

genutzt werden. Diese fügt wahlweise einen graduellen Übergang zu schwarzem Rand auf den äußeren 25 Pixeln des Bildes hinzu oder ergänzt das Bild um einen 10 Pixel breiten schwarzen Rand. Um additives weißes gaußsches Rauschen zu einem Bild hinzuzufügen, kann die Methode

```
./utils/functions/images/addNoise2image.m
```

genutzt werden, welche dafür die Matlab-Funktion `awgn` verwendet. Um diese Methode ausführen zu können, muss die `Communications System Toolbox` installiert sein.

## 6.2 Programmablauf

In diesem Abschnitt betrachten wir eine beispielhafte Ausführung des Programms, bei der alle relevanten Methoden genutzt werden. In der Benchmark-Datei für den Programmablauf müssen dafür, neben dem Eingangssignal  $f \in H_0^1(\Omega)$ , die Lösung  $u \in H_0^1(\Omega)$  von Problem 3.1 mit Eingangssignal  $f$ , die exakte Energie  $E(u)$  und der schwache Gradient von  $f$  gegeben sein. Damit diese Informationen tatsächlich genutzt werden, müssen zusätzlich einige Parameter aus Tabelle 6.4 passend gewählt werden. Diese Wahlen sind

- `useImage = false`,
- `inSiGradientKnown = true`,

Parametername	Beschreibung
<code>showPlots</code>	<code>true</code> , wenn während des Programmablaufs Plots angezeigt werden sollen, sonst <code>false</code>
<code>plotModeGrayscale</code>	<code>true</code> , wenn während des Programmablaufs Plots von Funktionen als Graufarbenbilder mit Blick aus positiver z-Richtung auf die xy-Ebene angezeigt werden sollen, sonst <code>false</code> (Wahl irrelevant, wenn <code>showPlots==false</code> )
<code>showProgress</code>	<code>true</code> , wenn während der primalen-dualen Iteration Informationen über den aktuellen Iterationsschritt angezeigt werden sollen, sonst <code>false</code>
<code>degree4Integrate</code>	algebraischer Exaktheitsgrad für Aufrufe der Methode <code>integrate</code> [Car+10, Abschnitt 1.8.2] des AFEM-Software-pakets
<code>plotGivenFunctions</code>	<code>true</code> , wenn Plots des Eingangssignals $f$ und, falls angegeben, der Lösung $u$ von Problem 3.1 erstellt und gespeichert werden sollen, sonst <code>false</code>
<code>refinementLevel4Plots</code>	Anzahl der Rotverfeinerungen der geladenen Geometrie, um Plots des Eingangssignal $f$ und, falls angegeben, der Lösung $u$ von Problem 3.1 zu erstellen (Wahl irrelevant, falls <code>plotGivenFunctions==false</code> )
<code>debugIfError</code>	<code>true</code> , wenn Matlab beim Auftreten eines Fehlers den Debug-Modus starten soll, sonst <code>false</code>

Tabelle 6.1: Diverse Parameter zur Kontrolle des Programmablaufs

Parametername	Beschreibung
<code>geometry</code>	Name der Geometrie auf deren Triangulierung der AFEM-Algorithmus angewendet werden soll (wird <code>'Square'</code> gesetzt, falls <code>useImage</code> aus Tabelle 6.4 den Wert <code>true</code> hat)
<code>initialRefinementLevel</code>	Anzahl der Rotverfeinerungen, die auf die Triangulierung der Geometrie <code>geometry</code> angewendet werden sollen vor Start des AFEM-Algorithmus (cf. <code>loadGeometry</code> in [Car+10, Abschnitt 1.9.1])
<code>parTheta</code>	Bulk-Parameter $\theta$ für den Mark-Schritt des AFEM-Algorithmus ( $\theta \in (0, 1)$ für adaptive und $\theta = 1$ für uniforme Netzverfeinerung)
<code>minNrDof</code>	Anzahl der Freiheitsgrade der Triangulierung eines Levels, die mindestens erreicht werden soll, bevor der AFEM-Algorithmus abbricht
<code>useProlongation</code>	<code>true</code> , wenn eine Prolongation der Lösung der primalen-dualen Iteration als Startwert für die Iteration des nächsten Levels genutzt werden soll, sonst <code>false</code>
<code>parGamma</code>	Parameter $\gamma$ aus Definition 4.10
<code>d</code>	Dimension $d$ aus Definition 4.10 (muss für diese Implementierung stets 2 sein)

Tabelle 6.2: Parameter für den AFEM-Algorithmus

## 6 Implementierung

Parametername	Beschreibung
<code>u0Mode</code>	Startwert $u_0$ für Algorithmus 5.1 auf dem ersten Level und, falls <code>useProlongation==false</code> , für die Iteration auf allen Leveln der AFEM-Routine (' <code>zeros</code> ' für $u_0 = 0$ und ' <code>interpolationInSi</code> ', falls für $u_0$ die Crouzeix-Raviart-Interpolation des Eingangssignals $f$ genutzt werden soll)
<code>epsStop</code>	Parameter $\varepsilon_{\text{stop}}$ für das Abbruchkriterium (5.3) von Algorithmus 5.1
<code>parTau</code>	Parameter $\tau$ aus Algorithmus 5.1
<code>maxIter</code>	Anzahl der Iterationsschritte der primalen-dualen Iteration, die auf jedem Level maximal durchgeführt werden sollen

Tabelle 6.3: Parameter für die primale-duale Iteration

Parametername	Beschreibung
<code>useImage</code>	<code>true</code> , wenn als Eingangssignal $f$ ein quadratisches Graufarbenbild anstelle eines <code>function_handle</code> genutzt werden soll, sonst <code>false</code>
<code>imageName</code>	Name (mit Dateiendung) eines quadratischen Graufarbenbildes im Ordner <code>./utils/functions/images/</code> , welches als Eingangssignal $f$ genutzt werden soll (Wahl irrelevant, falls <code>useImage==false</code> )
<code>parAlpha</code>	Parameter $\alpha$ aus Problem 3.1 und Problem 4.1
<code>parBeta</code>	Parameter $\beta$ einiger Experimente aus Kapitel 7
<code>inSiGradientKnown</code>	<code>true</code> , wenn ein <code>function_handle</code> des schwachen Gradienten $\nabla f$ eines Eingangssignals $f \in H_0^1(\Omega)$ gegeben ist und zur Berechnung der garantierten unteren Energieschranke aus Theorem 4.9 genutzt werden soll, sonst <code>false</code> (wird <code>false</code> gesetzt, falls <code>useImage==true</code> )
<code>exactSolutionKnown</code>	<code>true</code> , wenn die Lösung $u$ von Problem 3.1 bekannt ist und zur Berechnung von exakten Fehlern zu Iteraten von Algorithmus 5.1 genutzt werden soll, sonst <code>false</code> (wird <code>false</code> gesetzt, falls <code>useImage==true</code> )
<code>useExactEnergy</code>	<code>true</code> , wenn $E(u)$ für die Lösung $u$ von Problem 3.1 bekannt ist und in Auswertungen der Ergebnisse genutzt werden soll, sonst <code>false</code> (wird <code>false</code> gesetzt, falls <code>exactSolutionKnown==false</code> )
<code>exactEnergy</code>	Wert $E(u)$ für die Lösung $u$ von Problem 3.1 (Wahl irrelevant, falls <code>useExactEnergy==false</code> )

Tabelle 6.4: Parameter des Experiments

Parametername	Beschreibung
<code>expName</code>	Name des Ordners, der in <code>./../results/nonconforming/</code> zum Speichern der Ergebnisse von Durchläufen des Programms für ein Experiment erstellt werden soll
<code>dirInfoName</code>	Name des Ordners, der in <code>./../results/nonconforming/expName/</code> zum Speichern der Ergebnisse eines Durchlaufs des Programms erstellt werden soll

Tabelle 6.5: Parameter zum Speichern der Ergebnisse

Parametername	Beschreibung
<code>inputSignal</code>	<code>function_handle</code> des Eingangssignals $f$ (Wahl irrelevant, falls <code>useImage</code> aus Tabelle 6.4 den Wert <code>true</code> hat)
<code>gradientInputSignal</code>	<code>function_handle</code> des schwachen Gradienten $\nabla f$ des Eingangssignals $f$ (Wahl irrelevant, falls <code>inSiGradientKnown</code> aus Tabelle 6.4 den Wert <code>false</code> hat)
<code>exactSolution</code>	<code>function_handle</code> der Lösung $u$ von Problem 3.1 (Wahl irrelevant, falls <code>exactSolutionKnown</code> aus Tabelle 6.4 den Wert <code>false</code> hat)

Tabelle 6.6: Definierbare `function_handles`

- `exactSolutionKnown = true`,
- `useExactEnergy = true`.

*Bemerkung 6.1.* Wird `useImage = true` gewählt, dann erzeugt die Methode

```
./utils/functions/image2function.m
```

unter Nutzung der Matlab-Funktionen `imread` und `im2double` ein `function_handle`, welches das durch `imageName` aus Tabelle 6.4 gegebene quadratische Graufarbenbild beschreibt. Dabei wird genutzt, dass dieses Bild durch eine Funktion  $g : [0, 1]^2 \rightarrow [0, 1]$  repräsentiert werden kann. Um die Ergebnisse des Programms unter Berücksichtigung der in Kapitel 1 beschriebenen Anwendung des ROF-Modellproblems interpretieren zu können, wird dann mit dem Parameter  $\alpha$  aus Problem 3.1 ein `function_handle` der Funktion  $f = \alpha g$  zurückgegeben und als Eingangssignal gewählt.

Noch in der Benchmark-Datei werden mit den gewählten Einstellungen die benötigten geometrischen Datenstrukturen für die Triangulierung  $\mathcal{T}$  initialisiert. Diese werden mit allen weiteren gegebenen Informationen als Felder eines Structure Arrays gespeichert und an das Hauptprogramm übergeben. Dort werden zunächst Structure Arrays mit Feldern für die benötigten Daten während der AFEM-Routine sowie deren Ergebnisse erstellt. Insbesondere wird der Startwert  $u_0$  für die primale-duale Iteration auf dem ersten Level des AFEM-Algorithmus initialisiert. Dieser hängt von der Wahl des Parameters `u0Mode` aus Tabelle 6.3 ab. Falls dabei die Crouzeix-Raviart-Interpolation des Eingangssignals  $f$  bezüglich  $\mathcal{T}$  genutzt werden soll, wird diese in

```
./nonconforming/common/interpolationCR.m
```

ermittelt. Nach dem Berechnen einiger weiterer Informationen beginnt die AFEM-Schleife.

Zu Beginn der Schleife werden die für das aktuelle Level benötigten Daten ermittelt, zum Beispiel die Flächeninhalte der Dreiecke und die Freiheitsgrade, das heißt in dieser Implementierung die inneren Kanten, der aktuellen Triangulierung  $\mathcal{T}$ . Insbesondere werden die auf dem Level mehrmals benötigten Gradienten der lokalen Crouzeix-Raviart-Basisfunktionen aus Abschnitt 2.2 auf allen Dreiecken berechnet (cf. [Car+10, Abschnitt 1.4.2]) in

```
./nonconforming/common/computeGradsCR4e.m.
```

Mit diesen wird der stückweise Gradient des Startwerts  $u_0$  aus Algorithmus 5.1 berechnet in

```
./nonconforming/common/gradientCR.m.
```

## 6 Implementierung

Mit  $\nabla_{\text{NC}} u_0$  wird anschließend  $\Lambda_0$  aus Algorithmus 5.1 auf allen  $T \in \mathcal{T}$  initialisiert als

$$\Lambda_0|_T := \begin{cases} \frac{\nabla u_0|_T}{|\nabla u_0|_T}, & \text{falls } \nabla u_0|_T \neq 0, \\ 0, & \text{falls } \nabla u_0|_T = 0. \end{cases} \quad (6.1)$$

Außerdem werden die Steifigkeits- und Massenmatrix aus Bemerkung 5.2 aufgestellt in

```
./nonconforming/common/computeFeMatricesCR.m.
```

Dabei benutzten wir für jedes  $T \in \mathcal{T}$ , dass die lokale Massenmatrix auf  $T$  mithilfe der Gleichungen (2.2) und (2.3) berechnet werden kann durch

$$M_T = \frac{|T|}{3} I_3. \quad (6.2)$$

Das Aufstellen der lokalen Steifigkeitsmatrizen sowie die Berechnung der globalen Steifigkeits- und Massenmatrix aus ihren lokalen Versionen geschieht analog zu [Car+10, Abschnitt 1.4.2]). Danach werden in

```
./nonconforming/common/integralsWithInSi.m
```

mit den Crouzeix-Raviart-Basisfunktionen  $\psi_F$ ,  $F \in \mathcal{E}$ , für jedes Dreieck  $T \in \mathcal{T}$  mit Kanten  $\mathcal{E}(T) = \{E_1, E_2, E_3\}$  die Integrale  $\int_T f \psi_{E_k}|_T dx$ ,  $k \in \{1, 2, 3\}$ , berechnet. Dies geschieht mithilfe der `integrate` Methode [Car+10, Abschnitt 1.8.2] des AFEM-Softwarepaket. Die Berechnung der dafür benötigten lokalen Crouzeix-Raviart-Basisfunktionen  $\psi_{E_k}|_T$ ,  $k \in \{1, 2, 3\}$ , diskutieren wir in Abschnitt 6.4.1. Anschließend kann für alle  $F \in \mathcal{E}$  auch das Integral

$$\int_{\Omega} f \psi_F dx = \sum_{T \in \mathcal{T}} \int_T f \psi_F|_T dx = \sum_{T \in \mathcal{T}(F)} \int_T f \psi_F|_T dx$$

bestimmt werden. Damit sind alle notwendigen Daten ermittelt, um die primale-duale Iteration für den Solve-Schritt des aktuellen Levels zu starten. Genutzt wird dafür die Methode

```
./nonconforming/main/solvePrimalDualFormulation.m.
```

Details dazu diskutieren wir in Abschnitt 6.3. Mit dem letzten Iterat  $u_{\text{CR}, \mathcal{T}} \in \text{CR}_0^1(\mathcal{T})$  als Approximation der Lösung von Problem 4.1 auf dem Level berechnen wir unter anderem  $|E(u) - E_{\text{NC}}(u_{\text{CR}, \mathcal{T}})|$  und, mithilfe der Methode `error4eCRL2` [Car+10, Abschnitt 1.8.3] des AFEM-Softwarepaket, den Fehler  $\|u - u_{\text{CR}, \mathcal{T}}\|$ . Außerdem werden Informationen über die Iteration und ihre Ergebnisse gespeichert und gegebenenfalls ausgegeben. Zur Berechnung der garantierten oberen Energieschranke  $E_{\text{GUEB}, \mathcal{T}} = E_{\text{NC}}(J_{1, \mathcal{T}} u_{\text{CR}, \mathcal{T}})$  aus (4.15) nutzen wir anschließend die Methoden

```
./nonconforming/common/computeJ1.m,
```

die  $J_{1, \mathcal{T}} u_{\text{CR}, \mathcal{T}} \in P_1(\mathcal{T}) \cap C_0(\Omega) \subset \text{CR}_0^1(\mathcal{T})$  mit dem Operator  $J_{1, \mathcal{T}}$  aus (4.14) bestimmt,

```
./nonconforming/common/courant2CR.m,
```

welche die Koeffizienten von  $J_{1, \mathcal{T}} u_{\text{CR}, \mathcal{T}}$  bezüglich der Crouzeix-Raviart-Basisfunktionen aus Abschnitt 2.2 ermittelt, und schließlich

```
./nonconforming/common/computeDiscreteEnergyCR.m,
```

welche  $E_{NC}(J_{1,\mathcal{T}} u_{CR,\mathcal{T}})$ , wie zum Abschluss des nächsten Abschnitts 6.3 beschrieben, berechnet. Zum Bestimmen der garantierten unteren Energieschranke aus Theorem 4.9 und des Verfeinerungsindikators aus Definition 4.10 benötigen wir dann für alle  $T \in \mathcal{T}$  den Term  $\|f - \alpha u_{CR,\mathcal{T}}\|_{L^2(T)}^2$ . Zur Berechnung dieser Terme nutzen wir die Methode

```
./nonconforming/common/computeNormDiffInSiSolCrSquared4e.m
```

wie folgt. Sei  $T \in \mathcal{T}$  mit  $\mathcal{E}(T) = \{E_1, E_2, E_3\}$  und  $u_{CR,\mathcal{T}}|_T$  habe mit  $v_1, v_2, v_3 \in \mathbb{R}$  die Darstellung  $u_{CR,\mathcal{T}}|_T = \sum_{k=1}^3 v_k \psi_{E_k}|_T$ . Dann gilt

$$\begin{aligned} \|f - \alpha u_{CR,\mathcal{T}}\|_{L^2(T)}^2 &= \|f\|_{L^2(T)}^2 - 2\alpha (f, u_{CR,\mathcal{T}})_{L^2(T)} + \alpha^2 \|u_{CR,\mathcal{T}}\|_{L^2(T)}^2 \\ &= \|f\|_{L^2(T)}^2 - 2\alpha \sum_{k=1}^3 v_k \int_T f \psi_{E_k}|_T dx \\ &\quad + \alpha^2 \sum_{k=1}^3 \sum_{\ell=1}^3 v_k (\psi_{E_k}|_T, \psi_{E_\ell}|_T)_{L^2(T)} v_\ell. \end{aligned}$$

Um den ersten Term zu berechnen, nutzen wir die `integrate` Methode des AFEM-Softwarepaket und für die Berechnung des zweiten Terms merken wir an, dass die Integrale  $\int_T f \psi_{E_k}|_T dx$ ,  $k \in \{1, 2, 3\}$ , bereits ermittelt wurden. Der letzte Term vereinfacht sich durch Nutzung der expliziten Darstellung (6.2) der lokalen Massenmatrix  $M_T$  zu

$$\alpha^2 \sum_{k=1}^3 \sum_{\ell=1}^3 v_k (\psi_{E_k}|_T, \psi_{E_\ell}|_T)_{L^2(T)} v_\ell = \alpha^2 \sum_{k=1}^3 \sum_{\ell=1}^3 v_k M_{T,k\ell} v_\ell = \alpha^2 \frac{|T|}{3} \sum_{k=1}^3 v_k^2.$$

Unter erneuter Zuhilfenahme der `integrate` Methode wird anschließend in

```
./nonconforming/common/computeGleb.m
```

die garantierte untere Energieschranke berechnet. Zur Berechnung des Verfeinerungsindikators benötigen wir noch für alle  $F \in \mathcal{E}$  den Term  $\|[u_{CR,\mathcal{T}}]_F\|_{L^1(F)}$ . Diese ermitteln wir in der Methode

```
./nonconforming/common/computeL1NormOfJump4s.m.
```

Details dazu diskutieren wir in Abschnitt 6.4.3. Danach werden der Verfeinerungsindikator und seine Anteile  $\eta_{V,\mathcal{T}}$  und  $\eta_{J,\mathcal{T}}$  berechnet in der Methode

```
./nonconforming/estimate/estimateErrorCR4e.m.
```

Es folgt die Ausgabe einiger relevanter Informationen über das Level und das Abspeichern dieser und weiterer Ergebnisse in den durch die Parameter in Tabelle 6.5 gewählten Ordner mittels der Methode

```
./nonconforming/misc/saveResultsCR.m.
```

Endet nun der AFEM-Algorithmus nach Überprüfung des Abbruchkriteriums nicht, das heißt, ist die Anzahl der Freiheitsgrade der Triangulierung des aktuellen Levels geringer als die Zahl `minNrDof` aus Tabelle 6.2, so werden mithilfe des Verfeinerungsindikators  $\eta_{\mathcal{T}}$  die Dreiecke markiert, die verfeinert werden sollen. Dafür nutzen wir die folgenden Methoden des AFEM-Softwarepaket, die in [Car+10, Abschnitt 1.6] beschriebenen sind. Gilt für den Bulk-Parameter  $\theta$ , gewählt durch `parTheta` aus Tabelle 6.2, dass  $\theta = 1$ , so nutzen wir die Methode `markUniform`. Gilt  $\theta \in (0, 1)$ , so nutzen wir die Methode `markBulk`. Zum Verfeinern der Triangulierung wird anschließend die Methode `refineRGB` [Car+10, S. 1.7.2] des AFEM-Softwarepaket verwendet.

## 6 Implementierung

Zum Abschluss werden bereits Informationen für das nächste Level vorbereitet. Dazu gehören die Datenstrukturen, welche die verfeinerte Triangulierung  $\hat{\mathcal{T}}$  beschreiben, sowie einige weitere Eigenschaften von  $\hat{\mathcal{T}}$ , zum Beispiel die Anzahl und Längen der Seiten der Triangulierung. Soll als Startwert für das nächste Level eine Prolongation  $\hat{u}_{\text{CR},\mathcal{T}} \in \text{CR}_0^1(\hat{\mathcal{T}})$  von  $u_{\text{CR},\mathcal{T}}$  auf  $\hat{\mathcal{T}}$  genutzt werden, festgelegt durch den Parameter `useProlongation` aus Tabelle 6.2, wird dessen Berechnung in der Methode

```
./nonconforming/common/prolongationJ1.m
```

realisiert. Als Prolongation von  $u_{\text{CR},\mathcal{T}}$  wird ebenda mit dem Operator  $J_{1,\mathcal{T}}$  aus (4.14) die Funktion  $\hat{u}_{\text{CR},\mathcal{T}} := J_{1,\mathcal{T}} u_{\text{CR},\mathcal{T}} \in P_1(\mathcal{T}) \cap C_0(\Omega) \subseteq P_1(\hat{\mathcal{T}}) \cap C_0(\Omega) \subseteq \text{CR}_0^1(\hat{\mathcal{T}})$  gewählt. Die Berechnung der dabei benötigten lokalen Knotenwerte von  $u_{\text{CR},\mathcal{T}}$  geschieht wie in Abschnitt 6.4.2 beschrieben.

## 6.3 Realisierung der primalen-dualen Iteration

In diesem Abschnitt beschreiben wir die Methode

```
./nonconforming/main/solvePrimalDualFormulation.m.
```

Dabei benutzen wir die Bezeichnungen aus Algorithmus 5.1, der in dieser Methode realisiert wird. Wie in Abschnitt 6.2 beschrieben, sind vor Aufruf der Methode bereits alle benötigten Informationen über die aktuelle Triangulierung  $\mathcal{T}$  sowie die Steifigkeits- und Massenmatrix aus Bemerkung 5.2 und die Integrale  $\int_{\Omega} f \psi_F dx$  für alle Crouzeix-Raviart-Basisfunktionen  $\psi_F$ ,  $F \in \mathcal{E}$ , aus Abschnitt 2.2 bekannt. Mit der Steifigkeits- und Massenmatrix kann die Koeffizientenmatrix des Gleichungssystems (5.4) ohne Weiteres berechnet werden. Für die Implementierung der rechten Seite  $\bar{b}$  dieses Systems nutzen wir die folgende Beobachtung. Für alle  $k \in \{1, 2, \dots, |\mathcal{E}|\}$  und alle  $j \in \mathbb{N}$  gilt, da  $\nabla_{\text{NC}} u_{j-1}$ ,  $\Lambda_j$  und  $\nabla_{\text{NC}} \psi_{E_k}$  stückweise konstant sind, dass

$$\begin{aligned} \left( \frac{1}{\tau} \nabla_{\text{NC}} u_{j-1} - \Lambda_j, \nabla_{\text{NC}} \psi_{E_k} \right) &= \sum_{T \in \mathcal{T}} \left( \frac{1}{\tau} \nabla_{\text{NC}} u_{j-1} - \Lambda_j, \nabla_{\text{NC}} \psi_{E_k} \right)_{L^2(T)} \\ &= \sum_{T \in \mathcal{T}(E_k)} \frac{|T|}{\tau} (\nabla_{\text{NC}} u_{j-1} - \Lambda_j)|_T \cdot \nabla \psi_{E_k}|_T. \end{aligned}$$

Da die Terme  $(f, \psi_{E_k}) = \int_{\Omega} f \psi_{E_k} dx$  bereits berechnet wurden, können wir somit die rechte Seite  $\bar{b}$  implementieren. Es ist direkt ersichtlich wie der verbleibende Teil von Algorithmus 5.1 mit Abbruchkriterium (5.3) realisiert werden kann.

Es bleibt anzumerken, dass in `solvePrimalDualFormulation`, neben Algorithmus 5.1, der Ausgabe des Fortschritts sowie der Übergabe von Ergebnissen der Iteration, auch die Berechnung der nichtkonformen Energien  $E_{\text{NC}}(u_j)$  der Iterate realisiert wird. Diese wird ausgeführt in

```
./nonconforming/common/computeDiscreteEnergyCR.m.
```

Dabei werden die Massenmatrix  $M$  und die Integrale  $\int_{\Omega} f \psi_F dx$ ,  $F \in \mathcal{E}$ , wie folgt genutzt. Sei  $v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$  und  $v \in \mathbb{R}^{|\mathcal{E}|}$  enthalte die Koordinaten von  $v_{\text{CR}}$  bezüglich der Basis

$\{\psi_{E_1}, \psi_{E_2}, \dots, \psi_{E_{|\mathcal{E}|}}\}$  von  $\text{CR}^1(\mathcal{T})$ , das heißt,  $v_{\text{CR}} = \sum_{k=1}^{|\mathcal{E}|} v_k \psi_{E_k}$ . Dann gilt

$$\begin{aligned} E_{\text{NC}}(v_{\text{CR}}) &= \frac{\alpha}{2} \|v_{\text{CR}}\|^2 + \|\nabla_{\text{NC}} v_{\text{CR}}\|_{L^1(\Omega)} - \int_{\Omega} f v_{\text{CR}} \, dx \\ &= \frac{\alpha}{2} \sum_{k=1}^{|\mathcal{E}|} \sum_{\ell=1}^{|\mathcal{E}|} v_k (\psi_{E_k}, \psi_{E_\ell}) v_\ell + \sum_{T \in \mathcal{T}} \int_T |\nabla_{\text{NC}} v_{\text{CR}}| \, dx - \sum_{k=1}^{|\mathcal{E}|} v_k \int_{\Omega} f \psi_{E_k} \, dx \\ &= \frac{\alpha}{2} v \cdot M v + \sum_{T \in \mathcal{T}} |T| |(\nabla_{\text{NC}} v_{\text{CR}})_T| - \sum_{k=1}^{|\mathcal{E}|} v_k \int_{\Omega} f \psi_{E_k} \, dx. \end{aligned}$$

## 6.4 Mathematische Grundlagen ausgewählter Methoden

### 6.4.1 Berechnung lokaler Crouzeix-Raviart-Basisfunktionen

In der Methode

`./nonconforming/common/integralsWithInSi.m`

werden auf allen Dreiecken einer Triangulierung  $\mathcal{T}$  die lokalen Crouzeix-Raviart-Basisfunktionen aus Abschnitt 2.2 benötigt. Sei  $T \in \mathcal{T}$  mit  $T = \text{conv}\{P_1, P_2, P_3\}$ . Da nach Gleichung (2.2) die lokalen Crouzeix-Raviart-Basisfunktionen  $\psi_F|_T$ ,  $F \in \mathcal{E}(T)$ , darstellbar sind durch die baryzentrischen Koordinaten  $\lambda_1, \lambda_2, \lambda_3 \in P_1(T)$  aus Abschnitt 2.2, genügt es hier, deren Berechnung zu diskutieren. Dafür betrachten wir das Referenzdreieck, das definiert ist durch

$$T_{\text{ref}} := \text{conv} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\},$$

und mit der Matrix  $B := (P_1 - P_3, P_2 - P_3)$  die affine Transformation  $p : T_{\text{ref}} \rightarrow T$ ,  $x \mapsto Bx + P_3$ . Mithilfe der Umkehrabbildung von  $p$ , die gegeben ist durch  $p^{-1} : T \rightarrow T_{\text{ref}}$ ,  $x \mapsto B^{-1}(x - P_3)$ , können die baryzentrischen Koordinaten für alle  $x \in T$  berechnet werden durch

$$\begin{pmatrix} \lambda_1(x) \\ \lambda_2(x) \end{pmatrix} = p^{-1}(x) \quad \text{und} \quad \lambda_3(x) = 1 - \lambda_1(x) - \lambda_2(x).$$

### 6.4.2 Berechnung lokaler Knotenwerte einer Crouzeix-Raviart-Funktion

Sei  $v_{\text{CR}} \in \text{CR}^1(\mathcal{T})$ . Um  $J_{1,\mathcal{T}} v_{\text{CR}}$  für den Operator  $J_{1,\mathcal{T}}$  aus Gleichung (4.14) zu bestimmen und für die Berechnung der  $L^1$ -Norm der Kantensprünge von  $v_{\text{CR}}$  benötigen wir für jedes Dreieck  $T \in \mathcal{T}$  die Werte von  $v_{\text{CR}}$  in den Knoten von  $T$ . Dazu sei  $T = \text{conv}\{P_1, P_2, P_3\}$  mit den Kanten  $E_1 = \text{conv}\{P_1, P_2\}$ ,  $E_2 = \text{conv}\{P_2, P_3\}$  und  $E_3 = \text{conv}\{P_3, P_1\}$ . Die Funktion  $v := v_{\text{CR}}|_T$  habe in den Mittelpunkten der Kanten die Werte  $v_j := v(\text{mid}(E_j))$  für alle  $j \in \{1, 2, 3\}$ . Gesucht sind die Werte  $v(P_1)$ ,  $v(P_2)$  und  $v(P_3)$ .

Da  $v_{\text{CR}} \in \text{CR}^1(\mathcal{T})$ , ist  $v$  affin. Damit gilt für eine Kante  $E = \text{conv}\{P, Q\} \in \{E_1, E_2, E_3\}$ , dass  $v(\text{mid}(E))$  gegeben ist durch das arithmetische Mittel von  $v(P)$  und  $v(Q)$ . Somit erhalten wir die drei Gleichungen

$$v_1 = \frac{v(P_1) + v(P_2)}{2}, \quad v_2 = \frac{v(P_2) + v(P_3)}{2}, \quad v_3 = \frac{v(P_3) + v(P_1)}{2}.$$

Sind  $v_1$ ,  $v_2$  und  $v_3$  bekannt, können wir dieses Gleichungssystem nach  $v(P_1)$ ,  $v(P_2)$  und  $v(P_3)$  lösen und erhalten die gesuchten Werte

$$v(P_1) = v_1 + v_3 - v_2, \quad v(P_2) = v_1 + v_2 - v_3, \quad v(P_3) = v_2 + v_3 - v_1.$$

Wir realisieren diese Berechnung in der Methode

`./nonconforming/common/computeNodeValuesCR4e.m`.

### 6.4.3 Berechnung von Sprungtermen

Insbesondere für die Berechnung des Verfeinerungsindikators aus Definition 4.10 benötigen wir eine Methode, die für eine Crouzeix-Raviart-Funktion  $v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$  für jede Kante  $F \in \mathcal{E}$  den Term  $\| [v_{\text{CR}}]_F \|_{L^1(F)}$  bestimmt.

Da  $v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$ , ist  $[v_{\text{CR}}]_F$  affin und es gilt  $[v_{\text{CR}}]_F(\text{mid}(F)) = 0$  für alle Kanten  $F \in \mathcal{E}$ . Betrachten wir nun eine beliebige Kante  $F \in \mathcal{E}$  mit  $F = \text{conv}\{P_1, P_2\}$ . Wir definieren eine Parametrisierung  $\xi : [0, 2] \rightarrow \mathbb{R}^2$  von  $F$  durch  $\xi(t) := \frac{t}{2}(P_2 - P_1) + P_1$ . Es gilt  $|\xi'| \equiv \frac{1}{2}|P_2 - P_1| = \frac{1}{2}|F|$ . Sei außerdem  $p(t) := [v_{\text{CR}}]_F(\xi(t))$ . Dann gilt

$$\begin{aligned} \| [v_{\text{CR}}]_F \|_{L^1(F)} &= \int_F |[v_{\text{CR}}]_F| \, ds = \int_0^2 |p(t)| |\xi'(t)| \, dt = \frac{|F|}{2} \int_0^2 |p(t)| \, dt \\ &= \frac{|F|}{2} \left( \int_0^1 |p(t)| \, dt + \int_1^2 |p(t)| \, dt \right). \end{aligned}$$

Da  $v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$ , ist  $|p(\bullet)|$  auf  $[0, 1]$  und  $[1, 2]$  jeweils ein Polynom vom Grad 1 mit  $|p(1)| = |[v_{\text{CR}}]_F(\text{mid}(F))| = 0$ . Somit können wir  $|p|$  explizit ausdrücken durch

$$\begin{aligned} |p(t)| &= (1-t)|p(0)| \quad \text{für alle } t \in [0, 1] \quad \text{und} \\ |p(t)| &= (t-1)|p(2)| \quad \text{für alle } t \in [1, 2]. \end{aligned}$$

Damit erhalten wir, aufgrund der Exaktheit der Mittelpunktsregel für Polynome vom Grad 1, dass

$$\begin{aligned} \int_0^1 |p(t)| \, dt &= (1-0) \left| p\left(\frac{1}{2}\right) \right| = \frac{|p(0)|}{2} \quad \text{und} \\ \int_1^2 |p(t)| \, dt &= (2-1) \left| p\left(\frac{3}{2}\right) \right| = \frac{|p(2)|}{2}. \end{aligned}$$

Insgesamt gilt also

$$\begin{aligned} \| [v_{\text{CR}}]_F \|_{L^1(F)} &= \frac{|F|}{2} \left( \frac{|p(0)|}{2} + \frac{|p(2)|}{2} \right) = \frac{|F|}{4} (|p(0)| + |p(2)|) \\ &= \frac{|F|}{4} (|[v_{\text{CR}}]_F(P_1)| + |[v_{\text{CR}}]_F(P_2)|). \end{aligned}$$

Realisiert wird dies in der Methode

```
./nonconforming/common/computeL1NormOfJump4s.m
```

Dabei werden für  $v_{\text{CR}} \in \text{CR}_0^1(\mathcal{T})$  für jede Kante  $F \in \mathcal{E}$  mit  $F = \text{conv}\{P_1, P_2\}$  die Terme  $|[v_{\text{CR}}]_F(P_1)|$  und  $|[v_{\text{CR}}]_F(P_2)|$  berechnet in

```
./nonconforming/common/computeAbsNodeJumps4s.m
```

## 7 Numerische Beispiele

In diesem abschließenden Kapitel möchten wir die Ergebnisse der in Kapitel 6 beschriebenen Realisierung von Algorithmus 5.1 mit Abbruchkriterium (5.3) im Solve-Schritt der AFEM-Schleife aus Abbildung 6.1 untersuchen, deren initiales Level dabei mit 0 nummeriert wird. Bezüglich einer Triangulierung  $\mathcal{T}$  schreiben wir hier  $u_{\text{CR},\mathcal{T}}$  und  $\bar{\Lambda}_{0,\mathcal{T}}$  für  $u_{\text{CR}}$  und  $\bar{\Lambda}_0$  aus Theorem 5.3 sowie  $(u_{j,\mathcal{T}}, \Lambda_{j,\mathcal{T}})_{j \in \mathbb{N}}$  für die Iterate  $(u_j, \Lambda_j)_{j \in \mathbb{N}}$  von Algorithmus 5.1. Zunächst möchten wir die Belegung für alle Parameter aus den Tabellen 6.1–6.4, die für ein Experiment relevant sind, festlegen. Die Wahlen  $\tau = 1$  für Algorithmus 5.1 und  $\varepsilon_{\text{stop}} = 10^{-4}$  für das Abbruchkriterium (5.3) begründen wir in Abschnitt 7.1 experimentell. Außerdem nutzen wir  $\gamma = 1$  für den Parameter  $\gamma$  des Verfeinerungsindikators aus Definition 4.10, wobei wir den Einfluss dieses Parameters auf das Programm in Abschnitt 7.3 untersuchen. Eine Übersicht über alle somit festgelegten Standardwerte ist in Tabelle 7.1 zu finden. Wir werden in einem Experiment die Wahl eines Parameters nur erwähnen, wenn diese vom Standardwert abweicht. In Tabelle 7.2 wird für alle verwendeten Benchmark-Dateien aufgelistet, für welche Abbildungen die Ergebnisse des entsprechenden Programmaufrufs genutzt wurden.

Als Startwert für die primale-duale Iteration des ersten Levels der AFEM-Routine wählen wir  $u_0 \equiv 0$ . Für die Iteration der darauffolgenden Level wählen wir die zum Ende von Abschnitt 6.2 beschriebene Prolongation des Ergebnisses des jeweils vorherigen Levels. In allen Experimenten hatte dadurch die Wahl des Startwerts für die Iteration des ersten Levels keinen nennenswerten Einfluss auf die Dauer des Experiments oder die Güte der Ergebnisse. Für jeden Aufruf der primalen-dualen Iteration auf einer Triangulierung  $\mathcal{T}$  wählen wir dann  $\Lambda_{0,\mathcal{T}}$  wie in Gleichung (6.1) angegeben. Eingangssignale für Experimente, bei denen die exakte Lösung bekannt ist, konstruieren wir nach Abschnitt 3.2. Die schwachen Ableitungen dieser Eingangssignale und der entsprechenden exakten Lösungen berechnen wir ebenfalls nach Abschnitt 3.2. Dabei sei das Argument  $r$  einer Funktion stets aus  $[0, \infty)$ . Als Geometrie für das erste Level des AFEM-Algorithmus nutzen wir in Experimenten mit bekannter exakter Lösung die Triangulierung aus Abbildung 7.1a und in Experimenten mit einem Graufarbenbild als Eingangssignal die Triangulierung aus Abbildung 7.1b. Wir nutzen adaptive Netzverfeinerungen und dafür den Bulk-Parameter  $\theta = 0.5$  für den Mark-Schritt des AFEM-Algorithmus und den Parameter  $\gamma = 1$  für den Verfeinerungsindikator aus Definition 4.10. Die maximale Iterationszahl wurde mit  $10^{12}$  so gewählt, dass diese nie Grund für das Beenden der primalen-dualen Iteration war. Dementsprechend ist das Abbruchkriterium aus (5.3) für den Abbruch der Iteration verantwortlich. Die minimale Anzahl der Freiheitsgrade, welche eine Triangulierung während der AFEM-Routine zum Abbrechen erreichen soll, ist mit  $10^8$  so gewählt, dass die verfügbaren Ressourcen des verwendeten Rechners in der Regel ausgenutzt werden und die Rechnungen manuell beendet werden müssen. Dies geschah in allen Experimenten bei ungefähr  $10^6$  Freiheitsgraden. Beispielrechnungen mit verschiedenen Eingangssignalen ergaben, dass deutlich höhere Integrationsgrade als 10 zu keinen veränderten Konvergenzraten führen. Diese Wahl des Integrationsgrads erscheint daher als ausreichend. Da außerdem die Methode `integrate` des AFEM-Softwarepaket [Car09a] nicht während der primalen-dualen Iteration aufgerufen werden muss, hat eine möglicherweise zu hohe Wahl des Integrationsgrads keinen signifikanten Einfluss auf die Programmlaufzeit. Nur die Methode `errorCRL2` des AFEM-Softwarepaket, die zur Berechnung des exakten  $L^2$ -Fehlers verwendet wird, nutzt den Integrationsgrad 12, da diese unverändert übernommen wurde.

Parametername	Standardwert
<code>degree4Integrate</code>	10
<code>geometry</code>	'BigSquare' für Experimente mit bekannter exakter Lösung 'Square' für Experimente mit Graufarbenbild als Eingangssignal
<code>initialRefinementLevel</code>	0
<code>parTheta, <math>\theta</math></code>	0.5
<code>minNrDof</code>	$10^8$
<code>useProlongation</code>	<code>true</code>
<code>parGamma, <math>\gamma</math></code>	1
<code>d, d</code>	2
<code>u0Mode</code>	<code>zeros</code>
<code>epsStop, <math>\varepsilon_{\text{stop}}</math></code>	$10^{-4}$
<code>parTau, <math>\tau</math></code>	1
<code>maxIter</code>	$10^{12}$
<code>parAlpha, <math>\alpha</math></code>	$\alpha$ für Experimente mit Eingangssignal $f_\alpha$ $10^4$ für Experimente mit Eingangssignal <code>cameraman</code>
<code>parBeta, <math>\beta</math></code>	1 für Experimente mit Eingangssignal $f_\alpha$

Tabelle 7.1: Wahl der relevanten Parameter für die Experimente

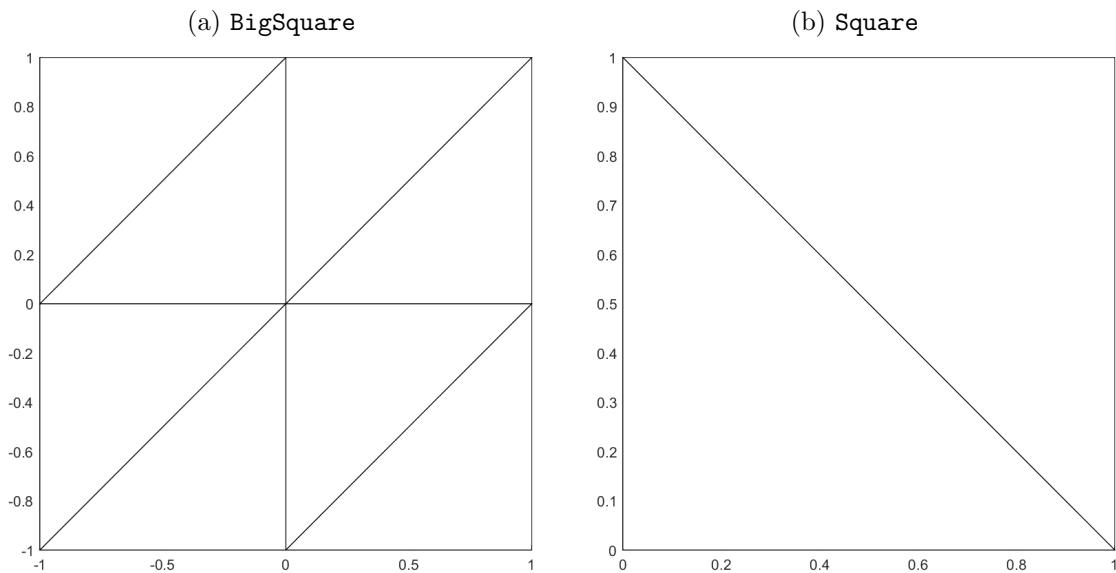


Abbildung 7.1: Initiale Triangulierung für Experimente mit bekannter exakter Lösung (a) und initiale Triangulierung für Experimente mit Graufarbenbild als Eingangssignal (b)

benchmark	Abbildungen
denoiseAlpha100	1.1c, 7.20, 7.21
denoiseAlpha1000	1.1d, 7.20, 7.21
denoiseAlpha2500	1.1e
denoiseAlpha5000	1.1f
denoiseAlpha10000	1.1g, 7.20, 7.21
denoiseAlpha50000	1.1h
paramsTau_f01_0Dot1	7.3a, 7.4, 7.8b
paramsTau_f01_0Dot5	7.3a, 7.4
standard_f01	7.3a, 7.4, 7.8a, 7.9, 7.10, 7.11, 7.12, 7.15, 7.16c, 7.18, 7.28b
paramsTau_cameraman_0Dot1	7.3b
paramsTau_cameraman_0Dot5	7.3b
standard_cameraman	7.3b, 7.19, 7.22, 7.28c
noTerminationTau_maxIter1e5	7.5
paramsEpsStop_1em2	7.6, 7.7a
paramsEpsStop_1em3	7.6
paramsEpsStop_1em4	7.6
paramsEpsStop_1em5	7.6, 7.7b
standardUniform_f01	7.11, 7.12, 7.28b
standard_f04	7.14, 7.18
standardUniform_f04	7.14
paramsGamma_0	7.15, 7.16a
paramsGamma_1em2	7.15
paramsGamma_0Dot5	7.15, 7.16b
standard_f01Alpha1e4	7.17, 7.18
standardUniform_f01Alpha1e4	7.17
standardUniform_cameraman	7.19, 7.28c
circleContinuousAdaptive	7.24a, 7.25a, 7.26a, 7.26b, 7.27, 7.28a
circleDiscontinuousAdaptive	7.24a
circleContinuousUniform	7.24b, 7.25b, 7.26c, 7.26d, 7.27, 7.28a
circleDiscontinuousUniform	7.24b

Tabelle 7.2: Eingabeparameter `benchmark` für `startAlgorithmCR`, um die Rechnung zu einem Experiment zu starten, und die Abbildungen, die auf den Ergebnissen dieses Experiments basieren

## 7 Numerische Beispiele

Nun möchten wir ein Eingangssignal definieren, welches in mehreren der folgenden Abschnitte verwendet wird. Sei  $\beta \geq 1/2$ , wobei wir stets  $\beta = 1$  wählen. Wir betrachten die Funktion

$$u(r) := \begin{cases} 1, & \text{falls } r \in [0, \frac{1}{6}], \\ 1 + (6r - 1)^\beta, & \text{falls } r \in (\frac{1}{6}, \frac{1}{3}], \\ 2, & \text{falls } r \in (\frac{1}{3}, \frac{1}{2}], \\ 2(\frac{5}{2} - 3r)^\beta, & \text{falls } r \in (\frac{1}{2}, \frac{5}{6}], \\ 0, & \text{falls } r \in (\frac{5}{6}, \infty), \end{cases}$$

und wählen

$$\operatorname{sgn}(\partial_r u(r)) := \begin{cases} 12r - 36r^2, & \text{falls } r \in [0, \frac{1}{6}], \\ 1, & \text{falls } r \in (\frac{1}{6}, \frac{1}{3}], \\ \cos(\pi(6r - 2)), & \text{falls } r \in (\frac{1}{3}, \frac{1}{2}], \\ -1, & \text{falls } r \in (\frac{1}{2}, \frac{5}{6}], \\ -\frac{1+\cos(\pi(6r-5))}{2}, & \text{falls } r \in (\frac{5}{6}, \infty). \end{cases}$$

Nach Gleichung (3.8) ist  $u$  mit dieser Wahl von  $\operatorname{sgn}(\partial_r u)$  die Lösung von Problem 3.1 mit Eingangssignal

$$f_\alpha(r) = \begin{cases} \alpha - 12(2 - 9r), & \text{falls } r \in [0, \frac{1}{6}], \\ \alpha(1 + (6r - 1)^\beta) - \frac{1}{r}, & \text{falls } r \in (\frac{1}{6}, \frac{1}{3}], \\ 2\alpha + 6\pi \sin(\pi(6r - 2)) - \frac{1}{r} \cos(\pi(6r - 2)), & \text{falls } r \in (\frac{1}{3}, \frac{1}{2}], \\ 2\alpha(\frac{5}{2} - 3r)^\beta + \frac{1}{r}, & \text{falls } r \in (\frac{1}{2}, \frac{5}{6}], \\ -3\pi \sin(\pi(6r - 5)) + \frac{1+\cos(\pi(6r-5))}{2r}, & \text{falls } r \in (\frac{5}{6}, \infty). \end{cases}$$

Dieses Eingangssignal für zwei Wahlen von  $\alpha$  und die exakte Lösung  $u$  von Problem 3.1 mit Eingangssignal  $f_\alpha$  können in Abbildung 7.2 betrachtet werden. Wir können anhand von Abbildung 7.2 ebenfalls den in Kapitel 1 beschriebenen Zusammenhang von Problem 3.1 und dem ROF-Modell sehen, denn für  $\alpha = 10^4$  scheint annähernd  $f_\alpha = \alpha u$  zu gelten. Die schwachen Gradienten von  $u$  und  $f_\alpha$  können bestimmt werden mithilfe der partiellen Ableitungen

$$\partial_r f_\alpha(r) = \begin{cases} 108, & \text{falls } r \in [0, \frac{1}{6}], \\ 6\alpha\beta(6r - 1)^{\beta-1} + \frac{1}{r^2}, & \text{falls } r \in (\frac{1}{6}, \frac{1}{3}], \\ (36\pi^2 + \frac{1}{r^2}) \cos(\pi(6r - 2)) + \frac{6\pi}{r} \sin(\pi(6r - 2)), & \text{falls } r \in (\frac{1}{3}, \frac{1}{2}], \\ -\left(6\alpha\beta(\frac{5}{2} - 3r)^{\beta-1} + \frac{1}{r^2}\right), & \text{falls } r \in (\frac{1}{2}, \frac{5}{6}], \\ -\left((18\pi^2 + \frac{1}{2r^2}) \cos(\pi(6r - 5)) + \frac{1}{2r^2} + \frac{3\pi}{r} \sin(\pi(6r - 5))\right), & \text{falls } r \in (\frac{5}{6}, \infty), \end{cases}$$

und

$$\partial_r u(r) = \begin{cases} 0, & \text{falls } r \in [0, \frac{1}{6}], \\ 6\beta(6r - 1)^{\beta-1}, & \text{falls } r \in (\frac{1}{6}, \frac{1}{3}], \\ 0, & \text{falls } r \in (\frac{1}{3}, \frac{1}{2}], \\ -6\beta(\frac{5}{2} - 3r)^{\beta-1}, & \text{falls } r \in (\frac{1}{2}, \frac{5}{6}], \\ 0, & \text{falls } r \in (\frac{5}{6}, \infty). \end{cases}$$

Durch Kenntnis des schwachen Gradienten von  $u$  erhalten wir für das Experiment mit Eingangssignal  $f_\alpha$  die Approximation  $E(u) \approx -2.05803$  für  $\alpha = 1$  und  $E(u) \approx -20580.34076$  für  $\alpha = 10^4$ . Ein weiteres Eingangssignal, welches wir in mehreren der folgenden Abschnitte nutzen werden, ist das Graufarbenbild `cameraman` aus Abb. 1.1a, aus dem nach Bemerkung 6.1 ein `function_handle` erzeugt wird.



Abbildung 7.2: Funktionen  $f_\alpha$  und  $u$  sowie deren Darstellungen entlang der x-Achse (blau) und der y-Achse (rot) für  $\alpha \in \{1, 10^4\}$

## 7.1 Wahl der Parameter für die primale-duale Iteration

Zunächst interessiert uns, wie der Parameter  $\tau$  aus Algorithmus 5.1 gewählt werden sollte. Durch Theorem 5.3 ist uns bereits bekannt, dass wir die Konvergenz der primalen-dualen Iteration nach ebendiesem Theorem nur für  $\tau \in (0, 1]$  garantieren können. In Abbildung 7.3 sehen wir, dass die Anzahlen der Iterationsschritte, und damit auch die Laufzeiten, der primalen-dualen Iterationen während der AFEM-Routine umso größer sind, je kleiner  $\tau$  gewählt wird. Da sich die betrachteten Graphen in Abbildung 7.4 für die verschiedenen Wahlen von  $\tau$  nicht sichtbar unterscheiden, vermuten wir, dass die ideale Wahl für  $\tau$  die maximale nach Theorem 5.3 zulässige ist. Deshalb wählen wir für unsere Experimente  $\tau = 1$ . Eine mögliche Erklärung für unsere Beobachtung liefert der Beweis von Theorem 5.3. Die darin bewiesene Ungleichung (5.14) impliziert

$$\sum_{j=1}^{\infty} \|u_{\text{CR},\tau} - u_{j,\tau}\|^2 \leq \frac{1}{2\alpha\tau} \left( \|u_{\text{CR},\tau} - u_{0,\tau}\|_{\text{NC}}^2 + \|\bar{\Lambda}_{0,\tau} - \Lambda_{0,\tau}\|^2 \right). \quad (7.1)$$

Die rechte Seite ist antiproportional zu  $\tau$ , womit womöglich die Folge  $(\|u_{\text{CR}} - u_{j,\tau}\|)_{j \in \mathbb{N}}$  schneller gegen 0 konvergiert und damit auch das Abbruchkriterium (5.3) nach einer geringeren Anzahl von Iterationen erfüllt ist. Allerdings schließt der Beweis von Theorem 5.3 die Konvergenz der primalen-dualen Iteration für  $\tau > 1$  nicht aus. Wie aber in Abbildung 7.5 zu sehen, haben wir schon für  $\tau = 1.2$  ein Beispiel gefunden, bei dem nicht davon ausgegangen werden kann, dass die primale-duale Iteration konvergiert. Dabei wurde die Iteration nach  $10^5$  Schritten abgebrochen, da kein anderes Verhalten mehr zu erwarten war. Zusätzlich legitimiert wird der Abbruch nach  $10^5$  Iterationen dadurch, dass in Abbildung 7.3a auf der gleichen Triangulierung für die suboptimale Wahl  $\tau = 0.1$  weniger als  $10^3$  Iterationen und für die Wahl  $\tau = 1$  sogar weniger als 10 Iterationen benötigt wurden. Weiterhin zeigt Abbildung 7.5a, dass auch für andere Varianten, den Unterschied zweier aufeinanderfolgender Iterate im Abbruchkriterium (5.3) zu messen, die Iteration in der Regel nicht abbricht. Es scheint nach Abbildung 7.5b außerdem festzustehen, dass die nichtkonformen Energien der Iterate nach etwa 100 Iterationen oszillierend Werte annehmen. Somit bleibt insgesamt festzuhalten, dass eine Konvergenzaussage im Wortlaut von Theorem 5.3 für  $\tau \in (0, 1.2]$  nicht mehr bewiesen werden kann.

Als Nächstes begründen wir die Wahl  $\varepsilon_{\text{stop}} = 10^{-4}$  für das Abbruchkriterium (5.3). Wie in Abbildung 7.6 zu sehen, stagniert der exakte  $L^2$ -Fehler für  $\varepsilon_{\text{stop}} = 10^{-2}$  ab circa  $10^4$  Freiheitsgraden und es lässt sich erahnen, dass dieser Effekt auch für  $\varepsilon_{\text{stop}} = 10^{-3}$  ab ungefähr  $6 \cdot 10^5$  Freiheitsgraden einsetzt. Dabei ist der erreichte Fehler für  $\varepsilon_{\text{stop}} = 10^{-3}$  geringer als für  $\varepsilon_{\text{stop}} = 10^{-2}$ . Da für  $\varepsilon_{\text{stop}} \in \{10^{-5}, 10^{-4}\}$  bis  $10^6$  Freiheitsgrade dieses Verhalten noch nicht zu sehen ist, scheint dies an einem zu frühen Abbruch der Iteration durch eine zu große Wahl von  $\varepsilon_{\text{stop}}$  zu liegen. Bei einer hohen Anzahl an Freiheitsgraden, das heißt bei kleinen Netzweiten, muss also  $\varepsilon_{\text{stop}}$  ausreichend klein sein, um die beobachtete Stagnation des exakten Fehlers zu verhindern. Da wir bei den Experimenten in dieser Arbeit  $10^6$  Freiheitsgrade nicht deutlich überschreiten und sich bis dahin die Ergebnisse für die Wahlen  $10^{-4}$  und  $10^{-5}$  für  $\varepsilon_{\text{stop}}$  in Abbildung 7.6 kaum unterscheiden, die Wahl  $\varepsilon_{\text{stop}} = 10^{-5}$  aber eine wesentlich längere Laufzeit des Programms verursacht, begründen wir damit unsere Wahl  $\varepsilon_{\text{stop}} = 10^{-4}$ . Allerdings bleibt anzumerken, dass der Verfeinerungsindikator  $\eta_{\tau}$  für alle getesteten Wahlen von  $\varepsilon_{\text{stop}}$  weiter fällt. Um dies zu untersuchen, betrachten wir die Definition 4.10 von  $\eta_{\tau}$ . Aus dieser folgt

$$\eta_{\tau} \leq \max_{T \in \mathcal{T}} |T| \|f - \alpha u_{\text{CR},\tau}\|^2 + 2 \max_{T \in \mathcal{T}} |T|^{\gamma/2} \sum_{F \in \mathcal{E}} \| [u_{\text{CR},\tau}]_F \|_{L^1(F)}. \quad (7.2)$$

Stagniert während der AFEM-Routine  $\|u - u_{\text{CR},\tau}\|$ , ist in Anbetracht der Interpretation des ROF-Modells aus Kapitel 1 nicht ausschließbar, dass  $\|f - \alpha u_{\text{CR},\tau}\|^2$  ebenfalls stagniert.

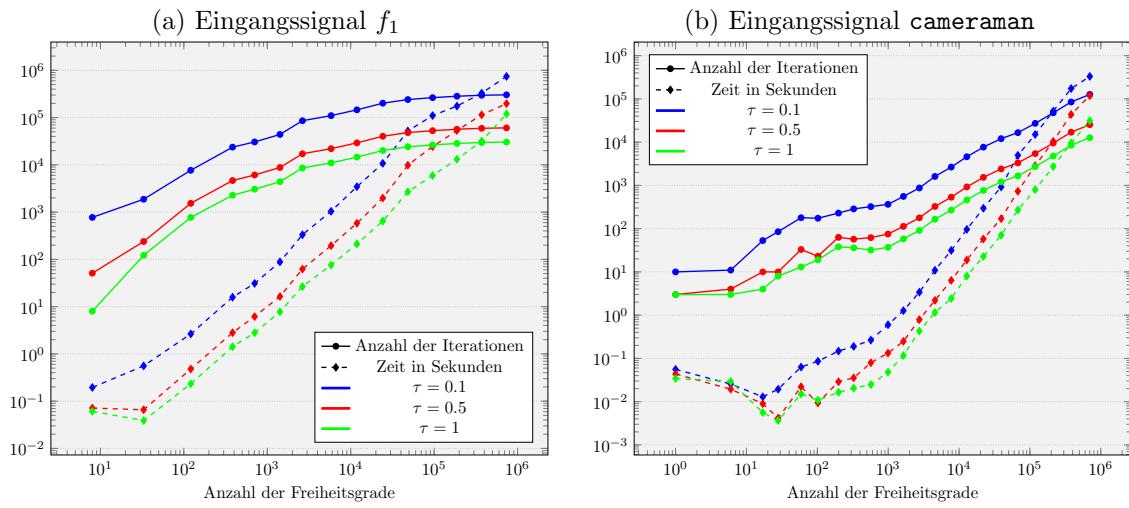


Abbildung 7.3: Anzahl der Iterationen und benötigte Zeit für die primalen-dualen Iterationen während des AFEM-Algorithmus mit verschiedenen Werten von  $\tau$  für die Eingangssignale  $f_1$  und **cameraman**

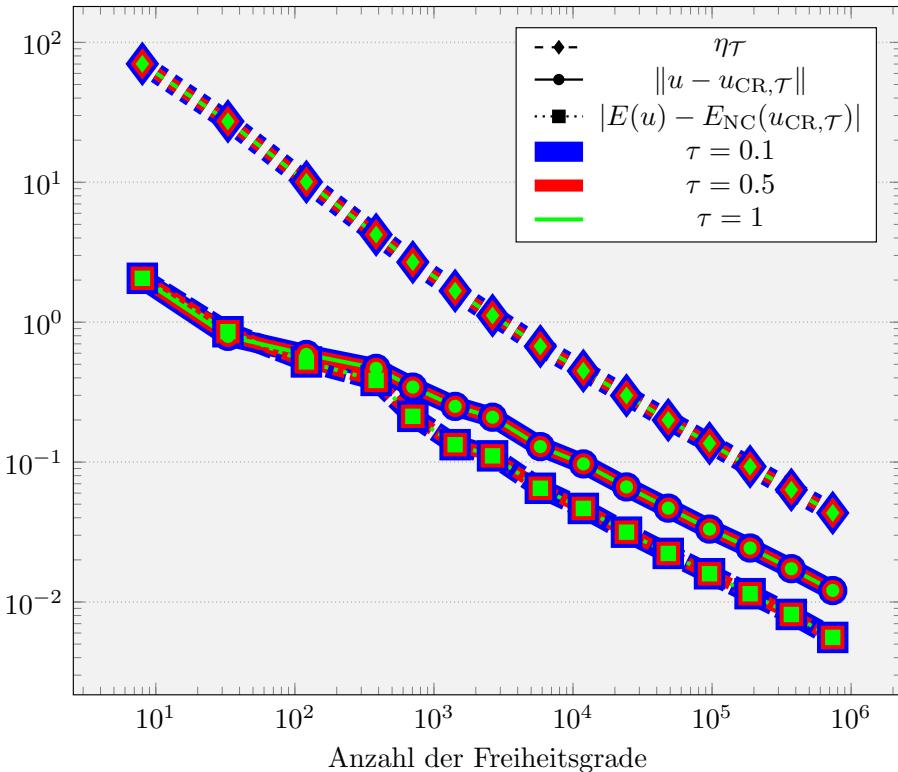


Abbildung 7.4: Ergebnisse des Programms mit verschiedenen Werten von  $\tau$  für das Eingangssignal  $f_1$

## 7 Numerische Beispiele

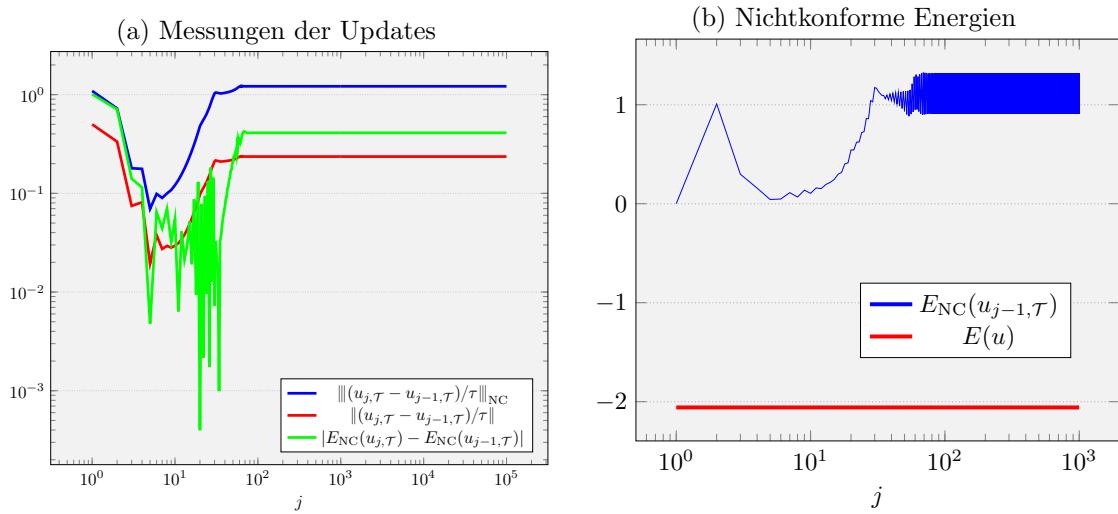


Abbildung 7.5: Verschiedene Messungen der Updates (a) und Verlauf der nichtkonformen Energien der ersten 1000 Iterate (b) der primalen-dualen Iteration auf der Triangulierung **BigSquare** aus Abbildung 7.1a mit  $\tau = 1.2$  für das Eingangssignal  $f_1$

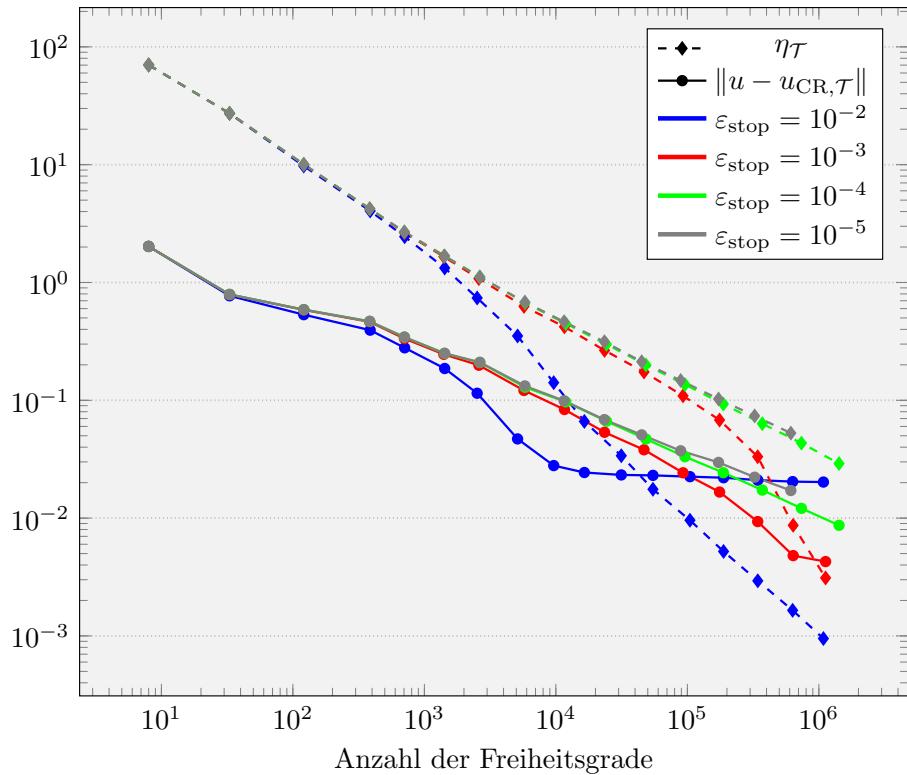


Abbildung 7.6: Ergebnisse des Programms mit verschiedenen Werten von  $\varepsilon_{stop}$  für das Eingangssignal  $f_1$

Außerdem können wir nach den Überlegungen zu Abbildung 7.12a in Abschnitt 7.3 davon ausgehen, dass wahrscheinlich auch in den Experimenten aus Abbildung 7.6 die Terme  $\sum_{F \in \mathcal{E}} \| [u_{\text{CR}, \mathcal{T}}]_F \|_{L^1(F)}$  im Verlauf des AFEM-Algorithmus nicht gegen 0 konvergieren. Dennoch ist nach Ungleichung (7.2) die Konvergenz von  $\eta_{\mathcal{T}}$  gegen 0 weiterhin möglich, solange die Flächen der Dreiecke  $T \in \mathcal{T}$  reduziert werden. Insbesondere kann davon ausgegangen werden, dass, nach Eintreten der beschriebenen Stagnation, der Flächeninhalt von Dreiecken beim Markieren durch  $\eta_{\mathcal{T}}$  zunehmend wichtiger wird. Diese Vermutung scheint von Abbildung 7.7 gestützt zu werden, in der wir sehen, dass bei einer vergleichbaren Anzahl von Freiheitsgraden der abgebildeten Triangulierungen die Fläche des größten Dreiecks in Abbildung 7.7a zu  $\varepsilon_{\text{stop}} = 10^{-2}$  kleiner ist als die des größten Dreiecks in Abbildung 7.7b zu  $\varepsilon_{\text{stop}} = 10^{-5}$ . Die vor Eintreten der Stagnation für  $\varepsilon_{\text{stop}} \in \{10^{-3}, 10^{-2}\}$  im Vergleich zu  $\varepsilon_{\text{stop}} \in \{10^{-5}, 10^{-4}\}$  größere Konvergenzgeschwindigkeit, die in Abbildung 7.6 zu sehen ist, lässt die Annahme zu, dass die Verfeinerung von Dreiecken mit großer Fläche kurzzeitig eine effektive Strategie ist, insgesamt aber nicht zu geringeren exakten Fehlern führt.

## 7.2 Eigenschaften der primalen-dualen Iteration

Nun möchten wir anhand einiger Experimente mit Eingangssignal  $f_1$  die primale-duale Iteration untersuchen. In Abbildung 7.8a sehen wir an einer Auswahl von Leveln der AFEM-Routine, dass die nichtkonforme Energie  $E_{\text{NC}}(\cdot)$  der Iterate von oben konvergiert. Dabei nimmt der Abstand des jeweiligen Grenzwerts zu der Energie der exakten Lösung  $E(u)$  mit zunehmender Anzahl von Freiheitsgraden ab. Anhand des Ansteigens der nichtkonformen Energie im ersten Iterationsschritt aller in Abbildung 7.8a dargestellten Level, mit Ausnahme von Level 3, ist aber auch klar erkennbar, dass die nichtkonformen Energien der Iterate nicht monoton fallend konvergieren. Dies können wir ebenfalls in Abbildung 7.8b sehen, in der, auf einem Level mit 33 Freiheitsgraden für das gleiche Experiment mit  $\tau = 10^{-1}$ , die nichtkonformen Energien der Iterate oszillierend fallen. Ebendieses Verhalten könnte ein Grund dafür sein, dass die bis zum Abbruch der primalen-dualen Iteration benötigte Anzahl an Iterationen für  $\tau = 10^{-1}$  deutlich höher ist als für  $\tau = 1$ , wie wir bereits in Abbildung 7.3a gesehen haben. Als Nächstes betrachten wir Abbildung 7.9a, in der die Entwicklung der Messung des Unterschieds zweier aufeinanderfolgender Iterate im Abbruchkriterium (5.3) für verschiedene Level während der AFEM-Routine zu sehen ist. Diese nimmt stets ohne Auffälligkeiten ab, bis sie den Wert  $\varepsilon_{\text{stop}}$  erreicht. Dabei ist wieder zu erkennen, dass die Anzahl der dazu benötigten Iterationen mit steigender Anzahl an Freiheitsgraden zunimmt. Zusätzlich sind in Abbildung 7.9b für das sechste Level derselben Rechnung die Verläufe zweier weiterer Varianten, den Unterschied von zwei aufeinanderfolgenden Iteraten zu messen, zu sehen. Deren Konvergenz gegen 0 ist zu erwarten, weshalb sie ebenfalls für das Abbruchkriterium in Betracht gezogen werden könnten. Dabei unterscheidet sich die Variante, bei der lediglich die  $L^2$ -Norm anstelle der Energienorm verwendet wird, nur um einen Faktor von etwa  $10^2$  von der im Programm benutzten. Dies passt zur bekannten Theorie, da die Norm  $\|\cdot\|_{\text{NC}}$  ungefähr so skaliert wie  $h^{-1} \|\cdot\|$  (cf. [Bar15b, Lemma 3.5, Lemma 3.7]), wobei für das in Abbildung 7.9b gezeigte Level ungefähr gilt  $h \approx 10^{-2}$ . Dementsprechend ist davon auszugehen, dass bei Wahl dieses Abbruchkriteriums nur die Toleranz  $\varepsilon_{\text{stop}}$  angepasst werden müsste. Die andere Variante zum Messen des Updates, das heißt der Betrag der Differenz der nichtkonformen Energien zweier Iterate, scheint sich aber, aufgrund der auftretenden Oszillationen, nur schlecht für ein Abbruchkriterium zu eignen. Abschließend können wir festhalten, dass sich demnach das von uns gewählte Abbruchkriterium (5.3) als sinnvoll erwiesen hat.

## 7 Numerische Beispiele

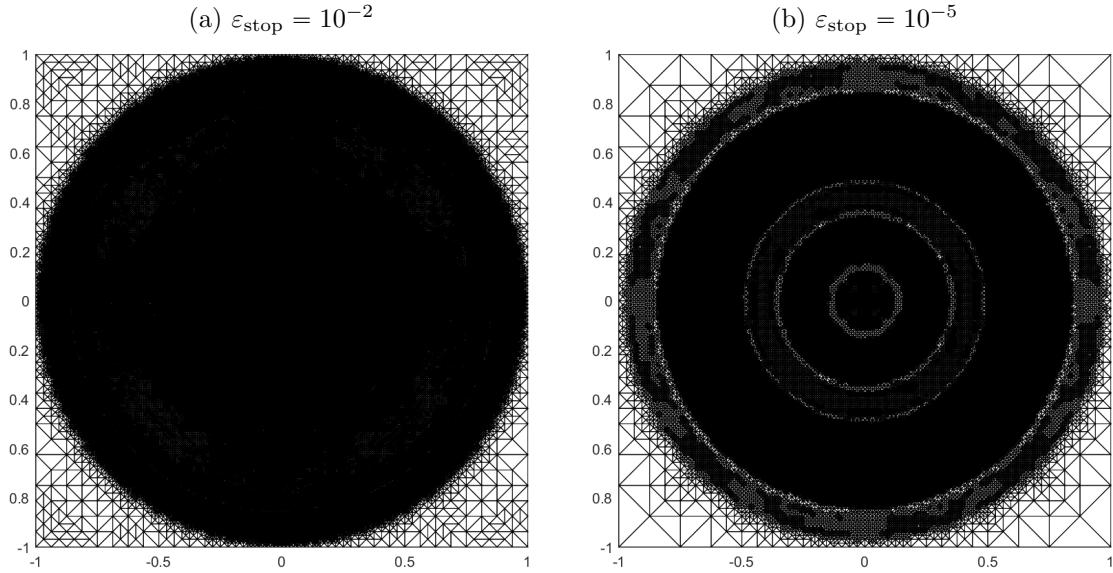


Abbildung 7.7: Triangulierungen nach den Rechnungen mit verschiedenen Werten von  $\varepsilon_{\text{stop}}$  nach jeweils circa 620 000 Freiheitsgraden für das Eingangssignal  $f_1$

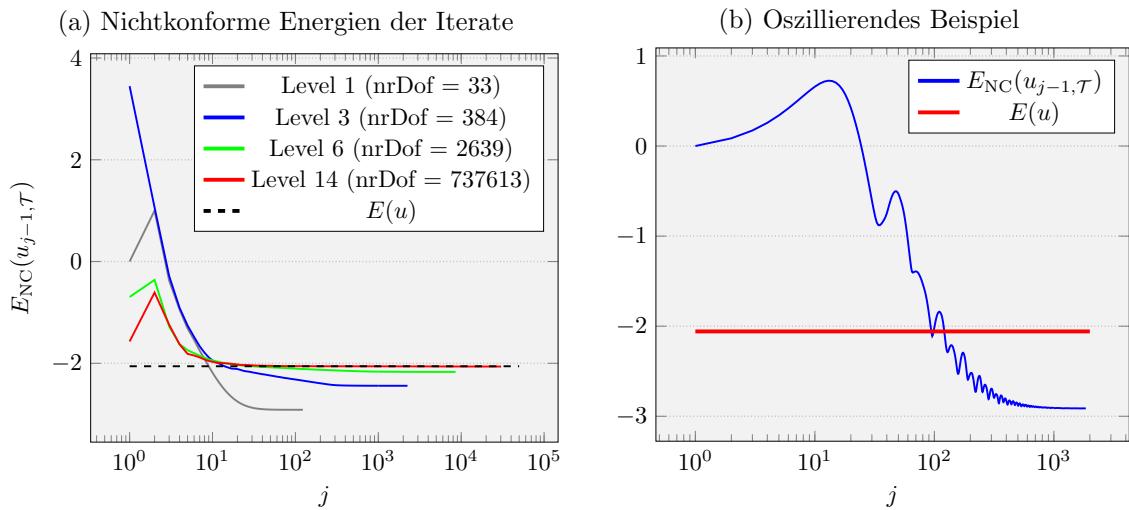


Abbildung 7.8: Entwicklung der nichtkonformen Energien der Iterate während der primalen-dualen Iteration auf verschiedenen Leveln der AFEM-Routine mit  $\tau = 1$  (a) und für ein Level mit 33 Freiheitsgraden und  $\tau = 10^{-1}$  (b) für das Eingangssignal  $f_1$

## 7.3 Experimente mit bekannter exakter Lösung

In diesem Abschnitt möchten wir zunächst die Ergebnisse der Experimente mit Eingangssignal  $f_1$  bei adaptiver und uniformer Netzverfeinerung untersuchen und einige Aussagen aus den theoretischen Kapiteln dieser Arbeit validieren. Die Lösung des Experiments mit adaptiver Netzverfeinerung ist in Abbildung 7.10 dargestellt und ähnelt erkennbar der exakten Lösung  $u$  aus Abbildung 7.2. Wie in Abbildung 7.11 zu sehen, unterscheiden sich die Konvergenzraten bei uniformer und adaptiver Netzverfeinerungen nur für den Anteil  $\eta_{V,\mathcal{T}}$  des Verfeinerungsindikators aus Definition 4.10 und für die Differenzen  $E_{\text{GUEB},\mathcal{T}} - E_{\text{GLEB},\mathcal{T}}$  sowie  $E(u) - E_{\text{GLEB},\mathcal{T}}$  mit den garantierten oberen und unteren Energieschranken  $E_{\text{GUEB},\mathcal{T}}$  und  $E_{\text{GLEB},\mathcal{T}}$  aus Abschnitt 4.3. Dabei können die Ratenunterschiede für diese Differenzen auf die für  $\eta_{V,\mathcal{T}}$  zurückgeführt werden, denn nach Definition von  $E_{\text{GLEB},\mathcal{T}}$  in Gleichung (4.12) und von  $\eta_{V,\mathcal{T}}$  ist zu erwarten, dass der Wert von  $E_{\text{GLEB},\mathcal{T}}$  umso kleiner ist, je größer der Wert von  $\eta_{V,\mathcal{T}}$  ist. Tatsächlich nimmt in Abbildung 7.11 bei adaptiver Netzverfeinerung  $\eta_{V,\mathcal{T}}$  größere Werte an als bei uniformer, weshalb sich die beiden Graphen, in denen  $E_{\text{GLEB},\mathcal{T}}$  subtrahiert wird, ebenso verhalten. Die Konvergenzrate von  $\eta_{V,\mathcal{T}}$  ist bei uniformer Netzverfeinerung etwa 1 und scheint bei adaptiver Netzverfeinerung kleiner als 1 aber womöglich größer als  $1/2$  zu sein. Allerdings wirken sich diese Ratenunterschiede von  $\eta_{V,\mathcal{T}}$  bei adaptiver und uniformer Netzverfeinerung nicht auf den Verfeinerungsindikator  $\eta_{\mathcal{T}}$  aus, denn dieser wird mit zunehmenden Freiheitsgraden deutlich von seinem Anteil  $\eta_{J,\mathcal{T}}$  dominiert. Damit hat  $\eta_{\mathcal{T}}$ , ebenso wie  $\eta_{J,\mathcal{T}}$ , die Rate  $1/2$ .

Durch Abbildung 7.8a war bereits zu vermuten, dass die nichtkonformen Energien der diskreten Lösungen  $E_{\text{NC}}(u_{\text{CR},\mathcal{T}})$  gegen die Energie der exakten Lösung  $E(u)$  konvergiert. Nun sehen wir, dass  $|E(u) - E_{\text{NC}}(u_{\text{CR},\mathcal{T}})|$  tatsächlich konvergiert mit einer Rate von etwa  $1/2$ . Weiterhin erwarten wir nach der Herleitung des diskreten Problems in Abschnitt 4.1, dass  $E(u_{\text{CR},\mathcal{T}})$ , im Gegensatz zu  $E_{\text{NC}}(u_{\text{CR},\mathcal{T}})$ , nicht gegen  $E(u)$  konvergiert. Dies wird experimentell bestätigt durch Abbildung 7.12a, in der wir sehen, dass in unserer nichtkonformen Formulierung Problem 4.1 der Term  $\sum_{F \in \mathcal{E}} \| [u_{\text{CR},\mathcal{T}}]_F \|_{L^1(F)} = E(u_{\text{CR},\mathcal{T}}) - E_{\text{NC}}(u_{\text{CR},\mathcal{T}})$  tatsächlich nicht minimiert wird. Zwar sind die jeweiligen Sprünge zwischen zwei Dreiecken bei einer hohen Anzahl von Freiheitsgraden klein, wie Abbildung 7.10 erahnen lässt, jedoch führt eine hohe Anzahl an Kanten zur Addition von vielen Sprüngen und scheint somit die Minimierung dieser Summe zu verhindern. Außerdem wollen wir anmerken, dass aus Theorem 3.7 und Abschnitt 4.1 folgt, dass

$$\frac{\alpha}{2} \|u - u_{\text{CR},\mathcal{T}}\|^2 \leq |E_{\text{NC}}(u_{\text{CR},\mathcal{T}}) - E(u)| + \sum_{F \in \mathcal{E}} \| [u_{\text{CR},\mathcal{T}}]_F \|_{L^1(F)}.$$

Nach Abbildung 7.11 gilt sogar die Aussage  $\frac{\alpha}{2} \|u - u_{\text{CR},\mathcal{T}}\|^2 \leq |E_{\text{NC}}(u_{\text{CR},\mathcal{T}}) - E(u)|$ , obwohl  $\sum_{F \in \mathcal{E}} \| [u_{\text{CR},\mathcal{T}}]_F \|_{L^1(F)}$  nicht minimiert wird.

Die Graphen von  $E_{\text{GUEB},\mathcal{T}} - E_{\text{GLEB},\mathcal{T}}$  sowie von  $E(u) - E_{\text{GLEB},\mathcal{T}}$  zeigen bei uniformer Netzverfeinerung die Konvergenzrate  $1/2$  und bei adaptiver Netzverfeinerung, wie bereits begründet, eine etwas geringere. Dabei gilt stets

$$\frac{\alpha}{2} \|u - u_{\text{CR},\mathcal{T}}\|^2 \leq E(u) - E_{\text{GLEB},\mathcal{T}} \leq E_{\text{GUEB},\mathcal{T}} - E_{\text{GLEB},\mathcal{T}}, \quad (7.3)$$

wie wir nach Theorem 4.9 und Ungleichung 4.15 erwartet haben. Die Gültigkeit der zweiten Ungleichung in (7.3) können wir deutlich an der Differenz  $E_{\text{GUEB},\mathcal{T}} - E(u)$  von  $E_{\text{GUEB},\mathcal{T}} - E_{\text{GLEB},\mathcal{T}}$  und  $E(u) - E_{\text{GLEB},\mathcal{T}}$  erkennen, die in Abbildung 7.12b dargestellt ist. Dabei konvergiert der quadrierte exakte Fehler  $\frac{\alpha}{2} \|u - u_{\text{CR},\mathcal{T}}\|^2$  mit Rate 1. Nach [Bar15b, S. 309, Theorem 10.7] ist für das dort betrachtete Problem, das heißt der Diskretisierung des ROF-Modells mit dem Courant-Finite-Elemente-Raum  $S^1(\mathcal{T})$  für ein Eingangssignal  $g \in L^\infty(\Omega)$ , für den quadrierten  $L^2$ -Fehler zwischen den Minimierern  $u_C \in S^1(\mathcal{T})$  und  $u \in \text{BV}(\Omega) \cap L^2(\Omega)$  des Funktionalen  $I$  aus Gleichung (1.1) in den entsprechenden

## 7 Numerische Beispiele

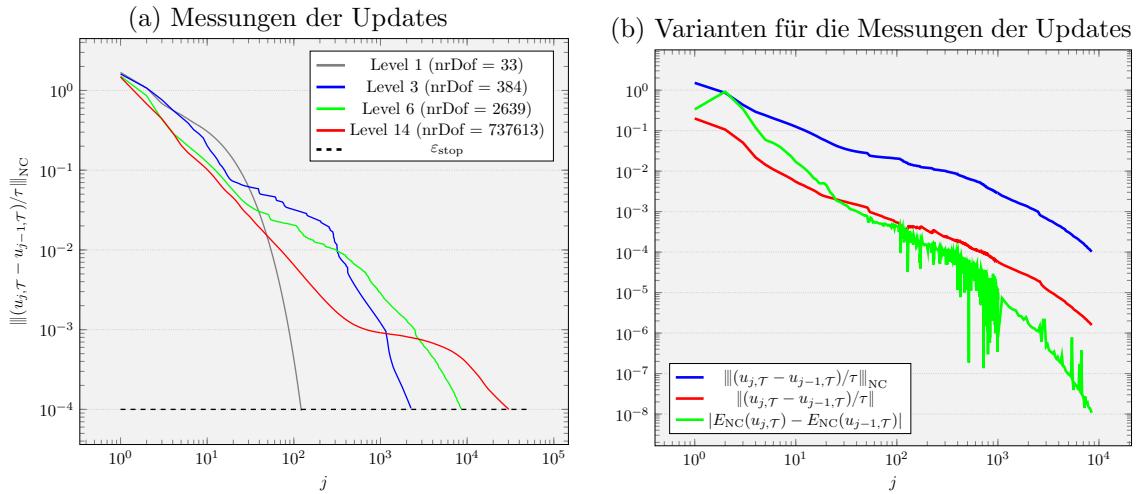


Abbildung 7.9: Verlauf der Messungen der Updates aus dem Abbruchkriterium (5.3) während der primalen-dualen Iteration auf verschiedenen Leveln der AFEM-Routine (a) und Verlauf von Varianten, diese Updates zu messen, für ein Level mit 2 639 Freiheitsgraden (b) für das Eingangssignal  $f_1$

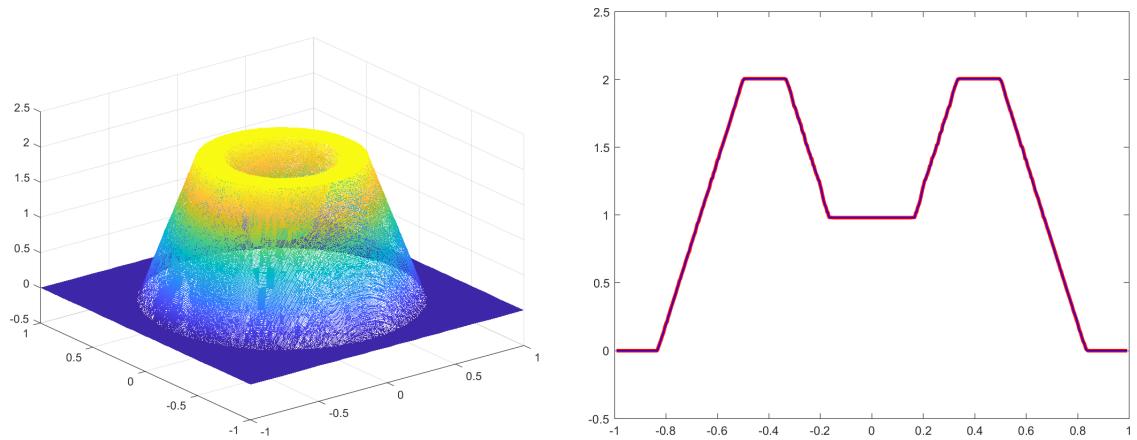


Abbildung 7.10: Lösung der Rechnung bei adaptiver Netzverfeinerung sowie deren Darstellungen entlang der x-Achse (blau) und der y-Achse (rot) nach etwa 740 000 Freiheitsgraden mit Eingangssignal  $f_1$

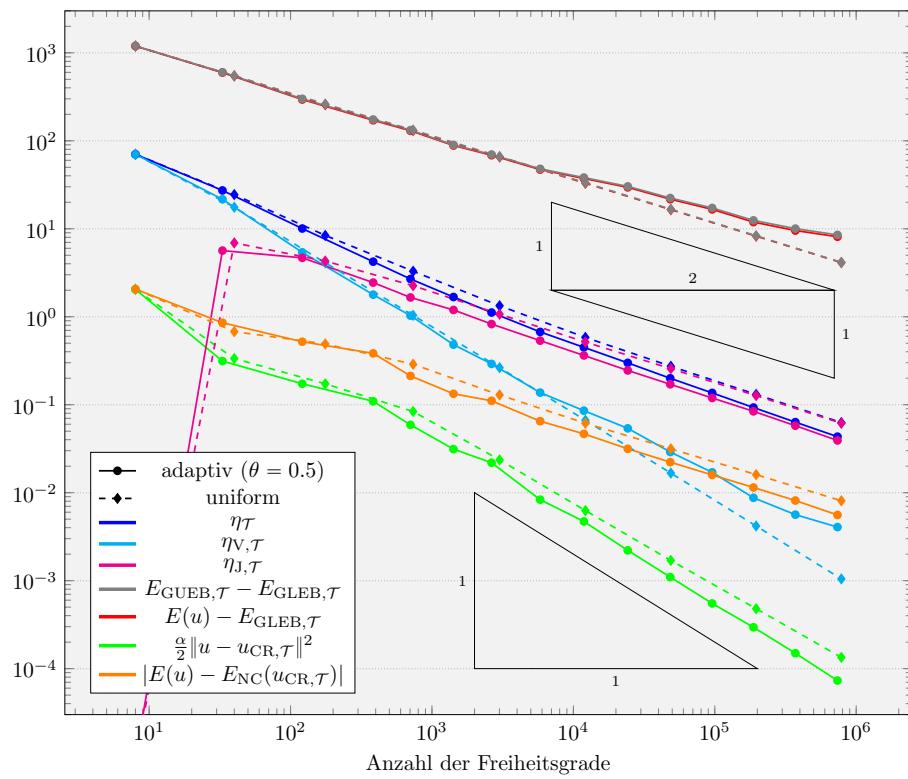


Abbildung 7.11: Ergebnisse der Rechnungen mit adaptiver und uniformer Netzverfeinerung für das Eingangssignal  $f_1$

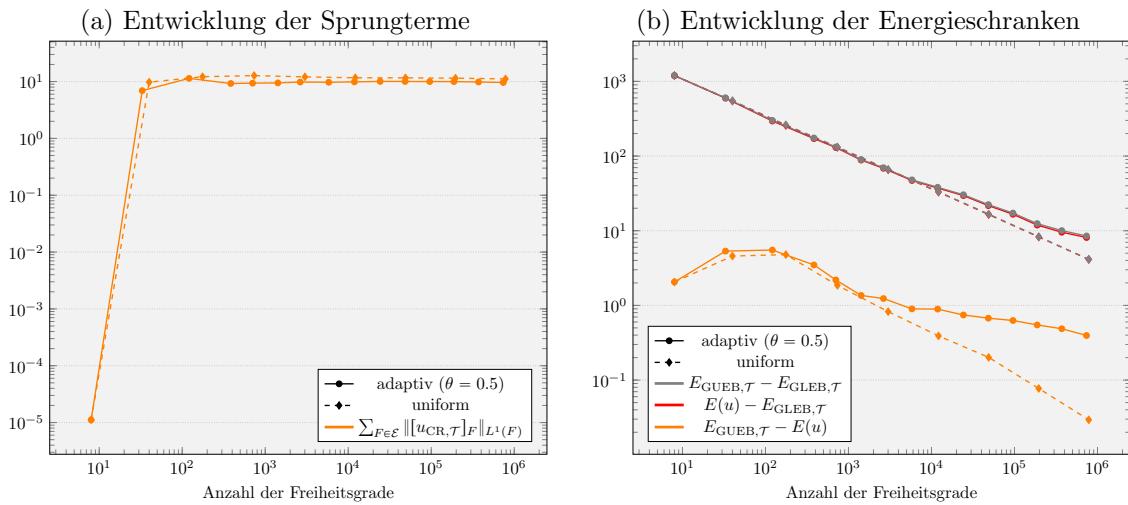


Abbildung 7.12: Zusätzliche Ergebnisse der Rechnungen mit adaptiver und uniformer Netzverfeinerung für das Eingangssignal  $f_1$

## 7 Numerische Beispiele

Räumen eine Rate von  $1/4$  garantiert. Obwohl wir eine alternative Formulierung des ROF-Modells betrachten und diese mit dem Crouzeix-Raviart-Finite-Elemente-Raum diskretisieren, können wir festhalten, dass die hier beobachtete Konvergenzrate für  $\frac{\alpha}{2} \|u - u_{\text{CR},\mathcal{T}}\|^2$  deutlich besser ist. Dabei bleibt aber noch hervorzuheben, dass für das Experiment in Abbildung 7.11 gilt  $u, f_1 \in H_0^1((-1, 1)^2)$ . Insbesondere ist unsere exakte Lösung sogar schwach differenzierbar und nicht nur eine Funktion von beschränkter Variation, weshalb die bessere Rate nicht unerwartet ist. Um ein anderes Eingangssignal für ein Experiment mit bekannter exakter Lösung zu betrachten und dabei zu untersuchen, ob noch stärkere Annahmen an die schwache Differenzierbarkeit der exakten Lösung und des Eingangssignals die Raten weiter verbessern können, betrachten wir nun die Funktion  $u_{\text{HR}} \in H_0^2((-1, 1)^2)$ , die gegeben ist durch

$$u_{\text{HR}}(r) := \begin{cases} 1, & \text{falls } r \in [0, \frac{1}{3}], \\ 54r^3 - 81r^2 + 36r - 4, & \text{falls } r \in (\frac{1}{3}, \frac{2}{3}], \\ 0, & \text{falls } r \in (\frac{2}{3}, \infty). \end{cases}$$

Mit der Wahl

$$\operatorname{sgn}(\partial_r u_{\text{HR}}(r)) := \begin{cases} -1458r^5 + 1215r^4 - 270r^3, & \text{falls } r \in [0, \frac{1}{3}], \\ -1, & \text{falls } r \in (\frac{1}{3}, \frac{2}{3}], \\ -243r^4 + 756r^3 - 864r^2 + 432r - 81, & \text{falls } r \in (\frac{2}{3}, \infty), \end{cases}$$

ist  $u_{\text{HR}}$  die exakte Lösung von Problem 3.1 mit Eingangssignal

$$f_{\text{HR}}(r) := \begin{cases} \alpha + 8748r^4 - 6075r^3 + 1080r^2, & \text{falls } r \in [0, \frac{1}{3}], \\ \alpha(54r^3 - 81r^2 + 36r - 4) + \frac{1}{r}, & \text{falls } r \in (\frac{1}{3}, \frac{2}{3}], \\ 1215r^3 - 3024r^2 + 2592r - 864 + \frac{81}{r}, & \text{falls } r \in (\frac{2}{3}, \infty), \end{cases}$$

für das ebenfalls gilt  $f_{\text{HR}} \in H_0^2((-1, 1)^2)$ . Die schwachen Ableitungen können wieder berechnet werden durch die partiellen Ableitungen

$$\partial_r f_{\text{HR}}(r) = \begin{cases} 34992r^3 - 18225r^2 + 2160r, & \text{falls } r \in [0, \frac{1}{3}], \\ \alpha(162r^2 - 162r + 36) - \frac{1}{r^2}, & \text{falls } r \in (\frac{1}{3}, \frac{2}{3}], \\ 3645r^2 - 6048r + 2592 - \frac{81}{r^2}, & \text{falls } r \in (\frac{2}{3}, \infty), \end{cases}$$

und

$$\partial_r u_{\text{HR}}(r) = \begin{cases} 0, & \text{falls } r \in [0, \frac{1}{3}], \\ 162r^2 - 162r + 36, & \text{falls } r \in (\frac{1}{3}, \frac{2}{3}], \\ 0, & \text{falls } r \in (\frac{2}{3}, \infty). \end{cases}$$

Wir wählen für das Experiment  $\alpha = 1$  und erhalten für die Energie der exakten Lösung  $E(u_{\text{HR}}) \approx -0.33411$ . Dieses Eingangssignal und die exakte Lösung sind in Abbildung 7.13 dargestellt. In Abbildung 7.14 stellen wir lediglich fest, dass im Vergleich zu Abbildung 7.11 alle Graphen um einen Faktor von etwa  $10^{1/2}$  nach unten verschoben sind. Insbesondere verändern sich die Konvergenzraten nicht. Eine noch stärkere Regularitätsannahme scheint die Raten also nicht weiter zu verbessern, es werden aber geringfügig kleinere Werte erreicht. Möglicherweise sind bei der hier benutzten Diskretisierung keine besseren Raten erreichbar und die Betrachtung einer Methode höherer Ordnung wäre dazu notwendig.

Nun betrachten wir noch einmal das Eingangssignal  $f_1$ , um den Einfluss der Wahl von  $\gamma \in (0, 1]$  aus der Definition 4.10 des Verfeinerungsindikators  $\eta_{\mathcal{T}}$  auf die Rechungen zu untersuchen. Je kleiner die Wahl von  $\gamma$ , desto dominanter sollte nach Definition der Einfluss

### 7.3 Experimente mit bekannter exakter Lösung

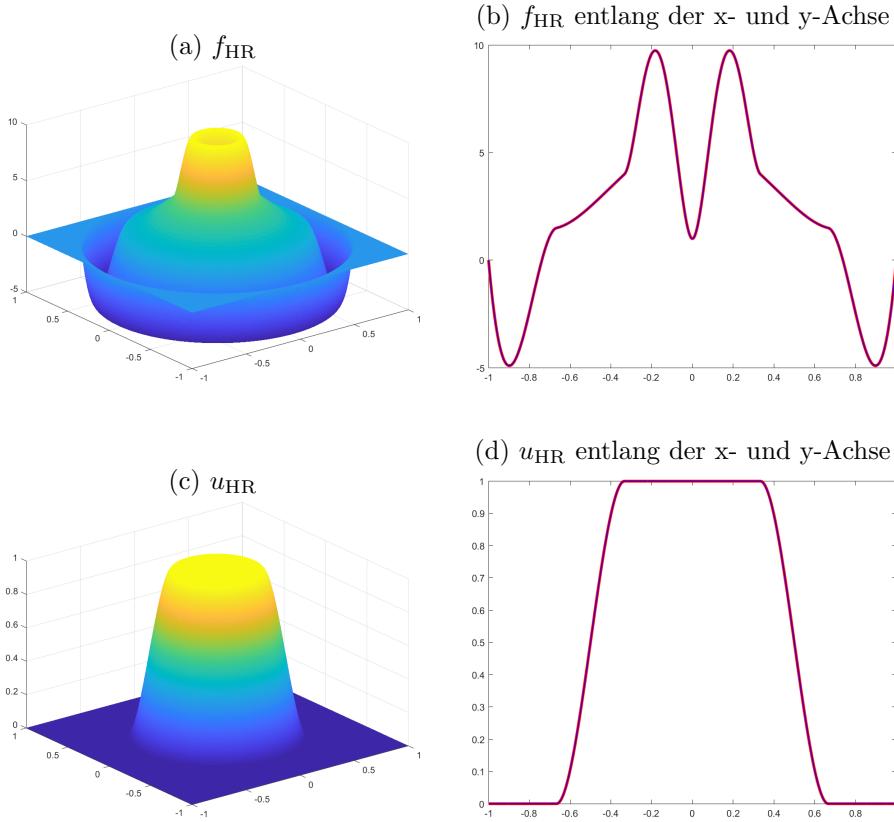


Abbildung 7.13: Funktionen  $f_{\text{HR}}$  und  $u_{\text{HR}}$  sowie deren Darstellungen entlang der x-Achse (blau) und der y-Achse (rot) für  $\alpha = 1$

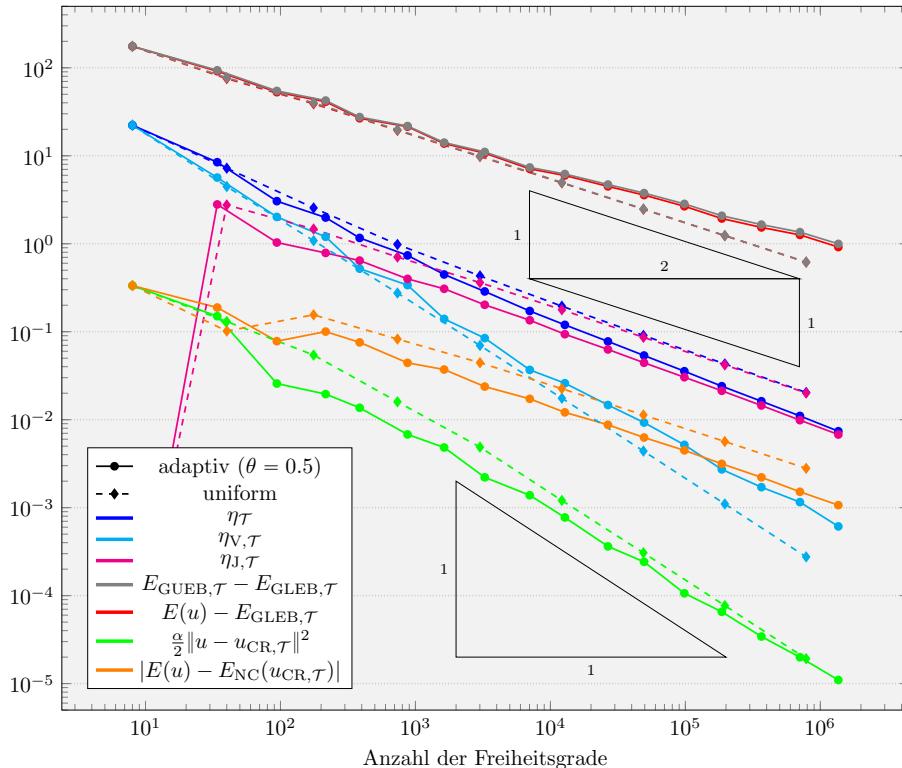


Abbildung 7.14: Ergebnisse der Rechnungen mit adaptiver und uniformer Netzverfeinerung für das Eingangssignal  $f_{\text{HR}}$

## 7 Numerische Beispiele

von  $\eta_{J,\tau}$  auf  $\eta_\tau$  sein und entsprechend dort stärker verfeinert werden, wo die Kanten sprünge größer sind. Dieser Effekt sollte auch für  $\gamma = 0$  zu beobachten sein, weshalb wir diese Wahl ebenfalls untersuchen. In Abbildung 7.15 ist tatsächlich zu erkennen, dass  $\eta_{J,\tau}$  mit kleinerem  $\gamma$  deutlich größere Werte annimmt, sodass sogar die Konvergenzrate von  $\eta_{J,\tau}$  abnimmt. Da  $\eta_{J,\tau}$  den Verfeinerungsindikator  $\eta_\tau$  dominiert, zeigt dieser das gleiche Konvergenzverhalten. Für  $\gamma = 0$  stagniert  $\eta_{J,\tau}$  sogar. Eine Erklärung für diese Beobachtungen ist, dass, wie bereits festgestellt, die Sprunsterme  $\sum_{F \in \mathcal{E}(\tau)} \| [u_{CR,\tau}]_F \|_{L^1(F)}$  nicht minimiert werden. Insbesondere für  $\gamma = 0$  wird somit im Verfeinerungsindikator der Anteil  $\eta_{J,\tau}$  stark gewichtet, der in diesem Fall nicht gegen 0 konvergiert. Weiterhin ist die Konvergenzrate von  $\eta_{V,\tau}$  etwas schlechter, wenn  $\gamma$  klein gewählt wird. Die Fehlergraphen  $\frac{\alpha}{2} \|u - u_{CR,\tau}\|^2$  und  $|E(u) - E_{NC}(u_{CR,\tau})|$  zeigen allerdings für die verschiedenen Wahlen von  $\gamma$  keine nennenswerten Unterschiede. Die für den Verfeinerungsindikator suboptimalen Wahlen von  $\gamma$  scheinen also keinen negativen Effekt auf diese zu haben. In Abbildung 7.16 ist die Wirkung des Verfeinerungsindikators gut zu erkennen. Die Kanten sprünge sind dort am größten, wo die Lösung  $u$ , und damit die meisten Iterate, nicht konstant sind. Dementsprechend wird mit kleinerem  $\gamma$  stark dort verfeinert, wo die Lösung  $u$  aus Abbildung 7.2 nicht konstant ist. Dies ist zu erkennen, obwohl die abgebildeten Triangulierungen für kleine  $\gamma$  deutlich mehr Freiheitsgrade haben als für  $\gamma = 1$ . Da wir im Experiment zu Abbildung 7.11 bereits gesehen haben, dass die Konvergenzraten bei adaptiver Netzverfeinerung nicht besser sind als bei uniformer Netzverfeinerung, ist die ausbleibende Verbesserung der Ergebnisse durch eine für  $\gamma < 1$  adaptivere Verfeinerung nicht unerwartet. Insgesamt halten wir fest, dass die zu Beginn des Kapitels getroffene Wahl von  $\gamma = 1$  weiterhin sinnvoll bleibt, da diese keinen Einfluss auf die Fehlergraphen zu haben scheint, aber bessere Konvergenzraten des Verfeinerungsindikators und seiner Anteile im Vergleich zu kleineren Werten von  $\gamma$  bewirkt. Auch hier möchten wir anmerken, dass Ungleichung (7.3) für alle Wahlen von  $\gamma$  erfüllt ist.

Als Nächstes wollen wir die Experimente mit Eingangssignal  $f_1$  aus Abbildung 7.11 noch mit den Experimenten mit  $\alpha = 10^4$  und dem Eingangssignal  $f_{10^4}$  vergleichen. Die Konvergenzraten für die Experimente mit Eingangssignal  $f_{10^4}$  in Abbildung 7.17 sind während eines vorasymptotischen Bereichs, der bei etwa  $10^5$  Freiheitsgraden endet, größer als bei den Experimenten mit Eingangssignal  $f_1$ . Danach sehen wir bei allen Graphen, außer dem für  $|E(u) - E_{NC}(u_{CR,\tau})|$ , die anscheinend gleichen Raten, die wir schon in Abbildung 7.11 beobachtet haben. Die Konvergenzrate von  $|E(u) - E_{NC}(u_{CR,\tau})|$  scheint nun 1 statt wie zuvor  $1/2$  zu sein. Ansonsten gibt es nur zwei weitere nennenswerte Unterschiede. Zum einen dominiert  $\eta_{V,\tau}$  hier lange den Verfeinerungsindikator, wobei ab etwa  $10^6$  Freiheitsgraden die Dominanz von  $\eta_{J,\tau}$  zu beginnen scheint, zum anderen erreichen alle Graphen, mit Ausnahme des quadrierten exakten Fehlers und  $\eta_{V,\tau}$ , geringere Werte nach  $10^6$  Freiheitsgraden als für das Eingangssignal  $f_1$ . Die Existenz des großen vorasymptotischen Bereichs wird möglicherweise dadurch verursacht, dass für  $\alpha = 10^4$  alle Graphen, abgesehen von  $\eta_{J,\tau}$ , zu Beginn des AFEM-Algorithmus deutlich größere Werte annehmen als für  $\alpha = 1$ . Das Eingangssignal  $f_{10^4}$  nimmt ebenfalls signifikant größere Werte an als  $f_1$ . Die Iterationen auf den ersten Leveln der AFEM-Routine erreichen dann jeweils starke Reduktionen der betrachteten Terme und bedingen somit den vorasymptotischen Bereich. Die Ungleichung (7.3) ist nach Abbildung 7.17 auch hier stets gültig.

Zum Abschluss dieses Abschnitts möchten wir für die betrachteten Experimente mit den Eingangssignalen  $f_1$ ,  $f_{10^4}$  und  $f_{HR}$  die Anzahl der Iterationen vergleichen. In Abbildung 7.18 ist zu erkennen, dass bis  $10^6$  Freiheitsgrade das Beispiel mit Eingangssignal  $f_{10^4}$ , welches im Gegensatz zu den anderen beiden Rechnungen den Parameter  $\alpha = 10^4$  nutzt, bei einer vergleichbaren Anzahl von Freiheitsgraden deutlich weniger Iterationen benötigt. Ähnlich zu Abschnitt 7.1 für den Parameter  $\tau$  könnte auch hier Ungleichung (7.1), deren rechte Seite antiproportional zu  $\alpha$  ist, der Grund für diese Beobachtung sein.

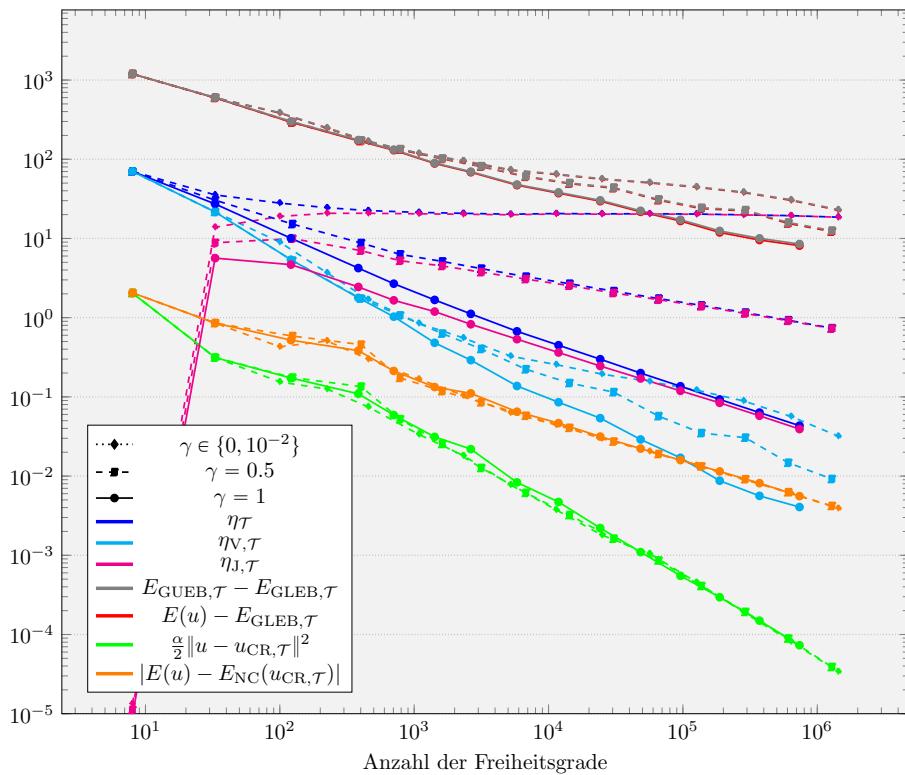


Abbildung 7.15: Ergebnisse des Programms mit verschiedenen Werten von  $\gamma$  für das Eingangssignal  $f_1$ , wobei die Graphen für  $\gamma \in \{0, 10^{-2}\}$  nicht sichtbar unterscheidbar sind und deshalb nur der Graph für  $\gamma = 0$  dargestellt ist

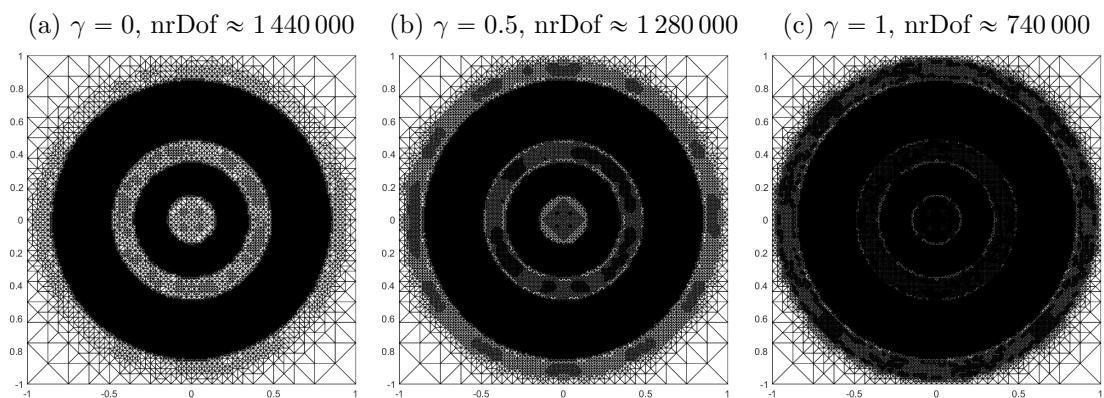


Abbildung 7.16: Triangulierungen des jeweils letzten Levels 14 der Rechnungen mit verschiedenen Werten von  $\gamma$  für das Eingangssignal  $f_1$

## 7 Numerische Beispiele

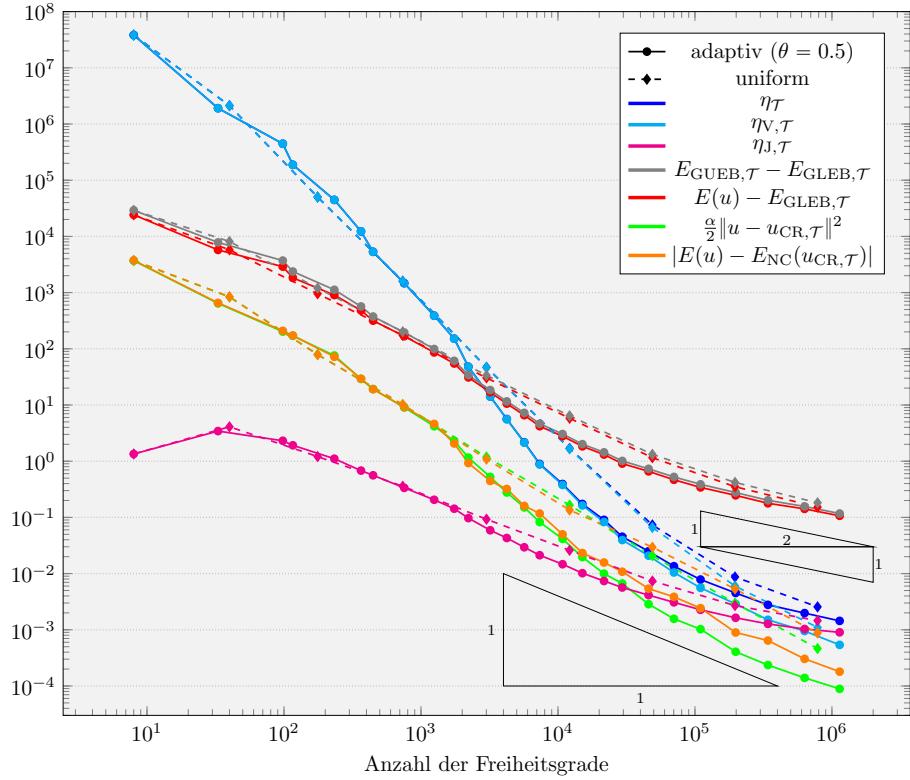


Abbildung 7.17: Ergebnisse der Rechnungen mit adaptiver und uniformer Netzverfeinerung für das Eingangssignal  $f_{10^4}$

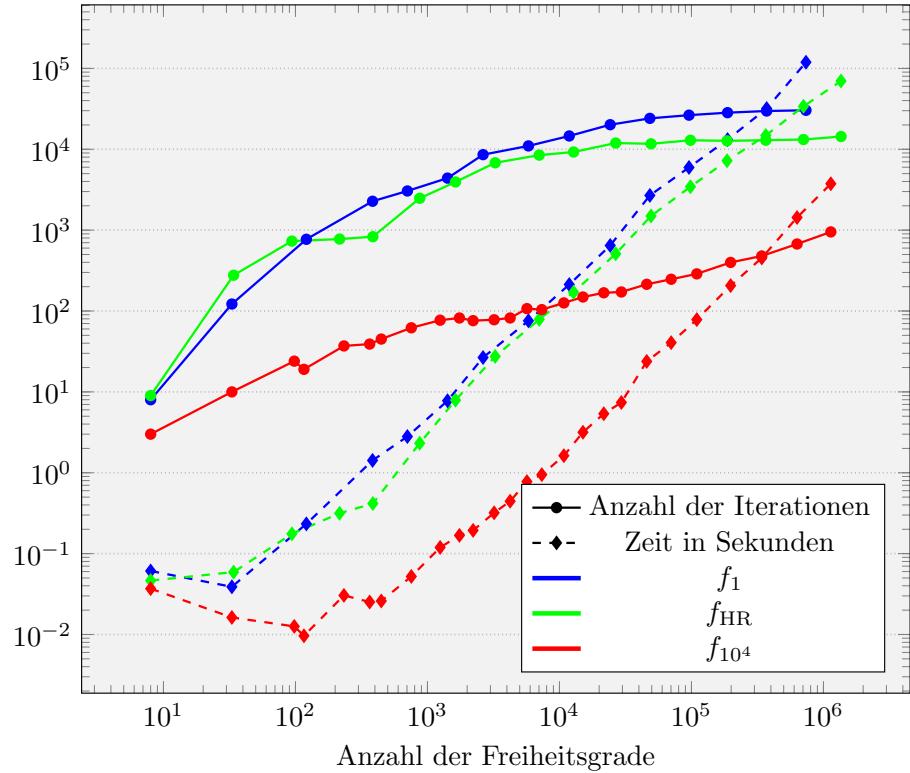


Abbildung 7.18: Anzahl der Iterationen und benötigte Zeit für die primalen-dualen Iterationen während des AFEM-Algorithmus bei den Rechnungen mit  $\alpha = 1$  für die Eingangssignale  $f_1$  und  $f_{\text{HR}}$  sowie bei der Rechnung mit  $\alpha = 10^4$  für das Eingangssignal  $f_{10^4}$

## 7.4 Graufarbenbilder als Eingangssignale

In diesem Abschnitt wollen wir Ergebnisse von Experimenten untersuchen, bei denen als Eingangssignal ein Graufarbenbild nach Bemerkung 6.1 gegeben ist. Für diese Rechnungen sind entsprechend weder exakte Lösung noch die schwache Ableitung des Eingangssignals bekannt. Zunächst betrachten wir das Eingangssignal `cameraman` aus Abbildung 1.1a. In Abbildung 7.19 sehen wir für den Verfeinerungsindikator  $\eta_{\mathcal{T}}$  und dessen Anteil  $\eta_{V,\mathcal{T}}$  die Konvergenzrate 1. Für den Anteil  $\eta_{J,\mathcal{T}}$  von  $\eta_{\mathcal{T}}$  ist ungefähr die Rate 1/2 zu erkennen, wobei  $\eta_{J,\mathcal{T}}$  bei uniformer Netzverfeinerung für das letzte Level einen etwas größeren Wert annimmt als für das vorletzte. Ansonsten ist bei den Konvergenzraten kein Unterschied zwischen uniformer und adaptiver Netzverfeinerung festzustellen. Die Raten von  $\eta_{V,\mathcal{T}}$  und  $\eta_{J,\mathcal{T}}$  stimmen mit den Raten überein, die schon beim Experiment mit Eingangssignal  $f_1$  in Abbildung 7.11 beobachtet wurden, wobei hier der Verfeinerungsindikator von  $\eta_{V,\mathcal{T}}$  dominiert wird und nicht, wie beim Experiment mit Eingangssignal  $f_1$ , von  $\eta_{J,\mathcal{T}}$ . Dementsprechend konvergiert hier  $\eta_{\mathcal{T}}$  auch nicht mit der Rate 1/2, sondern mit der Rate 1. Weiterhin nimmt  $\eta_{V,\mathcal{T}}$ , und somit auch  $\eta_{\mathcal{T}}$ , im Vergleich zur Rechnung mit Eingangssignal  $f_1$  zu Beginn der AFEM-Routine deutlich höhere Werte an. Diese sind vergleichbar mit den beim Experiment mit Eingangssignal  $f_{10^4}$  in Abbildung 7.17, bei dem ebenfalls  $\alpha = 10^4$  gewählt war, beobachteten Werten. Bei einer Rechnung mit einem Eingangssignal  $f$ , bei dem  $\alpha$  groß gewählt ist im Vergleich zu einem anderen Experiment, scheinen also auf den ersten Leveln der AFEM-Routine höhere Werte für  $\eta_{V,\mathcal{T}}$  aufzutreten als in einem Experiment mit einer kleineren Wahl von  $\alpha$ . Dies wird bekräftigt in Abbildung 7.20, in der wir die Konvergenzgraphen zu den Experimenten aus Abbildung 1.1 vergleichen für  $\alpha \in \{10^2, 10^3, 10^4\}$  und ebendieses Verhalten beobachten. Ein Grund dafür könnte sein, dass, wenn  $\alpha$  groß gewählt wird, die Terme  $\|f - \alpha u_{CR,\mathcal{T}}\|_{L^2(T)}^2$  im Anteil  $\eta_{V,\mathcal{T}}$  des Verfeinerungsindikators aus Definition 4.10 für viele Dreiecke  $T \in \mathcal{T}$  ebenfalls größere Werte annehmen als für ein Experiment mit einer kleineren Wahl für  $\alpha$ . Weiterhin können wir durch die benötigte Anzahl der Iterationen für jedes Level der Experimente zu Abbildung 1.1 in Abbildung 7.21 noch einmal unsere Hypothese aus Abschnitt 7.1 zur Ungleichung (7.1) stützen, denn auch in Abbildung 7.21 benötigt die primale-duale Iteration für eine größere Wahl von  $\alpha$  weniger Iterationsschritte.

Zum Abschluss dieser Abschnitts möchten wir anhand des Beispiels mit Eingangssignal `cameraman` die Wirkung des Verfeinerungsindikators veranschaulichen. Wir sehen in Abbildung 7.22a gut, dass dieser eine Verfeinerung zu den Unstetigkeiten bewirkt, insbesondere zu denen entlang deutlicher Farbkontraste. Außerdem ist zu erkennen, dass der hinzugefügte graduelle Übergang zu schwarzem Rand für das Eingangssignal den gewünschten Effekt hat, eine starke Verfeinerung zum Rand, die aufgrund der angenommenen Nullranddaten passieren würde, zu verhindern. Somit wird tatsächlich dort stark verfeinert, wo mehr Informationen benötigt werden, was am Rand nicht der Fall wäre. Auch in Abbildung 7.22c kann man, trotz der hohen Anzahl an Freiheitsgraden, noch erkennen, wo es im Bild kaum Farbkontraste gibt und eine Verfeinerung deshalb nicht unbedingt nötig ist. Die diskrete Lösung ähnelt dem Eingangssignal, was aufgrund der großen Wahl für  $\alpha$  nach der Interpretation des ROF-Modellproblems aus Kapitel 1 zu erwarten war.

## 7.5 Stetige Approximation eines unstetigen Eingangssignals

Im vorherigen Abschnitt 7.4 konnten wir mangels bekannter exakter Lösungen und schwacher Gradienten der Eingangssignale für die Experimente nur die Konvergenzraten des Verfeinerungsindikators und seiner Anteile  $\eta_{V,\mathcal{T}}$  und  $\eta_{J,\mathcal{T}}$  untersuchen. Da wir für  $\eta_{V,\mathcal{T}}$  und  $\eta_{J,\mathcal{T}}$  die gleichen Raten wie für die Experimente in Abschnitt 7.3 beobachten konnten, interessiert uns, ob andere Konvergenzraten ebenfalls vergleichbar sind zwischen Experi-

## 7 Numerische Beispiele

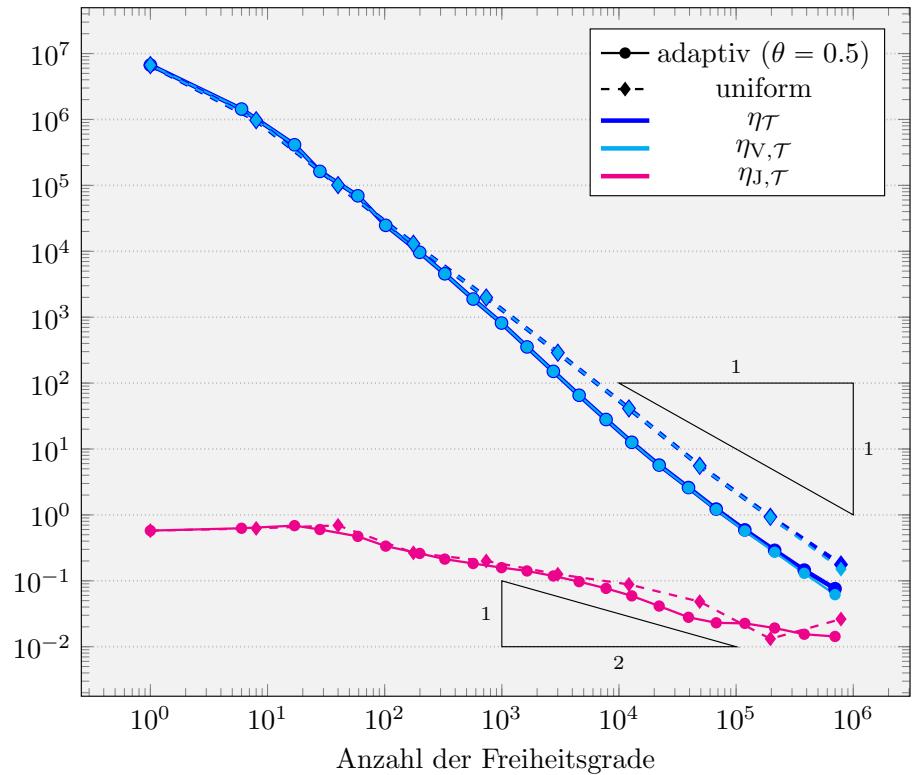


Abbildung 7.19: Ergebnisse der Rechnungen mit adaptiver und uniformer Netzverfeinerung für das Eingangssignal **cameraman**

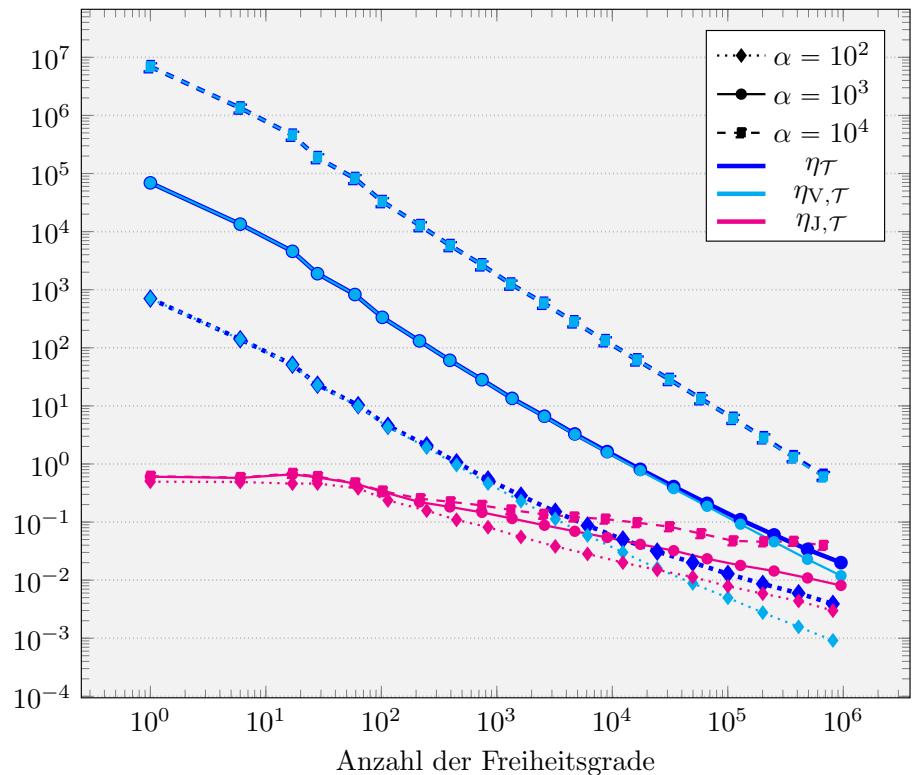


Abbildung 7.20: Ergebnisse der Rechnungen zu Abbildung 1.1 für drei verschiedene Werte von  $\alpha$

## 7.5 Stetige Approximation eines unstetigen Eingangssignals

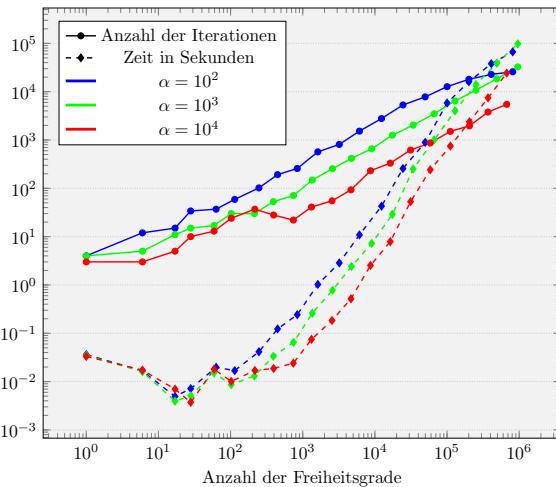


Abbildung 7.21: Anzahl der Iterationen und benötigte Zeit für die primalen-dualen Iterationen während des AFEM-Algorithmus bei den Rechnungen zu Abbildung 1.1 für drei verschiedene Werte von  $\alpha$



Abbildung 7.22: Triangulierung und diskrete Lösung, als Graufarbenplot aus der Draufsicht, nach der Rechnung mit Eingangssignal `cameraman` für Level 17 des AFEM-Algorithmus nach etwa 70 000 Freiheitsgraden und für Level 21 nach etwa 700 000 Freiheitsgraden

## 7 Numerische Beispiele

menten mit stetigen Eingangssignalen und solchen mit unstetigen Eingangssignalen. Deshalb möchten wir nun ein weiteres Beispiel mit einem unstetigen Eingangssignal, welches wir durch eine nach Abschnitt 3.2 konstruierte stetige Funktion approximieren können, betrachten. Für die Rechnungen mit dieser stetigen Approximation als Eingangssignal kennen wir die exakte Lösung sowie die schwachen Gradienten und können durch die Konvergenzraten der entsprechenden Graphen möglicherweise Rückschlüsse für die Raten der Rechnungen mit unstetigen Eingangssignalen ziehen. Wir nutzen für alle Experimente in diesem Abschnitt als initiale Triangulierung **BigSquare** aus Abbildung 7.1a. Als unstetiges Eingangssignal, das als Graufarbenbild interpretiert einem weißen Kreis mit Radius  $\frac{1}{2}$  entspricht, betrachten wir die Funktion

$$f_{DC}(r) := \begin{cases} 10^4, & \text{falls } r \in [0, \frac{1}{2}], \\ 0, & \text{falls } r \in (\frac{1}{2}, \infty). \end{cases}$$

Betrachten wir nun für einen Parameter  $\beta \in (0, 1)$  die Funktion

$$u_C(r) := \begin{cases} 1, & \text{falls } r \in [0, \frac{1-\beta}{2}], \\ -\frac{1}{\beta}r + \frac{1+\beta}{2\beta}, & \text{falls } r \in (\frac{1-\beta}{2}, \frac{1+\beta}{2}), \\ 0, & \text{falls } r \in (\frac{1+\beta}{2}, \infty), \end{cases}$$

so ist diese mit der Wahl

$$\operatorname{sgn}(\partial_r u_C(r)) := \begin{cases} \frac{4}{1-\beta}r \left( \frac{1}{1-\beta}r - 1 \right), & \text{falls } r \in [0, \frac{1-\beta}{2}], \\ -1, & \text{falls } r \in (\frac{1-\beta}{2}, \frac{1+\beta}{2}), \\ \frac{4}{(\beta-1)^3} (4r^3 - 3(\beta+3)r^2 + 6(\beta+1)r - 3\beta - 1), & \text{falls } r \in (\frac{1+\beta}{2}, \infty), \end{cases}$$

Lösung von Problem 3.1 mit Eingangssignal

$$f_C(r) := \begin{cases} \alpha - \frac{4}{1-\beta} \left( \frac{3}{1-\beta}r - 2 \right), & \text{falls } r \in [0, \frac{1-\beta}{2}], \\ -\frac{\alpha}{\beta} \left( r - \frac{1+\beta}{2} \right) + \frac{1}{r}, & \text{falls } r \in (\frac{1-\beta}{2}, \frac{1+\beta}{2}), \\ -\frac{4}{(\beta-1)^3} \left( 16r^2 - 9(\beta+3)r + 12(\beta+1) - \frac{3\beta+1}{r} \right), & \text{falls } r \in (\frac{1+\beta}{2}, \infty). \end{cases}$$

Die schwachen Gradienten von  $u_C$  und  $f_C$  können durch die partiellen Ableitungen

$$\partial_r f_C(r) = \begin{cases} -\frac{12}{(1-\beta)^2}, & \text{falls } r \in [0, \frac{1-\beta}{2}], \\ -\frac{\alpha}{\beta} - \frac{1}{r^2}, & \text{falls } r \in (\frac{1-\beta}{2}, \frac{1+\beta}{2}), \\ -\frac{4}{(\beta-1)^3} \left( 32r - 9(\beta+3) + \frac{3\beta+1}{r^2} \right), & \text{falls } r \in (\frac{1+\beta}{2}, \infty), \end{cases}$$

und

$$\partial_r u_C(r) = \begin{cases} 0, & \text{falls } r \in [0, \frac{1-\beta}{2}], \\ -\frac{1}{\beta}, & \text{falls } r \in (\frac{1-\beta}{2}, \frac{1+\beta}{2}), \\ 0, & \text{falls } r \in (\frac{1+\beta}{2}, \infty), \end{cases}$$

bestimmt werden. Wir wählen für unsere Experimente  $\alpha = 10^4$  und  $\beta = 10^{-3}$  und erhalten damit als Energie der exakten Lösung  $E(u) \approx -3924.37413$ . Augenscheinlich ist  $f_C$  für diese Wahl von  $\alpha$  und  $\beta$  tatsächlich eine stetige Approximation von  $f_{DC}$ , wie in Abbildung 7.23a zu sehen ist. Außerdem sehen wir in Abbildung 7.23 für diese große Wahl von  $\alpha$

## 7.5 Stetige Approximation eines unstetigen Eingangssignals

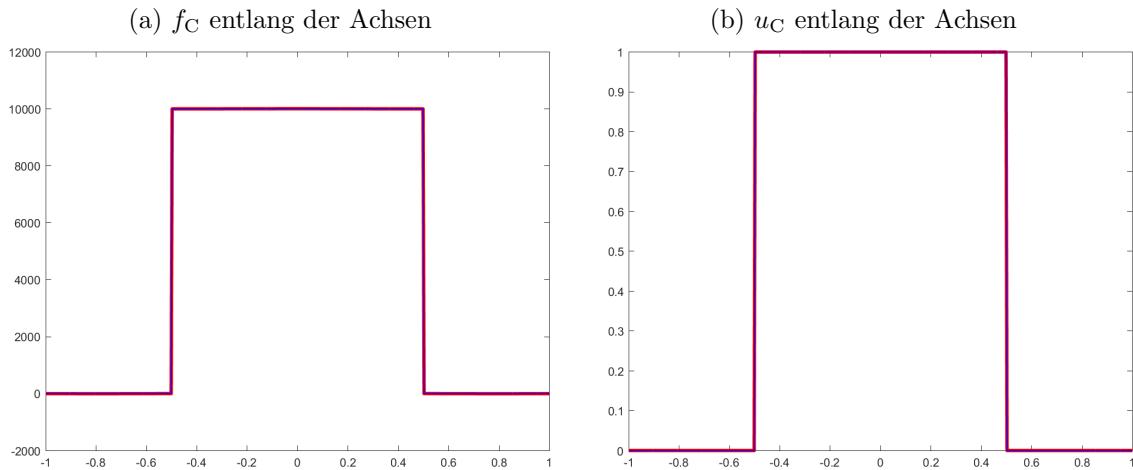


Abbildung 7.23: Funktionen  $f_C$  und  $u_C$  entlang der x-Achse (blau) und der y-Achse (rot)  
für  $\alpha = 10^4$  und  $\beta = 10^{-3}$

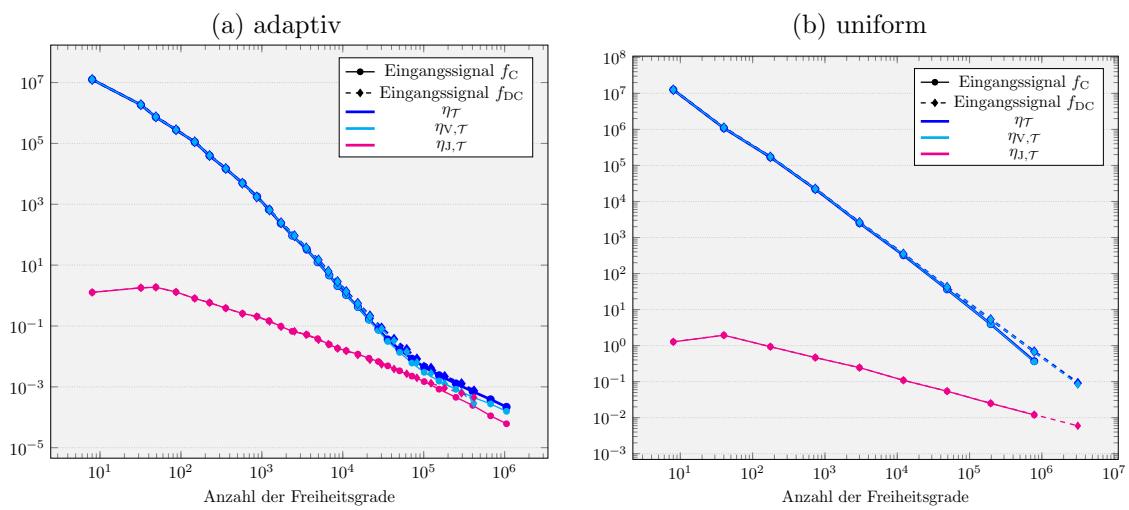


Abbildung 7.24: Vergleich des Verfeinerungsindikators und seiner Anteile zwischen den Rechnungen mit Eingangssignal  $f_C$  und  $f_{DC}$  bei adaptiver (a) und uniformer (b) Netzverfeinerung

noch einmal, dass das Eingangssignal  $f_C$  ungefähr gleich dem Produkt aus  $\alpha$  und der exakten Lösung  $u_C$  zu sein scheint, wie wir nach Kapitel 1 erwartet haben. Zunächst bestätigt Abbildung 7.24 sowohl für adaptive als auch uniforme Netzverfeinerung, dass es zwischen den Experimenten mit den Eingangssignalen  $f_C$  und  $f_{DC}$  bis etwa  $10^6$  Freiheitsgrade keine deutlichen Unterschiede in der Entwicklung des Verfeinerungsindikators und seiner Anteile gibt. In Abbildung 7.25 ist weiterhin zu sehen, dass der Verfeinerungsindikator bei den adaptiven Rechnungen mit den Eingangssignalen  $f_C$  und  $f_{DC}$  eine vergleichbare Netzverfeinerung zur Unstetigkeitsmenge der Funktion  $f_{DC}$  bewirkt. Außerdem ähneln die in Abbildung 7.26 dargestellten diskreten Lösungen der Rechnungen mit den Eingangssignalen  $f_C$  und  $f_{DC}$  beide der exakten Lösung für das Eingangssignal  $f_C$  aus Abbildung 7.23b. Insgesamt scheinen sich die vergleichbaren Ergebnisse der Experimente mit den Eingangssignalen  $f_C$  und  $f_{DC}$  tatsächlich kaum zu unterscheiden. Entsprechend möchten wir nun die Experimente mit Eingangssignal  $f_C$  untersuchen und gehen davon aus, dass die folgenden Beobachtungen ähnlich auch für Experimente mit Eingangssignal  $f_{DC}$  gültig sind. In Abbildung 7.27 ist zu sehen, dass bei diesem Experiment die AFEM-Routine für die Fehlergraphen und den Anteil  $\eta_{J,\mathcal{T}}$  des Verfeinerungsindikators mit adaptiver Netzverfeinerung bessere Konvergenzraten erzielt als mit uniformer Netzverfeinerung. Bei adaptiver Netzverfeinerung sind die entsprechenden Konvergenzraten etwa 1, wobei die Rate für die Fehlergraphen sogar etwas größer zu sein scheint, und bei uniformer Netzverfeinerung sind die Raten circa  $1/2$ . Dies wird bei diesem Beispiel daran liegen, dass eine starke Netzverfeinerung bei der Unstetigkeit hier sinnvoller ist als eine Verfeinerung dort, wo die Lösung konstant ist. Der Verfeinerungsindikator  $\eta_{\mathcal{T}}$  und sein Anteil  $\eta_{V,\mathcal{T}}$ , jeweils mit Rate 1, sowie die von  $EGLEB,\mathcal{T}$  abhängigen Graphen, beide mit Rate  $1/2$ , scheinen bei uniformer und adaptiver Netzverfeinerung zwar die gleichen Raten zu haben, aber bei adaptiver Netzverfeinerung deutlich geringere Werte zu erreichen. Ansonsten ähnelt das Konvergenzverhalten, insbesondere bei uniformer Netzverfeinerung, dem bereits beim Experiment mit Eingangssignal  $f_{10^4}$  in Abbildung 7.17 beobachteten. Dies scheint daran zu liegen, dass dort ebenfalls  $\alpha = 10^4$  gewählt wurde. Abschließend können wir betonen, dass Experimente existieren, bei denen adaptive Netzverfeinerung ein besseres Konvergenzverhalten bewirkt als uniforme. Allerdings möchten wir an dieser Stelle auch festhalten, dass der adaptive Algorithmus in den hier betrachteten Experimenten mehr Level durchläuft als der Algorithmus mit uniformer Netzverfeinerung und für Iterationen auf Triangulierungen mit vergleichbar vielen Freiheitsgraden in der Regel mehr Iterationsschritte ausführt. Insgesamt benötigt ein Experiment bei adaptiver Netzverfeinerung also mehr Zeit als bei uniformer Netzverfeinerung, wie in Abbildung 7.28 für drei Eingangssignale und insbesondere das Eingangssignal  $f_C$  zu sehen ist.

## 7.6 Fazit und Ausblick

Wir konnten in dieser Arbeit die Konvergenz der primalen-dualen Iteration nach [Bar15b] für eine nichtkonforme, mit Crouzeix-Raviart-Finite-Elemente-Funktionen diskretisierte Formulierung des ROF-Modellproblems beweisen, den AFEM-Algorithmus mit dieser Iteration im Solve-Schritt implementieren und in numerischen Experimenten untersuchen. Dabei konnten wir die garantierter Rate für das in [Bar15b] betrachte Problem übertreffen, auch wenn hierbei angemerkt werden muss, dass wir in den entsprechenden Beispielen regulärere Eingangssignale und exakte Lösungen betrachtet haben als lediglich Funktionen beschränkter Variation. Weiterhin konnten wir im Vergleich zu uniformer Netzverfeinerung die Verbesserung einiger Konvergenzraten bei adaptiver Netzverfeinerung für ein Problem feststellen, bei dem die Unstetigkeitsmenge des Eingangssignals entlang einer Teilmenge niedriger Dimension verlief, beziehungsweise ein lokal betragsgroßer schwacher Gradient in der stetigen Approximation des Eingangssignals auftrat, während dieses ansonsten nur

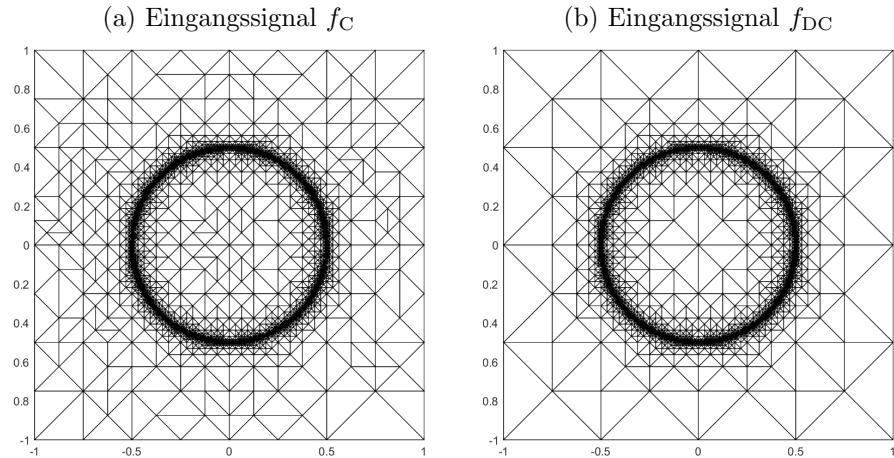


Abbildung 7.25: Triangulierungen nach den Rechnungen mit den Eingangssignalen  $f_C$  und  $f_{DC}$  für Level 17 des jeweiligen AFEM-Algorithmus nach etwa 15 000 Freiheitsgraden

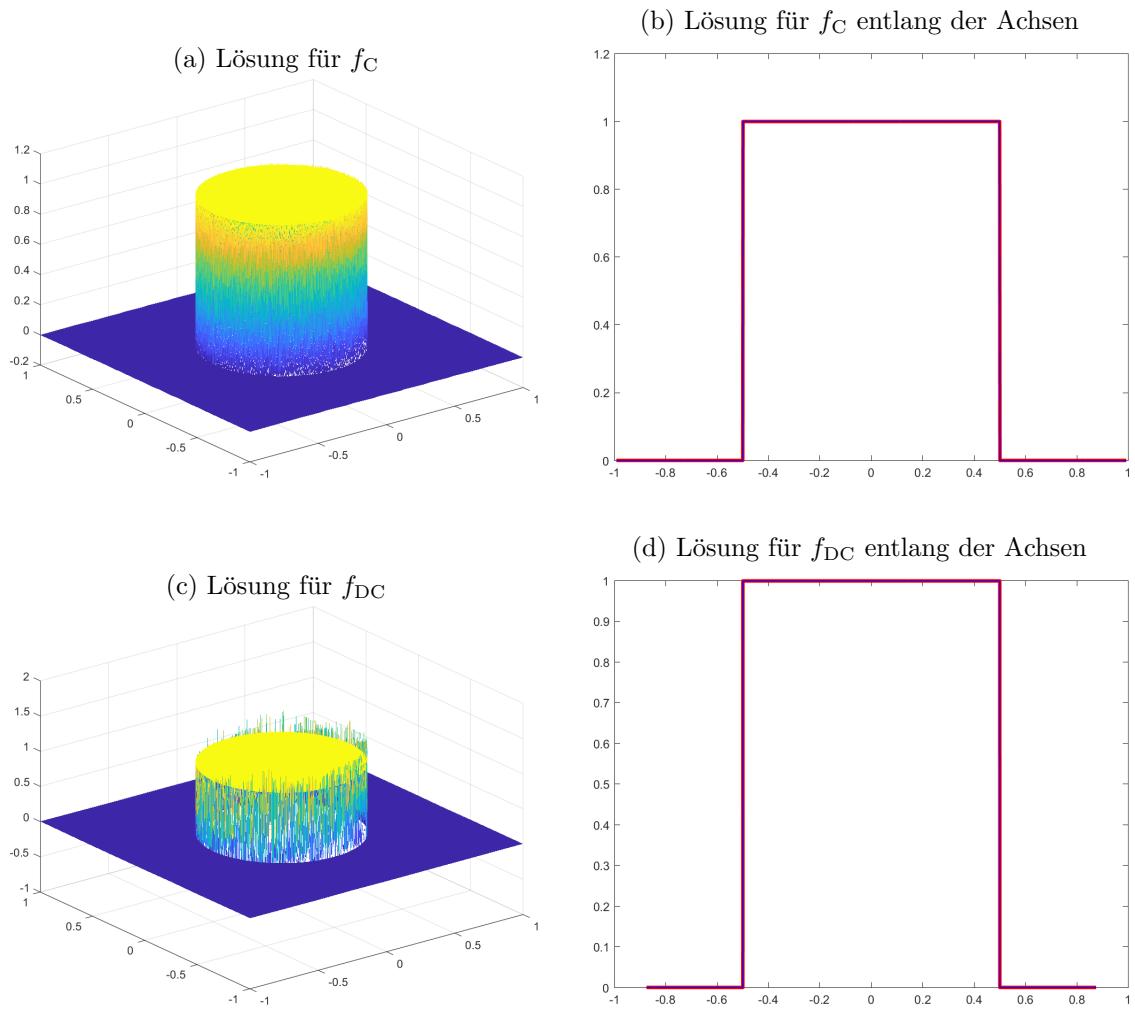


Abbildung 7.26: Diskrete Lösungen sowie deren Darstellungen entlang der x-Achse (blau) und der y-Achse (rot) nach den Rechnungen mit den Eingangssignalen  $f_C$  und  $f_{DC}$

## 7 Numerische Beispiele

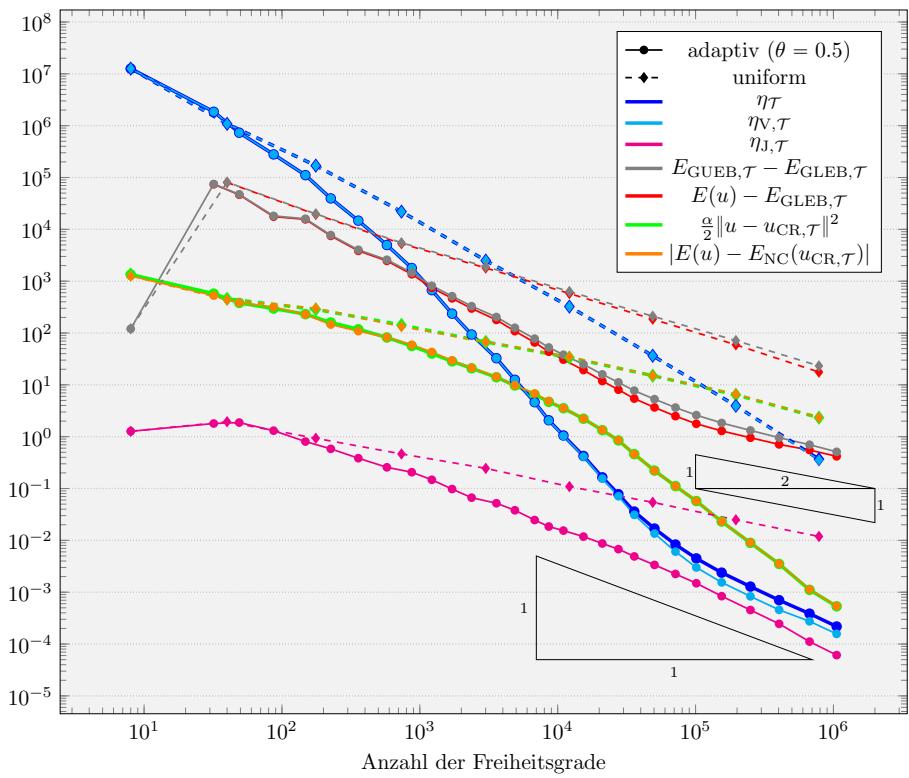


Abbildung 7.27: Ergebnisse der Rechnungen mit adaptiver und uniformer Netzverfeinerung für das Eingangssignal  $f_C$

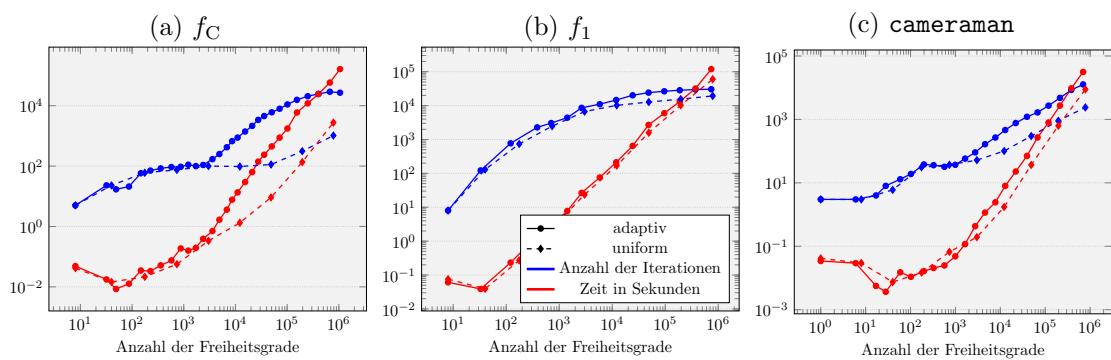


Abbildung 7.28: Anzahl der Iterationen und benötigte Zeit für die primalen-dualen Iterationen während des AFEM-Algorithmus bei den Rechnungen mit den Eingangssignalen  $f_C$ ,  $f_1$  und **cameraman**

konstante Werte annahm. Allerdings ging die Verbesserung der Ergebnisse bei adaptiver Netzverfeinerung mit einer erhöhten Rechenzeit einher. Ist Rechenzeit kein limitierender Faktor, dann ist dementsprechend davon auszugehen, dass bei solchen Experimenten adaptive Netzverfeinerung zu bevorzugen ist. Die durch den benutzten Verfeinerungsindikator entstandenen adaptiven Netze entsprachen unseren Erwartungen und für dessen Konvergenzverhalten konnten wir Erklärungen finden. Außerdem konnten wir die Gültigkeit der Interpretation des ROF-Modellproblems aus Kapitel 1 in mehreren Experimenten erkennen. Beim Test unserer Implementation entwickelten wir die Hypothese, dass die Größe der rechten Seite der Ungleichung (7.1) die benötigte Iterationzahl bis zum Erreichen des Abbruchkriteriums (5.3) für die primale-duale Iteration beeinflusst. Demnach führt eine größere Wahl von  $\tau$  und  $\alpha$ , wobei allerdings die Wahl von  $\alpha$  das betrachtete Experiment verändert, zu weniger Iterationen. Diese Hypothese wird von unseren Beobachtungen zur Größe der Parameter  $\tau$  und  $\alpha$  gestützt. Für die Wahl der Parameter  $\gamma$  und  $\varepsilon_{\text{stop}}$  konnten wir ebenfalls Empfehlungen formulieren. Weiterhin konnten theoretische Aussagen der Arbeit, insbesondere Ungleichung (7.3) zur Abschätzung des exakten Fehlers durch von den garantierten Energieschranken abhängigen Ausdrücken, experimentell bestätigt werden.

Zum Abschluss dieser Arbeit möchten wir noch mögliche nächste Schritte beschreiben. Zunächst kann das Programm erweitert werden, indem die Annahme, dass stets Nullranddaten vorliegen, aufgehoben wird. Dieses ist an einigen Stellen, zum Beispiel beim Erstellen der rechten Seite des Gleichungssystems während der primalen-dualen Iteration, noch unter ebendieser Annahme optimiert. Durch diese Erweiterung kann unter anderem bei Bildern als Eingangssignal auf das Hinzufügen eines Randes, um Nullranddaten zu garantieren, verzichtet werden. Weiterhin kann damit untersucht werden, ob für die Geometrie **Lshape** des AFEM-Softwarepaket ein Beispiel existiert, bei dem mit adaptiver Netzerfeinerung stark an der einspringenden Ecke verfeinert wird. In der aktuellen Implementation wird bei Eingangssignalen ohne Nullranddaten an allen Kanten gleichermaßen verfeinert. Als Nächstes kann die Implementierung der primalen-dualen Iteration für die konforme Formulierung des ROF-Modellproblems aus [Bar15b] angepasst werden, um deren Ergebnisse besser mit den Resultaten unserer Implementierung vergleichen zu können. Kann für die Implementierung in [Bar15b] außerdem ein Verfeinerungsindikator gefunden werden, so könnte diese ebenfalls im Rahmen des AFEM-Algorithmus realisiert werden, um das Konvergenzverhalten zu vergleichen. In der Theorie bleibt vor allem interessant, garantierte Konvergenzraten für unsere Implementierung zu beweisen. Zu diesem Zweck wäre die Konstruktion eines Beispiels mit bekannter exakter Lösung und einem Eingangssignal, welches von beschränkter Variation und nicht schwach differenzierbar ist, wichtig, um auch die Konvergenzraten für ein weniger reguläres Beispiel sehen zu können. Abschließend könnte auch die Wahl der Parameters  $\tau$  weiter untersucht werden, denn die Frage, ob es sinnvolle Möglichkeiten gibt,  $\tau$  nicht konstant zu wählen, bleibt offen. Damit könnte eventuell das Konvergenzverhalten der Iteration verbessert und somit die Anzahl der benötigten Iterationsschritte verringert werden. Außerdem könnte die Toleranz  $\varepsilon_{\text{stop}}$  für das Abbruchkriterium zum Beispiel abhängig von der Netzweite der Triangulierung gewählt werden, um sowohl die in einem Experiment dieser Arbeit beobachtete Stagnation des exakten Fehlers zu verhindern als auch die Iteration eines Levels nicht länger auszuführen als notwendig.

# Literatur

- [ABM14] Hedy Attouch, Giuseppe Buttazzo und Gérard Michaille. *Variational Analysis in Sobolev and BV Spaces. Applications to PDEs and Optimization*. Second Edition. Bd. 17. MOS-SIAM Series on Optimization. Philadelphia: Society for Industrial und Applied Mathematics, Mathematical Optimization Society, 2014. ISBN: 978-1-611973-47-1.
- [AK06] Gilles Aubert und Pierre Kornprobst. *Mathematical Problems in Image Processing. Partial Differential Equations and the Calculus of Variations*. Second Edition. Bd. 147. Applied Mathematical Sciences. New York: Springer, 2006. ISBN: 0-387-32200-0.
- [Bar12] Sören Bartels. „Total variation minimization with finite elements: convergence and iterative solution“. In: *SIAM Journal on Numerical Analysis* 50.3 (2012), S. 1162–1180. URL: <https://doi.org/10.1137/11083277X>.
- [Bar15a] Sören Bartels. „Error control and adaptivity for a variational model problem defined on functions of bounded variation“. In: *Mathematics of Computation* 84.293 (2015), S. 1217–1240. URL: <https://doi.org/10.1090/S0025-5718-2014-02893-7>.
- [Bar15b] Sören Bartels. *Numerical Methods for Nonlinear Partial Differential Equations*. Bd. 47. Springer Series in Computational Mathematics. Springer International Publishing, 2015. ISBN: 978-3-319-13796-4. DOI: [10.1007/978-3-319-13797-1](https://doi.org/10.1007/978-3-319-13797-1).
- [Bra98] Andrea Braides. *Approximation of free-discontinuity problems*. Bd. 1694. Lecture Notes in Mathematics. Berlin, Heidelberg, New York: Springer-Verlag, 1998. ISBN: 3-540-64771-6. URL: <https://doi.org/10.1007/BFb0097344>.
- [Car+10] Carsten Carstensen u. a. „AFEM-Dokumentation“. lokal verfügbar, 2010.
- [Car09a] Carsten Carstensen. *AFEM-Softwarepaket der Arbeitsgruppe Numerische Analysis*. Institut für Mathematik der Humboldt-Universität zu Berlin, 2009.
- [Car09b] Carsten Carstensen. „Yonsei Lectures at the WCU Department Computational Science and Engineering on Finite Element Method“. lokal verfügbar, 2009.
- [CGR12] Carsten Carstensen, Joscha Gedicke und Donsub Rim. „Explicit error estimates for Courant, Crouzeix-Raviart and Raviart-Thomas finite element methods“. In: *Journal of Computational Mathematics* 30.4 (2012), S. 337–353. ISSN: 0254-9409. DOI: [10.4208/jcm.1108-m3677](https://doi.org/10.4208/jcm.1108-m3677). URL: <https://doi.org/10.4208/jcm.1108-m3677>.
- [CH18] Carsten Carstensen und Friederike Hellwig. „Constants in Discrete Poincaré and Friedrichs Inequalities and Discrete Quasi-interpolation“. In: *Computational Methods in Applied Mathematics* 18.3 (2018), S. 433–450. ISSN: 1609-4840. DOI: [10.1515/cmam-2017-0044](https://doi.org/10.1515/cmam-2017-0044). URL: <https://doi.org/10.1515/cmam-2017-0044>.
- [CP10] Antonin Chambolle und Thomas Pock. „A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging“. In: *Journal of Mathematical Imaging and Vision* 40 (2010), S. 120–145. ISSN: 0924-9907. DOI: [10.1007/s10851-010-0251-1](https://doi.org/10.1007/s10851-010-0251-1). URL: <https://doi.org/10.1007/s10851-010-0251-1>.

- [Dac89] Bernard Dacorogna. *Direct Methods in the Calculus of Variations*. Bd. 78. Applied Mathematical Sciences. Berlin, Heidelberg: Springer-Verlag, 1989. ISBN: 978-3-642-51442-5.
- [EG92] Lawrence C. Evans und Ronald F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, 1992. ISBN: 0-8493-7157-0.
- [EM73] Martin A. Eisenberg und Lawrence E. Malvern. „On finite element integration in natural co-ordinates“. In: *International Journal for Numerical Methods in Engineering* 7 (4 1973), S. 574–575. URL: <https://doi.org/10.1002/nme.1620070421>.
- [Get12] Pascal Getreuer. „Rudin-Osher-Fatemi Total Variation Denoising using Split Bregman“. In: *Image Processing On Line* 2 (2012), S. 74–95. URL: <https://doi.org/10.5201/ipol.2012.g-tvd>.
- [Roc70] R. Tyrrell Rockafellar. *Convex Analysis*. New Jersey: Princeton University Press, 1970. ISBN: 0-691-08069-0.
- [ROF92] Leonid I. Rudin, Stanley Osher und Emad Fatemi. „Nonlinear total variation based noise removal algorithms“. In: *Physica D: Nonlinear Phenomena*. Bd. 60. 1-4. 1992, S. 259–268. DOI: 10.1016/0167-2789(92)90242-F. URL: [https://doi.org/10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F).
- [Zei85] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications. III: Variational Methods and Optimization*. New York: Springer Science+Business Media, LLC, 1985. ISBN: 978-1-4612-9529-7.
- [Zei86] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications. I: Fixed-Point Theorems*. New York, Berlin, Heidelberg, Tokyo: Springer-Verlag, 1986. ISBN: 0-387-90914-1.
- [Zei90a] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications. II/A: Linear Monotone Operators*. New York: Springer Science+Business Media, LLC, 1990. ISBN: 978-1-4612-6971-7.
- [Zei90b] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications. II/B: Nonlinear Monotone Operators*. New York: Springer Science+Business Media, LLC, 1990. ISBN: 978-1-4612-6969-4.

## Anhang

Datenträger mit dem Programm und einer digitalen Version dieser Arbeit

# **Selbstständigkeitserklärung**

Ich erkläre, dass ich die vorliegende Arbeit selbstständig verfasst und noch nicht für andere Prüfungen eingereicht habe. Sämtliche Quellen, einschließlich Internetquellen, die unverändert oder abgewandelt wiedergegeben werden, insbesondere Quellen für Texte, Grafiken, Tabellen und Bilder, sind als solche kenntlich gemacht. Mir ist bekannt, dass bei Verstößen gegen diese Grundsätze ein Verfahren wegen Täuschungsversuchs bzw. Täuschung eingeleitet wird.

Berlin, den 15. September 2021,