

# Adaptive Finite Elemente für Lineare Elastizität

Theo Koppenhöfer

Geboren am 9. November 2000 in Heidelberg

27. Juli 2022

Bachelorarbeit Mathematik

Betreuer: Prof. Dr. Joscha Gedicke

Zweitgutachter: Prof. Dr. Jochen Garcke

INSTITUT FÜR NUMERISCHE SIMULATION

MATHEMATISCH-NATURWISSENSCHAFTLICHE FAKULTÄT DER  
RHEINISCHEN FRIEDRICH-WILHELMS-UNIVERSITÄT BONN



# Inhaltsverzeichnis

<b>Zusammenfassung</b>	<b>1</b>
<b>1 Einleitung</b>	<b>1</b>
<b>2 Formulierung des kontinuierlichen Problems</b>	<b>4</b>
2.1 Verschiebung und Spannungstensor . . . . .	4
2.2 Randbedingungen . . . . .	5
2.3 Formulierung als Kräftegleichgewicht . . . . .	6
2.4 Differenzielle Formulierung . . . . .	6
2.5 Materialgesetze . . . . .	7
2.6 Variationelle Formulierung . . . . .	10
2.7 Energiebetrachtung . . . . .	11
<b>3 Existenz und Eindeutigkeit</b>	<b>12</b>
3.1 Lax-Milgram-Lemma . . . . .	12
3.2 Kornsche Ungleichung ohne Randbedingungen . . . . .	14
3.3 Kornsche Ungleichung mit Randbedingungen . . . . .	21
3.4 Elliptizität der Bilinearform $a$ . . . . .	23
<b>4 Diskretisierung mit P1 Finiten Elementen</b>	<b>24</b>
<b>5 A priori Fehlerabschätzung</b>	<b>26</b>
<b>6 A posteriori residualer Fehlerschätzer</b>	<b>27</b>
6.1 Konstruktion des residualen Fehlerschätzers . . . . .	27
6.2 Obere Abschätzung des Fehlers . . . . .	30
6.3 Untere Abschätzung des Fehlers . . . . .	31
<b>7 Implementierung der P1 Finiten Elemente</b>	<b>36</b>
7.1 Diskretisierung der Dirichletbedingungen . . . . .	37
7.2 Existenz und Eindeutigkeit des diskreten Problems . . . . .	38
7.3 Berechnung der Steifheitsmatrix . . . . .	39
<b>8 Numerische Experimente</b>	<b>41</b>
8.1 Phänomen des Lockings . . . . .	41
8.2 Benchmark: Quadratisches Gebiet . . . . .	42
8.3 Fehlerschätzer durch Mittelung . . . . .	44
8.4 Benchmark: L-förmiges Gebiet . . . . .	45
8.5 Fazit . . . . .	47
<b>Symbolverzeichnis, Literatur</b>	<b>49</b>

# Zusammenfassung

Viele Probleme aus der Statik werden mit Finite-Elemente-Methoden gelöst. Hierbei ist die Wahl des Gitters für den Fehler der Lösung oft entscheidend, weshalb a posteriori Fehlerschätzer eingesetzt werden, um das Gitter adaptiv zu verfeinern. In dieser Bachelorarbeit wird ein residualer a posteriori Fehlerschätzer für Probleme in der linearen Elastizität konstruiert und gezeigt, dass dieser zuverlässig und lokal effizient ist. Dazu werden grundlegende Begriffe aus der linearen Elastizität definiert, eine variationelle Formulierung hergeleitet und mithilfe der Kornischen Ungleichung und des Lax-Milgram-Lemmas Existenz und Eindeutigkeit gezeigt. Es wird ein P1 Finites Element besprochen und implementiert. Anschließend werden in numerischen Experimenten die Effizienz und Zuverlässigkeit des Schätzers bestätigt. Hierbei sieht man auch das Locking-Phänomen und die lineare Konvergenz für das P1 Finite Element.

## 1 Einleitung

Man stelle sich folgendes Problem vor: Ein quadratförmiges Werkstück wird an der linken Kante festgehalten und es wird an der rechten Kante gezogen, wie in Abbildung 1 dargestellt. Gesucht ist die Verformung des Werkstücks. Dies ist ein typisches Problem der Statik. Allerdings ist schon dieses simple Beispiel praktisch nur noch numerisch und nicht mehr analytisch lösbar. Für die numerische Lösung werden oft Finite-Elemente-Methoden verwendet. Hierbei

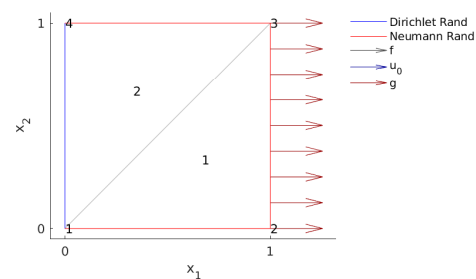


Abb. 1: Anfangsbedingungen

wird das Ausgangsobjekt mit einem Gitter in viele kleinere Teilgebiete, Elemente genannt, unterteilt. Anschließend wird für das Gitter eine Lösung des Problems berechnet. Eine solche Lösung ist für dieses Problem in Abbildung 2 dargestellt. Allerdings hängt der Fehler der berechneten Lösung von der Wahl des Gitters ab. Idealerweise wollen wir das Gitter an den Stellen stärker verfeinern, wo die Lösung besonders ungenau ist. Dafür benötigt man eine Methode, den Fehler der berechneten Lösung abzuschätzen ohne die exakte Lösung zu kennen. Dies führt auf den Begriff des a posteriori Fehlerschätzers. Man verwendet einen solchen Fehlerschätzer, um das Gitter adaptiv zu verfeinern, was für unser Problem in Abbildung 3 dargestellt ist.

Diese Arbeit soll einen kurzen Einblick in den Bereich der linearen Elastizitätstheorie liefern und einen residualen Fehlerschätzer für die adaptive Gitterverfeinerung konstruieren. Anschließend soll der Fehlerschätzer in numerischen Experimenten getestet werden.

Dazu werden im nachfolgenden Kapitel grundlegende Begriffe der linearen Elastizität eingeführt und mehrere Formulierungen des Problems motiviert. Das Kapitel beginnt mit der Definition der Begriffe „Verschiebung“ und „Spannungstensor“ und bespricht dann die Randbedingungen des Problems. Das Problem wird als Kräftegleichgewicht formuliert und daraus wird dann eine differenzielle Formulierung hergeleitet. Es folgt ein

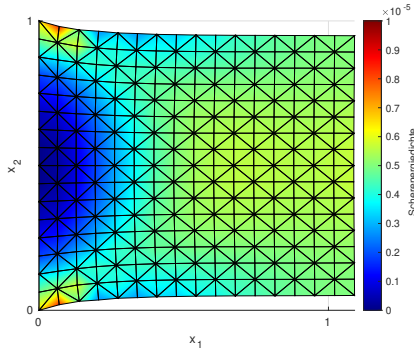


Abb. 2: uniforme Gitterverfeinerung

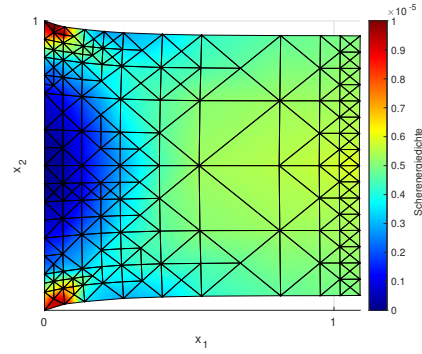


Abb. 3: adaptive Gitterverfeinerung

Intermezzo zu den Materialgesetzen, wobei wir uns im Wesentlichen auf lineare isotrope Materialien beschränken. Danach wird aus der differenziellen Formulierung eine variationelle Formulierung mit dem Prinzip der virtuellen Arbeit hergeleitet. Anschließend wird ein Ausdruck für die Energie hergeleitet, was eine Formulierung als Optimierungsproblem erlaubt, die äquivalent zur variationellen Formulierung ist.

Es folgt ein Kapitel über Existenz und Eindeutigkeit der schwachen Formulierungen. Hierfür nutzen wir das Lax-Milgram Lemma, was unter Anderem die Elliptizität der Bilinearform  $a$  voraussetzt. Die Elliptizität von  $a$  folgt im Wesentlichen aus den Kornschen Ungleichungen. Dazu wird zunächst die Kornsche Ungleichung ohne Randbedingungen auf dem Ganzraum gezeigt. Es wird dann ein Erweiterungsoperator von Epigraphen stetig differenzierbarer Funktionen auf den Ganzraum konstruiert. Dieser liefert zusammen mit einer Zerlegung der Eins die Kornsche Ungleichung ohne Randbedingungen auf  $C^1$ -Mengen. Es folgt ein Resultat für Starrkörperbewegungen, aus dem, zusammen mit der Kornschen Ungleichung ohne Randbedingungen, dann die Kornsche Ungleichung mit Randbedingungen gefolgert wird. In einem letzten Schritt wird hieraus die Elliptizität von  $a$  gefolgert. Dieses Vorgehen ist in Abbildung 4 dargestellt.

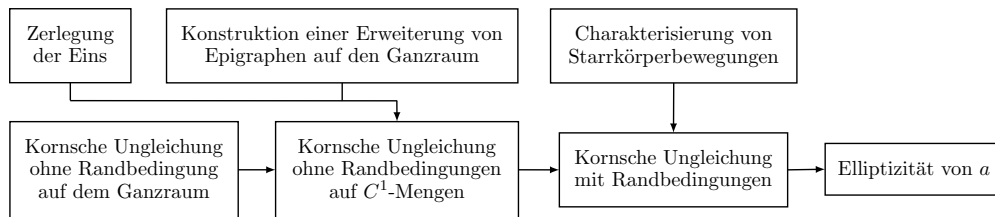


Abb. 4: Vorgehen beim Nachweis der Elliptizität

Dann wenden wir uns in Kapitel 4 dem diskreten Problem zu. Hier werden zunächst Triangulierungen besprochen und verschiedene Teilmengen der Kanten und Knoten definiert. Es wird anschließend eine diskrete Formulierung des Problems mit Finiten Elementen motiviert. Es folgt ein ähnlich kurzes Kapitel zum a priori Fehler dieser P1 Finiten Elemente. Hier wird der Fehler der Diskretisierung mit Céas Lemma und einem

geeignetem Interpolationsoperator nach oben abgeschätzt. Es wird gezeigt, dass unter gewissen Voraussetzungen der Fehler linear in  $h$  ist, wobei  $h$  die Gitterweite bezeichnet.

Im Kapitel über a posteriori Fehlerschätzer wird zunächst ein globaler und ein lokaler residualer Fehlerschätzer konstruiert. Danach wird mit Cléments Interpolationsoperator der Fehler nach oben beschränkt, also gezeigt, dass der Fehlerschätzer zuverlässig ist. Anschließend wird mit einem Hilfsresultat der lokale Fehler nach unten durch den lokalen Fehlerschätzer abgeschätzt.

Es folgt ein Kapitel zur Implementierung der P1 Finiten Elemente in der linearen Elastizität. Hier wird die Diskretisierung der rechten Seite und der Dirichletrandbedingungen besprochen und begründet, dass das resultierende lineare Gleichungssystem eine eindeutige Lösung besitzt. Es wird abschließend noch auf die Konstruktion der Steifheitsmatrix eingegangen.

Im letzten Kapitel wird die Implementierung an zwei numerischen Experimenten getestet. Da die a priori Abschätzung für die P1 Finiten Elemente von den Materialparametern  $\lambda$  und  $\mu$  kritisch abhängt, wird diese Abschätzung bei geeigneter Wahl der Materialparameter beliebig schlecht. Es stellt sich heraus, dass dies gerade für inkompressible Materialien der Fall ist. Dieses Phänomen wird als Locking bezeichnet und wird in einer ersten Versuchsreihe genauer untersucht. Hierbei wird die uniforme Triangulierung und ein Benchmark auf einem quadratischem Gebiet verwendet. Bei dieser Versuchsreihe sieht man das Locking und die vorausgesagte lineare Konvergenz in  $h$ . Man sieht in der Versuchsreihe auch die Effizienz und Zuverlässigkeit des residualen Fehlerschätzers. Im zweiten numerischen Experiment wird der Mehrgewinn durch die adaptive Gitterverfeinerung zu der uniformen Verfeinerung untersucht. Hierfür wird noch kurz ein zweiter a posteriori Fehlerschätzer eingeführt, der auf Mittelung basiert. Der verwendete Benchmark ist auf einem L-förmigen Gebiet definiert mit einer Singularität des Spannungstensors in der Null. Aufgrund dieser Singularität erhalten wir bei der uniformen Verfeinerung nicht mehr lineare Konvergenz in  $h$ . Die lineare Konvergenz wird allerdings mit der adaptiven Gitterverfeinerung mit dem residualen Fehlerschätzer erreicht. Auch hier ist der residuale Fehlerschätzer zuverlässig und effizient. Im Gegensatz dazu ist der Fehlerschätzer, der auf Mittelung basiert, nicht sonderlich zuverlässig oder effizient. Dieser verfeinert auch fast ausschließlich an der Singularität, was dazu führt, dass nach einer gewissen Anzahl an Verfeinerungen die Implementierung versagt.

Diese Bachelorarbeit orientiert sich maßgeblich an den Kapiteln III.§8 (A posteriori Abschätzungen), VI.§1 (Einführung in die Elastizitätstheorie) und VI.§3 (Lineare Elastizitätstheorie) in [3]. Die Implementierung orientiert sich an [1]. Dies sind auch die Hauptquellen der Arbeit. Im Kapitel zur Formulierung des kontinuierlichen Problems, wurden viele physikalische Argumente [14] und [8] entnommen. Aus [16] stammt im wesentlichen der Beweis für die Kornsche Ungleichung ohne Randbedingungen. Und aus [5] und [13] stammt die Folgerung der Kornschen Ungleichung mit Nebenbedingungen daraus. Im Abschnitt über a posteriori Fehlerschätzer wurden die Kapitel 1.1-1.6 und Kapitel 3.6 in [17] verwendet. Die numerischen Experimente stammen aus [4]. Die Implementierungen zur Bachelorarbeit wurden in MATLAB<sup>©</sup> verfasst, wofür mir Prof. Dr. Gedicke diverse Vorlagen gegeben hat.

## 2 Formulierung des kontinuierlichen Problems

Im folgenden Kapitel werden grundlegende Begriffe und Gleichungen aus der linearen Elastizität motiviert und definiert. Dazu werden die Verschiebung, der Spannungstensor und die Randbedingungen besprochen, woraus dann eine erste Formulierung des Problems als Kräftegleichgewicht motiviert wird. Hieraus wird eine differenzielle Formulierung gefolgert. Dann gibt es ein kleines Intermezzo zu den Materialgesetzen. Anschließend wird noch eine variationelle Formulierung und eine Formulierung als Optimierungsproblem hergeleitet. Die Definitionen und Herleitungen in diesem Kapitel entstammen im Wesentlichen [3, Kapitel VI.§1], [8, Kapitel III], [5] und einige physikalischere Argumentationen stammen aus [14].

### 2.1 Verschiebung und Spannungstensor

Das Objekt nehme in Referenzkonfiguration das Gebiet  $\bar{\Omega} \subseteq \mathbb{R}^d$  ein. Eine Deformation ist nach [3, S.276] eine Abbildung  $\chi: \bar{\Omega} \rightarrow \mathbb{R}^d$ , so dass  $\det \nabla \chi > 0$ . Eine Deformation ist also orientierungserhaltend, was gewährleistet, dass endliche Volumen nicht auf Singularitäten deformiert werden. Weiter definieren wir die Verschiebung  $u$  über  $\chi = \text{Id} + u$ . Die Menge der zulässigen Verschiebungen wird mit  $V$  bezeichnet. Wir bezeichnen mit  $v$  eine generische Verschiebung und mit  $u$  meistens die Verschiebung, welches unser Problem löst. Das deformierte Objekt nimmt die Menge  $\bar{\Omega}' = \chi(\bar{\Omega})$  ein. Man sieht in Abbildung 5 ein Beispiel einer Verschiebung. Im Folgenden gehen wir von kleinen Verschiebungen aus und setzen stillschweigend  $\Omega' = \Omega$ .

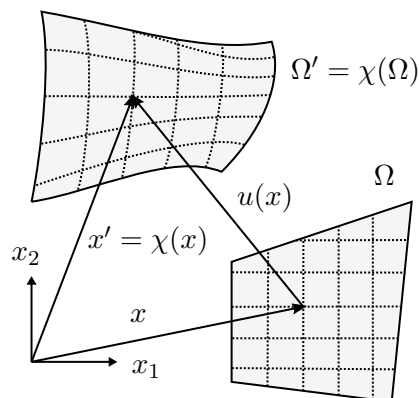


Abb. 5: Beispiel einer Deformation in 2D.

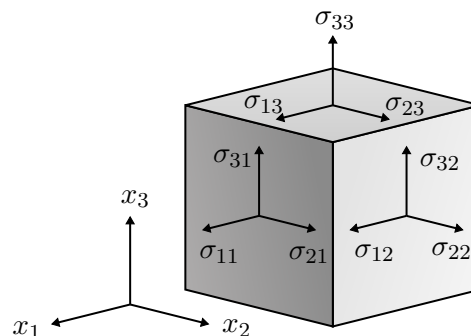


Abb. 6: Darstellung des dreidimensionalen Spannungstensors.

Weiter nehmen wir an, dass sich der deformierte Körper im Kräftegleichgewicht befindet. Wir betrachten also den statischen Fall. Auf den Körper wirken Volumenkräfte, repräsentiert durch eine Abbildung  $f: \Omega \rightarrow \mathbb{R}^d$ . Volumenkräfte können zum Beispiel Gravitationskräfte, Zentripetalkräfte oder elektromagnetische Kräfte sein. Außerdem wirken auf den Körper Oberflächenkräfte, repräsentiert durch eine Abbildung  $\sigma: \Omega \rightarrow \mathbb{R}^{d \times d}$ . Dabei bezeichnet  $\sigma_{ij}$  die Oberflächenkraft in Richtung  $i$ , die auf eine infinitesimale Oberfläche in Richtung  $j$  wirkt. Die Kraft, die auf eine in Richtung  $n$  liegenden infinitesimale

Oberfläche wirkt, lautet also

$$\sigma n = \sum_j \sigma_{ij} n_j e_i.$$

Wir fordern im Folgenden ein wenig Regularität für  $f$  und  $\sigma$ . Konkret nehmen wir  $f \in L^2(\Omega; \mathbb{R}^d)$  und  $\sigma \in L^2(\Omega; \mathbb{R}^{d \times d})$  an.

In dem Beispiel von der Einleitung wurde  $f = 0$  gesetzt und  $\sigma$  ist noch unbekannt. Allerdings wurden Randbedingungen gesetzt, die im Folgenden besprochen werden.

## 2.2 Randbedingungen

Wir bezeichnen  $\Gamma := \partial\Omega$  als den Rand des Gebietes,  $\Gamma_D \subseteq \Gamma$  als Dirichlet- und  $\Gamma_N \subseteq \Gamma$  als Neumann-Rand. Falls der Dirichlet-Rand  $\Gamma_D$  positives Flächenmaß hat,  $\Gamma_N = \Gamma \setminus \Gamma_D$  und die Randbedingungen

$$\begin{aligned} u &= w && \text{auf } \Gamma_D, \\ \sigma n &= g && \text{auf } \Gamma_N \end{aligned}$$

gelten, haben wir klassische Dirichlet-Neumann-Randbedingungen. Hierbei repräsentiert  $g: \Gamma_N \rightarrow \mathbb{R}^d$  Kräfte, die auf dem Neumann-Rand ausgeübt werden und  $w: \Gamma_D \rightarrow \mathbb{R}^d$  entspricht einer Anfangsverschiebung auf dem Dirichlet-Rand. In dem Beispiel von der Einleitung wurde als Dirichlet-Rand die linke Kante des Quadrats gewählt und der Rest als Neumann-Rand deklariert. Außerdem wurden  $w = 0$  und  $g = e_1$  auf  $\{1\} \times [0, 1]$  und  $g = 0$  sonst gewählt, wie in der Abbildung 1 dargestellt. Diese Randbedingungen können wir verallgemeinern zu gleitenden Randbedingungen

$$\begin{aligned} Mu &= w && \text{auf } \Gamma, \\ \sigma n &= g && \text{auf } \Gamma_N \end{aligned}$$

mit einer Matrix  $M: \Gamma \rightarrow \mathbb{R}^{d \times d}$ . Hierbei sind  $\Gamma_N$  und  $\Gamma_D$  sind nicht notwendigerweise disjunkt. Wir werden gemäß [1] die gleitenden Randbedingungen implementieren.

Im Folgenden fordern wir  $g \in L^2(\Gamma_N; \mathbb{R}^d)$ ,  $w \in H^1(\Omega; \mathbb{R}^d)$  sowie  $M \in L^\infty(\Gamma; \mathbb{R}^{d \times d})$ . Man beachte, dass  $w$  auf ganz  $\Omega$  definiert ist. Außerdem fordern wir, dass  $\Omega$  ein Lipschitzgebiet ist. Hierbei bezeichnen wir eine beschränkte offene Menge  $\Omega \subseteq \mathbb{R}^d$  als Lipschitz, falls es für jeden Randpunkt  $x \in \partial\Omega$  einen Radius  $r > 0$ , eine Lipschitz-stetige Funktion  $\psi: \mathbb{R}^{d-1} \rightarrow \mathbb{R}$  und eine affine Isometrie  $A: \mathbb{R}^d \rightarrow \mathbb{R}^d$  gibt, so dass

$$B_r(x) \cap \Omega = B_r(x) \cap A \operatorname{epi}(\psi). \quad (1)$$

Dabei ist für eine Funktion  $\psi: \mathbb{R}^{d-1} \rightarrow \mathbb{R}$  der Epigraph durch

$$\operatorname{epi}(\psi) := \{(x', x_d): x_d > \psi(x')\}$$

gegeben. In Abbildung 7 ist diese Definition bildlich dargestellt. Insbesondere sind polytope Gebiete, wie sie in der Implementierung auftauchen, Lipschitz.



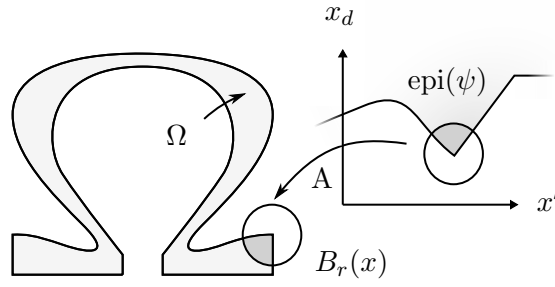


Abb. 7: Visualisierung eines Lipschitz-Gebiets.

### 2.3 Formulierung als Kräftegleichgewicht

Da sich der Körper im Kräftegleichgewicht befindet, muss sich für ein Gebiet  $\omega \subseteq \Omega$ , das regulär genug ist, die Gesamtkraft aufheben. Wir erhalten so die Gleichgewichtsbeziehung

$$\begin{array}{ll} \text{Volumenkräfte,} & \text{Oberflächenkräfte, die über die} \\ \text{die auf } \omega \text{ wirken} & \text{Oberfläche } \partial\omega \text{ auf } \omega \text{ wirken} \\ \downarrow & \downarrow \\ \int_{\omega} f \, dx + \int_{\partial\omega} \sigma n \, ds = 0 & \text{für alle } \omega \subseteq \Omega \text{ regulär genug} \end{array}$$

mit  $n: \Omega \rightarrow \mathbb{R}^d$  die äußere Normale. Die Gleichgewichtsbeziehung liefert eine erste Formulierung des Problems: Finde eine Deformation  $u \in V$ , so dass

$$\begin{array}{ll} - \int_{\partial\omega} \sigma n \, ds = \int_{\omega} f \, dx & \text{für alle } \omega \subseteq \Omega \text{ regulär genug,} \\ Mu = w & \text{auf } \Gamma, \\ \sigma n = g & \text{auf } \Gamma_N, \end{array}$$

wobei der Spannungstensor  $\sigma$  von der Verschiebung  $u$  abhängt und  $f, g$  als von  $u$  unabhängig angenommen werden.  $f$  und  $g$  werden dann auch als tote Lasten bezeichnet.

### 2.4 Differenzielle Formulierung

Wenn wir annehmen, dass  $\sigma$  differenzierbar ist, erhalten wir mit dem Satz von Gauß nach [14, S.5] die Beziehung

$$\begin{aligned} \int_{\omega} f \, dx &= - \int_{\partial\omega} \sigma n \, ds \\ &= - \int_{\partial\omega} \sum_{i,j} \sigma_{ij} n_j e_i \, ds \\ &= - \int_{\omega} \sum_{i,j} \partial_j \sigma_{ij} e_i \, dx. \end{aligned}$$

Da das für alle geeignete Gebiete  $\omega \subseteq \Omega$  gilt, folgt die Gleichgewichtsbeziehung in differenzieller Form

$$-\operatorname{Div} \sigma = - \sum_j \partial_j \sigma_{ij} e_i = f.$$

Wir erhalten also eine differenzielle Formulierung des Problems: Finde  $u \in V$ , so dass

$$\begin{aligned} -\operatorname{Div} \sigma &= f && \text{auf } \Omega, \\ Mu &= w && \text{auf } \Gamma, \\ \sigma n &= g && \text{auf } \Gamma_N. \end{aligned} \tag{2}$$

Da der Körper sich im Kräftegleichgewicht befindet, verschwindet auch der gesamte Drehmoment in beliebigen geeigneten Gebieten  $\omega \subseteq \Omega$ . In [14, S.5f.] wird hieraus mit dem Satz von Gauß die Symmetrie  $\sigma_{ij} = \sigma_{ji}$  herleiten.

## 2.5 Materialgesetze

Wir möchten nun einen Ausdruck für  $\sigma$  in Abhängigkeit der Verschiebung  $u$  motivieren und unterscheiden dazu im Folgenden zwischen deformiertem Zustand und Referenzkonfiguration. Sei  $z \in \mathbb{R}^d$  ein kleines Wegstück im Punkt  $x \in \mathbb{R}^d$ . Die Deformation transformiert den Punkt auf  $x' = \chi(x)$  und dieses Wegstück auf ein Wegstück

$$z' := \chi(x+z) - \chi(x) = z + u(x+z) - u(x).$$

Wir bezeichnen mit  $r = |z|$  und  $r' = |z'|$  die Weglängen. Wenn wir nun annehmen, dass  $\nabla u$  klein ist, erhalten wir nach [3, (1.3) in VI.§1] mit Taylor

$$\begin{aligned} r'^2 - r^2 &= |z'|^2 - r^2 \\ &= |z + \nabla u(x)z + o(r)|^2 - r^2 \\ &= |z|^2 + z^\top (\nabla u(x) + \nabla u(x)^\top) z + z^\top \nabla u(x)^\top \nabla u(x) z + o(r^2) - r^2 \\ &\approx 2z^\top \varepsilon(x) z + o(r^2), \end{aligned} \tag{3}$$

wobei wir den linearisierten Verzerrungstensor  $\varepsilon$  durch

$$2\varepsilon := \nabla u + \nabla u^\top$$

definiert haben. Der linearisierte Verzerrungstensor ist nach Gleichung (3) also ein Maß für die Weglängenänderung durch die Deformation.

Wir nennen ein Material linear-elastisch, falls der Spannungstensor linear vom linearisierten Verzerrungstensor abhängt, es also einen von  $v$  unabhängigen Hooke-Tensor  $C: \Omega \rightarrow \bigotimes_{i=1}^4 \mathbb{R}^d$  gibt, so dass das Hookesche Gesetz

$$\sigma_{ij} = \sum_{k,l} C_{ijkl} \varepsilon_{kl} \tag{4}$$

gilt. Dies ist eine Verallgemeinerung des eindimensionalen Hookeschen Gesetzes, welches besagt, dass die Kraft, die auf eine Feder wirkt, proportional zur Auslenkung aus der Ruhelage ist. Mit der Symmetrie von  $\sigma$  erhalten wir

$$\sum_{k,l} C_{ijkl} \varepsilon_{kl} = \sigma_{ij} = \sigma_{ji} = \sum_{k,l} C_{jikl} \varepsilon_{kl}.$$

Also gilt

$$C_{ijkl} = C_{jikl}. \quad (5)$$

Außerdem fordern wir noch die Symmetriebeziehung

$$C_{ijkl} = C_{klij}. \quad (6)$$

Wir möchten nun analog zu [15, S.120] einen Ausdruck für  $C$  in einem homogenen isotropen Material finden. Ein solches Material heißt auch St. Venant-Kirchhoff Material. Hodge hat in [12] gezeigt, dass die allgemeinste Form eines isotropen Tensors vierter Ordnung durch

$$C_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu \delta_{ik} \delta_{jl} + \nu \delta_{il} \delta_{jk}$$

gegeben ist. Dieses soll nun die Symmetrie (5) erfüllen. Es soll also gelten

$$\begin{aligned} 0 &= C_{ijkl} - C_{jikl} \\ &= \lambda \delta_{ij} \delta_{kl} + \mu \delta_{ik} \delta_{jl} + \nu \delta_{il} \delta_{jk} - (\lambda \delta_{ji} \delta_{kl} + \mu \delta_{jk} \delta_{il} + \nu \delta_{jl} \delta_{ik}) \\ &= (\mu - \nu) \delta_{ik} \delta_{jl} + (\nu - \mu) \delta_{il} \delta_{jk}, \end{aligned}$$

womit  $\mu = \nu$  folgt. Es gilt also für ein isotropes linear elastisches Material

$$C_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}), \quad (7)$$

wobei  $\lambda, \mu \in \mathbb{R}$  Lamé-Koeffizienten heißen. Der Parameter  $\mu$  wird auch Schubmodul genannt. Nach [14, S.11, S.14] zeigen experimentelle Befunde und theoretische Überlegungen, dass  $\lambda > 0$  und  $\mu > 0$  gilt. In Abbildung 8 sind einige durchschnittliche Materialparameter für gebräuchliche Materialien angegeben.

	$\lambda$ ( $10^7$ Pa)	$\mu$ ( $10^7$ Pa)	$\nu$
Stahl	10	8.2	0.28
Glas	2.2	2.2	0.25
Blei	4.6	0.63	0.44
Gummi	0.4	0.012	0.485

Abb. 8: Einige durchschnittliche Parameter nach [5, S.129]

Mit Gleichung (7) erhält man einen konkreteren Ausdruck für den Spannungstensor von St. Venant-Kirchhoff-Materialien

$$\begin{aligned}
\sigma_{ij} &= \sum_{k,l} C_{ijkl} \varepsilon_{kl} \\
&= \sum_k \lambda \delta_{ij} \varepsilon_{kk} + \mu (\varepsilon_{ij} + \varepsilon_{ji}) \\
&= \lambda \operatorname{Tr}(\varepsilon) \delta_{ij} + 2\mu \varepsilon_{ij}.
\end{aligned} \tag{8}$$

Man kann Relation (8) in  $d = 2$  Dimensionen nach [1, S.244] auch schreiben als

$$\tau(v) := \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{12} \end{bmatrix} = \underbrace{\begin{bmatrix} \lambda + 2\mu & \lambda \\ \lambda & \lambda + 2\mu \\ & \mu \end{bmatrix}}_{=:\hat{C}} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ 2\varepsilon_{12} \end{bmatrix} =: \hat{C}\gamma(v) \tag{9}$$

und nach [1, S.246] in  $d = 3$  Dimensionen als

$$\tau(v) := \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{13} \\ \sigma_{23} \end{bmatrix} = \underbrace{\begin{bmatrix} \lambda + 2\mu & \lambda & \lambda & & & \\ \lambda & \lambda + 2\mu & \lambda & & & \\ \lambda & \lambda & \lambda + 2\mu & & & \\ & & & \mu & & \\ & & & & \mu & \\ & & & & & \mu \end{bmatrix}}_{=:\hat{C}} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{12} \\ 2\varepsilon_{13} \\ 2\varepsilon_{23} \end{bmatrix} =: \hat{C}\gamma(v), \tag{10}$$

wobei wir die Voigt-Notation  $\gamma(v) \in \mathbb{R}^{d(d+1)/2}$  für den linearisierten Verzerrungstensor  $\varepsilon(v)$  und  $\tau(v) \in \mathbb{R}^{d(d+1)/2}$  für den Spannungstensor  $\sigma(v)$  eingeführt haben.

Im Folgenden wird angenommen, dass  $C \in L^\infty(\Omega; \bigotimes_{i=1}^4 \mathbb{R}^d)$  beschränkt und, dass  $C$  gemäß [8, S.103] elliptisch ist, also Gleichung (11) erfüllt.

**Proposition 1** (Elliptizität von  $C$ ). *Für ein St. Venant-Kirchhoff-Material ist der Hooke-Tensor  $C \in L^\infty(\Omega; \bigotimes_{i=1}^4 \mathbb{R}^d)$  elliptisch. Es gibt also  $c_c > 0$ , so dass*

$$\sum_{i,j,k,l} C_{ijkl} \varepsilon_{ij} \varepsilon_{kl} \geq c_c \sum_{i,j} \varepsilon_{ij}^2. \tag{11}$$

*Beweis.* Es gilt

$$\begin{aligned}
\sum_{i,j,k,l} C_{ijkl} \varepsilon_{ij} \varepsilon_{kl} &= \sum_{i,j} \sigma_{ij} \varepsilon_{ij} \\
&= \sum_{i,j} (\lambda \operatorname{Tr}(\varepsilon) \varepsilon_{ii} + 2\mu \varepsilon_{ij} \varepsilon_{ij}) \\
&= d\lambda \operatorname{Tr}(\varepsilon)^2 + 2\mu \sum_{i,j} \varepsilon_{ij}^2 \\
&\geq c_c \sum_{i,j} \varepsilon_{ij}^2
\end{aligned}$$

mit einer Konstante  $c_c := 2\mu$ . □

## 2.6 Variationelle Formulierung

Wir setzen  $V := H^1(\Omega; \mathbb{R}^d)$  als Menge der möglichen Verschiebungen. Außerdem definieren wir die Menge der kinematisch zulässigen Verschiebungen des homogenen Problems durch

$$V^0 := \{v \in V : Mv = 0 \text{ auf } \Gamma\}.$$

Unser Ziel ist es, einen Ausdruck für die Energie eines Zustandes zu bestimmen. Wir betrachten die (virtuelle) Arbeit, die verrichtet wird, wenn wir das Objekt im Endzustand  $u \in V$  um ein kleines  $v \in V^0$  verschieben. Wir erhalten nach [14, S.8] dafür den Ausdruck

$$\begin{aligned}
 \int_{\Gamma_N} g \cdot v \, ds + \int_{\Omega} f \cdot v \, dx &= \int_{\partial\Omega} (\sigma(u) n) \cdot v \, ds + \int_{\Omega} \sum_i f_i v_i \, dx \\
 &\quad \begin{array}{l} \nearrow \text{von } g \text{ auf } v \text{ verrichtete Arbeit} \\ \nearrow \text{von } f \text{ auf } v \text{ verrichtete Arbeit} \end{array} \\
 &\quad \text{Differenzielle Formulierung (2)} \\
 &= \int_{\partial\Omega} \sum_{i,j} \sigma_{ij} v_i n_j \, ds - \int_{\Omega} \sum_{i,j} (\partial_j \sigma_{ij}) v_i \, dx \\
 &\quad \text{Satz von Gauß} \\
 &= \int_{\Omega} \sum_{i,j} \partial_j (\sigma_{ij} v_i) \, dx - \int_{\Omega} \sum_{i,j} (\partial_j \sigma_{ij}) v_i \, dx \\
 &= \int_{\Omega} \sum_{i,j} \sigma_{ij} (\partial_j v_i) \, dx \\
 &\quad \text{Symmetrie von } \sigma \\
 &= \int_{\Omega} \sum_{i,j} \sigma_{ij} \frac{1}{2} (\partial_j v_i + \partial_i v_j) \, dx \\
 &= \int_{\Omega} \sum_{i,j} \sigma_{ij}(u) \varepsilon_{ij}(v) \, dx.
 \end{aligned}$$

Nun setzt man

$$a(u, v) := \int_{\Omega} \sigma(u) : \varepsilon(v) \, dx := \int_{\Omega} \sum_{i,j} \sigma_{ij}(u) \varepsilon_{ij}(v) \, dx, \quad (12)$$

sowie

$$\ell(v) := \int_{\Omega} f \cdot v \, dx + \int_{\Gamma_N} g \cdot v \, ds = \langle f, v \rangle_{0,\Omega} + \langle g, v \rangle_{0,\Gamma_N}. \quad (13)$$

Wir erhalten also eine variationelle Formulierung des Problems: Finde  $u \in V$ , so dass

$$\begin{aligned}
 a(u, v) &= \ell(v) \quad \text{für alle } v \in V^0, \\
 Mu &= w \quad \text{auf } \Gamma.
 \end{aligned} \quad (14)$$

## 2.7 Energiebetrachtung

Nun möchten wir einen Ausdruck für die potenzielle Energie  $W$  eines Zustandes  $u$  finden. Ändern wir den Zustand  $u$  um eine kleine Verrückung  $v$ , so soll sich die Energie gerade um die von  $v$  ausgeübte Arbeit auf das Material  $a(u, v)$  abzüglich der von der Umgebung auf  $v$  verrichtete Arbeit  $\ell(v)$  ändern. Also muss für die Gâteaux-Ableitung gelten

$$\partial_v W(u) = a(u, v) - \ell(v) = 0.$$

$u$  muss also ein kritischer Punkt von  $W$  sein. Proposition 2 zeigt, dass diese Eigenschaft unter gewissen Voraussetzungen an  $a$  und  $\ell$  von der Funktion

$$W(v) := \frac{1}{2}a(v, v) - \ell(v) \tag{15}$$

erfüllt wird.

**Proposition 2** (Zusammenhang variationelle Formulierung und Sattelpunktproblem). *Seien  $V$  ein linearer Raum,  $\ell: V \rightarrow \mathbb{R}$  linear,  $a: V \times V \rightarrow \mathbb{R}$  eine symmetrische elliptische bilineare Form und  $u \in V$ . Dann gilt*

$$a(u, v) = \ell(v)$$

*für alle  $v \in V$  genau dann, wenn  $u \in V$  eindeutiger Minimierer von*

$$W(v) = \frac{1}{2}a(v, v) - \ell(v)$$

*ist.*

*Beweis.* Siehe Satz 2.2 in [3, S.34]. □

Sofern die Voraussetzungen von Proposition 2 erfüllt sind, erhalten wir eine zur variationellen Formulierung (14) äquivalente Formulierung als Optimierungsproblem: Finde  $u \in V$ , so dass

$$\begin{aligned} u \text{ minimiert} \quad & W = \frac{1}{2}a(\cdot, \cdot) - \ell \\ \text{unter der Nebenbedingung} \quad & Mu|_{\Gamma} = w|_{\Gamma}. \end{aligned} \tag{16}$$

Zur Visualisierung der Lösungen im Kapitel zu den numerischen Experimenten plotten wir die elastische Scherenergiedichte, die sich nach [1, S.252] ergibt als

$$U = \frac{1}{4\mu} \left\| \sigma - \frac{1}{d} \text{Tr}(\sigma) \text{Id} \right\|_F^2,$$

wobei  $\|\cdot\|_F$  die Frobenius-Norm bezeichnet.

### 3 Existenz und Eindeutigkeit

Im folgenden Abschnitt nutzen wir das Lax-Milgram-Lemma, um Existenz und Eindeutigkeit unseres Problems (14), beziehungsweise (16) zu erhalten. Dafür zeigen wir zunächst hinreichende Kriterien dafür, dass  $a$  eine stetige symmetrische bilineare Form und  $\ell$  eine stetige lineare Abbildung ist. Das Lemma setzt allerdings auch Elliptizität von  $a$  voraus, deren Nachweis sich als recht involviert herausstellt. In Abbildung 4 aus der Einleitung wurde bereits das Vorgehen umrissen. Wir zeigen zunächst die Kornsche Ungleichung ohne Randbedingungen auf dem Ganzraum. Dann konstruieren wir einen Erweiterungsoperator, der  $v$  von Epigraphen stetig differenzierbarer Funktionen auf den Ganzraum erweitert. Mit einer Zerlegung der Eins ergibt sich hieraus die Kornsche Ungleichung ohne Randbedingungen für  $C^1$ -Mengen. Dann nutzen wir ein Resultat über Starrkörperbewegungen, um die Kornsche Ungleichung mit Randbedingungen zu folgern und nutzen dieses Resultat schließlich, um die Elliptizität von  $a$  zu zeigen. Der Beweis von der Kornschen Ungleichung ohne Randbedingungen folgt im wesentlichen der Struktur von [16]. Die Folgerung der Kornschen Ungleichung mit Randbedingungen und der Elliptizität folgt dagegen der Struktur der Beweise in [5], [13] und [8].

#### 3.1 Lax-Milgram-Lemma

Wir legen nun die Grundsteine zur Anwendung des Lax-Milgram-Lemmas.

**Proposition 3** (Bilinearität und Symmetrie). *Die Abbildung  $a: V \times V \rightarrow \mathbb{R}$  definiert in Gleichung (12) ist symmetrisch und bilinear.*

*Beweis.* Da  $\varepsilon(v)$  linear in  $v$  und folglich auch  $\sigma(v) = C\varepsilon(v)$  linear in  $v$  ist, folgt, dass

$$a(v_1, v_2) = \langle \sigma(v_1), \varepsilon(v_2) \rangle_{0,\Omega}$$

eine Bilinearform ist. Ferner folgt mit Gleichung (6)

$$\begin{aligned} a(v_1, v_2) &= \int_{\Omega} \sigma(v_1) : \varepsilon(v_2) \, dx \\ &= \int_{\Omega} \sum_{i,j,k,l} C_{ijkl} \varepsilon_{kl}(v_1) \varepsilon_{ij}(v_2) \, dx \\ &= \int_{\Omega} \sum_{i,j,k,l} C_{klij} \varepsilon_{ij}(v_2) \varepsilon_{kl}(v_1) \, dx \\ &= a(v_2, v_1) \end{aligned}$$

die Symmetrie. □

**Proposition 4** (Stetigkeit von  $a$ ). *Ist  $C \in L^\infty(\Omega; \bigotimes_{i=1}^4 \mathbb{R}^d)$ , dann ist  $a: V \times V \rightarrow \mathbb{R}$  stetig, d. h. es gibt ein  $c_A > 0$ , so dass für alle  $v_1, v_2 \in V$  gilt*

$$a(v_1, v_2) \leq c_A |v_1|_{1,\Omega} |v_2|_{1,\Omega} \leq c_A \|v_1\|_{1,\Omega} \|v_2\|_{1,\Omega}.$$

*Beweis.* Man sieht, dass für alle  $v \in H^1(\Omega; \mathbb{R}^d)$  gilt

$$\|\varepsilon(v)\|_{0,\Omega} = \frac{1}{2} \|\nabla v + \nabla v^\top\|_{0,\Omega} \leq \frac{1}{2} \left( \|\nabla v\|_{0,\Omega} + \|\nabla v^\top\|_{0,\Omega} \right) = |v|_{1,\Omega}. \quad (17)$$

Weiter sieht man mit Cauchy-Schwarz und der Dreiecksungleichung

$$\begin{aligned} \|\sigma(v)\|_{0,\Omega}^2 &= \sum_{i,j} \left\| \sum_{k,l} C_{ijkl} \varepsilon_{kl}(v) \right\|_{0,\Omega}^2 \\ &\leq \sum_{i,j} \left\| \left( \sum_{k,l} C_{ijkl}^2 \right)^{1/2} \left( \sum_{k,l} \varepsilon_{kl}^2 \right)^{1/2} \right\|_{0,\Omega}^2 \\ &\leq d^2 \|C\|_\infty^2 \sum_{i,j,k,l} \|\varepsilon_{kl}\|_{0,\Omega}^2 \\ &= d^4 \|C\|_\infty^2 \sum_{k,l} \|\varepsilon_{kl}\|_{0,\Omega}^2 \\ &= c_A^2 \|\varepsilon(v)\|_{0,\Omega}^2 \end{aligned} \quad (18)$$

mit einer Konstante  $c_A := d^2 \|C\|_\infty$ . Damit folgt erneut mit Cauchy-Schwarz

$$\begin{aligned} a(v_1, v_2) &= \langle \sigma(v_1), \varepsilon(v_2) \rangle_{0,\Omega} \\ &\leq \|\sigma(v_1)\|_{0,\Omega} \|\varepsilon(v_2)\|_{0,\Omega} \\ &\leq c_A \|\varepsilon(v_1)\|_{0,\Omega} \|\varepsilon(v_2)\|_{0,\Omega} \\ &\leq c_A |v_1|_{1,\Omega} |v_2|_{1,\Omega} \end{aligned}$$

wie behauptet.  $\square$

**Proposition 5.** *Die Abbildung  $\ell$  definiert in Gleichung (13) ist linear und beschränkt, also stetig.*

*Beweis.* Die Linearität ist klar. Die Beschränktheit folgt aus Cauchy-Schwarz und der Beschränktheit des Trace-Operators

$$\begin{aligned} \ell(v) &= \langle f, v \rangle_{0,\Omega} + \langle g, v \rangle_{0,\Gamma_N} \\ &\leq \|f\|_{0,\Omega} \|v\|_{0,\Omega} + \|g\|_{0,\Gamma_N} \|v\|_{0,\Gamma_N} \\ &\leq \|f\|_{0,\Omega} \|v\|_{1,\Omega} + c_F \|g\|_{0,\Gamma_N} \|v\|_{1,\Omega}. \end{aligned}$$

$\square$

Wir erinnern an das Lemma von Lax-Milgram.

**Satz 6** (Lax-Milgram Lemma). *Seien  $V^0$  ein Banach-Raum,  $\ell: V^0 \rightarrow \mathbb{R}$  eine stetige lineare Form und  $a: V^0 \times V^0 \rightarrow \mathbb{R}$  eine stetige symmetrische elliptische bilineare Form. Dann hat das Problem  $u \in V^0$  zu finden, so dass*

$$a(u, v) = \ell(v)$$



für alle  $v \in V^0$  eine eindeutige Lösung. Dieses ist dann auch eindeutige Lösung des Problems  $u \in V^0$  zu finden, so dass  $u$  das Funktional

$$W(v) = \frac{1}{2}a(v, v) - \ell(v)$$

minimiert.

*Beweis.* Siehe Theorem 6.3-2. in [5, S.288].  $\square$

Um Existenz und Eindeutigkeit zu zeigen, nutzen wir die Annahme, dass die Randbedingungen erfüllbar sind, also dass es ein  $u_\Gamma \in V$  gibt, so dass  $Mu_\Gamma = w$  auf  $\Gamma$ .

**Korollar 7** (Existenz und Eindeutigkeit des Problems). *Existiert  $u_\Gamma \in V$ , so dass  $Mu_\Gamma = w$  und erfüllen  $a$  und  $\ell$  auf dem Raum  $V^0$  die Voraussetzungen des Lax-Milgram-Lemmas, dann besitzt unser Problem (14), beziehungsweise (16) eine eindeutige Lösung.*

*Beweis.* Es ist  $u \in V$  genau dann eine Lösung von

$$\begin{aligned} a(u, v) &= \ell(v) & \text{für alle } v \in V^0, \\ Mu &= w & \text{auf } \Gamma, \end{aligned}$$

wenn  $u - u_\Gamma \in V^0$  eine Lösung ist von

$$a(u - u_\Gamma, v) = \ell(v) - a(u_\Gamma, v) \quad \text{für alle } v \in V^0.$$

Dieses Problem hat aber nach dem Lax-Milgram Lemma eine eindeutige Lösung.  $\square$

Die Proposition 5 zeigt, dass  $\ell$  die Voraussetzungen erfüllt und die Propositionen 3 und 4 zeigen, dass unter angemessenen Voraussetzungen  $a$  eine stetige symmetrische bilineare Form ist. Allerdings ist, wie schon angedeutet, der Nachweis der Elliptizität von  $a$  etwas subtiler und erfolgt über die Kornschen Ungleichungen.

### 3.2 Kornsche Ungleichung ohne Randbedingungen

Wir definieren nun für  $v \in H^1(\Omega; \mathbb{R}^d)$  die Norm

$$\|v\|_{K,\Omega}^2 := \|v\|_{0,\Omega}^2 + \|\varepsilon(v)\|_{0,\Omega}^2. \quad (19)$$

Da  $\varepsilon: H^1(\Omega; \mathbb{R}^d) \rightarrow L^2(\Omega; \mathbb{R}^d)$  linear ist, folgt die 1-Homogenität und die Dreiecksungleichung für  $\|\cdot\|_{K,\Omega}$ . Es ist nun unser Ziel die Positivdefinitheit von  $\|\cdot\|_{K,\Omega}$  nachzuweisen. Auf dem Ganzraum folgt sie aus folgendem Resultat.

**Lemma 8** (Kornsche Ungleichung ohne Randbedingungen auf dem Ganzraum). *Sei  $\Omega = \mathbb{R}^d$  und  $v \in H^1(\Omega; \mathbb{R}^d)$ . Dann gilt*

$$|v|_{1,\Omega} \leq \sqrt{2} \|\varepsilon(v)\|_{0,\Omega}.$$

*Beweis.* Wir folgen im wesentlichen dem Beweis aus [3, S.292]. Es gilt für  $v \in H^2(\Omega; \mathbb{R}^d)$

$$\begin{aligned}
& 2\varepsilon(v) : \varepsilon(v) - \nabla v : \nabla v \\
&= \frac{1}{2} \sum_{i,j} (\partial_i v_j + \partial_j v_i)^2 - \sum_{i,j} (\partial_i v_j)^2 \\
&= \frac{1}{2} \sum_{i,j} \cancel{(\partial_i v_j)^2} + \frac{1}{2} \sum_{i,j} \cancel{(\partial_j v_i)^2} + \sum_{i,j} (\partial_i v_j)(\partial_j v_i) - \sum_{i,j} \cancel{(\partial_i v_j)^2} \\
&= \sum_{i,j} (\partial_i v_j)(\partial_j v_i) + \sum_{i,j} v_j (\partial_i \partial_j v_i) - \sum_{i,j} v_i (\partial_i \partial_j v_j) \\
&\quad - \sum_{i,j} (\partial_j v_j)(\partial_i v_i) + \sum_{i,j} (\partial_j v_j)(\partial_i v_i) \\
&= \sum_{i,j} \partial_i (v_j (\partial_j v_i) - (\partial_j v_j) v_i) + \left( \sum_i (\partial_i v_i) \right)^2 \\
&\geq \sum_{i,j} \partial_i (v_j (\partial_j v_i) - (\partial_j v_j) v_i) .
\end{aligned}$$

Die Aussage folgt auf  $H^2(\Omega; \mathbb{R}^d)$  dann mit Hilfe von partieller Integration aus

$$\begin{aligned}
2\|\varepsilon(v)\|_{0,\Omega}^2 - |v|_{1,\Omega}^2 &= \int_{\Omega} 2\varepsilon(v) : \varepsilon(v) - \nabla v : \nabla v \, dx \\
&\geq \int_{\Omega} \sum_{i,j} \partial_i (v_j (\partial_j v_i) - (\partial_j v_j) v_i) \, dx \\
&= \lim_{R \rightarrow \infty} \int_{B_R(0)} \sum_{i,j} \partial_i (v_j (\partial_j v_i) - (\partial_j v_j) v_i) \, dx \\
&= \lim_{R \rightarrow \infty} \int_{\partial B_R(0)} \sum_{i,j} (v_j (\partial_j v_i) - (\partial_j v_j) v_i) n_i \, dx \\
&= 0 .
\end{aligned}$$

Da  $H^2(\Omega; \mathbb{R}^d)$  dicht in  $H^1(\Omega; \mathbb{R}^d)$  ist, folgt die Aussage auch auf  $H^1(\Omega; \mathbb{R}^d)$ .  $\square$

Unser Ziel ist es, diese Aussage auf mehr Gebiete  $\Omega$  zu erweitern. Dazu bezeichnen wir mit  $H_S^k(\Omega; \mathbb{R}^d)$  die Menge aller  $v \in H^k(\Omega; \mathbb{R}^d)$ , so dass  $v = 0$  auf einer Randmenge  $S \subseteq \Gamma$ . Mit dieser Notation erhalten wir das folgende Lemma.

**Lemma 9** (Erweiterung von Epigraphen). *Gegeben sei ein Ball  $B = B_r(0) \subseteq \mathbb{R}^d$ , eine stetig differenzierbare Funktion  $\psi: \mathbb{R}^{d-1} \rightarrow \mathbb{R}$  mit beschränkter Ableitung  $|\nabla \psi| \leq M$  und  $\psi(0) = 0$ . Diese definieren ein Gebiet  $\Omega := B \cap \text{epi}(\psi)$ . Wir definieren einen Radius  $\tilde{r}^2 := r^2 + (2M + 1)^2 r^2$  und einen Ball  $\tilde{B} := B_{\tilde{r}}(0)$ . Dann existieren eine vom Gebiet unabhängige Konstante  $c_E > 0$  und eine lineare Erweiterung  $\mathfrak{E}: H_{\partial B}^1(\Omega; \mathbb{R}^d) \rightarrow H_{\partial \tilde{B}}^1(\tilde{B}; \mathbb{R}^d)$  mit  $\mathfrak{E}v|_{\Omega} = v$  für  $v \in H_{\partial B}^1(\Omega; \mathbb{R}^d)$  und*

$$\|\varepsilon(\mathfrak{E}v)\|_{0,\tilde{B}}^2 \leq c_E (\|\varepsilon(v)\|_{0,\Omega}^2 + M^2 |v|_{1,\Omega}^2) . \quad (20)$$

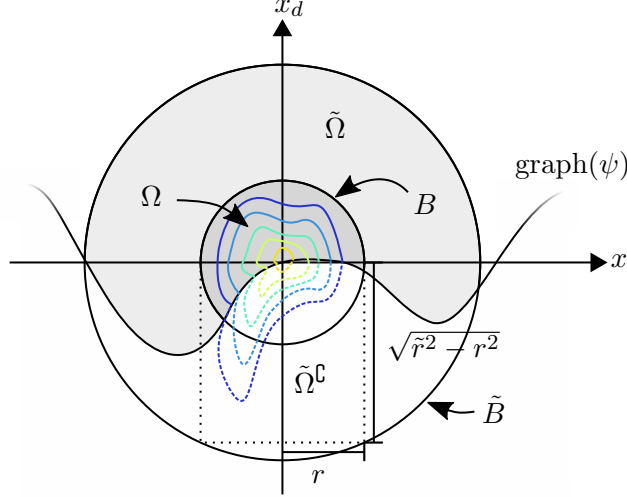


Abb. 9: Visualisierung der Erweiterung.

*Beweis.* Der Beweis folgt im Wesentlichen [16]. Man definiert das Komplement  $\tilde{\Omega}^c := \tilde{B} \setminus \tilde{\Omega}$  mit  $\tilde{\Omega} := \text{epi}(\psi) \cap \tilde{B}$ . Wir definieren eine Spiegelung  $\Phi: \tilde{\Omega}^c \rightarrow \text{epi}(\psi)$  durch

$$\Phi_t(x) := \begin{bmatrix} x' \\ \psi(x') + t(\psi(x') - x_d) \end{bmatrix},$$

wobei  $x' = [x_1 \dots x_{d-1}]^\top$ . Für  $t > 0$  ist  $t(\psi(x') - x_d) > 0$ , und es folgt  $\Phi_t(x) \in \text{epi}(\psi)$ . Wir definieren nun eine Erweiterung für  $v \in H_{\partial B}^1(\Omega; \mathbb{R}^d)$  durch

$$(\mathfrak{E}v(x))_j := \begin{cases} v_j(x) & \text{falls } x \in \tilde{\Omega} \\ a_j v_j(\Phi_1(x)) + b_j v_j(\Phi_2(x)) & \text{falls } x \in \tilde{\Omega}^c \end{cases},$$

wobei wir stillschweigend  $v_j = 0$  auf  $\text{epi}(\psi) \setminus B$  erweitert haben. Hierbei setzen wir  $a_d = -3$  und  $b_d = 4$  sowie  $a_i = 3$  und  $b_i = -2$  für  $i < d$ .

Wir zeigen zunächst, dass  $\mathfrak{E}v \in C^0(\tilde{B}; \mathbb{R}^d)$  für  $v \in C^0(\Omega; \mathbb{R}^d)$ . Es geht aus der Konstruktion hervor, dass  $\mathfrak{E}v|_{\tilde{\Omega}} \in C^0(\tilde{\Omega}; \mathbb{R}^d)$  und  $\mathfrak{E}v|_{\tilde{\Omega}^c} \in C^0(\tilde{\Omega}^c; \mathbb{R}^d)$ . Es verbleibt also die Stetigkeit an der Naht  $\text{graph}(\psi) = \{(x', \psi(x')) : x' \in \mathbb{R}^d\}$  zu zeigen. Sei dazu  $x^k = (x', x_d^k) \in \tilde{\Omega}^c$  eine Folge, die gegen  $x^* = (x', \psi(x')) \in \text{graph}(\psi)$  konvergiert. Dann gilt  $\psi(x') - x_d^k \rightarrow 0$  für  $k \rightarrow \infty$  und es folgt  $\Phi_t(x^k) \rightarrow x^*$ . Damit ergibt sich aufgrund der Wahl der  $a_i$  und  $b_i$

$$\begin{aligned} \lim_{k \rightarrow \infty} (\mathfrak{E}v(x^k))_j &= \lim_{k \rightarrow \infty} (a_j v_j(\Phi_1(x^k)) + b_j v_j(\Phi_2(x^k))) \\ &= a_j v_j(x^*) + b_j v_j(x^*) = v_j(x^*), \end{aligned}$$

wie erwünscht. Da  $C^0(\Omega; \mathbb{R}^d)$  dicht in  $H^1(\Omega; \mathbb{R}^d)$  und  $\mathfrak{E}$  stetig ist, folgt, dass  $\mathfrak{E}v \in H^1(\tilde{B}; \mathbb{R}^d)$  für alle  $v \in H_{\partial B}^1(\Omega; \mathbb{R}^d)$ .

Wir zeigen nun  $\mathfrak{E}v \in H_{\partial \tilde{B}}^1(\tilde{B}; \mathbb{R}^d)$ , also dass  $\mathfrak{E}v(x) = 0$  für alle Randpunkte  $x \in \partial \tilde{B}$ . Für den Fall, dass  $|x'| > r$ ,  $x_d > 0$  oder  $x \in \partial \tilde{\Omega}$ , ist  $\mathfrak{E}v(x) = 0$  nach Konstruktion. Falls

nun  $|x'| \leq r$ ,  $x_d \leq 0$  und  $x_d \in \partial\tilde{\Omega}^{\mathbb{C}}$ , so folgt

$$\tilde{r}^2 = |x|^2 = |x'|^2 + x_d^2 \leq r^2 + x_d^2$$

und wegen  $x_d \leq 0$  ist

$$x_d \leq -(\tilde{r}^2 - r^2)^{1/2} = -2Mr - r.$$

Es folgt aus dem Hauptsatz der Integralrechnung angewandt auf Linienintegrale

$$\psi(x') = \int_0^1 \nabla \psi(tx') \cdot x' dt \geq - \int_0^1 Mr dt = -Mr,$$

wobei wir  $\psi(0) = 0$  verwendet haben. Damit erhält man

$$x_d \leq -2Mr - r \leq 2\psi(x') - r$$

und folglich wegen  $x_d \in \partial\tilde{\Omega}^{\mathbb{C}}$

$$r \leq 2\psi(x') - x_d \leq \psi(x') + t(\psi(x') - x_d).$$

Es ist also  $\Phi_t(x) \in \text{epi}(\psi) \setminus B$ , also  $v_j(\Phi_t(x)) = 0$  und somit auch  $\mathfrak{E}v(x) = 0$ . Dies war auch die Motivation für die Wahl von  $\tilde{r}$ .

Die Linearität des Operators folgt aus der Konstruktion. Es verbleibt also die Eigenschaft (20) zu zeigen. Es gilt mit der Kettenregel für Sobolev-Funktionen (siehe beispielsweise [2, S.125]) für  $i < d$

$$\begin{aligned} \partial_i(v_j \circ \Phi_t) &= \nabla v_j|_{\Phi_t} \partial_i \Phi_t \\ &= \partial_i v_j|_{\Phi_t} + \partial_d v_j|_{\Phi_t} \partial_i(\psi(x') + t(\psi(x') - x_d)) \\ &= \partial_i v_j|_{\Phi_t} + (1+t) \partial_i \psi \partial_d v_j|_{\Phi_t} \end{aligned}$$

und für  $i = d$

$$\begin{aligned} \partial_d(v_j \circ \Phi_t) &= \nabla v_j|_{\Phi_t} \partial_d \Phi_t \\ &= \partial_d v_j|_{\Phi_t} \partial_d(\psi(x') + t(\psi(x') - x_d)) \\ &= -t \partial_d v_j|_{\Phi_t}. \end{aligned}$$

Wir erhalten auf  $\tilde{\Omega}^{\mathbb{C}}$  für  $i, j < d$

$$\begin{aligned} |2\varepsilon_{ij}(\mathfrak{E}v)| &= |\partial_i(a_j v_j(\Phi_1(x)) + b_j v_j(\Phi_2(x))) + \partial_j(a_i v_i(\Phi_1(x)) + b_i v_i(\Phi_2(x)))| \\ &= |a_j(\partial_i v_j|_{\Phi_1} + 2\partial_i \psi \partial_d v_j|_{\Phi_1}) + b_j(\partial_i v_j|_{\Phi_2} + 3\partial_i \psi \partial_d v_j|_{\Phi_2}) \\ &\quad + a_i(\partial_j v_i|_{\Phi_1} + 2\partial_j \psi \partial_d v_i|_{\Phi_1}) + b_i(\partial_j v_i|_{\Phi_2} + 3\partial_j \psi \partial_d v_i|_{\Phi_2})| \\ &\leq |a_j \partial_i v_j|_{\Phi_1} + a_i \partial_j v_i|_{\Phi_1}| + |a_j 2\partial_i \psi \partial_d v_j|_{\Phi_1}| + |b_j \partial_i v_j|_{\Phi_2} + b_i \partial_j v_i|_{\Phi_2}| \\ &\quad + |b_j 3\partial_i \psi \partial_d v_j|_{\Phi_2}| + |a_i 2\partial_j \psi \partial_d v_i|_{\Phi_1}| + |b_i 3\partial_j \psi \partial_d v_i|_{\Phi_2}| \\ &\leq 8|\partial_i v_j|_{\Phi_1} + \partial_j v_i|_{\Phi_1}| + 8M|\partial_d v_j|_{\Phi_1}| + 8|\partial_i v_j|_{\Phi_2} + \partial_j v_i|_{\Phi_2}| \\ &\quad + 8M|\partial_d v_j|_{\Phi_2}| + 8M|\partial_d v_i|_{\Phi_1}| + 8M|\partial_d v_i|_{\Phi_2}| \\ &\leq 2^4 \left( |\varepsilon_{ij}(v)|_{\Phi_1}| + M|\nabla v|_{\Phi_1}| + |\varepsilon_{ij}(v)|_{\Phi_2}| + M|\nabla v|_{\Phi_2}| \right), \end{aligned}$$

da wir  $a_j = a_i = 3$  und  $b_j = b_i = -2$  gewählt haben. Es gilt für  $j = d$  und  $i < d$

$$\begin{aligned}
|2\varepsilon_{id}(\mathfrak{E}v)| &= |\partial_i(a_d v_d(\Phi_1(x)) + b_d v_d(\Phi_2(x))) + \partial_d(a_i v_i(\Phi_1(x)) + b_i v_i(\Phi_2(x)))| \\
&= |a_d(\partial_i v_d|_{\Phi_1} + 2\partial_i \psi \partial_d v_d|_{\Phi_1}) + b_d(\partial_i v_d|_{\Phi_2} + 3\partial_i \psi \partial_d v_d|_{\Phi_2}) \\
&\quad - a_i \partial_d v_i|_{\Phi_1} - b_i 2\partial_d v_i|_{\Phi_2}| \\
&\leq |a_d \partial_i v_d|_{\Phi_1} - a_i \partial_d v_i|_{\Phi_1}| + |a_d 2\partial_i \psi \partial_d v_d|_{\Phi_1}| \\
&\quad + |b_d \partial_i v_d|_{\Phi_2} - b_i 2\partial_d v_i|_{\Phi_2}| + |b_d 3\partial_i \psi \partial_d v_d|_{\Phi_2}| \\
&\leq 8|\partial_i v_d|_{\Phi_1} + \partial_d v_i|_{\Phi_1}| + 2^4 M |\partial_d v_d|_{\Phi_1}| \\
&\quad + 8|\partial_i v_d|_{\Phi_2} + \partial_d v_i|_{\Phi_2}| + 2^4 M |\partial_d v_d|_{\Phi_2}| \\
&\leq 2^4 (|\varepsilon_{id}(v)|_{\Phi_1}| + M |\nabla v|_{\Phi_1}| + |\varepsilon_{id}(v)|_{\Phi_2}| + M |\nabla v|_{\Phi_2}|),
\end{aligned}$$

da wir  $a_d = -a_i = -3$  und  $b_d = -2b_i = 4$  gewählt haben. Außerdem gilt

$$\begin{aligned}
|2\varepsilon_{dd}(\mathfrak{E}v)| &= 2|\partial_d(a_d v_d(\Phi_1(x)) + b_d v_d(\Phi_2(x)))| \\
&= 2| -a_d \partial_d v_d|_{\Phi_1} - 2b_d \partial_d v_d|_{\Phi_2}| \\
&\leq 2(8|\partial_d v_d|_{\Phi_1}| + 8|\partial_d v_d|_{\Phi_2}|) \\
&= 2^4 (|\varepsilon_{dd}(v)|_{\Phi_1}| + |\varepsilon_{dd}(v)|_{\Phi_2}|).
\end{aligned}$$

Es folgt also in jedem Fall

$$|\varepsilon_{ij}(\mathfrak{E}v)| \leq 2^3 (F \circ \Phi_1 + F \circ \Phi_2)$$

für

$$F := \max_{i,j} |\varepsilon_{ij}(v)| + M |\nabla v|.$$

Damit folgt

$$\begin{aligned}
\|\varepsilon_{ij}(\mathfrak{E}v)\|_{0,\tilde{\Omega}\mathfrak{E}}^2 &\leq 2^6 \|F \circ \Phi_1 + F \circ \Phi_2\|_{0,\tilde{\Omega}\mathfrak{E}}^2 \\
&= 2^6 (\|F \circ \Phi_1\|_{0,\tilde{\Omega}\mathfrak{E}}^2 + \|F \circ \Phi_2\|_{0,\tilde{\Omega}\mathfrak{E}}^2 + 2\langle F \circ \Phi_1, F \circ \Phi_2 \rangle_{0,\tilde{\Omega}\mathfrak{E}}) \\
&\leq 2^6 (\|F \circ \Phi_1\|_{0,\tilde{\Omega}\mathfrak{E}}^2 + \|F \circ \Phi_2\|_{0,\tilde{\Omega}\mathfrak{E}}^2 + 2\|F \circ \Phi_1\|_{0,\tilde{\Omega}\mathfrak{E}} \|F \circ \Phi_2\|_{0,\tilde{\Omega}\mathfrak{E}}) \\
&= 2^6 (\|F \circ \Phi_1\|_{0,\tilde{\Omega}\mathfrak{E}} + \|F \circ \Phi_2\|_{0,\tilde{\Omega}\mathfrak{E}})^2.
\end{aligned}$$

Es folgt mit dem Transformationssatz und  $\text{supp } F \subseteq \overline{\Omega}$

$$\begin{aligned}
\|F \circ \Phi_t\|_{0,\tilde{\Omega}\mathfrak{E}}^2 &= \int_{\tilde{\Omega}\mathfrak{E}} |F \circ \Phi_t|^2 dx \\
&= \int_{\Phi_t^{-1}(\tilde{\Omega}\mathfrak{E})} |F|^2 |\det \nabla \Phi_t| dx \\
&\leq t \int_{\Omega} |F|^2 dx \\
&\leq 2\|F\|_{0,\Omega}^2.
\end{aligned}$$

Also erhalten wir

$$\begin{aligned}
\|\varepsilon_{ij}(\mathfrak{E}v)\|_{0,\tilde{\Omega}\mathfrak{C}}^2 &\leq 2^6 (\|F \circ \Phi_1\|_{0,\tilde{\Omega}\mathfrak{C}} + \|F \circ \Phi_2\|_{0,\tilde{\Omega}\mathfrak{C}})^2 \\
&\leq 2^7 (\|F\|_{0,\Omega} + \|F\|_{0,\Omega})^2 \\
&= 2^9 \|F\|_{0,\Omega}^2 \\
&= 2^9 \left\| \max_{i,j} |\varepsilon_{ij}(v)| + M|\nabla v| \right\|_{0,\Omega}^2 \\
&= 2^9 \left( \left\| \max_{i,j} |\varepsilon_{ij}(v)| \right\|_{0,\Omega}^2 + \|M|\nabla v|\|_{0,\Omega}^2 + 2 \left\langle \max_{i,j} |\varepsilon_{ij}(v)|, M|\nabla v| \right\rangle_{0,\Omega} \right) \\
&\leq 2^9 \left( \left\| \max_{i,j} |\varepsilon_{ij}(v)| \right\|_{0,\Omega}^2 + \|M|\nabla v|\|_{0,\Omega}^2 + 2 \left\| \max_{i,j} |\varepsilon_{ij}(v)| \right\|_{0,\Omega} \|M|\nabla v|\|_{0,\Omega} \right) \\
&= 2^9 \left( \left\| \max_{i,j} |\varepsilon_{ij}(v)| \right\|_{0,\Omega} + M\|v\|_{1,\Omega} \right)^2 \\
&\leq c_1 \left( \sum_{i,j} \|\varepsilon_{ij}(v)\|_{0,\Omega}^2 + M^2 \|v\|_{1,\Omega}^2 \right) \\
&= c_1 (\|\varepsilon(v)\|_{0,\Omega}^2 + M^2 \|v\|_{1,\Omega}^2),
\end{aligned}$$

wobei wir im vorletzten Schritt die Normäquivalenz von  $|\cdot|_1$  und  $|\cdot|_\infty$  auf dem  $\mathbb{R}^{d \times d}$  und die Normäquivalenz von  $|\cdot|_1$  und  $|\cdot|_2$  auf dem  $\mathbb{R}^{d^2+1}$  ausgenutzt haben. Durch Aufsummieren der einzelnen Komponenten ergibt sich

$$\begin{aligned}
\|\varepsilon(\mathfrak{E}v)\|_{0,\mathbb{R}^d}^2 - \|\varepsilon(v)\|_{0,\Omega}^2 &= \|\varepsilon(\mathfrak{E}v)\|_{0,\tilde{\Omega}\mathfrak{C}}^2 \\
&= \sum_{i,j} \|\varepsilon_{ij}(\mathfrak{E}v)\|_{0,\tilde{\Omega}\mathfrak{C}}^2 \\
&\leq d^2 c_1 (\|\varepsilon(v)\|_{0,\Omega}^2 + M^2 \|v\|_{1,\Omega}^2),
\end{aligned}$$

womit dann die Behauptung folgt.  $\square$

Wir bezeichnen eine Lipschitz-Menge als  $C^1$ -Menge, falls für Randpunkt  $x \in \Gamma$  die lokale Parametrisierung des Randes  $\psi: \mathbb{R}^{d-1} \rightarrow \mathbb{R}$  in Gleichung (1) stetig differenzierbar ist. Für diese Mengen gilt das folgende Lemma.

**Lemma 10** (Zerlegung der Eins). *Sei  $\Omega \subseteq \mathbb{R}^d$  eine  $C^1$ -Menge und  $M > 0$ . Dann gibt es*

- *Bälle  $B_{r_i}(x^i)$  um Randpunkte  $x^1, \dots, x^N \in \partial\Omega$ , die den Rand  $\Gamma \subseteq \bigcup_i B_{r_i}(x^i)$  überdecken, und  $B_{\tilde{r}_i}(x^i)$  mit Radii  $\tilde{r}_i^2 := r^2 + (2M+1)^2 r^2$ .*
- *Stetig differenzierbare lokale Parametrisierungen des Randes  $\psi_i: \mathbb{R}^{d-1} \rightarrow \mathbb{R}$  mit uniform beschränkten Ableitungen  $|\nabla \psi_i| \leq M$  und affinen Isometrien  $A_i: \mathbb{R}^d \rightarrow \mathbb{R}^d$ , so dass  $B_{\tilde{r}_i}(x^i) \cap \Omega = B_{\tilde{r}_i}(x^i) \cap A_i \text{ epi}(\psi_i)$  und  $\psi_i(0) = 0$ .*
- *Abschneidefunktionen  $\theta_i \in C_c^\infty(B_{r_i}(x^i); [0; 1])$  und  $\theta_0 \in C_c^\infty(\Omega; [0; 1])$  mit der Eigenschaft  $\sum_i \theta_i = 1$  auf  $\bar{\Omega}$ .*

*Beweis.* Da  $\Omega$  eine  $C^1$ -Menge ist, gibt es endlich viele Bälle  $B_{r_i}(x^i)$  mit  $x^i \in \Gamma$ , welche den Rand  $\Gamma$  überdecken, sowie  $\psi_i$  und  $A_i$ , so dass  $B_{r_i}(x^i) \cap \Omega = B_{r_i}(x^i) \cap A_i \operatorname{epi}(\psi_i)$  und  $\psi_i(0) = 0$  gelten. Wir können nach möglicher Verfeinerung annehmen, dass auch  $B_{\tilde{r}_i}(x^i) \cap \Omega = B_{\tilde{r}_i}(x^i) \cap A_i \operatorname{epi}(\psi_i)$  und, dass  $|\nabla \psi_i| \leq M$ . Da die  $B_{r_i}(x_i)$  die kompakte Menge  $\overline{\Omega}$  überdecken, liefert die Folgerung einer Proposition in [2, S.118f.] eine Zerlegung der Eins bezüglich der Überdeckung  $B_{r_i}(x^i)$ , so dass  $\sum_i \theta_i = 1$  auf  $\overline{\Omega}$  erfüllt ist.  $\square$

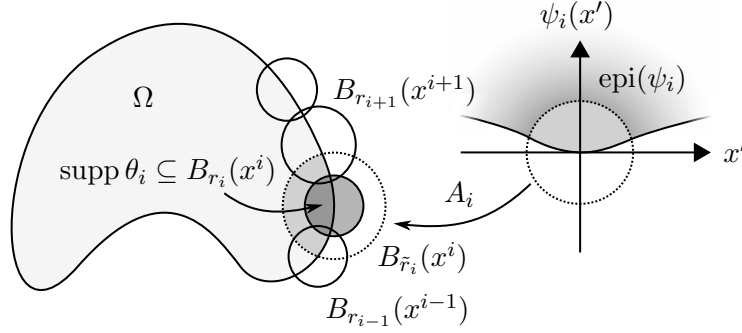


Abb. 10: Visualisierung zur Zerlegung der Eins.

Zusammen mit der Kornschen Ungleichung auf dem Ganzraum und der in Lemma 9 konstruierten Erweiterung erhalten wir das folgende Resultat.

**Satz 11** (Kornsche Ungleichung ohne Randbedingungen für  $C^1$ -Gebiete). *Sei  $\Omega \subseteq \mathbb{R}^d$  eine  $C^1$ -Menge, dann gibt es ein  $c_{K1} > 0$ , so dass für alle  $v \in H^1(\Omega; \mathbb{R}^d)$  gilt*

$$\|v\|_{1,\Omega} \leq c_{K1} \|v\|_{K,\Omega}. \quad (21)$$

*Beweis.* Wir folgen dem Beweis aus [16, S.242]. Setze  $M^2 := 1/(4c_E)$  und seien  $x^i$ ,  $r_i$ ,  $\tilde{r}_i$ ,  $\theta_i$ ,  $A_i$  und  $\psi_i$  wie im Lemma zur Partition der Eins. Wir definieren  $B_i := B_{r_i}(0)$ ,  $\tilde{B}_i := B_{\tilde{r}_i}(0)$  und  $\Omega_i := B_i \cap \operatorname{epi}(\psi_i)$ . Es gibt nach Lemma 9 stetige lineare Erweiterungen

$$\mathfrak{E}_i: H^1_{\partial B_i}(\Omega_i; \mathbb{R}^d) \rightarrow H^1_{\partial \tilde{B}_i}(\tilde{B}_i; \mathbb{R}^d),$$

die Gleichung (20) erfüllen. Wir bezeichnen mit  $\mathfrak{E}_0$  die Erweiterung von  $v \in H^1_{\partial \Omega}(\Omega; \mathbb{R}^d)$  mit 0 außerhalb von  $\Omega_0 := \Omega$  und setzen  $A_0 := \operatorname{Id}$ . Es gilt für  $x \in \Omega$

$$\mathfrak{E}_i((\theta_i v) \circ A_i) \circ A_i^{-1}(x) = (\theta_i v) \circ A_i(A_i^{-1}(x)) = (\theta_i v)(x),$$

womit dann mit Lemma 8 und Gleichung (20) folgt

$$\begin{aligned} |\theta_i v|_{1,\Omega}^2 &\leq |\mathfrak{E}_i((\theta_i v) \circ A_i) \circ A_i^{-1}|_{1,\mathbb{R}^d}^2 \\ &= |\mathfrak{E}_i((\theta_i v) \circ A_i)|_{1,\mathbb{R}^d}^2 \\ &\leq 2\|\varepsilon(\mathfrak{E}_i((\theta_i v) \circ A_i))\|_{0,\mathbb{R}^d}^2 \\ &\leq 2c_E(\|\varepsilon((\theta_i v) \circ A_i)\|_{0,\Omega_i}^2 + M^2|(\theta_i v) \circ A_i|_{1,\Omega_i}^2) \\ &\leq 2c_E\|\varepsilon(\theta_i v)\|_{0,\Omega}^2 + \frac{1}{2}|\theta_i v|_{1,\Omega}^2 \end{aligned}$$

Man erhält also

$$|\theta_i v|_{1,\Omega} \leq c_1 \|\varepsilon(\theta_i v)\|_{0,\Omega}$$

für ein  $c_1 > 0$ . Es gilt nun

$$\begin{aligned} \|\varepsilon_{jk}(\theta_i v)\|_{0,\Omega} &= \|\theta_i \varepsilon_{jk}(v) + \frac{1}{2} v_j \partial_k \theta_i + \frac{1}{2} v_k \partial_j \theta_i\|_{0,\Omega} \\ &\leq \|\theta_i \varepsilon_{jk}(v)\|_{0,\Omega} + \frac{1}{2} \|v_j \partial_k \theta_i\|_{0,\Omega} + \frac{1}{2} \|v_k \partial_j \theta_i\|_{0,\Omega} \\ &\leq \|\varepsilon_{jk}(v)\|_{0,\Omega} + \frac{1}{2} \|\nabla \theta_i\|_{\infty} \|v\|_{0,\Omega} + \frac{1}{2} \|\nabla \theta_i\|_{\infty} \|v\|_{0,\Omega} \\ &\leq \|\varepsilon(v)\|_{0,\Omega} + c_2 \|v\|_{0,\Omega} \end{aligned}$$

mit  $c_2 := \max_i \|\nabla \theta_i\|_{\infty}$ . Damit folgt

$$|\theta_i v|_{1,\Omega} \leq c_1 \|\varepsilon(\theta_i v)\|_{0,\mathbb{R}^d} \leq c_3 (\|\varepsilon(v)\|_{0,\Omega}^2 + \|v\|_{0,\Omega}^2)^{1/2}$$

für ein  $c_3 > 0$ . Schließlich erhalten wir

$$\begin{aligned} |v|_{1,\Omega} &= \left| \sum_i \theta_i v \right|_{1,\Omega} \leq \sum_i |\theta_i v|_{1,\Omega} \\ &\leq \sum_i c_3 (\|\varepsilon(v)\|_{0,\Omega}^2 + \|v\|_{0,\Omega}^2)^{1/2} \\ &= c_4 (\|\varepsilon(v)\|_{0,\Omega}^2 + \|v\|_{0,\Omega}^2)^{1/2} \end{aligned}$$

für ein  $c_4 > 0$ , womit die Behauptung folgt.  $\square$

Man bemerke, dass die Voraussetzung, dass  $\Omega$  ein  $C^1$ -Gebiet ist, eine recht starke Voraussetzung ist, die zum Beispiel Polyeder nicht erfüllen. Allerdings gilt diese Aussage auch allgemeiner für Lipschitz-Mengen. Hier wurde der Übersichtlichkeit halber dieser Beweis nicht geführt. Es wird beispielsweise in [16] mit derselben Beweisidee das Resultat für Lipschitz-Mengen gezeigt. Einen anderen Beweis von Lemma 11 liefert [13, S.106ff.]. Dieser zeigt das Resultat für Lipschitz-Gebiete auf dem Sobolevraum  $W^{1,p}(\Omega)$  für  $1 < p < \infty$  unter entsprechender Anpassung der Normen in Gleichung (21).

### 3.3 Kornsche Ungleichung mit Randbedingungen

Wir bezeichnen eine offene Zusammenhängende Menge  $\Omega \subseteq \mathbb{R}^d$  als Gebiet. Ein Gebiet, welches auch eine  $C^1$ -Menge ist, bezeichnen wir als  $C^1$ -Gebiet. Wir benötigen ein Lemma.

**Lemma 12** (Eine Charakterisierung von Starrkörperbewegungen). *Seien  $d = 3$  und  $\Omega \subseteq \mathbb{R}^d$  ein Gebiet, dann sind für  $v \in H^1(\Omega; \mathbb{R}^d)$  äquivalent*

1.  $\varepsilon(v) = 0$ .
2. Es gibt  $a, b \in \mathbb{R}^d$ , so dass  $v(x) = a \times x + b$ .



*Beweis.* Wir folgen den Beweisen von [5, S.293] und [3, S.295]. Da  $C^\infty(\Omega; \mathbb{R}^d)$  dicht ist in  $H^1(\Omega; \mathbb{R}^d)$ , reicht es, die Behauptung für  $v \in C^\infty(\Omega; \mathbb{R}^d)$  zu zeigen. Wir nehmen zunächst 1. an. Da  $\Omega$  zusammenhängend ist und für alle zweiten Ableitungen gilt

$$\begin{aligned}\partial_j \partial_k v_i &= \frac{1}{2}(\partial_j \partial_k v_i + \cancel{\partial_j \partial_i v_k}) + \frac{1}{2}(\cancel{\partial_k \partial_i v_j} + \partial_k \partial_j v_i) - \frac{1}{2}(\cancel{\partial_i \partial_j v_k} + \cancel{\partial_i \partial_k v_j}) \\ &= \partial_j \varepsilon_{ki}(v) + \partial_k \varepsilon_{ij}(v) - \partial_i \varepsilon_{jk}(v) = 0,\end{aligned}$$

gibt es  $A \in \mathbb{R}^{d \times d}$  und  $b \in \mathbb{R}^d$ , so dass  $v(x) = Ax + b$ . Es folgt aus

$$2\varepsilon(v) = \nabla v + \nabla v^\top = A + A^\top \quad (22)$$

und  $\varepsilon(v) = 0$ , dass  $A$  antisymmetrisch ist. Wegen

$$\begin{bmatrix} 0 & A_{12} & A_{13} \\ -A_{12} & 0 & A_{23} \\ -A_{13} & -A_{23} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} A_{12}x_2 + A_{13}x_3 \\ -A_{12}x_1 + A_{23}x_3 \\ -A_{13}x_1 - A_{23}x_2 \end{bmatrix} = \begin{bmatrix} -A_{23} \\ A_{13} \\ -A_{12} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (23)$$

gibt es ein  $a \in \mathbb{R}^3$ , so dass  $v(x) = a \times x + b$ , also folgt die 2. Charakterisierung. Ist andererseits 2. gegeben, so definiert die Beziehung (23) ein antisymmetrisches  $A \in \mathbb{R}^{3 \times 3}$ , so dass  $v(x) = Ax + b$ . Dann folgt aus Gleichung (22) und der Antisymmetrie von  $A$ , dass  $\varepsilon(v) = 0$ , also, dass 1. gilt.  $\square$

Nach [13, S.115] gilt Lemma 12 auch für Dimensionen  $d = 2$  und  $d = 1$  unter entsprechender Interpretation des Kreuzproduktes. Nun können wir die Kornsche Ungleichung mit Randbedingungen folgern.

**Satz 13** (Kornsche Ungleichung mit Randbedingungen). *Seien  $d = 3$ ,  $\Omega \subseteq \mathbb{R}^d$  ein  $C^1$ -Gebiet und die Randmenge  $S \subseteq \Gamma$  habe positives Flächenmaß. Dann gibt es  $c_{K2} > 0$ , so dass für alle  $v \in H_S^1(\Omega; \mathbb{R}^d)$  gilt*

$$\|v\|_{1,\Omega} \leq c_{K2} \|\varepsilon(v)\|_{0,\Omega}. \quad (24)$$

*Beweis.* Der Beweis folgt im wesentlichen [5, S.294]. Angenommen (24) gilt nicht, dann gibt es eine Folge  $\tilde{v}: \mathbb{N} \rightarrow H_S^1(\Omega; \mathbb{R}^d)$  mit

$$k \|\varepsilon(\tilde{v}_k)\|_{0,\Omega} < \|\tilde{v}_k\|_{1,\Omega}.$$

Da nach Reskalierung  $v_k := \tilde{v}_k / \|\tilde{v}_k\|_{1,\Omega} \in H_S^1(\Omega; \mathbb{R}^d)$  beschränkt ist, gibt es eine schwach konvergente Teilfolge  $v_{k_l}$ , die schwach gegen  $v \in H_S^1(\Omega; \mathbb{R}^d)$  konvergiert. Die Seminorm  $\|\varepsilon(\cdot)\|_{0,\Omega}$  ist konvex und stetig, also schwach unterhalbstetig. Es folgt aus

$$k \|\varepsilon(v_k)\|_{0,\Omega} < \|v_k\|_{1,\Omega} = 1$$

und aus der schwachen Unterhalbstetigkeit, dass

$$0 \leq \|\varepsilon(v)\|_{0,\Omega} \leq \liminf_{l \rightarrow \infty} \|\varepsilon(v_{k_l})\|_{0,\Omega} = 0.$$

Also ist  $|\varepsilon(v)| = 0$  fast überall und es folgt aus Lemma 12, dass  $v(x) = a \times x + b$  für ein  $a, b \in \mathbb{R}^3$ . Da  $v = 0$  auf  $S$  ist, haben wir  $a = b = 0$  und somit  $v = 0$ . Nach Rellich konvergiert  $v_{k_l}$  bezüglich  $\|\cdot\|_{0,\Omega}$  gegen  $v$ , also strebt  $\|v_{k_l}\|_{0,\Omega} \rightarrow 0$ . Nun folgt aber aus der Kornschen Ungleichung ohne Randbedingungen

$$1 = \|v_{k_l}\|_{1,\Omega} \leq c_{K1} (\|v_{k_l}\|_{0,\Omega}^2 + \|\varepsilon(v_{k_l})\|_{0,\Omega}^2)^{1/2} \xrightarrow{l \rightarrow \infty} 0,$$

ein Widerspruch. □

### 3.4 Elliptizität der Bilinearform $a$

Damit haben wir alle Voraussetzungen, um die Elliptizität von  $a$  zu folgern.

**Proposition 14** (Elliptizität von  $a$ ). *Seien  $d = 3$ ,  $\Omega \subseteq \mathbb{R}^d$  ein  $C^1$ -Gebiet mit elliptischem Hooke-Tensor  $C \in L^\infty(\Omega; \bigotimes_{i=1}^4 \mathbb{R}^d)$  und  $V^0 \subseteq H_S^1(\Omega; \mathbb{R}^d)$  für eine Randmenge  $S \subseteq \Gamma$  mit positivem Flächenmaß. Dann ist  $a: V^0 \times V^0 \rightarrow \mathbb{R}$  auch elliptisch, d. h. es gibt ein  $c_a > 0$ , so dass für alle  $v \in V^0$  gilt*

$$a(v, v) \geq c_a \|v\|_{1,\Omega}^2.$$

*Beweis.* Mit der Kornschen Ungleichung mit Randbedingungen folgt

$$\begin{aligned} a(v, v) &= \int_{\Omega} \sigma(v) : \varepsilon(v) \, dx \\ &= \int_{\Omega} \sum_{i,j,k,l} C_{ijkl} \varepsilon_{kl}(v) \varepsilon_{ij}(v) \, dx \\ &\geq c_c \int_{\Omega} \sum_{i,j} \varepsilon_{ij}(v) \varepsilon_{ij}(v) \, dx \\ &\geq \frac{c_c}{c_{K2}} \|v\|_{1,\Omega}^2, \end{aligned}$$

womit sich die Behauptung mit  $c_a := c_c/c_{K2}$  ergibt. □

Man bemerke, dass die einzige Stelle, an der verwendet wurde, dass  $\Omega$  ein  $C^1$  Gebiet ist, im Beweis der Kornschen Ungleichung ohne Randbedingungen war. Da diese Ungleichung, wie vorhin schon bemerkt auch auf Lipschitz-Mengen gilt, erhält man eine analoge Aussage zu Proposition 14, indem wir  $C^1$ -Gebiet mit Lipschitz-Gebiet ersetzt. Da wir die Voraussetzung  $d = 3$  ausschließlich für das Lemma 12 verwendet haben und dieses Lemma analog für  $d \leq 3$  gilt, kann man die Aussage 14 auch für  $d \leq 3$  verallgemeinern. Wir haben also gezeigt, dass unter gewissen Annahmen Existenz und Eindeutigkeit des kontinuierlichen Problems folgen. Damit ist das kontinuierliche Problem ausreichend besprochen und wir können uns dem diskreten Problem widmen.

## 4 Diskretisierung mit P1 Finiten Elementen

Wir diskretisieren in diesem Abschnitt das kontinuierliche Problem (14). Dazu definieren wir zunächst Begriffe aus [3, Kapitel II.§5] und führen anschließend das P1 Finite Element ein. Im Fall von  $d = 3$  Dimensionen soll man im Folgenden die Begriffe „Dreiecke“ und „Kanten“ als „Tetraeder“ und „Seitenflächen“ interpretieren.

Sei  $\Omega$  ein polytopes Lipschitz-Gebiet. Eine gemäß [3, Definition II.5.1] zulässige Triangulierung  $\mathcal{T}$  von  $\Omega$  habe die  $n$  Knotenpunkte  $x^1, \dots, x^n \in \mathbb{R}^d$ . Man definiert:

- $\mathcal{E} = \{E: E \text{ ist Kante von } T \text{ für ein } T \in \mathcal{T}\}$  ist die Menge aller Kanten
- $\mathcal{E}_D = \{E \in \mathcal{E}: E \subseteq \Gamma_D\}$  ist die Menge der Dirichlet-Kanten
- $\mathcal{E}_N = \{E \in \mathcal{E}: E \subseteq \Gamma_N\}$  ist die Menge der Neumann-Kanten
- $\mathcal{E}_\Gamma = \mathcal{E}_D \cup \mathcal{E}_N$  ist die Menge der Rand-Kanten
- $\mathcal{E}_\Omega := \mathcal{E} \setminus \mathcal{E}_\Gamma$  ist die Menge der inneren Kanten
- $\mathcal{N} = \{x^i\}_{i=1}^n$  ist die Menge der Knoten
- $\mathcal{N}_\Gamma = \{x^{ij} \in \mathcal{N}: x^{ij} \in \Gamma\}$  ist die Menge der Rand-Knoten

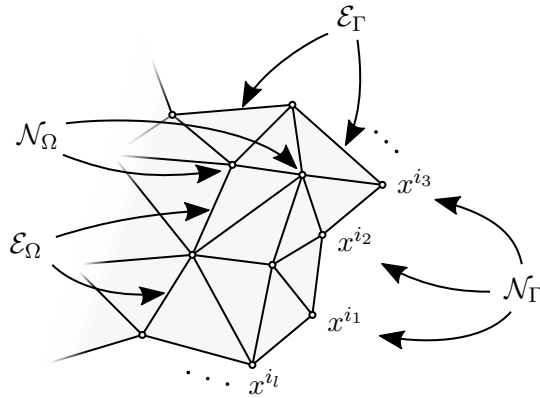


Abb. 11: Eine zulässige Triangulierung.

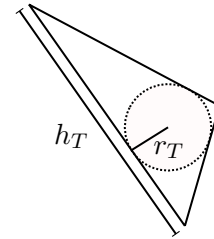


Abb. 12: Der Durchmesser  $h_T$  und der Innenkreisradius  $r_T$  des Dreiecks  $T$ .

Weiter bezeichnen wir mit  $h_T := \text{diam}(T)$  den Durchmesser eines Dreiecks  $T$ . Eine Familie an zulässigen Triangulierungen  $\mathcal{T}_h$  wird uniform genannt, falls es einen Regularitätsparameter  $\kappa > 0$  gibt, so dass jedes Dreieck  $T \in \mathcal{T}_h$  einen Kreis mit Radius  $r_T$  enthält, so dass gilt

$$r_T \geq \frac{h_T}{\kappa}.$$

Wir definieren  $\varphi_i \in H^1(\Omega)$  für  $i \in \{1, \dots, n\}$  als die nodale Basis in einer Dimension. Sie ist gegeben durch  $\varphi_i(x^j) = \delta_{ij}$ , wobei  $\varphi_i|_T$  auf jedem Dreieck  $T \in \mathcal{T}$  linear interpoliert

wird. In Abbildung 13 ist ein solcher Basisvektor dargestellt. Mit der eindimensionalen nodalen Basis definiert man die  $d$ -dimensionale nodale Basis durch

$$[\phi_1 \cdots \phi_{dn}] = [\varphi_1 e_1 \cdots \varphi_1 e_d \quad \cdots \quad \varphi_n e_1 \cdots \varphi_n e_d] .$$

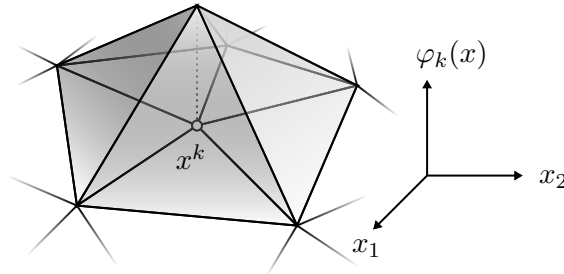


Abb. 13: Darstellung eines Basisvektors der nodalen Basis

Wir diskretisieren die Menge der zulässigen Verschiebungen  $V = H^1(\Omega; \mathbb{R}^d)$  zu

$$V_h := \text{Span}\{\phi_i\}_i \subseteq V . \quad (25)$$

Dieses Element wird P1-Element oder auch Courant-Element genannt. Nun diskretisieren wir die Menge  $V^0$  zu

$$V_h^0 := V_h \cap V^0 . \quad (26)$$

Damit  $V_h^0$  nichtleer ist, sollten wir annehmen, dass  $M$  stückweise linear ist, also, dass  $M|_E \in P^1(E; \mathbb{R}^{d \times d})$  auf jeder Rand-Kante  $E \in \mathcal{E}_\Gamma$ . Damit die Randbedingung

$$Mu = w \quad \text{auf } \Gamma$$

erfüllbar ist, müssen wir annehmen, dass auch  $w$  stückweise linear ist, also  $w|_E \in P^1(E; \mathbb{R}^d)$  auf allen Neumann-Kanten  $E \in \Gamma_N$ . Die diskrete Formulierung vom Problem (14) lautet nun: Finde  $u_h \in V_h$ , so dass

$$\begin{aligned} a(u_h, v_h) &= \ell(v_h) \quad \text{für alle } v_h \in V_h^0, \\ Mu_h &= w \quad \text{auf } \Gamma. \end{aligned} \quad (27)$$

Falls  $a$  und  $\ell$  die Voraussetzungen des Lax-Milgram Lemmas erfüllen, liefert Korollar 7 die Existenz und Eindeutigkeit des diskreten Problems (27).

## 5 A priori Fehlerabschätzung

Seien im Folgenden  $u$  Lösung des kontinuierlichen Problems (14) und  $u_h$  Lösung des diskreten Problems (27). Wir definieren den Fehler

$$e := u - u_h.$$

Unser Ziel ist es im folgenden Abschnitt, den Fehler  $\|e\|_{1,\Omega}$  für das P1 Finite Element durch  $u$  und der Gitterweite  $h := \max_{T \in \mathcal{T}} h_T$  abzuschätzen, also ohne die diskrete Lösung  $u_h$  zu verwenden. Dazu folgen wir der Argumentationsstruktur von [3, Satz II.7.3]. Wir benötigen zunächst ein Resultat.

**Proposition 15** (Céas Lemma). *Es sei  $a$  eine elliptische Bilinearform. Dann erfüllt der Approximationsfehler*

$$\|e\|_{1,\Omega} \leq \frac{c_A}{c_a} \inf_{v_h \in V_h^0} \|u - v_h\|_{1,\Omega}.$$

*Beweis.* Siehe Satz 4.2 in [3, S.53] für das Resultat in einer Komponente, d. h. auf  $H^1(\Omega)$ . Der Fall mit mehreren Komponenten folgt analog.  $\square$

Wir benötigen noch ein weiteres Resultat.

**Proposition 16** (Interpolation, Approximationssatz). *Seien  $d = 2$  und  $\mathcal{T}$  eine uniforme Triangulierung. Dann gibt es einen Interpolationsoperator  $I_h: H^2(\Omega; \mathbb{R}^d) \rightarrow V_h$  und ein  $c_{I1} > 0$ , welches von  $\Omega$  und  $\kappa$  abhängt, so dass für alle  $v \in H^2(\Omega; \mathbb{R}^d)$*

$$\|v - I_h v\|_{1,\Omega} \leq c_{I1} h |v|_{2,\Omega}.$$

*Beweis.* Satz 6.4 in [3, S.75] mit  $t = 2$  und  $m = 1$  liefert die Aussage in einer Komponente. Man definiert  $I_h$  komponentenweise und das Resultat folgt dann durch Aufsummieren der beiden Komponenten.  $\square$

Wenn man diese beiden Aussagen kombiniert, erhält man

**Satz 17** (A priori Fehler). *Seien  $d = 2$ ,  $\mathcal{T}$  eine uniforme Triangulierung und  $u \in H^2(\Omega; \mathbb{R}^d)$ . Dann gibt es ein  $c > 0$ , welches von  $\Omega$  und  $\kappa$  abhängt, so dass*

$$\|e\|_{1,\Omega} \leq c \frac{c_A}{c_a} h |u|_{2,\Omega}.$$

*Beweis.* Dies folgt aus

$$\begin{aligned} \|e\|_{1,\Omega} &\leq \frac{c_A}{c_a} \inf_{v_h \in V_h^0} \|u - v_h\|_{1,\Omega} \\ &\leq \frac{c_A}{c_a} \|u - I_h u\|_{1,\Omega} \\ &\leq c_{I1} \frac{c_A}{c_a} h |u|_{2,\Omega}. \end{aligned}$$

$\square$

Die diskrete Lösung  $u_h$  konvergiert also bezüglich der Norm  $\|\cdot\|_{1,\Omega}$  in linearer Ordnung in  $h$  gegen die kontinuierliche Lösung  $u$ , falls  $u$  regulär genug ist.

## 6 A posteriori residualer Fehlerschätzer

Seien  $u, u_h$  und  $e$  wie im vorhergehenden Abschnitt. Unser Ziel ist, den Fehler  $\|e\|_{1,\Omega}$  anhand von  $u_h, f$  und  $g$ , also ohne die exakte Lösung  $u$  zu verwenden, nach oben und nach unten abzuschätzen. Dies führt auf den Begriff des a posteriori Fehlerschätzers. Zunächst konstruieren wir einen residualen Fehlerschätzer und zeigen mit einer Interpolation vom Clément-Typ, dass dieser den Fehler nach oben beschränkt, also zuverlässig ist. Anschließend zeigen wir mit einem Hilfsresultat, dass der Fehlerschätzer den Fehler lokal nach unten beschränkt, also auch lokal effizient ist. Hierbei folgen wir der Struktur von [3, Kapitel III.§8]. Außerdem verwenden wir Kapitel 1.1-1.6 und Kapitel 3.6 in [17]. Der Einfachheit halber werden wir im Folgenden von klassischen Dirichlet-Neumann-Randbedingungen ausgehen.

### 6.1 Konstruktion des residualen Fehlerschätzers

Wir definieren zunächst die flächenbezogenen Residuen  $R_T \in L^2(T; \mathbb{R}^d)$  durch

$$R_T := (f + \operatorname{Div} \sigma(u_h))|_T \quad (28)$$

für alle Dreiecke  $T \in \mathcal{T}$ . Für eine innere Kante  $E \in \mathcal{E}_\Omega$  gibt es angrenzende Dreiecke  $T_1, T_2 \in \mathcal{T}$ , so dass  $E = T_1 \cap T_2$ . Wir definieren die kantenbezogenen Sprünge durch

$$[\![\sigma(u_h)n_E]\!] := \sigma(u_h)n_{T_1}|_E + \sigma(u_h)n_{T_2}|_E.$$

Hierbei bezeichnen  $\cdot|_E: H^1(T_i; \mathbb{R}^d) \rightarrow L^2(E; \mathbb{R}^d)$  den (eingeschränkten) Trace Operator und  $n_{T_i}$  die äußeren Normalen von  $T_i$  für  $i \in \{1, 2\}$ . In der Implementation des Fehlerschätzers wird verwendet, dass  $n_{T_1} = -n_{T_2}$  auf der Kante  $E$  gilt. Man definiert nun die kantenbezogenen Residuen  $R_E \in L^2(E; \mathbb{R}^d)$  durch

$$R_E = \begin{cases} -[\![\sigma(u_h)n_E]\!] & \text{auf inneren Kanten } E \in \mathcal{E}_\Omega, \\ g - \sigma(u_h)n_E & \text{auf Neumann-Kanten } E \in \mathcal{E}_N, \\ 0 & \text{sonst,} \end{cases} \quad (29)$$

für Kanten  $E \in \mathcal{E}$ . Wir zeigen jetzt zwei Resultate für den Fehler.

**Proposition 18** (Galerkin Orthogonalität). *Für alle Testfunktionen  $v_h \in V_h^0$  gilt*

$$a(e, v_h) = 0.$$

*Beweis.* Wegen  $V_h^0 \subseteq V^0$  gilt

$$a(e, v_h) = a(u, v_h) - a(u_h, v_h) = \ell(v_h) - \ell(v_h) = 0$$

für alle Testfunktionen  $v_h \in V_h^0$ . □

**Lemma 19.** *Für alle Testfunktionen  $v \in V^0$  gilt*

$$a(e, v) = \sum_{T \in \mathcal{T}} \langle R_T, v \rangle_{0,T} + \sum_{E \in \mathcal{E}} \langle R_E, v \rangle_{0,E}.$$

*Beweis.* Man rechnet

$$\begin{aligned}
a(u_h, v) &= \int_{\Omega} \sigma(u_h) : \varepsilon(v) \, dx \\
&= \sum_{T \in \mathcal{T}} \int_T \sum_{i,j} \frac{1}{2} \sigma_{ij}(u_h) (\partial_i v_j + \partial_j v_i) \, dx \\
&= \sum_T \frac{1}{2} \int_T \sum_{i,j} (\sigma_{ij}(\partial_i v_j) + \sigma_{ij}(\partial_j v_i)) \, dx \\
&\stackrel{\text{Leibniz Regel für schwache Ableitungen}}{=} \sum_T \frac{1}{2} \int_T \sum_{i,j} (\partial_i(\sigma_{ij} v_j) + \partial_j(\sigma_{ij} v_i) \\
&\quad - (\partial_i \sigma_{ij}) v_j - (\partial_j \sigma_{ij}) v_i) \, dx \\
&\stackrel{\text{Symmetrie von } \sigma}{=} \sum_T \left( \int_T \sum_{i,j} (\partial_i(\sigma_{ij} v_j) \, dx - \int_T \sum_{i,j} (\partial_i \sigma_{ji}) v_j \, dx \right) \\
&\stackrel{\text{Schwacher Satz von Gauß}}{=} \sum_T \int_{\partial T} \sum_{i,j} \sigma_{ji} v_j n_{T,i} \, ds - \sum_T \int_T (\text{Div } \sigma) \cdot v \, dx \\
&= \sum_{T \in \mathcal{T}} \int_{\partial T} \sigma(u_h) n_T \cdot v \, ds - \sum_{T \in \mathcal{T}} \langle \text{Div } \sigma(u_h), v \rangle_{0,T}
\end{aligned}$$

mit  $n_T: \partial T \rightarrow \mathbb{R}^d$  die äußere Normale. In der ersten Summe wird über die inneren Kanten  $\mathcal{E}_{\Omega}$  doppelt und die Rand-Kanten  $\mathcal{E}_{\Gamma}$  einfach integriert und wir erhalten nach Umsortieren der Summanden

$$\begin{aligned}
\sum_{T \in \mathcal{T}} \int_{\partial T} \sigma(u_h) n_T \cdot v \, ds &= \sum_{E \in \mathcal{E}_{\Omega}} \int_E \llbracket \sigma(u_h) n_E \rrbracket \cdot v \, ds + \sum_{E \in \mathcal{E}_N} \int_E \sigma(u_h) n_E \cdot v \, ds \\
&\quad + \sum_{E \in \mathcal{E}_{\Gamma} \setminus \mathcal{E}_N} \int_E \underbrace{\sigma(u_h) n_E \cdot v}_{=0, \text{ da } v|_E=0} \, ds \\
&= - \sum_{E \in \mathcal{E}_{\Omega}} \int_E R_E \cdot v \, ds - \sum_{E \in \mathcal{E}_N} \int_E R_E \cdot v \, ds \\
&\quad - \sum_{E \in \mathcal{E}_{\Gamma} \setminus \mathcal{E}_N} \int_E R_E \cdot v \, ds + \sum_{E \in \mathcal{E}_N} \int_E g \cdot v \, ds \\
&= - \sum_{E \in \mathcal{E}} \langle R_E, v \rangle_{0,E} + \langle g, v \rangle_{0,\Gamma_N}.
\end{aligned}$$

Hierbei haben wir ausgenutzt, dass, falls eine Kante  $E \in \mathcal{E}$  zwischen zwei Dreiecken  $T_1, T_2 \in \mathcal{T}$  liegt,  $v|_{\partial T_1} = v|_{\partial T_2}$  auf  $E$  gilt, sowie

$$\llbracket \sigma(u_h) n_E \rrbracket = \sigma(u_h) n_{T_1}|_E + \sigma(u_h) n_{T_2}|_E.$$

Es ergibt sich also

$$\begin{aligned} -a(u_h, v) &= - \sum_{T \in \mathcal{T}} \int_{\partial T} \sigma(u_h) n_T \cdot v \, ds + \sum_{T \in \mathcal{T}} \langle \operatorname{Div} \sigma(u_h), v \rangle_{0,T} \\ &= \sum_{E \in \mathcal{E}} \langle R_E, v \rangle_{0,E} - \langle g, v \rangle_{0,\Gamma_N} + \sum_{T \in \mathcal{T}} \langle \operatorname{Div} \sigma(u_h), v \rangle_{0,T}. \end{aligned}$$

Damit erhalten wir wie erwünscht

$$\begin{aligned} a(e, v) &= a(u, v) - a(u_h, v) \\ &= \ell(v) - a(u_h, v) \\ &= \langle f, v \rangle_{0,\Omega} + \langle g, v \rangle_{0,\Gamma_N} + \sum_{E \in \mathcal{E}} \langle R_E, v \rangle_{0,E} - \langle g, v \rangle_{0,\Gamma_N} + \sum_{T \in \mathcal{T}} \langle \operatorname{Div} \sigma(u_h), v \rangle_{0,T} \\ &= \sum_{T \in \mathcal{T}} \langle f, v \rangle_{0,T} + \sum_{E \in \mathcal{E}} \langle R_E, v \rangle_{0,E} + \sum_{T \in \mathcal{T}} \langle \operatorname{Div} \sigma(u_h), v \rangle_{0,T} \\ &= \sum_{T \in \mathcal{T}} \langle R_T, v \rangle_{0,T} + \sum_{E \in \mathcal{E}} \langle R_E, v \rangle_{0,E}. \end{aligned}$$

□

Tatsächlich ist Lemma 19 die Motivation für die Definition der flächenbezogenen Residuen  $R_T$  und der kantenbezogenen Residuen  $R_E$ . Für einen lokalen residualen Schätzer wird der Fehler auf einem Dreieck  $T$  nun mit  $R_T$  und  $R_E$  für  $E \in \partial T$  abgeschätzt. Konkret definieren wir einen lokalen residualen Fehlerschätzer durch

$$\eta_{R,T}^2 := h_T^2 \|R_T\|_{0,T}^2 + \frac{1}{2} \sum_{E \in \mathcal{E}_\Omega \cap \partial T} h_E \|R_E\|_{0,E}^2 + \sum_{E \in \mathcal{E}_\Gamma \cap \partial T} h_E \|R_E\|_{0,E}^2, \quad (30)$$

wobei  $h_E := \operatorname{diam}(E)$  den Durchmesser von  $E$  bezeichnet. Wir setzen den lokalen Fehlerschätzer zu einem globalen Fehlerschätzer zusammen

$$\begin{aligned} \eta_R^2 &:= \sum_{T \in \mathcal{T}} \eta_{R,T}^2 \\ &= \sum_{T \in \mathcal{T}} h_T^2 \|R_T\|_{0,T}^2 + \sum_{T \in \mathcal{T}} \frac{1}{2} \sum_{E \in \mathcal{E}_\Omega \cap \partial T} h_E \|R_E\|_{0,E}^2 \\ &\quad + \sum_{T \in \mathcal{T}} \sum_{E \in \mathcal{E}_\Gamma \cap \partial T} h_E \|R_E\|_{0,E}^2 \\ &= \sum_{T \in \mathcal{T}} h_T^2 \|R_T\|_{0,T}^2 + \sum_{E \in \mathcal{E}} h_E \|R_E\|_{0,E}^2, \end{aligned} \quad (31)$$

wobei wir ausgenutzt haben, dass wir über alle inneren Kanten doppelt und über alle Randkanten nur einfach summiert haben.



## 6.2 Obere Abschätzung des Fehlers

Wir zeigen im Folgenden, dass  $\eta_R$  den Fehler nach oben beschränkt. Hierfür benötigen wir ein Resultat und eine dazugehörige Definition aus [17, S.6f.].

Wir definieren eine Umgebung  $\tilde{\omega}_T$  von  $T$  als die Menge aller Dreiecke  $T' \in \mathcal{T}$ , so dass  $T$  und  $T'$  mindestens einen Knoten gemeinsam haben. Der Einfachheit halber behandeln wir im Folgenden nur den Fall, dass wir  $d = 2$  Dimensionen haben. Allerdings wird in [17, S.108ff.] gezeigt, dass sich das folgende Resultat auch auf andere Dimensionen verallgemeinern lässt.

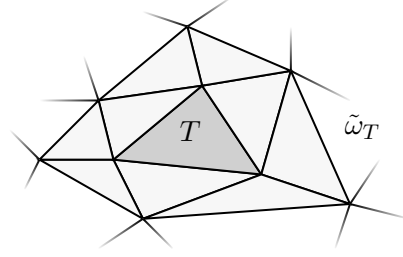


Abb. 14: Die Umgebung  $\tilde{\omega}_T$  von  $T$ .

**Proposition 20** (Interpolation vom Clément-Typ). *Sei  $d = 2$  und  $\mathcal{T}$  eine uniforme Triangulierung. Dann existiert eine Abbildung  $I_h: V^0 \rightarrow V_h^0$  und eine von  $\Omega$  und dem Regularitätsparameter abhängigen Konstante  $c_{I2}$  mit der Eigenschaft*

$$\begin{aligned} \|v - I_h v\|_{0,T} &\leq c_{I2} h_T |v|_{1,\tilde{\omega}_T}, \\ \|v - I_h v\|_{0,E} &\leq c_{I2} h_E^{1/2} |v|_{1,\tilde{\omega}_T} \end{aligned}$$

für alle  $v \in V^0$ , Kanten  $E \in \partial T$  und Dreiecke  $T \in \mathcal{T}$ .

*Beweis.* In [17, S.8] wird dieses Resultat auf  $H_{\Gamma_D}^1(\Omega)$  gezeigt. Indem wir  $I_h$  dann Komponentenweise definieren, folgen die Abschätzungen dann durch Aufsummieren der einzelnen Komponenten.  $\square$

**Satz 21** (Obere Abschätzung des Fehlers, Zuverlässigkeit). *Sei  $d = 2$ ,  $\mathcal{T}$  eine uniforme Triangulierung von  $\Omega$  und sei  $a$  elliptisch. Dann gibt es eine vom Gebiet, von  $c_a$  und vom Regularitätsparameter abhängige Konstante  $c$ , so dass für den Fehler gilt*

$$\|e\|_{1,\Omega} \leq c \eta_R.$$

*Beweis.* Im Folgenden bezeichnet  $I_h$  den Operator aus Proposition 20. Damit haben wir

$$\begin{aligned} &\text{Elliptizität von } a \\ &\downarrow \\ c_a \|e\|_{1,\Omega}^2 &\leq a(e, e) \\ &\downarrow \text{Galerkin-Orthogonalität} \\ &= a(e, e - I_h e) \\ &\downarrow \text{Proposition 19} \\ &= \sum_{T \in \mathcal{T}} \langle R_T, e - I_h e \rangle_{0,T} + \sum_{E \in \mathcal{E}} \langle R_E, e - I_h e \rangle_{0,E} \\ &\downarrow \text{Cauchy-Schwarz Ungleichung} \\ &\leq \sum_T \|R_T\|_{0,T} \|e - I_h e\|_{0,T} + \sum_E \|R_E\|_{0,E} \|e - I_h e\|_{0,E} \end{aligned}$$

und man erhält weiter

$$\begin{aligned}
c_a \|e\|_{1,\Omega}^2 &\leq \sum_T \|R_T\|_{0,T} \|e - I_h e\|_{0,T} + \sum_E \|R_E\|_{0,E} \|e - I_h e\|_{0,E} \\
&\stackrel{\text{Proposition 20}}{\leq} \sum_T c_{I2} h_T \|R_T\|_{0,T} \|e\|_{1,\tilde{\omega}_T} + \sum_E c_{I2} h_E^{1/2} \|R_E\|_{0,E} \|e\|_{1,\tilde{\omega}_T} \\
&\stackrel{\text{Cauchy-Schwarz Ungleichung}}{\leq} c_{I2} \left( \sum_{T \in \mathcal{T}} h_T^2 \|R_T\|_{0,T}^2 + \sum_{E \in \mathcal{E}} h_E \|R_E\|_{0,E}^2 \right)^{1/2} \left( \sum_T \|e\|_{1,\tilde{\omega}_T}^2 + \sum_E \|e\|_{1,\tilde{\omega}_T}^2 \right)^{1/2} \\
&\leq c_{I2} (d+2)^{1/2} \eta_R \left( \sum_T \|e\|_{1,\tilde{\omega}_T}^2 \right)^{1/2} \\
&\stackrel{(*)}{\leq} c_U \eta_R \left( \sum_{T \in \mathcal{T}} \|e\|_{1,T}^2 \right)^{1/2} \\
&\leq c_U \eta_R \|e\|_{1,\Omega}.
\end{aligned}$$

Hierbei haben wir in  $(*)$  ausgenutzt, dass aufgrund der uniformen Triangulation die Anzahl der Dreiecke in der Umgebung  $\tilde{\omega}_T$  nach oben beschränkt ist. Die Konstante  $c_U$  hängt also vom Gebiet und vom Regularitätsparameter ab. Nach Division durch  $c_a \|e\|_{1,\Omega}$  ergibt sich die Behauptung mit  $c = c_U/c_a$ .  $\square$

Man bemerke, dass die Konstante der oberen Abschätzung von der Elliptizitätskonstante  $c_a$  abhängt, also insbesondere auch vom Materialparameter  $\mu$ .

### 6.3 Untere Abschätzung des Fehlers

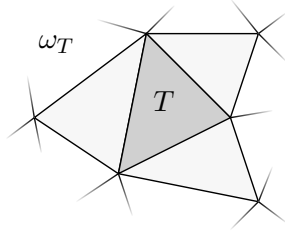


Abb. 15: Die Umgebung  $\omega_T$  von  $T$ .

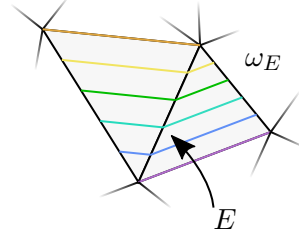


Abb. 16: Die Umgebung  $\omega_E$  von  $E$ .

Die folgenden Definitionen und die folgende Proposition sind [3, S.170f.] und [17, S.83] entnommen. Wir bezeichnen die Menge aller Dreiecke  $T' \in \mathcal{T}$ , sodass  $T$  und  $T'$  eine Kante gemeinsam haben, als Umgebung  $\omega_T$ . Für eine Kante  $E \in \mathcal{E}$  bezeichnen wir mit  $\omega_E$  die Menge aller Dreiecke  $T' \in \mathcal{T}$ , sodass  $E$  eine Kante von  $T'$  ist. Wir definieren außerdem eine Erweiterung  $\mathfrak{E}: L^2(E) \rightarrow L^2(\omega_E)$ , indem wir  $\mathfrak{E}v(x') = v(x)$  setzen, wobei

$x' \in T' \in \omega_E$  und  $x \in E$  mit  $\text{dist}(x, E') = \text{dist}(x', E')$  für eine Kante  $E'$  von  $T'$  mit  $E \neq E'$ . Diese Konstruktion ist in Abbildung 16 angedeutet.

Für ein Dreieck  $T \in \mathcal{T}$  mit Knoten  $x^1, \dots, x^{d+1}$  definieren wir die Blasenfunktion  $\theta_T \in H^1_\Gamma(\Omega)$  durch

$$\theta_T|_T := (d+1)^{d+1} \prod_{i=1}^{d+1} \varphi_i|_T \in P^{d+1}(T)$$

und setzen  $\theta_T = 0$  außerhalb von  $T$ . Man sieht dann, dass  $\text{supp } \theta_T = T$ ,  $0 \leq \theta_T \leq 1$  und  $1 = \max \theta_T$ . Für eine Kante  $E$  mit Knoten  $x^1, \dots, x^d$  konstruieren wir eine Abschneidefunktion  $\theta_E \in H^1(\Omega)$  durch

$$\theta_E|_{\omega_E} := d^d \prod_{i=1}^d \varphi_i|_{\omega_E} \in P^d(T)$$

und setzen  $\theta_E = 0$  außerhalb von  $\omega_E$ . Auch hier sieht man, dass  $\text{supp } \theta_T = \omega_E$  und  $0 \leq \theta_E \leq 1 = \max \theta_E$ .

**Proposition 22.** *Sei  $\mathcal{T}$  eine uniforme Triangulierung. Dann gilt für  $T \in \mathcal{T}$  und  $E \in \mathcal{E}$  mit  $E \subseteq \partial T$*

$$c_1 \|p\|_{0,T} \leq \|\theta_T^{1/2} p\|_{0,T} \quad \text{für } p \in P^k(T; \mathbb{R}^d), \quad (32)$$

$$|\theta_T p|_{1,T} \leq c_2 h_T^{-1} \|\theta_T p\|_{0,T} \quad \text{für } p \in P^k(T; \mathbb{R}^d), \quad (33)$$

$$c_3 \|q\|_{0,E} \leq \|\theta_E^{1/2} q\|_{0,E} \quad \text{für } q \in P^k(E; \mathbb{R}^d), \quad (34)$$

$$|\theta_E \mathfrak{E} q|_{1,T} \leq c_4 h_E^{-1/2} \|q\|_{0,E} \quad \text{für } q \in P^k(E; \mathbb{R}^d), \quad (35)$$

$$\|\theta_E \mathfrak{E} q\|_{0,T} \leq c_5 h_E^{1/2} \|q\|_{0,E} \quad \text{für } q \in P^k(E; \mathbb{R}^d), \quad (36)$$

wobei die Konstanten  $c_1, \dots, c_5$  nur vom Regularitätsparameter der Triangulierung und vom Polynomgrad  $k$  abhängen.

*Beweis.* Die Aussage folgt für die Komponenten  $p_i \in P^2(T)$  und  $q_i \in P^2(E)$  aus der Proposition in [17, S.112], welches durch Transformation auf ein Referenz-Simplex bewiesen wird. Das Resultat folgt dann durch Aufsummieren der einzelnen Komponenten.  $\square$

Wir bezeichnen mit  $\mathcal{L}^0(\mathcal{T}; \mathbb{R}^d)$  alle Abbildungen  $f_h: \Omega \rightarrow \mathbb{R}^d$ , so dass auf allen Dreiecken  $T \in \mathcal{T}$  gilt, dass  $f_h|_T$  konstant ist. Analog bezeichnen wir mit  $\mathcal{L}^0(\mathcal{E}_N; \mathbb{R}^d)$  alle Abbildungen  $g_h: \bigcup \mathcal{E}_N \rightarrow \mathbb{R}^d$ , so dass für alle Kanten  $E \in \mathcal{E}_N$  gilt  $g_h|_E$  ist konstant. Damit erhalten wir schließlich folgende Abschätzung für P1 Finite Elemente.

**Satz 23** (Untere Abschätzung des Fehlers, lokale Effizienz). *Seien  $d = 2$ ,  $\mathcal{T}$  eine uniforme Triangulierung und  $a$  stetig. Dann gibt es eine von  $c_A$  und vom Regularitätsparameter abhängige Konstante  $c > 0$ , so dass für alle  $f_h \in \mathcal{L}^0(\mathcal{T}; \mathbb{R}^d)$  und  $g_h \in \mathcal{L}^0(\mathcal{E}_N; \mathbb{R}^d)$  gilt*

$$\eta_{R,T}^2 \leq c \left( \sum_{T \in \omega_T} \|e\|_{0,T}^2 + \sum_{T \in \omega_T} h_T^2 \|f_h - f\|_{0,T}^2 + \sum_{E \in \mathcal{E}_N \cap \partial T} h_E \|g_h - g\|_{0,E}^2 \right).$$

*Beweis.* Wir folgen im Beweis [3, S.173f.]. Da

$$R_{T,h} := \text{Div } \sigma_h + f_h$$

polynomiell ist, können wir Proposition 22 anwenden und erhalten

$$\begin{aligned}
& \overset{|\theta_T| \leq 1}{\curvearrowright} c_1^2 \|\theta_T R_{T,h}\|_{0,T} \|R_{T,h}\|_{0,T} \leq c_1^2 \|R_{T,h}\|_{0,T}^2 \\
& \overset{\text{Gleichung (32)}}{\curvearrowright} \leq \|\theta_T^{1/2} R_{T,h}\|_{0,T}^2 \\
& = \langle R_{T,h}, \theta_T R_{T,h} \rangle_{0,T} \\
& \overset{R_{T,h} = R_T + f_h - f}{\curvearrowright} = \langle R_T, \theta_T R_{T,h} \rangle_{0,T} + \langle f_h - f, \theta_T R_{T,h} \rangle_{0,T} \\
& \overset{\text{supp } \theta_T \subseteq T \text{ und } \theta_T|_E = 0 \text{ für alle } E \in \mathcal{E}}{\curvearrowright} = \sum_{T \in \mathcal{T}} \langle R_T, \theta_T R_{T,h} \rangle_{0,T} + \sum_{E \in \mathcal{E}} \langle R_E, \theta_T R_{T,h} \rangle_{0,E} \\
& \quad + \langle f_h - f, \theta_T R_{T,h} \rangle_{0,T} \\
& \overset{\text{Proposition 19}}{\curvearrowright} = a(e, \theta_T R_{T,h}) + \langle f_h - f, \theta_T R_{T,h} \rangle_{0,T} \\
& \overset{\text{Stetigkeit von } a, \text{ supp } \theta_T \subseteq T \text{ und Cauchy-Schwarz}}{\curvearrowright} \leq c_A |e|_{1,T} |\theta_T R_{T,h}|_{1,T} + \|f_h - f\|_{0,T} \|\theta_T R_{T,h}\|_{0,T} \\
& \overset{\text{Gleichung (33)}}{\curvearrowright} \leq c_6 h_T^{-1} \|e\|_{1,T} \|\theta_T R_{T,h}\|_{0,T} + \|f_h - f\|_{0,T} \|\theta_T R_{T,h}\|_{0,T} \\
& \leq c_7 h_T^{-1} (\|e\|_{1,T} + h_T \|f_h - f\|_{0,T}) \|\theta_T R_{T,h}\|_{0,T}.
\end{aligned}$$

Es folgt nach Multiplikation mit  $h_T / (c_1^2 \|\theta_T R_{T,h}\|_{0,T})$

$$h_T \|R_{T,h}\|_{0,T} \leq c_8 (\|e\|_{1,T} + h_T \|f_h - f\|_{0,T})$$

und mit der Dreiecksungleichung weiter

$$\begin{aligned}
h_T \|R_T\|_{0,T} &= h_T \|f + \text{Div } \sigma_h\|_{0,T} \\
&\leq h_T (\|f_h + \text{Div } \sigma_h\|_{0,T} + \|f_h - f\|_{0,T}) \\
&= h_T (\|R_{T,h}\|_{0,T} + \|f_h - f\|_{0,T}) \\
&\leq c_9 (\|e\|_{1,T} + h_T \|f_h - f\|_{0,T}) \\
&\leq c_{10} (\|e\|_{1,T}^2 + h_T^2 \|f_h - f\|_{0,T}^2)^{1/2}.
\end{aligned} \tag{37}$$

Nun zu den Kanten. Wir nutzen aus, dass für P1 Finite Elemente

$$R_{E,h} := \begin{cases} g_h - \sigma(u_h) n_E & \text{auf Neumann-Kanten } E \in \mathcal{E}_N, \\ R_E & \text{sonst} \end{cases}$$

polynomiell ist. Man sieht mit Gleichung (34), dass

$$c_3^2 \|R_{E,h}\|_{0,E}^2 \leq \|\theta_E^{1/2} R_{E,h}\|_{0,E}^2 = \langle R_{E,h}, \theta_E R_{E,h} \rangle_{0,E}.$$

Weiter sehen wir

$$\begin{aligned}
\langle R_E, \theta_E R_{E,h} \rangle_{0,E} & \stackrel{\theta_E = 0 \text{ auf } \bigcup \mathcal{E} \setminus E}{=} \sum_{E \in \mathcal{E}} \langle R_E, \theta_E \mathfrak{E} R_{E,h} \rangle_{0,E} \\
& \stackrel{\text{Proposition 19}}{=} a(e, \theta_E \mathfrak{E} R_{E,h}) - \langle R_{\mathcal{T}}, \theta_E \mathfrak{E} R_{E,h} \rangle_{0,\Omega} \\
& \stackrel{\text{Stetigkeit von } a, \text{ supp } \theta_E \subseteq \omega_E, \text{ Cauchy-Schwarz-Ungleichung}}{\leq} c_A |e|_{1,\omega_E} |\theta_E \mathfrak{E} R_{E,h}|_{1,\omega_E} + \|R_{\mathcal{T}}\|_{0,\omega_E} \|\theta_E \mathfrak{E} R_{E,h}\|_{0,\omega_E} \\
& \stackrel{\text{Gleichungen (35) und (36)}}{\leq} (c_A c_4 h_E^{-1/2} |e|_{1,\omega_E} + c_5 h_E^{1/2} \|R_{\mathcal{T}}\|_{0,\omega_E}) \|R_{E,h}\|_{0,E} \\
& \leq c_{11} h_E^{-1/2} \sum_{T \in \omega_E} (\|e\|_{1,T} + h_T \|R_T\|_{0,T}) \|R_{E,h}\|_{0,E} \\
& \leq c_{12} h_E^{-1/2} \left( \sum_{T \in \omega_E} (\|e\|_{1,T}^2 + h_T^2 \|R_T\|_{0,T}^2) \right)^{1/2} \|R_{E,h}\|_{0,E} \\
& \stackrel{\text{Gleichung (37)}}{\leq} c_{13} h_E^{-1/2} \left( \sum_{T \in \omega_E} (\|e\|_{1,T}^2 + h_T^2 \|f - f_h\|_{0,T}^2) \right)^{1/2} \|R_{E,h}\|_{0,E},
\end{aligned}$$

wobei wir die Notation  $R_{\mathcal{T}} := \sum_{T \in \mathcal{T}} R_T$  verwendet haben. Falls  $E$  keine Neumann-Kante ist, erhalten wir

$$\begin{aligned}
c_3^2 \|R_{E,h}\|_{0,E}^2 & \leq \langle R_{E,h}, \theta_E R_{E,h} \rangle_{0,E} \\
& = \langle R_E, \theta_E R_{E,h} \rangle_{0,E} \\
& \leq c_{13} h_E^{-1/2} \left( \sum_{T \in \omega_E} (\|e\|_{1,T}^2 + h_T^2 \|f - f_h\|_{0,T}^2) \right)^{1/2} \|R_{E,h}\|_{0,E}
\end{aligned}$$

und nach Multiplikation mit  $h_E^{1/2}/(c_3^2 \|R_{E,h}\|_{0,E})$

$$h_E^{1/2} \|R_{E,h}\|_{0,E} \leq c_{14} \left( \sum_{T \in \omega_E} (\|e\|_{1,T}^2 + h_T^2 \|f - f_h\|_{0,T}^2) \right)^{1/2}. \quad (38)$$

Falls dagegen  $E$  eine Neumann-Kante ist, erhalten wir

$$\begin{aligned}
c_3^2 \|R_{E,h}\|_{0,E}^2 &\leq \langle R_{E,h}, \theta_E R_{E,h} \rangle_{0,E} \\
&\stackrel{\curvearrowright}{=} \langle R_E, \theta_E R_{E,h} \rangle_{0,E} + \langle g_h - g, \theta_E R_{E,h} \rangle_{0,E} \\
&\stackrel{\curvearrowright}{\leq} c_{13} h_E^{-1/2} \left( \sum_{T \in \omega_E} (\|e\|_{1,T}^2 + h_T^2 \|f - f_h\|_{0,T}^2) \right)^{1/2} \|R_{E,h}\|_{0,E} \\
&\quad + \|g_h - g\|_{0,E} \|R_{E,h}\|_{0,E} \\
&\stackrel{\curvearrowright}{\leq} c_{15} h_E^{-1/2} \left( \sum_{T \in \omega_E} (\|e\|_{1,T}^2 + h_T^2 \|f - f_h\|_{0,T}^2) + h_E \|g_h - g\|_{0,E}^2 \right)^{1/2} \|R_{E,h}\|_{0,E}
\end{aligned}$$

und nach Multiplikation mit  $h_E^{1/2}/(c_3^2 \|R_{E,h}\|_{0,E})$

$$h_E^{1/2} \|R_{E,h}\|_{0,E} \leq c_{16} \left( \sum_{T \in \omega_E} (\|e\|_{1,T}^2 + h_T^2 \|f - f_h\|_{0,T}^2) + h_E \|g_h - g\|_{0,E}^2 \right)^{1/2}. \quad (39)$$

Die Behauptung folgt dann mit Gleichungen (37), (38) und (39) aus

$$\begin{aligned}
\eta_{R,T}^2 &= h_T^2 \|R_T\|_{0,T}^2 + \frac{1}{2} \sum_{E \in \mathcal{E}_\Omega \cap \partial T} h_E \|R_E\|_{0,E}^2 + \sum_{E \in \mathcal{E}_T \cap \partial T} h_E \|R_E\|_{0,E}^2 \\
&\leq c_{10}^2 (\|e\|_{1,T}^2 + h_T^2 \|f_h - f\|_{0,T}^2) + c_{16}^2 \sum_{E \in \mathcal{E}_N \cap \partial T} h_E \|g_h - g\|_{0,E}^2 \\
&\quad + \max \{c_{14}^2/2, c_{14}^2, c_{16}^2\} \sum_{\substack{T \in \omega_E \\ E \in \mathcal{E} \cap \partial T}} (\|e\|_{1,T}^2 + h_T \|f_h - f\|_{0,T}^2) \\
&\leq c_{17} \left( \sum_{T \in \omega_T} \|e\|_{1,T}^2 + \sum_{T \in \omega_T} h_T^2 \|f_h - f\|_{0,T}^2 + \sum_{E \in \mathcal{E}_N \cap \partial T} h_E \|g_h - g\|_{0,E}^2 \right).
\end{aligned}$$

□

Die Terme

$$\sum_{T \in \omega_T} h_T^2 \|f_h - f\|_{0,T}^2 \quad \text{und} \quad \sum_{E \in \mathcal{E}_N \cap \partial T} h_E \|g_h - g\|_{0,E}^2$$

werden auch als oszillatorische Anteile bezeichnet und hängen von der Regularität der Anfangsdaten  $f$  und  $g$  ab. Da im Beweis lediglich die Eigenschaft von P1 Finiten Elementen verwendet wurde, dass  $R_E$  für alle  $E \in \mathcal{E} \setminus \mathcal{E}_N$  polynomiell ist, folgt die Aussage auch für andere Finite Elemente Methoden mit dieser Eigenschaft. Schließlich bemerke man noch, dass die Konstante  $c$  von der Stetigkeitskonstante  $c_A$  abhängt, die Abschätzung also insbesondere auch von  $\lambda$  und  $\mu$  abhängt.

## 7 Implementierung der P1 Finiten Elemente

In diesem Kapitel möchten wir das diskretisierte Problem (27) auf eine implementierbare Form bringen. Dazu diskretisieren wir zunächst  $\ell$ , diskretisieren dann die Nebenbedingungen und schreiben das Problem als lineares Gleichungssystem. Es wird Existenz und Eindeutigkeit dieses Gleichungssystems gezeigt und anschließend wird die Konstruktion der Steifheitsmatrix besprochen. Im gesamten Abschnitt folgen wir in großen Teilen [1].

Da wir im Abschnitt über A priori Fehlerschätzer gesehen haben, dass  $u_h$  gegen  $u$  linear konvergiert, können wir nach [1, S.249f.] unsere rechte Seite  $\ell(v_h)$  diskretisieren, indem wir die Mittelpunktsregel auf jedes Dreieck und jede Kante separat anwenden. Dies verschlechtert nicht das Konvergenzverhalten. Man erhält also

$$\begin{aligned}\ell(v_h) &= \langle f, v_h \rangle_{0,\Omega} + \langle g, v_h \rangle_{0,\Gamma_N} \\ &\approx \sum_{T \in \mathcal{T}} |T| f(x_T) v_h(x_T) + \sum_{E \in \mathcal{E}_N} |E| g(x_E) v_h(x_E) \\ &=: \ell_H(v_h)\end{aligned}$$

mit Dreiecksmittelpunkten  $x_T$  und Kantenmittelpunkten  $x_E$ . Unser Problem lautet dann: Finde  $u_H \in V_h$ , so dass

$$\begin{aligned}u_H &\text{ minimiert} & W_H(v_H) &= \frac{1}{2} a(v_H, v_H) - \ell_H(v_H) \\ \text{unter der Nebenbedingung} & & M u_H|_{\Gamma} &= w|_{\Gamma}.\end{aligned}\tag{40}$$

Seien nun  $u_H, v_H \in V_h$ , dann gibt es Koeffizienten  $\hat{u}, \hat{v} \in \mathbb{R}^{nd} =: \hat{V}$ , so dass

$$u_H = \sum_i \hat{u}_i \phi_i \quad \text{und} \quad v_H = \sum_j \hat{v}_j \phi_j.$$

Wir schreiben  $a$  als

$$\begin{aligned}a(u_H, v_H) &= a\left(\sum_i \hat{u}_i \phi_i, \sum_j \hat{v}_j \phi_j\right) \\ &= \sum_{i,j} a(\phi_i, \phi_j) \hat{u}_i \hat{v}_j \\ &= \sum_{i,j} A_{ij} \hat{u}_i \hat{v}_j \\ &= \hat{u}^\top A \hat{v}\end{aligned}\tag{41}$$

mit der Steifheitsmatrix

$$A_{ij} := a(\phi_i, \phi_j).$$

Nun können wir  $\ell_H$  schreiben als

$$\ell_H(v_H) = \ell_H\left(\sum_j \hat{v}_j \phi_j\right) = \sum_j \ell_H(\phi_j) \hat{v}_j = \hat{\ell}_j^\top \hat{v}\tag{42}$$

mit Load-Vektor

$$\hat{\ell}_j := \ell_H(\phi_j),$$

wobei wir für die Konkrete Berechnung bemerken, dass  $\phi_j(x_T) = 1/3$  und  $\phi(x_E) = 1/2$ .

## 7.1 Diskretisierung der Dirichletbedingungen

Bezeichnen nun  $x^{i_j} \in \mathcal{N}_\Gamma$ ,  $j \in \{1, \dots, l\}$ , die Randknoten. Dann lautet die Randbedingung

$$M(x^{i_j})u(x^{i_j}) = w(x^{i_j}) \quad \text{für alle } x^{i_j} \in \mathcal{N}_\Gamma.$$

Wir definieren  $M^{i_j} := M(x^{i_j}) \in \mathbb{R}^{d \times d}$  und  $\hat{w}^{i_j} := w(x^{i_j}) \in \mathbb{R}^d$ , sowie  $\hat{u}^i := u_h(x^i) \in \mathbb{R}^d$ . Damit ergibt sich die Randbedingung zu

$$M^{i_j} \hat{u}^{i_j} = \hat{w}^{i_j} \quad \text{für alle } j \in \{1, \dots, l\}.$$

Dieses Gleichungssystem können wir in Matrixform schreiben. Dazu konstruieren wir nach [1, S.250f.] eine Matrix  $B \in \mathbb{R}^{ld \times nd}$  durch

$$B := \begin{bmatrix} * & & 0 \\ & M^{i_j} & \leftarrow \\ 0 & \uparrow & * \end{bmatrix} \begin{array}{l} \\ j\text{-ter Zeilenblock} \\ i_j\text{-ter Spaltenblock} \end{array}$$

und sehen, dass

$$B\hat{u} = \begin{bmatrix} * & & 0 \\ & M^{i_j} & \\ 0 & & * \end{bmatrix} \begin{bmatrix} \hat{u}^1 \\ \vdots \\ \hat{u}^j \\ \vdots \\ \hat{u}^n \end{bmatrix} = \begin{bmatrix} M^{i_1} \hat{u}^{i_1} \\ \vdots \\ M^{i_l} \hat{u}^{i_l} \end{bmatrix} = \begin{bmatrix} \hat{w}^{i_1} \\ \vdots \\ \hat{w}^{i_l} \end{bmatrix} =: \hat{w}.$$

Unsere Randbedingungen lassen sich also schreiben als

$$\begin{bmatrix} B_1 \\ \vdots \\ B_{ld} \end{bmatrix} \hat{u} := B\hat{u} = \begin{bmatrix} \hat{w}_1 \\ \vdots \\ \hat{w}_{ld} \end{bmatrix},$$

Seien  $B_{i_j}$  für  $j = 1, \dots, k$  genau die Zeilen von  $B$ , die nicht null sind. Wir streichen alle Nullzeilen aus  $B$  und setzen

$$B' := \begin{bmatrix} B_{i_1} \\ \vdots \\ B_{i_k} \end{bmatrix}.$$



Es sollte für alle Nullzeilen  $B_i$  von  $B$  auch  $\hat{w}_i = 0$  gelten, denn sonst kann es keine Lösung des Problems geben. Damit kann man die Randbedingungen schreiben als

$$B' \hat{u} = \begin{bmatrix} B_{i_1} \\ \vdots \\ B_{i_k} \end{bmatrix} \hat{u} = \begin{bmatrix} \hat{w}_{i_1} \\ \vdots \\ \hat{w}_{i_k} \end{bmatrix} =: \hat{w}'. \quad (43)$$

Mit Gleichungen (41), (42) und (43), können wir das Problem (40) jetzt schreiben als: Finde  $\hat{u} \in \hat{V}$ , so dass

$$\begin{aligned} \hat{u} \text{ minimiert} \quad & \hat{W}(\hat{v}) = \frac{1}{2} \hat{v}^\top \hat{A} \hat{v} - \hat{\ell}^\top \hat{v} \\ \text{unter der Nebenbedingung} \quad & B' \hat{u} = \hat{w}'. \end{aligned}$$

## 7.2 Existenz und Eindeutigkeit des diskreten Problems

Falls  $a$  symmetrisch und elliptisch ist, so ist  $A$  nach Konstruktion symmetrisch positiv definit. Es sei nun an ein Resultat aus der quadratischen Programmierung erinnert.

**Proposition 24** (Quadratische Programme mit Gleichheitsrestriktionen). *Die Steifheitsmatrix  $A$  sei symmetrisch positiv definit. Dann gilt*

1.  $\hat{u} \in \hat{V}$  ist genau dann Lösung von

$$\begin{aligned} \hat{u} \text{ minimiert} \quad & \hat{W}(\hat{v}) = \frac{1}{2} \hat{v}^\top \hat{A} \hat{v} - \hat{\ell}^\top \hat{v} \\ \text{unter der Nebenbedingung} \quad & B' \hat{u} = \hat{w}', \end{aligned} \quad (44)$$

wenn das Paar  $(\hat{u}, \hat{p})$  KKT-Punkt von (44) ist.

2. Das Paar  $(\hat{u}, \hat{p})$  ist genau dann KKT-Punkt von (44), wenn es das Gleichungssystem

$$\begin{bmatrix} A & B'^\top \\ B' & 0 \end{bmatrix} \begin{bmatrix} \hat{u} \\ \hat{p} \end{bmatrix} = \begin{bmatrix} \hat{b} \\ \hat{w}' \end{bmatrix} \quad (45)$$

löst.

3. Sind die Zeilen von  $B'$  linear unabhängig (d.h.  $B'$  ist surjektiv), so besitzt das Gleichungssystem (45) genau eine Lösung.

*Beweis.* 1. Siehe Korollar 2.47 in [11, S.59]

2. Siehe Satz 5.1 in [11, S.198]

3. Siehe Kapitel 2, Satz 19 in [9]

□

Im Programm `ECFEM.m` wird das System (45) aufgestellt und gelöst. Nach dem Resultat oben hat dieses Gleichungssystem eine eindeutige Lösung, sofern  $a$  symmetrisch und elliptisch ist. Die Matrix  $B'$  und der Load-Vektor  $\hat{\ell}$  werden wie oben angegeben aufgestellt. Der 3. Punkt ist die Begründung dafür, dass wir in der Konstruktion von  $B'$  alle Nullzeilen gestrichen haben. Die Steifheitsmatrix  $A$  aufzustellen ist ein wenig aufwendiger und wird im folgenden Abschnitt beschrieben.

### 7.3 Berechnung der Steifheitsmatrix

Das Aufstellen der Steifheitsmatrix macht einen signifikanten Teil der Funktion `ECFEM.m` in der Implementation aus. Hierbei folgen wir [1, S.244-247]. Wir definieren die Abbildung  $\pi: \mathcal{T} \times \{1, \dots, d(d+1)\} \rightarrow \{1, \dots, nd\}$  dadurch, dass  $\pi(T, \cdot)$  die geordneten Indizes aller  $\phi_i$  gibt, die auf  $T$  nicht verschwinden. Genauer, seien  $x^{i_1}, \dots, x^{i_{d+1}}$  die Knoten von  $T \in \mathcal{T}$ , so bezeichnen wir

$$\begin{bmatrix} \phi_{\pi(T,1)} \\ \vdots \\ \phi_{\pi(T,d)} \\ \vdots \\ \phi_{\pi(T,d^2+1)} \\ \vdots \\ \phi_{\pi(T,d(d+1))} \end{bmatrix} = \begin{bmatrix} \varphi_{i_1} e_1 \\ \vdots \\ \varphi_{i_1} e_d \\ \vdots \\ \varphi_{i_{d+1}} e_1 \\ \vdots \\ \varphi_{i_{d+1}} e_d \end{bmatrix}.$$

Für  $k \in \{1, 2\}$  und  $j \in \{1, 2, 3\}$  erhält man mit der in den Gleichungen (9) und (10) eingeführten Voigt-Notation für den Fall  $d = 2$

$$\gamma(\varphi_{i_j} e_k) = \begin{bmatrix} \partial_1 \varphi_{i_j} \delta_{k1} \\ \partial_2 \varphi_{i_j} \delta_{k2} \\ \partial_2 \varphi_{i_j} \delta_{k1} + \partial_1 \varphi_{i_j} \delta_{k2} \end{bmatrix},$$

womit dann folgt

$$\begin{aligned} \gamma(\phi_{\pi(T,j)}) \Big|_T &= \begin{bmatrix} \gamma(\varphi_{i_1} e_1) & \gamma(\varphi_{i_1} e_2) & \cdots & \gamma(\varphi_{i_3} e_1) & \gamma(\varphi_{i_3} e_2) \end{bmatrix} e_j \\ &= \underbrace{\begin{bmatrix} \partial_1 \varphi_{i_1} & \partial_2 \varphi_{i_1} & \partial_1 \varphi_{i_2} & \partial_2 \varphi_{i_2} & \partial_1 \varphi_{i_3} & \partial_2 \varphi_{i_3} \\ \partial_2 \varphi_{i_1} & \partial_1 \varphi_{i_1} & \partial_2 \varphi_{i_2} & \partial_1 \varphi_{i_2} & \partial_2 \varphi_{i_3} & \partial_1 \varphi_{i_3} \end{bmatrix}}_{=:R} e_j. \end{aligned}$$

Analog folgt für  $k \in \{1, 2, 3\}$  und  $j \in \{1, 2, 3, 4\}$  im Fall  $d = 3$  aus

$$\gamma(\varphi_{i_j} e_k) = \begin{bmatrix} \partial_1 \varphi_{i_j} \delta_{k1} \\ \partial_2 \varphi_{i_j} \delta_{k2} \\ \partial_3 \varphi_{i_j} \delta_{k3} \\ \partial_2 \varphi_{i_j} \delta_{k1} + \partial_1 \varphi_{i_j} \delta_{k2} \\ \partial_3 \varphi_{i_j} \delta_{k1} + \partial_1 \varphi_{i_j} \delta_{k3} \\ \partial_3 \varphi_{i_j} \delta_{k2} + \partial_2 \varphi_{i_j} \delta_{k3} \end{bmatrix},$$

dass

$$\begin{aligned} \gamma(\phi_{\pi(T,j)}) \Big|_T &= \begin{bmatrix} \gamma(\varphi_{i_1} e_1) & \cdots & \gamma(\varphi_{i_1} e_3) & \cdots & \gamma(\varphi_{i_4} e_1) & \cdots & \gamma(\varphi_{i_4} e_3) \end{bmatrix} e_j \\ &= \underbrace{\begin{bmatrix} \partial_1 \varphi_{i_1} & & & & \partial_1 \varphi_{i_4} & & \\ & \partial_2 \varphi_{i_1} & & \cdots & & \partial_2 \varphi_{i_4} & \\ & & \partial_3 \varphi_{i_1} & & & & \partial_3 \varphi_{i_4} \\ \partial_2 \varphi_{i_1} & \partial_1 \varphi_{i_1} & & & \partial_2 \varphi_{i_4} & \partial_1 \varphi_{i_4} & \\ \partial_3 \varphi_{i_1} & & \partial_1 \varphi_{i_1} & \cdots & \partial_3 \varphi_{i_4} & & \partial_1 \varphi_{i_4} \\ & \partial_3 \varphi_{i_1} & \partial_2 \varphi_{i_1} & & \partial_3 \varphi_{i_4} & \partial_2 \varphi_{i_4} & \end{bmatrix}}_{=:R} e_j. \end{aligned}$$

Damit folgt

$$\begin{aligned} (\tilde{A}(T))_{kl} &:= \int_T \sigma(\phi_{\pi(T,k)}) : \varepsilon(\phi_{\pi(T,l)}) \, dx \\ &= \int_T \tau(\phi_{\pi(T,k)})^\top \gamma(\phi_{\pi(T,l)}) \, dx \\ &= \int_T \gamma(\phi_{\pi(T,k)})^\top \hat{C} \gamma(\phi_{\pi(T,l)}) \, dx \\ &= \int_T (Re_k)^\top \hat{C} Re_l \, dx \\ &= |T| (R^\top \hat{C} R)_{kl}. \end{aligned}$$

Man erhält schließlich

$$\begin{aligned} A_{ij} &= a(\phi_i, \phi_j) \\ &= \int_\Omega \sigma(\phi_i) : \varepsilon(\phi_j) \, dx \\ &= \sum_{T \in \mathcal{T}} \int_T \sigma(\phi_i) : \varepsilon(\phi_j) \, dx \\ &= \sum_{\substack{T \in \mathcal{T} \\ k, \text{ s.d. } \pi(T,k)=i \\ l, \text{ s.d. } \pi(T,l)=j}} (\tilde{A}(T))_{kl} \end{aligned}$$

als Ausdruck für die Steifheitsmatrix.

## 8 Numerische Experimente

Wir möchten nun die Implementation und die Resultate der vorhergehenden Abschnitte anhand von zwei Benchmarks aus [4, Abschnitt 3] und [1, S.255f.] überprüfen. Im ersten Benchmark wird das Phänomen des Lockings untersucht. Dazu wird zunächst kurz erläutert, wann Locking auftritt. Anschließend werden das Benchmark beschrieben und die Ergebnisse interpretiert. Im zweiten Benchmark wird adaptive Verfeinerung mit dem Mittelungsfehlerschätzer und dem residualem Fehlerschätzer mit der uniformen Verfeinerung verglichen. Dazu ist es notwendig, die Konstruktion des Mittelungsfehlerschätzers nach [1, S.253f.] zu erläutern. Anschließend werden das Benchmark beschrieben und die Ergebnisse interpretiert. Zum Schluss werden die Ergebnisse der Experimente in einem Fazit zusammengefasst.

Wir werden für den Fehler in den numerischen Experimenten die Norm  $\|\sigma(\cdot)\|_{0,\Omega}$  statt  $\|\cdot\|_{1,\Omega}$  verwenden. Nach der folgenden Proposition gilt für diese Norm:

**Proposition 25** (Normäquivalenz). *Seien  $a$  und  $C$  elliptisch und erfülle  $a$  (18). Dann sind die Normen  $|\cdot|_{1,\Omega}$ ,  $\|\cdot\|_{1,\Omega}$ ,  $\|\varepsilon(\cdot)\|_{0,\Omega}$ ,  $a(\cdot, \cdot)$  und  $\|\sigma(\cdot)\|_{0,\Omega}$  äquivalent.*

*Beweis.* Es folgt mit Gleichung (11) und Cauchy-Schwarz für  $v \in V$

$$\langle \sigma(v), \varepsilon(v) \rangle_{0,\Omega} \leq \|\sigma(v)\|_{0,\Omega} \langle \varepsilon(v), \varepsilon(v) \rangle_{0,\Omega}^{1/2} \leq \frac{1}{c_c} \|\sigma(v)\|_{0,\Omega} \langle \sigma(v), \varepsilon(v) \rangle_{0,\Omega}^{1/2}$$

und nach Division durch  $\langle \sigma(v), \varepsilon(v) \rangle_{0,\Omega}^{1/2}$

$$a(v, v) = \langle \sigma(v), \varepsilon(v) \rangle_{0,\Omega} \leq \frac{1}{c_c^2} \|\sigma(v)\|_{0,\Omega}^2$$

und damit weiter mit Gleichungen (18), (17), und der Elliptizität von  $a$

$$\begin{aligned} \frac{1}{c_A^2} \|\sigma(v)\|_{0,\Omega}^2 &\leq \|\varepsilon(v)\|_{0,\Omega}^2 \\ &\leq |v|_{1,\Omega}^2 \\ &\leq \|v\|_{1,\Omega}^2 \\ &\leq \frac{1}{c_a} a(v, v) \\ &\leq \frac{1}{c_a c_c^2} \|\sigma(v)\|_{0,\Omega}^2, \end{aligned}$$

womit die Behauptung folgt □

### 8.1 Phänomen des Lockings

Das erste Experiment handelt vom Phänomen des Lockings, welches im Folgenden erörtert werden soll. Man definiert die Poissonzahl durch

$$\nu := \frac{\lambda}{2(\lambda + \mu)}.$$

Es gilt nach [14, S.14] in der Praxis  $\nu \in [0, \frac{1}{2}]$ . Falls  $\nu \rightarrow 1/2$  und  $\mu$  fest ist, haben wir  $\lambda \rightarrow \infty$ . Aber dann folgt mit der Wahl der Konstanten der Propositionen 1, 4 und 14

$$\frac{c_A}{c_a} = \frac{c_{K2} d^2 \|C\|_\infty}{c_c} = \frac{c_{K2} d^2 (\lambda + 2\mu)}{2\mu} \rightarrow \infty.$$

Die a priori Abschätzung unseres Fehlers in Satz 17 wird also beliebig schlecht. Die Lösung mit dem P1 Finitem Element zeigt dann bis zu sehr kleinen  $h$  suboptimale Konvergenzeigenschaften. Dieses Phänomen wird als Locking bezeichnet und tritt beim Grenzübergang zu inkompressiblen Materialien auf. Nach der Tabelle auf Seite 8 ist beispielsweise Gummi mit  $\nu \approx 4.85$  fast inkompressibel. Das Locking wird im nachfolgenden Benchmark genauer untersucht.

## 8.2 Benchmark: Quadratisches Gebiet

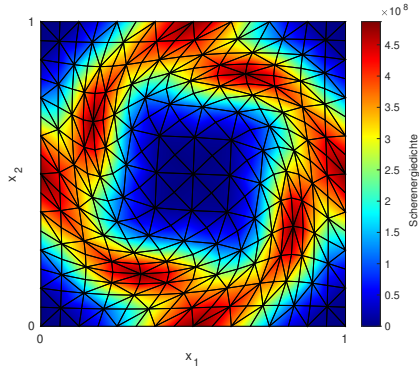


Abb. 17: Mögliche Deformation

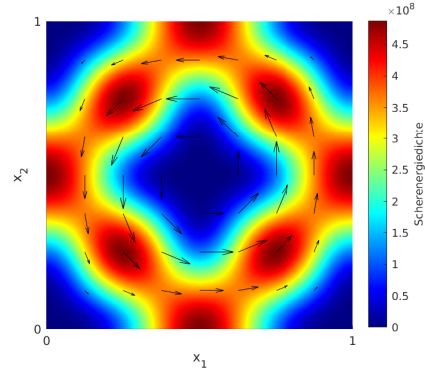


Abb. 18: Lösung

Das folgende Benchmark stammt aus [4, Abschnitt 3.1]. Wir haben das Gebiet  $\Omega = [0, 1]^2 \subseteq \mathbb{R}^2$  mit reinem Dirichlet-Rand  $\Gamma_D = \partial\Omega$ , einen festen Parameter  $\mu = 10^7$  und Funktionen

$$u(x) = \pi \begin{bmatrix} \cos(\pi x_2) \sin^2(\pi x_1) \sin(\pi x_2) \\ -\cos(\pi x_1) \sin(\pi x_1) \sin^2(\pi x_2) \end{bmatrix}$$

und

$$f(x) = 2\mu\pi^3 \begin{bmatrix} -\cos(\pi x_2) \sin(\pi x_2) (2 \cos(2\pi x_1) - 1) \\ \cos(\pi x_1) \sin(\pi x_1) (2 \cos(2\pi x_2) - 1) \end{bmatrix}.$$

Die Anfangsbedingungen sind in Abbildung 19 dargestellt. Die exakte Lösung  $u$  und die Scherenergiedichte für  $\nu = 0.2$  ist in den Plots 17 und 18 zu sehen. Ziel des Experiments ist es, das Lockings nachzuweisen und den residualen Fehlerschätzer zu prüfen. Dazu lassen wir 3 Versuchesreihen mit unterschiedlichen Poissonzahlen  $\nu \in \{0.2, 0.49, 0.499\}$  laufen, in der in jeder Iteration das Gitter uniform verfeinert wird.

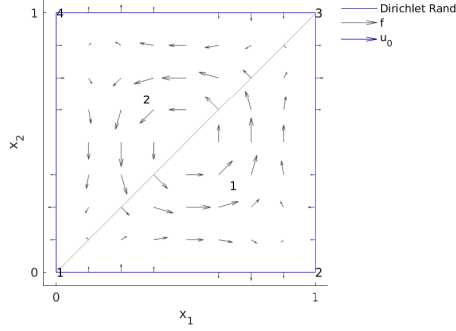


Abb. 19: Anfangsbedingungen

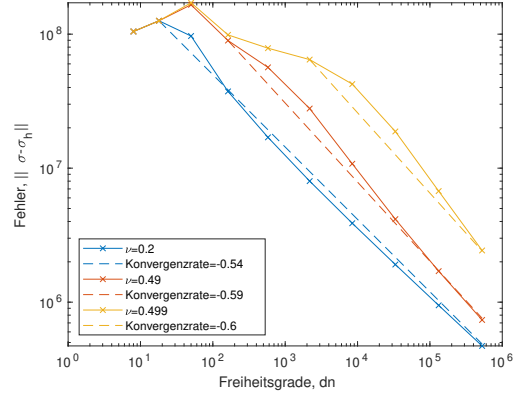


Abb. 20: Fehler

In Abbildung 20 sieht man eine Näherung des Fehlers  $\|\sigma(e)\|_{0,\Omega}$  in Abhängigkeit der Freiheitsgrade  $dn$  geplottet. Dadurch, dass die Mittelpunktsregel bei der Auswertung der Integrale in  $\|\sigma(e)\|_{0,\Omega}$  verwendet wurde und  $\ell$  durch  $\ell_H$  approximiert wurde, sind die Daten erst für feinere Gitter aussagekräftig. Allerdings sieht man hier dennoch, dass die Konvergenzrate suboptimal ist für  $\nu$  nahe an  $1/2$ , also, dass Locking auftritt. Man sieht auch, dass die in der Satz 17 vorhergesagten linearen Konvergenzrate annähernd erreicht wird. Das Locking macht sich auch in der Energie  $a(u, u)/2$  in Abbildung 21 und in einer Näherung der Kondition durch **condest** in Abbildung 22 bemerkbar. Man sieht an der Energie, dass die Lösungen auf dem groben Gitter näher bei der Nulllösung sind, wenn  $\nu$  näher am kritischen Wert  $1/2$  ist. Man beachte dabei, dass die ersten beiden berechneten Lösungen in jedem Fall aufgrund der Wahl des Anfangsgitters in Abbildung 19 Nulllösungen sind. In der Tat wird beim ersten Verfeinern nur der Mittelpunkt des Quadrats hinzugefügt, wo aufgrund der Symmetrie der Volumenkräfte  $f$  um diesen Punkt die Verschiebung Null ist.

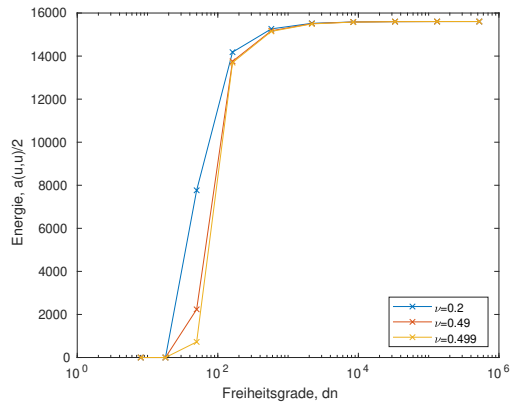


Abb. 21: Energie

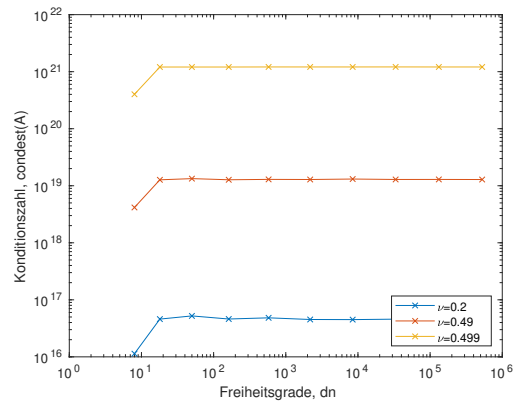


Abb. 22: Kondition von  $A$

In Abbildung 23 sieht man, dass die in Satz 23 vorausgesagte lokale Effizienz und die in Satz 21 vorhergesagte Zuverlässigkeit erfüllt sind. Der Effizienzindex  $\eta_R/\|\sigma(e)\|_{0,\Omega}$  ist nach oben und nach unten beschränkt und man sieht die Abhängigkeit dieser Schranken vom Materialparameter. Auch hier sind die ersten Messwerte wegen Näherungen der Integrale des Fehlerschätzers und des Fehlers durch die Mittelpunktsregel nicht aussagekräftig.

Ein völlig anderes Verhalten zeigt dagegen der nachfolgend vorgestellte Fehlerschätzer durch Mittelung in Abbildung 24. Man beachte das Spektrum an Werten, die der Quotient  $\eta_M/\|\sigma(e)\|_{0,\Omega}$  annimmt.

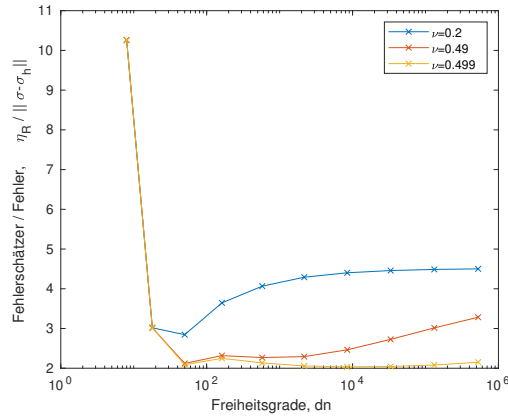


Abb. 23: Effizienz

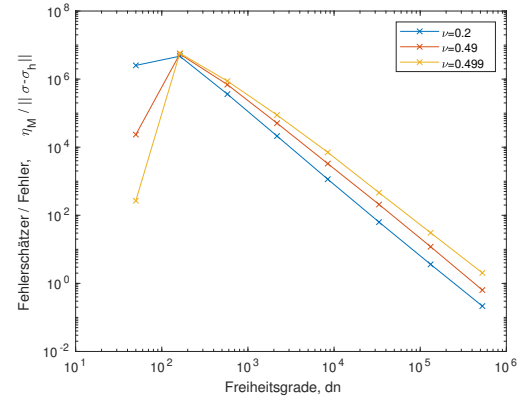


Abb. 24: Ineffizienz

### 8.3 Fehlerschätzer durch Mittelung

Wir vergleichen den residualen Fehlerschätzer mit einem Fehlerschätzer durch Mittelung, wie er in [1, S.253f.] beschrieben wird. Man definiert eine stetige Approximation  $\tilde{\sigma}_h$  an  $\sigma_h$ , indem man an den Knoten den Wert von  $\tilde{\sigma}_h$  auf das gewichtete Mittel von  $\sigma_h|_T$  der angrenzenden Dreiecke  $T \in \mathcal{T}$  setzt und diesen dann linear auf den Dreiecken interpoliert. Genauer: Für einen Knoten  $x^i \in \mathcal{N}$  definiert man die Umgebung  $\omega_{x^i}$  als die Menge aller Dreiecke  $T \in \mathcal{T}$ , so dass  $x^i$  Knoten von  $T$  ist. Man setzt

$$\tilde{\sigma}_h(x^i) = \frac{1}{|\omega_{x^i}|} \sum_{T \in \omega_{x^i}} |T| \cdot \sigma_h|_T$$

und interpoliert  $\tilde{\sigma}_h$  dann linear auf den Dreiecken. Dies liefert den Fehlerschätzer

$$\eta_{M,T} := \|\tilde{\sigma}_h - \sigma_h\|_{0,T}.$$

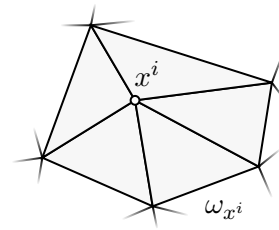


Abb. 25: Die Umgebung  $\omega_{x^i}$  von  $x^i$ .

## 8.4 Benchmark: L-förmiges Gebiet

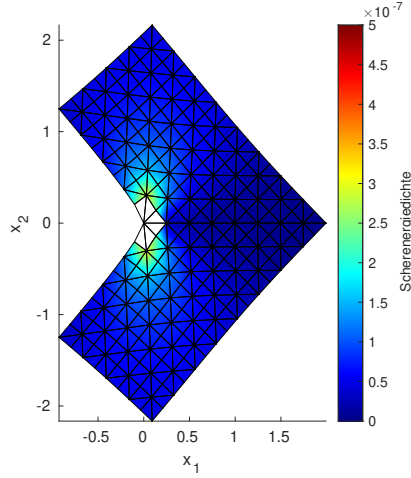


Abb. 26: Mögliche Deformation

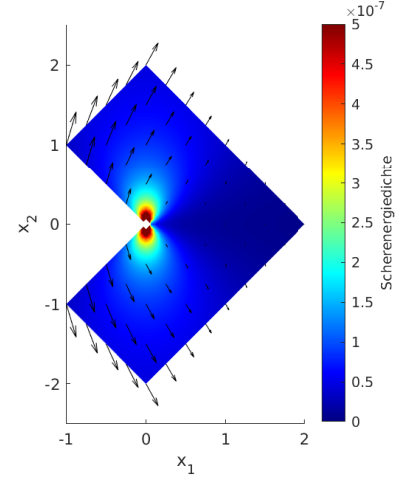


Abb. 27: Lösung

Das folgende Benchmark stammt aus [4, Abschnitt 3.4] und [1, S.255f.]. Wir verwenden ein L-förmiges Gebiet mit Dirichlet- und Neumann-Rand wie in Abbildung 28 dargestellt. Wir setzen weiter  $\mu = 10^7$  und  $\nu = 0.3$ . Die Verschiebung  $u$  ist in Polarkoordinaten gegeben durch

$$\begin{aligned} u_r(r, \varphi) &= \frac{r^\alpha}{2\mu} \left( -(\alpha + 1) \cos((\alpha + 1)\varphi) \right. \\ &\quad \left. + (C_2 - \alpha - 1)C_1 \cos((\alpha - 1)\varphi) \right) \\ u_\varphi(r, \varphi) &= \frac{r^\alpha}{2\mu} \left( (\alpha + 1) \sin((\alpha + 1)\varphi) \right. \\ &\quad \left. + (C_2 + \alpha - 1)C_1 \sin((\alpha - 1)\varphi) \right) \end{aligned}$$

mit Konstanten

$$C_1 := -\frac{\cos((\alpha + 1)\omega)}{\cos((\alpha - 1)\omega)}, \quad C_2 := 2\frac{\lambda + 2\mu}{\lambda + \mu}, \quad \alpha \approx 0.544483736782, \quad \omega := \frac{3\pi}{4}$$

nach [4, Abschnitt 3.4]. Weiter setzen wir  $f = 0$  und  $g = 0$ . Die exakte Lösung  $u$  sieht man in Abbildungen 26 und 27. In diesem Experiment wird die Art der Verfeinerung zwischen uniform und adaptiv mit Fehlerschätzern  $\eta_R$  und  $\eta_M$  variiert. In Abbildung 29 ist die Routine für die adaptive Gitterverfeinerung dargestellt. Das markieren in der Funktion `MARK.m` erfolgt nach der Methode von Dörfler in [7]. Hierbei werden für einen Parameter  $\zeta \in (0, 1)$  solange die Dreiecke  $T \in \mathcal{T}$  mit den größten Fehlerschätzungen  $\eta_{T,R}$  markiert, bis

$$\sum_{T \text{ markiert}} \eta_{R,T}^2 \geq \zeta \eta_R^2$$



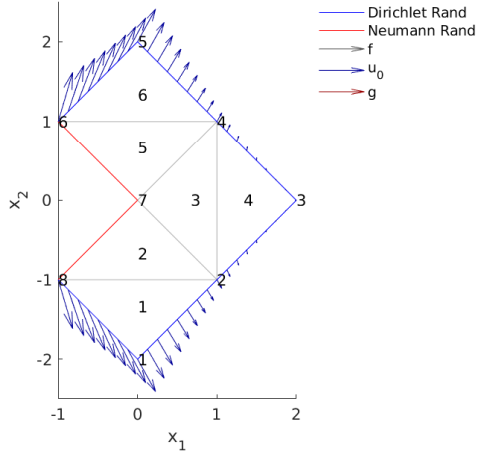


Abb. 28: Anfangsbedingungen

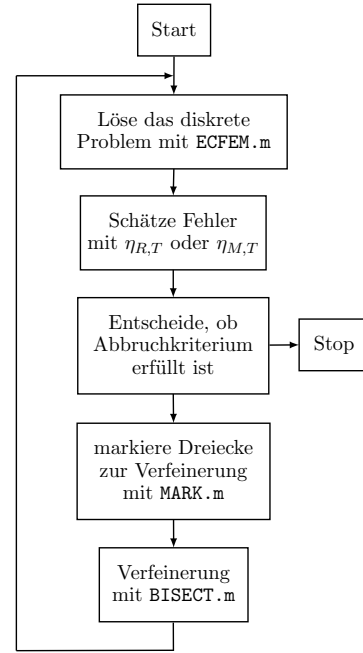


Abb. 29: Adaptiven Gitterverfeinerung

erfüllt ist. In diesem Experiment wurde  $\zeta = 0.7$  gesetzt. Für die Gitterverfeinerung wird in `BISECT.m` ein Bisektionsalgorithmus verwendet, der auf dem Zerteilen der längsten Kante basiert.

Da dieses Problem eine Singularität im Ursprung hat und wir  $u \in H^2(\Omega; \mathbb{R}^d)$  für die a priori Abschätzung in Satz 17 benötigen, ist lineare Konvergenz des Fehlers in  $h$  nicht mehr garantiert. In der Tat konvergiert die uniforme Verfeinerung nicht mehr linear, wie in Abbildung 30 zu erkennen ist. Die adaptive Verfeinerung mit dem residualen Fehlerschätzer liefert dagegen wieder die optimale Konvergenzrate für das Verfahren. Bei der adaptiven Verfeinerung mit dem Fehlerschätzer  $\eta_M$  wird die Steifheitsmatrix bei  $d \cdot n \approx 3000$  Freiheitsgraden singulär, was sich interessanterweise nicht in der Kondition in Abbildung 31 bemerkbar macht. Der Grund für dieses Verhalten liegt darin, dass der Mittelungsfehlerschätzer erfolgreich eine Singularität im Ursprung erkennt, diese aber zu stark gewichtet, so dass fast ausschließlich die Dreiecke am Ursprung verfeinert werden. Dieses Verhalten sieht man in den Abbildungen 34a-34d.

In Abbildung 33 sieht man, dass der Fehlerschätzer durch Mittelung auch bei diesem Problem nicht zuverlässig ist. Der residuale Fehlerschätzer  $\eta_R$  weist dagegen wie im ersten Benchmark in Abbildung 32 Effizienz und Zuverlässigkeit auf.

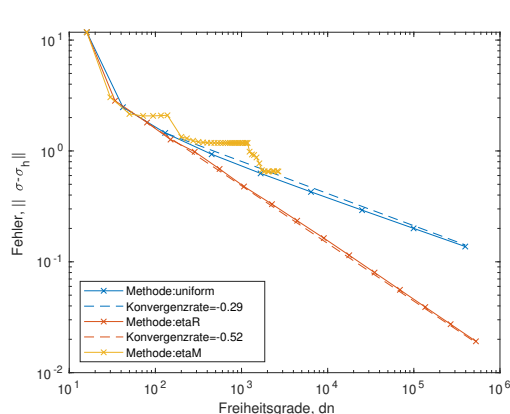


Abb. 30: Fehler

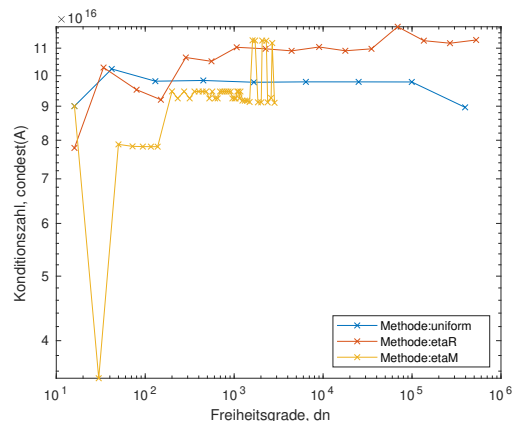


Abb. 31: Kondition

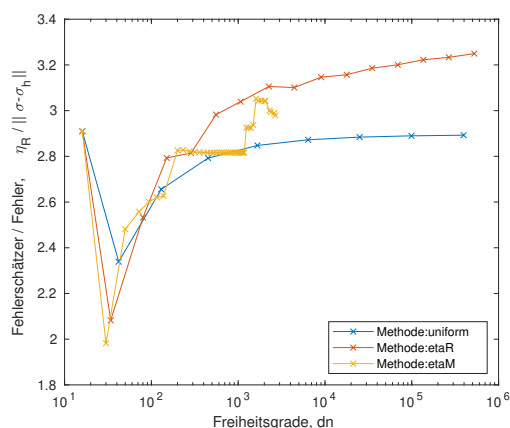


Abb. 32: Effizienz

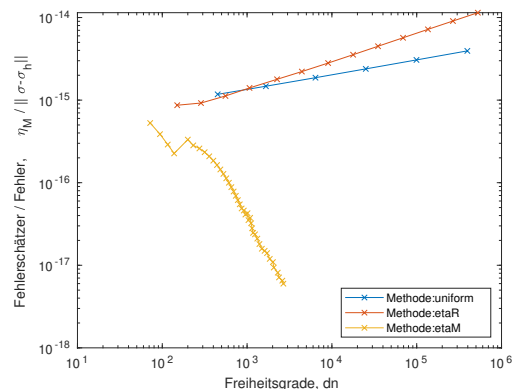
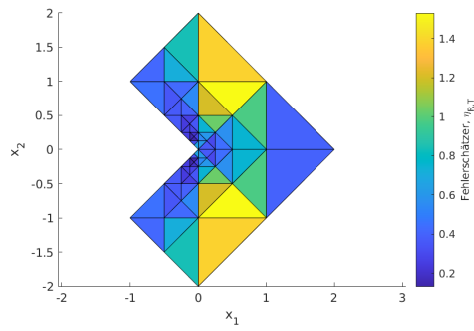


Abb. 33: Ineffizienz

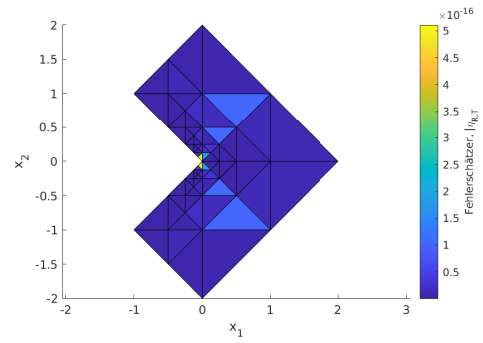
## 8.5 Fazit

Im ersten numerischen Experiment haben wir gesehen, dass die in Satz 17 a priori vorhergesagte lineare Konvergenz in der Praxis vom P1 Finiten Element erfüllt wird, sofern die Verschiebung  $u$  regulär genug ist. Allerdings tritt unter anderem bei ungünstigen Materialparametern mit dieser Methode Locking auf.

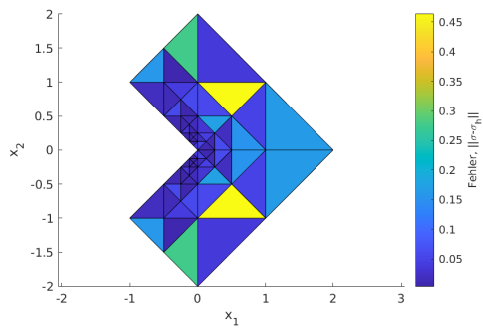
Die Konvergenz im zweiten Experiment mit uniformer Gitterverfeinerung ist aufgrund einer Singularität im Ursprung nicht mehr linear. Wir haben gesehen, dass adaptive Gitterverfeinerung mit dem Fehlerschätzer durch Mittelung hier keinen Mehrgewinn bringt, da dieser Fehlerschätzer nicht zuverlässig ist. Wir können dagegen mit der adaptiven Gitterverfeinerung mit dem residualen Fehlerschätzer die für das Verfahren optimale lineare Konvergenz zurückgewinnen. In beiden Experimenten wurde nochmal bestätigt, was im Abschnitt über den residualen Fehlerschätzer gezeigt wurde, nämlich, dass dieser zuverlässig und effizient ist.



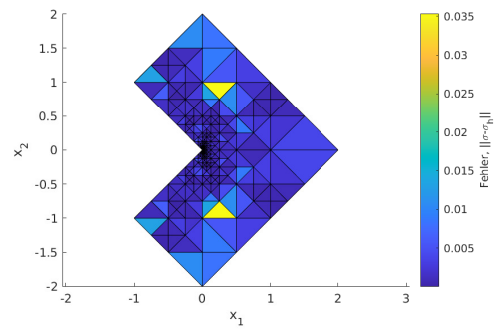
(a) lokaler Fehlerschätzer  $\eta_{R,T}$



(b) lokaler Fehlerschätzer  $\eta_{M,T}$

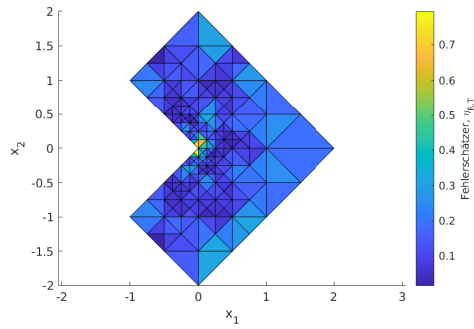


(c) lokaler Fehler  $\|\sigma - \sigma_h\|$

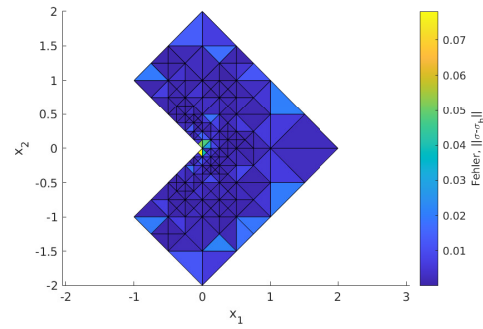


(d) lokaler Fehler  $\|\sigma - \sigma_h\|$

Abb. 34: Adaptive Gitterverfeinerung mit dem Mittelungfehlerschätzer  $\eta_M$  und  $d \cdot n = 94$  Freiheitsgraden, bzw.  $d \cdot n = 1610$  Freiheitsgraden in Abbildung 34d.



(a) lokaler Fehlerschätzer  $\eta_{R,T}$



(b) lokaler Fehler  $\|\sigma - \sigma_h\|$

Abb. 35: Adaptive Gitterverfeinerung mit dem residualen Fehlerschätzer  $\eta_R$  und  $d \cdot n = 286$  Freiheitsgraden.

## Symbolverzeichnis

$\Omega$	eine Menge im $\mathbb{R}^d$
$\Gamma$	Rand von $\Omega$
$\chi$	Deformation
$V, V^0$	Mengen von Verschiebungen $H^1(\Omega; \mathbb{R}^{d \times d})$ und $\{v \in V : Mv = 0\}$
$u, v$	Verschiebungen
$f, g$	Volumenkraftdichte und Oberflächenkraftdichte
$M$	Matrix für gleitende Randbedingungen, definiert auf Seite 5
$\text{Div}$	Divergenz auf $H^1(\Omega; \mathbb{R}^d)$ , definiert durch $\text{Div } \sigma := \sum_j \partial_j \sigma_{ij} e_i$
$\varepsilon, \sigma$	linearisierter Verzerrungstensor und Spannungstensor
$\gamma, \tau$	Voigtrepresentation, definiert in (9) und (10)
$\lambda, \mu, \nu$	Lamé-Parameter und Poissonzahl
$C$	Hooke-Tensor, definiert in (4)
$a$	Bilinearform, definiert in (12)
$\ell$	Funktional, definiert in (13)
$W$	Potenzielle Energie, definiert in (15)
$H^k(\Omega), H^k(\Omega; \mathbb{R}^d)$	Sobolevräume $W^{k,p}(\Omega)$ und $W^{k,p}(\Omega; \mathbb{R}^d)$ mit $p = 2$
$\langle \cdot, \cdot \rangle_{k,\Omega}$	inneres Produkt auf $H^k(\Omega), H^k(\Omega; \mathbb{R}^d)$ und $H^k(\Omega; \mathbb{R}^{d \times d})$
$\ \cdot\ _{k,\Omega}$	vom inneren Produkt induzierte Sobolevnorm
$ \cdot _{1,\Omega}$	Seminorm $\ \nabla \cdot\ _{0,\Omega}$
$\ \cdot\ _\infty$	Supremumsnorm auf Sobolevräumen
$\ \cdot\ _F$	Frobenius-Norm, induziert vom inneren Produkt „:“ auf $\mathbb{R}^{d \times d}$
$\ \cdot\ _K$	Kornsche Norm, definiert in (19)
$\text{epi}$	Epigraph einer Funktion
$\psi$	lokale Parametrisierung von $\Omega$
$A$	affine Transformation oder Steifheitsmatrix
$\theta$	Abschneidefunktionen oder Zerlegung der Eins
$\mathfrak{E}$	Erweiterungsoperator
$c_A, c_a$	Stetigkeits- und Elliptizitätskonstante aus Propositionen 4 und 14
$c_c$	Elliptizitätskonstante von $C$ , definiert in (11)
$c_E$	Stetigkeitskonstanten der Erweiterung $\mathfrak{E}$ aus Lemma 9
$c_{K1}, c_{K2}$	Konstanten der Kornschen Ungleichungen aus Satz 11 und 13
$c_{I1}, c_{I2}$	Konstanten der Interpolationsoperatoren in Propositionen 16 und 20
$\varphi, \phi$	nodale Basis in einer und $d$ Dimensionen
$V_h, V_h^0$	P1-Finite-Elemente-Räume, definiert in Gleichungen (25) und (26)
$u_h, v_h$	Verschiebungen des diskreten Problems
$\mathcal{T}$	zulässige Triangulierung
$\mathcal{E}, \mathcal{E}_\Gamma, \mathcal{E}_N, \mathcal{E}_\Omega$	Mengen von Kanten einer Triangulierung, definiert auf Seite 24
$\mathcal{N}, \mathcal{N}_\Gamma$	Mengen von Knoten einer Triangulierung, definiert auf Seite 24
$h_T, h_E$	Durchmesser $\text{diam}(T)$ und Kantenlänge $\text{diam}(E)$
$n$	Anzahl der Knoten von $\mathcal{T}$ oder äußere Normale

$e$	Fehler $u - u_h$ oder Einheitsvektor $e_i = \delta_{\cdot i}$
$R_T, R_E$	flächen- und kantenbezogene Residuen, definiert in (28) und (29)
$\eta_{R,T}, \eta_R$	residualer Fehlerschätzer, definiert in (30) und (31)
$\eta_{M,T}, \eta_M$	Fehlerschätzer durch Mittelung, definiert auf Seite 44
$\tilde{\omega}_T, \omega_T$	Umgebungen von $T$ , definiert auf Seite 30 und 31
$\omega_E$	Umgebung von $E$ , definiert auf Seite 31
$\omega_{x^i}$	Umgebung von $x^i$ , definiert auf Seite 44

## Literatur

- [1] J. Albery, C. Carstensen, S. A. Funken, and R. Klose. Matlab implementation of the finite element method in elasticity. *Computing*, 69(3):239–263, 2002.
- [2] Hans Wilhelm Alt. *Linear functional analysis. An application-oriented introduction*. Universitext. Translated from the German edition by Robert Nürnberg. Springer-Verlag London, Ltd., London, 2016.
- [3] Dietrich Braess. *Finite Elemente. Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*. Springer Verlag, Berlin, fourth edition, 2007.
- [4] C. Carstensen, M. Eigel, and J. Gedicke. Computational competition of symmetric mixed FEM in linear elasticity. *Comput. Methods Appl. Mech. Engrg.*, 200(41-44):2903–2915, 2011.
- [5] Philippe G. Ciarlet. *Mathematical elasticity. Vol. I. Three-dimensional elasticity*, volume 20 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1988.
- [6] Sergio Conti. Einführung in die Funktionalanalysis. Lecture notes. Universität Bonn, February 2022.
- [7] Willy Dörfler. A convergent adaptive algorithm for Poisson’s equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.
- [8] G. Duvaut and J.-L. Lions. *Inequalities in mechanics and physics*, volume 219 of *Grundlehren der Mathematischen Wissenschaften*. Translated from the French by C. W. John. Springer-Verlag, Berlin-New York, 1976.
- [9] Joscha Gedicke. Einführung in die Numerische Mathematik. Lecture notes. Universität Bonn, June 2021.
- [10] Joscha Gedicke. Wissenschaftliches Rechnen I. Lecture notes. Universität Bonn, February 2022.
- [11] C. Geiger and C. Kanzow. *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer-Verlag, Berlin-Heidelberg, first edition, 2002.

- [12] Philip G. Hodge, Jr. Classroom Notes: On Isotropic Cartesian Tensors. *Amer. Math. Monthly*, 68(8):793–795, 1961.
- [13] N. Kikuchi and J. T. Oden. *Contact problems in elasticity: a study of variational inequalities and finite element methods*, volume 8 of *SIAM Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1988.
- [14] L. D. Landau and E. M. Lifshitz. *Theory of elasticity*. Course of Theoretical Physics, Vol. 7. Translated by J. B. Sykes and W. H. Reid. Pergamon Press, London-Paris-Frankfurt; Addison-Wesley Publishing Company, Inc., Reading, Mass., 1959.
- [15] Zdeněk Martinec. Continuum mechanics. Lecture notes. Charles University, Prague, January 2011.
- [16] J. A. Nitsche. On Korn’s second inequality. *RAIRO Anal. Numér.*, 15(3):237–248, 1981.
- [17] Rüdiger Verfürth. *A posteriori error estimation techniques for finite element methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.