Humboldt-Universität zu Berlin
Mathematisch-Naturwissenschaftliche Fakultät
Institut für Mathematik

Least-Squares Methoden in der 2D-Elastizitätstheorie

Bachelorarbeit

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

Name:	Franz Friedrich August Bethke
Geburtsdatum:	31.10.1991
Geburtsort:	Berlin
Gutachter:	Prof. Dr. C. Carstensen
eingereicht am:	8. Juli 2016

Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit der numerischen Lösung der partiellen Differentialgleichungen, welche die Verformung von Festkörpern beschreiben, mit Hilfe von Least-Squares-Finite-Elemente-Methoden. Dabei geht es insbesondere um den Vergleich verschiedener kontinuierlicher Formulierungen und deren Diskretisierungen. Im ersten Abschnitt werden die physikalischen Vorgänge modelliert und die Annahmen an die Problemsituation aufgezeigt. Der zweite Abschnitt behandelt eine einheitliche Notation für verschiedene Least-Squares-Formulierungen des Systems erster Ordnung und den Beweis über deren Wohlgestelltheit, sowie einen alternativen Ansatz in dem die Kontinuitätsgleichung exakt über eine Nebenbedingung gelöst wird.

Für diese Formulierungen werden im dritten Abschnitt die natürliche Diskretisierung niedrigster Ordnung vorgestellt und die Berechnung von Steifigkeitsmatrizen zur numerischen Lösung beschrieben. Außerdem wird auf das Locking-Verhalten konformer Diskretisierungen niedriger Ordnung eingegangen, welches ein Materialparameter abhängiges, vorasymptotisches, verringertes Konvergenzverhalten beschreibt. Daraus motiviert wird noch die nicht-konforme Kouhia-Stenberg Diskretisierung vorgestellt und die Steifigkeitsmatrizen berechnet.

Der vierte Abschnitt enthält die wichtigsten Informationen zum Aufbau und Nutzung der beiliegenden Programme. Im letzten Abschnitt werden die numerischen Experimente präsentiert in denen die verschiedenen Löser zu den verschiedenen Formulierungen verglichen werden.

Inhaltsverzeichnis

1 Modellierung linearer Elastizitätsprobleme

1.1 Herleitung der Differentialgleichungen

Die Elastizitätstheorie beschäftigt sich mit der Verformung von Feststoffen unter Krafteinwirkungen. Die partiellen Differentialgleichungen, die eine solche Verformung beschreiben, sollen hier kurz motiviert und die zur Modellierung nötigen Annahmen und Vereinfachungen herausgearbeitet werden. Dabei sind die Argumente und Schlüsse denen aus [EGK02, Kapitel 5.10] nachempfunden.

In Feststoffen sind die Atome oder Moleküle in einem Gitter angeordnet. Eine elastische Verformung soll die durch das Gitter gegebenen Nachbarschaften dieser nicht verändern. Dadurch ist es sinnvoll eine solche elastische Verformung in Lagrangeschen Koordinaten anzugeben, welche die Position eines Materiepunktes in einer Referenzkonfiguration beschreiben. Dabei ist die Verformung an einem Punkt x im Gebiet Ω zu einem Zeitpunkt t soll durch die Funktion $\varphi : [t_0, \infty) \times \Omega \to \Omega(t)$ gegeben sein. Die zugehörige Verschiebung u betrachtet den Unterschied zwischen der Referenzkonfiguration und der Verformung, das heißt

$$u(t, x) = \varphi(t, x) - x.$$

Die ersten Annahmen, die im Rahmen dieser Arbeit gemacht werden sollen, sind die Beschränkung auf zweidimensionale Gebiete $\Omega \subset \mathbb{R}^2$, sowie die zeitliche Unabhängikeit der Verformung und damit auch der Verschiebung, also

$$\forall x \in \Omega, \qquad \varphi(t, x) = \varphi(x) \text{ und } u(t, x) = u(x).$$

Die Annahme der zeitlichen Unabhängigkeit kann als Grenzwert einer Verformung für $t \to \infty$ bei sich nicht ändernden äußeren Einwirkungen verstanden werden.

Durch den Impulserhaltungssatz in Lagrangschen Koordinaten für eine volumenbezogene Kraftdichte $f : \Omega \to \mathbb{R}^2$ und den Spannungstensor $\sigma : \Omega \to \mathbb{R}^{2\times 2}$

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega \rho \partial_t u \,\mathrm{d}x = \int_\Omega f \,\mathrm{d}x + \int_\Gamma \sigma \nu \,\mathrm{d}s$$

folgt mit dieser Annahme und dem Satz von Gauß dann

$$\int_\Omega f + \operatorname{div} \sigma \,\mathrm{d}x = 0.$$

Aus der Betrachtung dieser Gleichung auf allen messbaren Teilgebieten von Ω und geeigneten Regularitätsannahmen an f und σ erhält man die Formulierung

$$\operatorname{div} \sigma = -f,$$

welche die Kontinuitätsgleichung des Systems der partiellen Differentialgleichungen zur Beschreibung linearer elastischer Vorgänge darstellt. Zur vollständigen Beschreibung wird noch eine konstitutive Gleichung benötigt, welche den Zusammenhang zwischen der Verschiebung u und dem Spannungstensor σ herstellt.

Bei Betrachtung zweier Punkte x und $x+a$ mit einem kleinen Abstand a in der Referenzkonfiguration kann der durch die Verformung geänderte Abstand dieser Punkte mittels des Verformungsgradienten $D\varphi(x)$ angenähert werden, das heißt

$$|\varphi(x+a)-\varphi(x)| \sim |D\varphi(x)a| = (a^T D\varphi(x)^T D\varphi(x)a)^{1/2}.$$

Der Greensche Verzerrungstensor G, gegeben durch

$$G := (D\varphi^T D\varphi - I)/2 = (Du + Du^T + Du^T Du)/2,$$

beschreibt die lokalen Längenänderungen. Um ein lineares Modell der elastischen Vorgänge zu erhalten wird für G erneut linearisiert, sodass die quadratischen Terme in Du vernachlässigt werden. Diese Linearisierung soll mit ε bezeichnet werden, also ist

$$G \sim (Du + Du^T)/2 =: \varepsilon(u).$$

Der gesuchte Zusammenhang zwischen der Verschiebung und der Spannung ist dann gegeben durch das Hooksche Gesetz

$$\sigma_{i,j}(u) = \sum_{k,l=1}^{2} a_{i,j,k,l}\varepsilon_{k,l}(u) \quad \text{für } i,j \in \{1,2\}.$$

Wird das Material als isotrop und homogen angenommen, so ergeben sich für die Koeffizienten des Hookschen Tensors $a_{i,j,k,l}$ die Symetrieeingenschaften $a_{i,j,k,l} = a_{j,i,k,l} = a_{k,l,i,j}$. Und durch den Satz von Rivlin-Ericksen [EGK02, Satz 5.13] ergeben sich die Koeffizienten

$$a_{i,j,k,l} = \lambda\delta_{i,j}\delta_{k,l} + \mu(\delta_{i,k}\delta_{j,l} + \delta_{i,j}\delta_{j,k}).$$

Dabei sind die Lamé-Konstanten λ und μ vom Material abhängig. Der zweite Lamé-Parameter μ wird häufig auch als Schubmodul bezeichnet. Der Hooksche Tensor kann dann als linearer Operator $\mathbb{C}$ geschrieben werden,

$$\mathbb{C}\varepsilon(u) = 2\mu\varepsilon(u) + \lambda \operatorname{tr}(\varepsilon(u))I_{2\times 2}.$$

Für die konstitutive Gleichung ergibt sich somit

$$\sigma = \mathbb{C}\varepsilon(u).$$

Für die eindeutige Lösbarkeit müssen noch Randdaten an die Verschiebung u auf dem Rand Γ_D und die normalen Komponenten der Spannungen $\sigma\nu$ auf dem Rand Γ_N vorgegeben werden. Dabei soll gelten $\Gamma_D \cap \Gamma_N = \varnothing$ und $\Gamma_D \cup \Gamma_N = \Gamma = \partial\Omega$. Dadurch ergibt

sich das System

$$\begin{aligned} \operatorname{div}\sigma &= -f && \text{in } \Omega, \\ \sigma &= \mathbb{C}\varepsilon(u) && \text{in } \Omega, \\ u &= g && \text{auf } \Gamma_D, \\ \sigma\nu &= t && \text{auf } \Gamma_N. \end{aligned} \tag{FOS}$$

Die im nächsten Abschnitt folgenden Modifikationen des Systems erster Ordnung sind im Wesentlichen aus [Hel14b, Abschnitt 7.5] entnommen, und sollen hier lediglich kurz zusammengefasst werden.

1.2 Skalierung der Gebietsgröße

Die im Rahmen dieser Arbeit implementierten Methoden sind zum Teil nicht unabhängig von der Größe des Gebietes Ω, auf dem eine Lösung zu den Gleichungen der linearen Elastizität errechnet wird. Daher ist es von Vorteil alle Gebiete auf eine Referenzgröße zu skalieren. Dazu sei $c \in \mathbb{R}$, sodass $\tilde{\Omega} := \Omega/c \subseteq [-1,1]^2$. Dies sei durch eine Transformation $\Phi = c\,\mathrm{id}\ : \tilde{\Omega} \to \Omega$ organisiert. Die Transformationen der restlichen Größen sind dann gegeben durch

$$\tilde{u} = u \circ \Phi, \quad \tilde{\sigma} = \sigma \circ \Phi, \quad \tilde{f} = f \circ \Phi, \quad \tilde{g} = g \circ \Phi, \quad \text{und} \quad \tilde{t} = t \circ \Phi.$$

Mit der Kettenregel folgt durch die Differentialoperatoren $\operatorname{div}\tilde{\sigma} = c \operatorname{div}\sigma$ und $\varepsilon(\tilde{u}) = c\,\varepsilon(u)$. Somit ergibt sich das skalierte System

$$\begin{aligned} \operatorname{div}\tilde{\sigma} &= -c\tilde{f} && \text{in } \tilde{\Omega} \\ \tilde{\sigma} &= \mathbb{C}\varepsilon(\tilde{u})/c && \text{in } \tilde{\Omega} \\ \tilde{u} &= \tilde{g} && \text{auf } \tilde{\Gamma}_D \\ \tilde{\sigma}\tilde{\nu} &= \tilde{t} && \text{auf } \tilde{\Gamma}_N. \end{aligned}$$

Durch die Substitution $\hat{f} = c^2\tilde{f}$, $\hat{\sigma} = c\tilde{\sigma}$ und $\hat{t} = c\tilde{t}$ kann die Form des ursprünglichen Systems wieder hergestellt werden. Um also auf einem beliebigen Gebiet zu lösen, werden erst die gegeben Daten f, g und t in $\hat{f}, \tilde{g}$ und $\hat{t}$ transformiert, dann die Lösungen $\hat{\sigma}$ und $\tilde{u}$ ermittelt und diese anschließend zurück transformiert.

1.3 Skalierung der Materialparameter

Eine Skalierung des Lamé-Parameters μ kann durch weitere Substitutionen erreicht werden. Dazu werden zunächst die zweite und dritte Gleichung des Systems erster Ordnung durch die Linearität des symmetrischen Gradienten und des Spur-Operators zu

$$\begin{aligned} \sigma &= \varepsilon(2\mu u) + \lambda/(2\mu)\,\mathrm{tr}(\varepsilon(2\mu u))I_{2\times 2} && \text{in } \Omega \\ 2\mu u &= 2\mu g && \text{auf } \Gamma_D \end{aligned}$$

umgeformt. Durch die Substitutionen

$$\tilde{\mu} := 1/2, \qquad \tilde{\lambda} := \lambda/(2\mu), \qquad \tilde{u} := 2\mu u, \qquad \tilde{g} := 2\mu g$$

wird das System wieder in die ursprüngliche Form zurückgeführt. Die eventuell bestehenden Abhängigkeiten der Funktionen f, $\tilde{g}$, und t von den Materialparametern λ und μ bleiben dabei erhalten. Die Lösung u kann aus der Lösung $\tilde{u}$ zu den Daten $\tilde{\mu}, \tilde{\lambda}$ und $\tilde{g}$ durch $u = 1/(2\mu)\tilde{u}$ gewonnen werden.

An dieser Stelle sei noch kurz erwähnt, dass in der linearen Elastizitätstheorie statt den oben beschriebenen Lamé-Parametern λ und μ häufig auch das Elastizitätsmodul (Youngscher Modul) E und die Querkontraktionszahl (Poission-Zahl) ν zur Beschreibung eines Materials genutzt werden. Die Zusammenhänge zwischen diesen Größen lassen sich durch

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad \text{und} \quad \mu = \frac{E}{2(1+\nu)}$$

beschreiben. Diese Arbeit wird alle Aussagen in Termen der Lamé-Parameter machen, die zugehörige Software ist aber auch in der Lage Eingaben eines Elastizitätsmoduls und einer Querkontraktionszahl behandeln.

2 Herleitung der Least-Squares-Formulierungen

Das Ziel des ersten Teils dieses Abschnittes ist es, einige Eigenschaften des Elastizitätstensors $\mathbb{C}$ zusammenzutragen, und eine einheitliche Notation für verschiedene Formulierungen des Systems erster Ordnung zu schaffen.

Der zweite Teil dieses Abschnittes formuliert die Least-Squares Ansätze mit ihren Variationen, und trägt die nötigen Resultate für den Beweis der Wohlgestelltheit dieser Formulierungen zusammen. Der Beweis selbst versucht, die in den Abschätzungen entstehenden Konstanten möglichst genau zu betrachten und damit Rückschlüsse auf die Güte der verschiedenen Least-Squares-Methoden zu ziehen.

Vorab sollen noch einige gängige Notationen eingeführt werden. Für generische Konstanten $C \in \mathbb{R}$ und Abschätzungen der Form $A \leqslant CB$ wird $A \lesssim B$, und für den Fall $A \lesssim B$ und $B \lesssim A$, $A \approx B$ geschrieben. Des Weiteren ist für $A, B \in \mathbb{R}^{n\times n}$ durch

$$A : B := \sum_{i,j=1}^{n} A_{i,j} B_{i,j}$$

ein Skalarprodukt auf dem Raum der Matrizen gegeben. Außerdem wird die Standardnotation für Sobolevräume $H^k(\Omega; \mathbb{R}^N)$ mit den Skalarprodukten $\langle \bullet, \bullet \rangle_{H^k(\Omega;\mathbb{R}^N)}$ sowie den induzierten Normen $\| \bullet \|_{H^k(\Omega;\mathbb{R}^N)}$ für k-fach schwach differenzierbare L^2-Funktionen mit Werten im $\mathbb{R}^N$ verwendet. Für $k = 0$ ist $H^0(\Omega) = L^2(\Omega)$ und das Skalarprodukt wird nur mit $\langle \bullet, \bullet \rangle$ bezeichnet.

2.1 Eigenschaften des Elastitätstensors

Bemerkung 2.1. *Der lineare Operator $\mathbb{C} : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ ist diagonalisierbar, da durch*

$$E_1 := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad E_2 := \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad E_3 := \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad E_4 := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

eine Basis von Eigenmatrizen des Operators $\mathbb{C}$ angegeben werden kann. Insbesondere existieren also lineare Operatoren $\mathbb{C}^s : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ für alle $s \in \mathbb{R}$ und für $s_1, s_2 \in \mathbb{R}$ gilt $\mathbb{C}^{s_1} \circ \mathbb{C}^{s_2} = \mathbb{C}^{s_1+s_2}$. Alle diese Operatoren sind Automorphismen auf der Gruppe der symmetrischen Matrizen $\mathbb{S}^{2\times 2}$.

Besonders relevante Beispiele sind der inverse Operator $\mathbb{C}^{-1}$ zu $\mathbb{C}$, sowie $\mathbb{C}^{1/2}$ und dessen inverser Operator $\mathbb{C}^{-1/2}$. Die Wirkungen dieser Operatoren auf ein $A \in \mathbb{R}^{2\times 2}$ sind gegeben

durch die Vorschriften

$$\begin{aligned}
\mathbb{C}A &= (2\mu)A + \lambda \operatorname{tr}(A) I_{2\times 2} \\
\mathbb{C}^{-1}A &= (2\mu)^{-1}A + 2^{-2}((\lambda+\mu)^{-1} - \mu^{-1}) \operatorname{tr}(A) I_{2\times 2} \\
\mathbb{C}^{1/2}A &= (2\mu)^{1/2}A + 2^{-1/2}((\lambda+\mu)^{1/2} - \mu^{1/2}) \operatorname{tr}(A) I_{2\times 2} \\
\mathbb{C}^{-1/2}A &= (2\mu)^{-1/2}A + 2^{-3/2}((\lambda+\mu)^{-1/2} - \mu^{-1/2}) \operatorname{tr}(A) I_{2\times 2}.
\end{aligned}$$

Für eine kürzere und allgemeine Notation eines allgemeinen Operators $\mathbb{C}^s$ werden die Konstanten $\mathfrak{c}_1(s,\mu) := (2\mu)^s$ und $\mathfrak{c}_2(s,\lambda,\mu) := 2^{s-1}((\lambda+\mu)^s - \mu^s)$ definiert. Damit kann der Operator $\mathbb{C}^s$ dargestellt werden durch

$$\mathbb{C}^s A = \mathfrak{c}_1(s,\mu)A + \mathfrak{c}_2(s,\lambda,\mu) \operatorname{tr}(A) I_{2\times 2}.$$

Für die Richtigkeit dieser Darstellung muss für $s_1, s_2 \in \mathbb{R}$ $\mathbb{C}^{s_1}\mathbb{C}^{s_2}A = \mathbb{C}^{s_1+s_2}A$ nachgerechnet werden, was mit einer expliziten Rechnung und ohne weitere Argumente gelingt.

Bemerkung 2.2. *Für $A, B \in \mathbb{R}^{2\times 2}$ gilt $\operatorname{tr}(A) = A : I_{2\times 2}$ und damit*

$$\operatorname{tr}(A)\operatorname{tr}(B) = \operatorname{tr}(A) I_{2\times 2} : B = A : \operatorname{tr}(B) I_{2\times 2}.$$

Lemma 2.3. *Für Matrizen $A, B \in \mathbb{R}^{2\times 2}$ und $s_1, s_2 \in \mathbb{R}$ gilt*

$$\mathbb{C}^{s_1}A : \mathbb{C}^{s_2}B = \mathbb{C}^{s_1+s_2}A : B = \mathfrak{c}_1(s_1+s_2)A : B + \mathfrak{c}_2(s_1+s_2)\operatorname{tr}(A)\operatorname{tr}(B).$$

Beweis. Die Behauptung ist eine direkte Folgerung aus Bemerkung 2.2 und Bemerkung 2.1. Der Übersichtlichkeit halber wird im ersten Teil die Substitution $\tilde{A} := \mathbb{C}^{s_1}A$ vorgenommen und die Abhängigkeiten von $\mathfrak{c}_1$ und $\mathfrak{c}_2$ von den Materialparametern nicht ausgeschrieben. Es gilt also

$$\begin{aligned}
\mathbb{C}^{s_1}A : \mathbb{C}^{s_2}B &= \tilde{A} : \mathbb{C}^{s_2}B \\
&= \mathfrak{c}_1(s_2)\tilde{A} : B + \mathfrak{c}_2(s_2)\operatorname{tr}(B) I_{2\times 2} : \tilde{A} \\
&= \mathfrak{c}_1(s_2)\tilde{A} : B + \mathfrak{c}_2(s_2)\operatorname{tr}(\tilde{A}) I_{2\times 2} : B \\
&= \mathbb{C}^{s_2}\tilde{A} : B \\
&= \mathbb{C}^{s_1+s_2}A : B \\
&= \mathfrak{c}_1(s_1+s_2)A : B + \mathfrak{c}_2(s_1+s_2)\operatorname{tr}(A) I_{2\times 2} : B \\
&= \mathfrak{c}_1(s_1+s_2)A : B + \mathfrak{c}_2(s_1+s_2)\operatorname{tr}(A)\operatorname{tr}(B).
\end{aligned}$$

□

Bemerkung 2.4. *Der Operator $\mathbb{C}$ besitzt den dreifachen Eigenwert 2μ zu den Eigenmatrizen $E_1, \ldots, E_3$ und den einfachen Eigenwert $(2\mu + 2\lambda)$ zur Eigenmatrix E_4. Folglich besitzt der Operator $\mathbb{C}^s$ den dreifachen Eigenwert $(2\mu)^s$ und den einfachen Eigenwert*

$(2\mu + 2\lambda)^s$. *Durch*

$$(2\lambda + 2\mu)^s = 2^s((\lambda + \mu)^s - \mu^s) + (2\mu)^s = 2\mathbb{c}_2(s, \lambda, \mu) + \mathbb{c}_1(s, \mu)$$

ergibt sich die Darstellung der Eigenwerte als $\mathbb{c}_1(s, \mu)$ *und* $2\mathbb{c}_2(s, \mu, \lambda) + \mathbb{c}_1(s, \mu)$. *Werden* λ *und* μ *positiv angenommen, so ist jedes* $\mathbb{C}^s$ *damit ein symmetrisch positiv-definiter Operator.*

Bezüglich seiner Eigenwerte kann jeder Operator $\mathbb{C}^s$ in einen deviatorischen und einen Spuranteil zerlegt werden. Der Deviator einer Matrix A ist dabei durch $A = \operatorname{dev} A + 1/2 \operatorname{tr}(A) I_{2\times 2}$ gegeben.

Bemerkung 2.5. *Durch die Eigenschaft* $\operatorname{tr}(\operatorname{dev} A) = 0$ *des Deviators einer Matrix, folgt die Zerlegung bezüglich der Eigenwerte durch*

$$\begin{aligned}\mathbb{C}^s A &= \mathbb{C}^s(\operatorname{dev} A + 1/2 \ \operatorname{tr}(A) I_{2\times 2}) \\ &= \mathbb{c}_1(s, \mu) \operatorname{dev} A + (2\mathbb{c}_2(s, \lambda, \mu) + \mathbb{c}_1(s, \mu))/2 \operatorname{tr}(A) I_{2\times 2}.\end{aligned}$$

Durch Bemerkung 2.2 gilt auch $\operatorname{dev} A : \operatorname{tr}(A) I_{2\times 2} = \operatorname{tr}(\operatorname{dev} A) I_{2\times 2} : A = 0$. *Damit besitzt auch* $\|\mathbb{C}^s A\|^2$ *eine Darstellung bezüglich der Eigenwerte von* $\mathbb{C}^s$ *als*

$$\|\mathbb{C}^s A\|^2 = \mathbb{c}_1(s, \mu)^2 \|\operatorname{dev} A\|^2 + (2\mathbb{c}_2(s, \lambda, \mu) + \mathbb{c}_1(s, \mu))^2 \|\operatorname{tr} A\|^2.$$

Bemerkung 2.6. *Um die Eigenwerte von* $\mathbb{C}^s$ *in Relation zueinander stellen zu können, ist eine Untersuchung des Vorzeichens von* $\mathbb{c}_2(s, \lambda, \mu)$ *notwendig. Da offensichtlich*

$$2\mathbb{c}_2(s, \lambda, \mu) + \mathbb{c}_1(s, \mu) \leqslant \mathbb{c}_1(s, \mu) \quad \Leftrightarrow \quad \mathbb{c}_2(s, \lambda, \mu) \leqslant 0$$

gilt. Wird $\lambda, \mu > 0$ *angenommen, so ergibt sich*

$$\operatorname{sign}(\mathbb{c}_2(s, \lambda, \mu)) = \operatorname{sign}(2^{s-1}((\lambda + \mu)^s - \mu^s)) = \operatorname{sign}((\lambda + \mu)^s - \mu^s) = \operatorname{sign}(s).$$

Lemma 2.7. *Wird mit* $\mathbb{c}_{max}(s, \lambda, \mu)$ *der größte Eigenwert von* $\mathbb{C}^s$ *bezeichnet, d.h.*

$$\mathbb{c}_{max}(s, \lambda, \mu) := \begin{cases} \mathbb{c}_1(s, \mu) & \text{für } s \leqslant 0, \\ 2\mathbb{c}_2(s, \lambda, \mu) + \mathbb{c}_1(s, \mu) & \text{für } s > 0 \end{cases}$$

und sei $\tau \in L^2(\Omega; \mathbb{R}^{2\times 2})$, *so gelten folgende Abschätzungen*

(i) $\|\mathbb{C}^s \tau\|_{L^2(\Omega)} \leqslant \mathbb{c}_{max}(s, \lambda, \mu) \|\tau\|_{L^2(\Omega)}$, *(ii)* $\|\tau\|_{L^2(\Omega)} \leqslant \mathbb{c}_{max}(s, \lambda, \mu) \|\mathbb{C}^{-s} \tau\|_{L^2(\Omega)}$.

Beweis. Die Abschätzungen in $L^2(\Omega; \mathbb{R}^{2\times 2})$ folgen durch punktweise Anwendung von Abschätzungen für $A \in \mathbb{R}^{2\times 2}$.

(i) Für eine Matrix $A \in \mathbb{R}^{2\times 2}$ gilt $|\mathbb{C}^s A| \leqslant \|\mathbb{C}^s\||A|$, wobei $\|\mathbb{C}^s\|$ die Spektralnorm von $\mathbb{C}^s$ bezeichnet. Die Spektralnorm ist gleich dem größten Eigenwert und damit nach 2.6 und der obigen Definition gleich $\mathbb{c}_{max}(s,\lambda,\mu)$. Es folgt also $|\mathbb{C}^s A| \leqslant \mathbb{c}_{max}(s,\lambda,\mu)|A|$.

(ii) Folgt aus (i) mit $|A| = |\mathbb{C}^s\mathbb{C}^{-s}A| \leqslant \mathbb{c}_1(s,\mu)|\mathbb{C}^{-s}A|$. □

Bemerkung 2.8. *Durch die obige Definition von* $\mathbb{c}_{max}$ *gilt für* $s_1, s_2 \in \mathbb{R}$ *mit gleichen Vorzeichen, also* $\operatorname{sign}(s_1 s_2) = 1$*, dass* $\mathbb{c}_{max}(s_1,\lambda,\mu)\mathbb{c}_{max}(s_2,\lambda,\mu) = \mathbb{c}_{max}(s_1+s_2,\lambda,\mu)$*.*

Werden beide Abschätzungen aus Lemma 2.7 benötigt, so gibt es offensichtlich keine Wahl von $s \neq 0$, die immer λ-unabhänige Abschätzungen liefert. Aus diesem Grund wird im Folgenden noch eine weitere Abschätzung der Form $\|\sigma\|_{L^2(\Omega)} \leqslant C\|\mathbb{C}^s\sigma\|_{L^2(\Omega)}$ entwickelt, welche im wesentlichen auf der Darstellung von $\|\mathbb{C}^s\tau\|^2$ aus Bemerkung 2.5 und dem folgenden Lemma, beruht.

Lemma 2.9 (tr-dev-div-Lemma)**.** *Für alle* $\tau \in H(\operatorname{div},\Omega,\mathbb{R}^{2\times 2})$ *mit* $\int_\Omega \operatorname{tr}\tau \,\mathrm{d}x = 0$ *existiert eine Konstante* c_{tdd}*, sodass*

$$\|\tau\|_{L^2(\Omega;\mathbb{R}^{2\times 2})} \leqslant c_{\mathrm{tdd}} \left(\|\operatorname{dev}\tau\|_{L^2(\Omega;\mathbb{R}^{2\times 2})} + \|\operatorname{div}\tau\|_{L^2(\Omega;\mathbb{R}^2)} \right).$$

Beweis. Ein Beweis kann in [BBF13, Proposition 9.1.1] gefunden werden. □

Es folgt nun unmittelbar eine Abschätzung der oben gewünschten Form, wenn zusätzlich gegen $\|\operatorname{div}\tau\|$ abgeschätzt wird. Für die späteren Anwendungen ist dies aber keineswegs ein Hindernis, allerdings muss bemerkt werden, dass die verwendete Konstante c_{tdd} in keiner der genannten Quellen explizit angegeben wird.

Lemma 2.10. *Für alle* $\tau \in H(\operatorname{div},\Omega,\mathbb{R}^{2\times 2})$ *mit* $\int_\Omega \operatorname{tr}\tau \,\mathrm{d}x = 0$*, gilt*

$$\|\tau\|_{L^2(\Omega;\mathbb{R}^{2\times 2})} \leqslant c_{\mathrm{tdd}}/\mathbb{c}_1(s,\mu)\|\mathbb{C}^s\tau\|_{L^2(\Omega;\mathbb{R}^{2\times 2})} + c_{\mathrm{tdd}}\|\operatorname{div}\tau\|_{L^2(\Omega;\mathbb{R}^2)}.$$

Beweis. Nach Bemerkung 2.5 gilt

$$\begin{aligned}
&\|\mathbb{C}^s\tau\|^2 = \mathbb{c}_1(s,\mu)^2\|\operatorname{dev}\tau\|^2 + (2\mathbb{c}_2(s,\lambda,\mu) + \mathbb{c}_1(s,\mu))^2\|\operatorname{tr}\tau\|^2 \\
\Leftrightarrow\quad &\|\operatorname{dev}\tau\|^2 = \|\mathbb{C}^s\|^2/\mathbb{c}_1(s,\mu)^2 - (2\mathbb{c}_2(s,\lambda,\mu) + \mathbb{c}_1(s,\mu))^2/\mathbb{c}_1(s,\mu)^2\|\operatorname{tr}\tau\|^2 \\
\Rightarrow\quad &\|\operatorname{dev}\tau\| \leqslant \|\mathbb{C}^s\tau\|/\mathbb{c}_1(s,\mu).
\end{aligned}$$

Wird diese Abschätzung nun in die des Lemmas 2.9 eingesetzt, so erhält man die Behauptung

$$\begin{aligned}
\|\tau\|_{L^2(\Omega;\mathbb{R}^{2\times 2})} \leqslant& c_{\mathrm{tdd}} \left(\|\operatorname{dev}\tau\|_{L^2(\Omega;\mathbb{R}^{2\times 2})} + \|\operatorname{div}\tau\|_{L^2(\Omega;\mathbb{R}^2)} \right) \\
\leqslant& c_{\mathrm{tdd}}/\mathbb{c}_1(s,\mu)\|\mathbb{C}^s\tau\|_{L^2(\Omega;\mathbb{R}^{2\times 2})} + c_{\mathrm{tdd}}\|\operatorname{div}\tau\|_{L^2(\Omega;\mathbb{R}^2)}.
\end{aligned}$$

□

2.2 Wohlgestelltheit verschiedener Problemformulierungen

In diesem Abschnitt wird von homogenen Dirichletranddaten ausgegangen. Der natürlich zu betrachtende Funktionenraum zur Approximation der Verschiebungsvariablen u ist dann

$$H^1_D(\Omega;\mathbb{R}^2) := \{v \in H^1(\Omega;\mathbb{R}^2) \mid v = 0 \text{ auf } \Gamma_D\},$$

da für u im System erster Ordnung eine schwache Ableitung in Form des symmetrischen Gradienten $\varepsilon(u)$ existieren muss und Randdaten auf dem Dirichletrand gesetzt sind. Zur Beschreibung des Approximationsraumes für die Stressvariable σ, werde für den Augenblick angenommen, dass der Neumannrand $\Gamma_N = \varnothing$, also $\Gamma_D = \partial\Omega$. Durch Anwendung der Spur auf die zweite Gleichung des Systems erster Ordnung [FOS] und eine partielle Integration folgt dann

$$\int_\Omega \operatorname{tr}\sigma \,\mathrm{d}x = \int_\Omega \operatorname{tr}\mathbb{C}\varepsilon(u)\,\mathrm{d}x = \int_\Omega (2\mu+2\lambda)\operatorname{div} u\,\mathrm{d}x = (2\mu+2\lambda)\int_{\partial\Omega} u\cdot\nu\,\mathrm{d}x.$$

Da aber $u = 0$ auf dem gesamten Rand $\Gamma_D = \partial\Omega$ gilt, folgt

$$\int_\Omega \operatorname{tr}\sigma\,\mathrm{d}x = 0.$$

Diese Gleichheit zu fordern stellt also keine Einschränkung an den Approximationsraum von σ dar, solange der Neumannrand kein positives Oberflächenmaß besitzt. Werden wie oben homogene Randbedingungen betrachtet, so ergibt also

$$H_N(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}) := \begin{cases} \{\tau \in H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}) \mid \int_\Omega \operatorname{tr}\tau\,\mathrm{d}x = 0\}, & \text{falls } \Gamma_N = \varnothing \\ \{\tau \in H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}) \mid \tau\cdot\nu = 0 \text{ auf } \Gamma_N\} & \text{sonst.} \end{cases}$$

Für die Stressvariable wird dabei nur die Existenz von schwachen Divergenzen gefordert. Die für Lemma 2.9 wichtige Eigenschaft $\int_\Omega \operatorname{tr}\tau\,\mathrm{d}x = 0$ ist mit dieser Definition also nicht immer erfüllt. In [CS03] wird aber eine Verallgemeinerung des Lemmas 2.9 für alle $t \in H_N(\operatorname{div},\Omega;\mathbb{R}^{2\times 2})$ gezeigt. Von nun an werden daher die Konstante c_{tdd}, sowie das Lemma 2.10 im Sinne dieser Verallgemeinerung verstanden.

Als letztes wird der Produktraum

$$X := H_N\left(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}\right) \times H^1_D\left(\Omega;\mathbb{R}^2\right)$$

definiert, in dem dann Lösungstupel (σ,u) gefunden werden können. Der Raum X wird mit der Norm $\|(\tau,v)\|^2_X := \|\tau\|^2_{H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2})} + \|v\|^2_{H^1(\Omega;\mathbb{R}^2)}$ versehen.

Betrachtet man nun das System erster Ordnung [FOS] und wendet auf die zweite Gleichung den Operator $\mathbb{C}^s$ an, so erhält man für jedes $s \in \mathbb{R}$ eine äquivalente Formulierung der Problemstellung.

Problem 1 (klassische Formulierung)**.** *Zu* $f \in L^2(\Omega;\mathbb{R}^2)$ *finde* $(\sigma,u) \in H_N(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}_{sym}) \times H^1_D(\Omega,\mathbb{R}^2)$*, sodass*

$$\begin{aligned} \operatorname{div}\sigma &= -f && \text{f.ü. in } \Omega \\ \mathbb{C}^s\sigma &= \mathbb{C}^{1+s}\varepsilon(u) && \text{f.ü. in } \Omega. \end{aligned}$$

Zu jeder dieser Formulierungen können Least-Squares-Funktionale

$$\begin{aligned} \mathrm{LS}_{\operatorname{div}}(f;\sigma) :=& \|f+\operatorname{div}\sigma\|^2_{L^2(\Omega)}, \\ \mathrm{LS}_{\mathbb{C},s}(f;\sigma,u) :=& \|\mathbb{C}^s\sigma - \mathbb{C}^{1+s}\varepsilon(u)\|^2_{L^2(\Omega)} \text{ und} \\ \mathrm{LS}_s(f;\sigma,u) :=& \mathrm{LS}_{\operatorname{div}}(f;\sigma) + \mathrm{LS}_{\mathbb{C},s}(f;\sigma,u) \\ =& \|f+\operatorname{div}\sigma\|^2_{L^2(\Omega)} + \|\mathbb{C}^s\sigma - \mathbb{C}^{1+s}\varepsilon(u)\|^2_{L^2(\Omega)} \end{aligned}$$

als Residuen der Gleichheiten konstruiert werden. Offensichtlich sind Lösungen des Systems erster Ordnung [FOS] Minimalstellen der Least-Squares-Funktionale. Dies motiviert die folgende Formulierung des Problems.

Problem 2 (Least-Squares-Formulierung)**.** *Zu* $f \in L^2(\Omega;\mathbb{R}^2)$ *finde* $(\sigma,u) \in H_N(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}_{sym}) \times H^1_D(\Omega,\mathbb{R}^2)$*, sodass*

$$(\sigma,u) = \operatorname*{argmin}_{\substack{(\tau,v)\in H_N(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}_{sym}) \\ \times H^1_D(\Omega,\mathbb{R}^2)}} \mathrm{LS}_s(f;\tau,v).$$

Die Minimalstellen dieser Least-Squares-Funktionale sind ihrerseits Nullstellen der ersten Variationen. Dies liefert eine variationelle Formulierung des Problems.

Problem 3 (variationelle Formulierung)**.** *Zu* $f \in L^2(\Omega;\mathbb{R}^2)$ *finde* $(\sigma,u) \in H_N(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}_{sym}) \times H^1_D(\Omega,\mathbb{R}^2)$*, sodass für*

$$\begin{aligned} \mathcal{B}\big((\sigma,u),(\tau,v)\big) :=& \int_\Omega \operatorname{div}\sigma\cdot\operatorname{div}\tau\,\mathrm{d}x \\ &+ \int_\Omega \big(\mathbb{C}^s\sigma - \mathbb{C}^{1+s}\varepsilon(u)\big) : \big(\mathbb{C}^s\tau - \mathbb{C}^{1+s}\varepsilon(v)\big)\,\mathrm{d}x, \\ F\big((\tau,v)\big) :=& \int_\Omega f\cdot\operatorname{div}\tau\,\mathrm{d}x. \end{aligned}$$

und für alle $(\tau,v) \in H_N(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}_{sym}) \times H^1_D(\Omega,\mathbb{R}^2)$ *gilt,*

$$\mathcal{B}\big((\sigma,u),(\tau,v)\big) = -F\big((\tau,v)\big).$$

Lemma 2.11. *Die variationelle Formulierung ist in dem Sinne äquivalent zur Least-Squares-Formulierung, dass Lösungen der einen auch Lösungen der anderen sind.*

Beweis. Eine Minimalstelle des Funktionals LS_s eine Nullstelle der ersten Variation. Es genügt die beiden Variationen der Summanden von LS_s einzeln zu betrachten. Beginnend mit $\mathrm{LS}_{\mathrm{div}}$ ergibt sich die Bedingung

$$0 \stackrel{!}{=} \mathrm{D}\,\mathrm{LS}_{\mathrm{div}}(\sigma)(\tau) := \left[\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{LS}_{\mathrm{div}}(\sigma + t\tau)\right]_{t=0} = \left[\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega (f + \operatorname{div}(\sigma + t\tau))^2\,\mathrm{d}x\right]_{t=0}.$$

Da $\sigma, \tau \in H_N(\operatorname{div}, \Omega; \mathbb{R}^{2\times 2})$ und $f \in L^2(\Omega; \mathbb{R}^2)$ ist der gesamte Integrand L^2-integrierbar und der Satz von Lebesque liefert

$$\begin{aligned}
0 \stackrel{!}{=}& \left[\frac{d}{dt}\int_\Omega (f + \operatorname{div}(\sigma + t\tau))^2\,\mathrm{d}x\right]_{t=0} \\
=& \left[\int_\Omega 2t(\operatorname{div}\tau)^2 + 2(\operatorname{div}\sigma \cdot \operatorname{div}\tau) + 2(f \cdot \operatorname{div}\tau)\,\mathrm{d}x\right]_{t=0} \\
=& 2\int_\Omega \operatorname{div}\sigma \cdot \operatorname{div}\tau\,\mathrm{d}x + 2\int_\Omega f \cdot \operatorname{div}\tau\,\mathrm{d}x.
\end{aligned}$$

Aus dieser Rechnung ergibt sich also die Äquivalenz der Bedingungen

$$\mathrm{D}\,\mathrm{LS}_{\mathrm{div}}(\sigma) = 0 \quad \Longleftrightarrow \quad \int_\Omega \operatorname{div}\sigma \cdot \operatorname{div}\tau\,\mathrm{d}x = -\int_\Omega f \cdot \operatorname{div}\tau\,\mathrm{d}x.$$

Für das zweite Funktional folgt durch eine ähnliche Rechnung,

$$\begin{aligned}
0 \stackrel{!}{=} \mathrm{D}\,\mathrm{LS}_{\mathbb{C},s}(\sigma, u)(\tau, v) :=& \left[\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{LS}_{\mathbb{C},s}(\sigma + t\tau, u + tv)\right]_{t=0} \\
=& \left[\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega (\mathbb{C}^s(\sigma + t\tau) - \mathbb{C}^{1+s}\varepsilon(u + tv))^2\,\mathrm{d}x\right]_{t=0}.
\end{aligned}$$

Wiederum lassen sich die Differentiation und die Integration vertauschen und es gilt

$$\begin{aligned}
0 \stackrel{!}{=}& \left[\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega (\mathbb{C}^s(\sigma + t\tau) - \mathbb{C}^{1+s}\varepsilon(u + tv))^2\,\mathrm{d}x\right]_{t=0} \\
=& \Bigg[2\int_\Omega \mathbb{C}^s\sigma : \mathbb{C}^s\tau - \mathbb{C}^s\sigma : \mathbb{C}^{1+s}\varepsilon(v) - \mathbb{C}^s\tau : \mathbb{C}^{1+s}\varepsilon(u) + \mathbb{C}^{1+s}\varepsilon(u) : \mathbb{C}^{1+s}\varepsilon(v) \\
&+ t\big((\mathbb{C}^s\tau)^2 - 2(\mathbb{C}^s\tau : \mathbb{C}^{1+s}\varepsilon(v)) + (\mathbb{C}^{1+s}\varepsilon(v))^2\big)\,\mathrm{d}x\Bigg]_{t=0} \\
=& 2\int_\Omega \big(\mathbb{C}^s\sigma - \mathbb{C}^{1+s}\varepsilon(u)\big) : \big(\mathbb{C}^s\tau - \mathbb{C}^{1+s}\varepsilon(v)\big)\,\mathrm{d}x.
\end{aligned}$$

Als zweite Bedingung an eine Nullstelle ergibt sich also

$$\mathrm{D}\,\mathrm{LS}_{\mathbb{C},s}(\sigma,u)(\tau,v)=0 \quad \Longleftrightarrow \quad \int_\Omega \big(\mathbb{C}^s\sigma-\mathbb{C}^{1+s}\varepsilon(u)\big):\big(\mathbb{C}^s\tau-\mathbb{C}^{1+s}\varepsilon(v)\big)\,\mathrm{d}x=0.$$

Insgesamt liefern diese beiden Bedingungen die variationelle Formulierung

$$\int_\Omega \operatorname{div}\sigma\cdot\operatorname{div}\tau\,\mathrm{d}x\int_\Omega+\big(\mathbb{C}^s\sigma-\mathbb{C}^{1+s}\varepsilon(u)\big):\big(\mathbb{C}^s\tau-\mathbb{C}^{1+s}\varepsilon(v)\big)\,\mathrm{d}x=-\int_\Omega f\cdot\operatorname{div}\tau\,\mathrm{d}x.\,\square$$

Das vorgeführte Vorgehen beim herleiten variationeller Formulierungen benötigt unbedingt Systeme von partiellen Differentialgleichungen erster Ordnung. Für Differentialgleichungen zweiter, oder höherer Ordnung enstehen durch den Least-Squares-Ansatz zusätzliche Regularitätsanforderungen an die Lösung, welche zu schwerer zu Konstruierenden Approximationsräumen und schließlich zu schlechteren Konditionszahlen der numerischen Lösungsmethoden führen. Diese Problematik wird in [BG09, Kapitel 2.2.2] an einem einfachen Beispiel verdeutlicht. Folgend werden noch problemspezifische Normen $\|\bullet\|_{\mathbb{C}^k}$ definiert, welche durch die Skalarprodukte

$$(\sigma,\tau)_{\mathbb{C}^k}:=\int_\Omega \sigma:\mathbb{C}^k\tau\,\mathrm{dx},\qquad \text{für alle } \sigma,\tau\in L^2(\Omega;\mathbb{R}^{2\times2})$$

induziert werden.

Bemerkung 2.12. *Die Zusammenhänge zwischen der Notation in diesen Normen und den Problemformulierungen mit beliebigem $s\in\mathbb{R}$ können leicht als*

$$\|\mathbb{C}^s\tau-\mathbb{C}^{1+s}\varepsilon(v)\|_{\mathbb{C}^k}=\|\mathbb{C}^{(k/2)+s}\tau-\mathbb{C}^{(k/2)+1+s}\varepsilon(v)\|_{L^2}$$

für $(\tau,v)\in H(\operatorname{div},\Omega;\mathbb{R}^{2\times2})\times L^2(\Omega;\mathbb{R}^2)$ nachgerechnet werden.

Speziell durch

$$|||v|||^2:=\|\varepsilon(v)\|_{\mathbb{C}}^2=\int_\Omega \varepsilon(v):\mathbb{C}\varepsilon(v)\,\mathrm{d}x=2\mu\|\varepsilon(v)\|^2_{L^2(\Omega;\mathbb{R}^{2\times2})}+\lambda\|\operatorname{div}v\|^2_{L^2(\Omega;\mathbb{R}^2)}$$

ist dann die Norm der elastischen Energie einer Verschiebung v gegeben. Ein entscheidendes Resultat stellt in diesem Zusammenhang die Kornsche Ungleichung dar.

Satz 2.13 (Kornsche Ungleichung)**.** *Es existiert eine positive Konstante c_{Korn} mit*

$$\|v\|_{H^1(\Omega;\mathbb{R}^2)}\leqslant c_{\mathrm{Korn}}\|\varepsilon(v)\|_{L^2(\Omega;\mathbb{R}^{2\times2})},\qquad \textit{für alle } v\in H^1(\Omega;\mathbb{R}^2).$$

Ein Beweis kann in [BS02, Theorem 11.2.16+Korollar 11.2.22] gefunden werden.

Bemerkung 2.14. *Durch die Kornsche Ungleichung ist dann die Energienorm äquivalent zur H^1-Norm d.h.,*

$$(2(\mu+\lambda))^{-1/2}|||v||| \leqslant \|v\|_{H^1} \leqslant c_{\mathrm{Korn}}(2\mu)^{-1/2}|||v|||, \qquad \textit{für alle } v \in H^1(\Omega;\mathbb{R}^2).$$

Beweis. Die erste Ungleichung folgt direkt aus $\|\varepsilon(v)\|_{L^2} \leqslant \|\operatorname{D} v\|_{L^2}$ und $\|\operatorname{div} v\|_{L^2} \leqslant 2\|\operatorname{D} v\|_{L^2}$ durch

$$\begin{aligned}|||v|||^2 &= 2\mu\|\varepsilon(v)\|_{L^2}^2 + \lambda\|\operatorname{div} v\|_{L^2}^2 \leqslant 2\mu\|\operatorname{D} v\|_{L^2}^2 + 2\lambda\|\operatorname{D} v\|_{L^2}^2 = 2(\mu+\lambda)\|\operatorname{D} v\|_{L^2}^2 \\ &\leqslant 2(\mu+\lambda)(\|\operatorname{D} v\|_{L^2}^2 + \|v\|_{L^2}^2) = 2(\mu+\lambda)\|v\|_{H^1}^2.\end{aligned}$$

Die zweite Abschätzung folgt aus der Kornschen Ungleichung und der Anwendung des Lemmas 2.7 mit $s = -1/2$

$$\|v\|_{H^1} \leqslant c_{\mathrm{Korn}}\|\varepsilon(v)\|_{L^2} \leqslant c_{\mathrm{Korn}}(2\mu)^{-1/2}\|\mathbb{C}^{1/2}\varepsilon(v)\|_{L^2} = c_{\mathrm{Korn}}(2\mu)^{-1/2}|||v|||. \qquad \square$$

Der nachfolgende Beweis über die Wohlgestelltheit der Formulierung 2 basiert im Wesentlichen auf den Ideen von [CS03] und [CS04, Theorem 3.1]. Die dort durchgeführten Beweise sind die Spezialfälle für $s = -1/2$ und $s = -1$. Die hier vorliegende Version des Beweises für allgemeine $s \in \mathbb{R}$ soll sowohl die Wahl von [CS04] für λ-unabhängige Konstanten bestätigen, als auch zeigen, dass sich für alle $s \in \mathbb{R}$ dieselben Ideen anwenden lassen.

Satz 2.15. *Es existieren eine positive Stetigkeitskonstante $\alpha(s,\lambda,\mu)$ und eine positive Elliptizitätskonstante $\beta(s,\lambda,\mu)$, sodass für alle $(\tau, v) \in H_N(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}_{sym}) \times H^1_D(\Omega,\mathbb{R}^2)$ gilt*

$$\beta(s,\lambda,\mu)^2\|(\tau,v)\|_X^2 \leqslant \mathrm{LS}_s(0;\tau,v) \leqslant \alpha(s,\lambda,\mu)^2\|(\tau,v)\|_X^2.$$

Beweis. Aus Gründen der Übersichtlichkeit werden im gesammten Beweis die Notationen der Normen verkürzt. Außerdem werden die Abhänigkeiten von $\mathfrak{c}_{max}(s,\lambda,\mu)$ von λ und μ, sowie die Abhänigkeiten von $\mathrm{LS}_{\mathbb{C},s}(\tau,v)$ und $\mathrm{LS}_{\mathrm{div}}(0;\tau)$ von τ und v nicht explitzit notiert.

Stetigkeit. Da in der zu zeigenden Abschätzung nicht gegen $\|\operatorname{div}\tau\|_{L^2}$ abgeschätzt werden soll, kann das tr-dev-div-Lemma 2.10 nicht verwendet werden. Es verbleibt nur die Abschätzung über die Eigenwerte aus Lemma 2.7. Um von λ unabhänige Konstanten zu erhalten muss also $s \in \mathbb{R}$ geeignet gewählt werden.

Mit dem Lemma 2.7 und der Tatsache, dass $\|\varepsilon(v)\|_{L^2} \leqslant \|\mathrm{D}v\|_{L^2}$ folgt die Abschätzung durch

$$\begin{aligned}\mathrm{LS}_s &\leqslant \|\operatorname{div}\tau\|_{L^2}^2 + 2\mathfrak{c}_{max}(s)^2\|\tau\|_{L^2}^2 + 2\mathfrak{c}_{max}(s+1)^2\|Dv\|_{L^2}^2 \\ &\leqslant \max\{1, 2\mathfrak{c}_{max}(s)^2, 2\mathfrak{c}_{max}(1+s)^2\}\|(\tau,v)\|_X^2.\end{aligned}$$

Die Konstante $\alpha(s,\lambda,\mu) = \max\{1, 2\mathfrak{c}_{max}(s)^2, 2\mathfrak{c}_{max}(1+s)^2\}^{1/2}$ kann also explizit angegeben werden und ist nur für $s \leqslant -1$ von λ unabhägnig, da nur dann $\mathfrak{c}_{max}(1+s) = \mathfrak{c}_1(1+s,\mu)$ ist. Diese Einschränkung an s muss im zweiten Teil des Beweises berücksichtigt werden.

Elliptizität. Der Beweis ist komplexer und erfolgt daher in mehreren Schritten. Nach jedem Schritt werden die enstehenden Kostanten zusammengefasst und die Bedingungen an $s \in \mathbb{R}$ diskutiert. Da in dieser Richtung gegen $\|\operatorname{div}\tau\|_{L^2}$ abgeschätzt wird, kann nun zusätzlich zu den Eigenwertabschätzungen aus Lemma 2.7 auch das tr-dev-div-Lemma 2.10 verwendet werden.

Schritt 1. Im ersten Schritt wird die Behauptung auf Abschätzungen von $\|\mathbb{C}^{-1/2}\tau\|_{L^2}^2$ und $\|v\|_{H^1}$ zurückgeführt. Die Wahl von $\mathbb{C}^{-1/2}$ scheint dabei zunächst unmotiviert, ist aber für spätere Schritte essentiell. Die Abschätzung mit dem tr-dev-div-Lemma 2.10, ergibt

$$\begin{aligned}\|(\tau,v)\|_X^2 &= \|\operatorname{div}\tau\|_{L^2}^2 + \|\tau\|_{L^2}^2 + \|v\|_{H^1}^2\\ &\leqslant \|\operatorname{div}\tau\|_{L^2}^2 + 2\mathfrak{c}_1(-1/2)^2/c_{\mathrm{tdd}}^2\|\mathbb{C}^{-1/2}\tau\|_{L^2}^2 + 2c_{\mathrm{tdd}}^2\|\operatorname{div}\tau\|_{L^2}^2 + \|v\|_{H^1}^2\\ &= (1+2c_{\mathrm{tdd}}^2)\,\mathrm{LS}_{\mathrm{div}} + 2\mathfrak{c}_1(-1)/c_{\mathrm{tdd}}^2\|\mathbb{C}^{-1/2}\tau\|_{L^2}^2 + \|v\|_{H^1}^2.\end{aligned}$$

Die Konstanten $c_{11} := (1+2c_{\mathrm{tdd}}^2)$ und $c_{12} := 2\mathfrak{c}_1(-1)/c_{\mathrm{tdd}}^2$ sind offensichtlich für alle $s \in \mathbb{R}$ von λ unabhängig. Mit diesen Konstanten folgt

$$\|(\tau,v)\|_X^2 \leqslant c_{11}\,\mathrm{LS}_{\mathrm{div}} + c_{12}\|\mathbb{C}^{-1/2}\tau\|_{L^2}^2 + \|v\|_{H^1}^2.$$

Schritt 2. Die Kornsche Ungleichung aus Satz 2.13, Lemma 2.7 und die Dreiecksungleichung, liefern eine Abschätzung für $\|v\|_{H^1}$ der Form

$$\begin{aligned}\|v\|_{H^1} &\leqslant c_{\mathrm{Korn}}\|\varepsilon(v)\|_{L^2}\\ &\leqslant c_{\mathrm{Korn}}\mathfrak{c}_{max}(-1-s)\|\mathbb{C}^{1+s}\varepsilon(v)\|_{L^2}\\ &\leqslant c_{\mathrm{Korn}}\mathfrak{c}_{max}(-1-s)\left(\|\mathbb{C}^{1+s}\varepsilon(v) - \mathbb{C}^s\tau\|_{L^2} + \|\mathbb{C}^s\tau\|_{L^2}\right)\\ &\leqslant c_{\mathrm{Korn}}\mathfrak{c}_{max}(-1-s)\left(\mathrm{LS}_{\mathbb{C},s}^{1/2} + \mathfrak{c}_{max}(s-1/2)\|\mathbb{C}^{-1/2}\tau\|_{L^2}\right).\end{aligned}$$

Während $\mathfrak{c}_{max}(s-1/2) = \mathfrak{c}_1(s-1/2)$ gilt, da bereits $s \leqslant -1$ ist aufgrund der Abschätzungen für die Stetigkeitskonstante, muss für $\mathfrak{c}_{max}(-1-s) = \mathfrak{c}_1(-1-s)$ noch $s \geqslant -1$ gefordert werden. Da im Allgemeinen nicht $\int_\Omega \operatorname{tr}\varepsilon(v) = 0$ gilt, kann diese Abschätzung mit den vorgestellten Mitteln nicht umgangen werden. Also muss $s = -1$ wie in [CS04] gesetzt werden. Die Konstanten werden in $c_{21} := c_{\mathrm{Korn}}\mathfrak{c}_{max}(-1-s)$ und $c_{22} := \mathfrak{c}_{max}(s-1/2)$ zusammengefasst, sodass sich

$$\|v\|_{H^1} \leqslant c_{21}\,\mathrm{LS}_{\mathbb{C},s}^{1/2} + c_{22}\|\mathbb{C}^{-1/2}\tau\|_{L^2}$$

ergibt.

Schritt 3. Im nächsten Schritt wird mit der Abschätzung von $\|\mathbb{C}^{-1/2}\tau\|_{L^2}^2$ begonnen.

Zunächst folgt aus der Cauchy-Schwarzschen Ungleichung und des Lemmas 2.7, dass

$$\begin{aligned}\|\mathbb{C}^{-1/2}\tau\|_{L^2}^2 &= \langle\mathbb{C}^{-1}\tau,\tau\rangle = \langle\mathbb{C}^{-1}\tau-\varepsilon(v),\tau\rangle + \langle\varepsilon(v),\tau\rangle\\ &\leqslant \|\mathbb{C}^{-1}\tau-\varepsilon(v)\|_{L^2}\|\tau\|_{L^2} + \langle\varepsilon(v),\tau\rangle\\ &\leqslant \mathbb{c}_{max}(-1-s)\|\mathbb{C}^s\tau-\mathbb{C}^{1+s}\varepsilon(v)\|_{L^2}\|\tau\|_{L^2} + \langle\varepsilon(v),\tau\rangle.\end{aligned}$$

Die Abschätzung mit $\mathbb{c}_{max}(-1-s)$ stellt bei der Wahl von $s=-1$ wiederum keine λ Abhänigkeit dar. Da $\|\mathbb{C}^s\tau-\mathbb{C}^{1+s}\varepsilon(v)\|_{L^2} = \mathrm{LS}_{\mathbb{C},s}^{1/2}$, kann dann mit dem tr-dev-div-Lemma 2.10 und den Konstanten $c_{31} := \mathbb{c}_{max}(-1-s)\mathbb{c}_1(-1/2)/c_{\mathrm{tdd}}$ und $c_{32} := \mathbb{c}_{max}(-1-s)c_{\mathrm{tdd}}$ gefolgert werden, dass

$$\begin{aligned}\|\mathbb{C}^{-1/2}\tau\|_{L^2}^2 &\leqslant \mathbb{c}_{max}(-1-s)\,\mathrm{LS}_{\mathbb{C},s}^{1/2}\left(\mathbb{c}_1(-1/2)/c_{\mathrm{tdd}}\|\mathbb{C}^{-1/2}\tau\|_{L^2} + c_{\mathrm{tdd}}\|\operatorname{div}\tau\|_{L^2}\right) + \langle\varepsilon(v),\tau\rangle\\ &= \mathbb{c}_{max}(-1-s)\mathbb{c}_1(-1/2)/c_{\mathrm{tdd}}\,\mathrm{LS}_{\mathbb{C},s}^{1/2}\,\|\mathbb{C}^{-1/2}\tau\|_{L^2}\\ &\quad + \mathbb{c}_{max}(-1-s)c_{\mathrm{tdd}}\,\mathrm{LS}_{\mathbb{C},s}^{1/2}\,\mathrm{LS}_{\mathrm{div}}^{1/2} + \langle\varepsilon(v),\tau\rangle\\ &= c_{31}\,\mathrm{LS}_{\mathbb{C},s}^{1/2}\,\|\mathbb{C}^{-1/2}\tau\|_{L^2} + c_{32}\,\mathrm{LS}_s + \langle\varepsilon(v),\tau\rangle.\end{aligned}$$

Schritt 4. Der nächste Schritt schätzt $\langle\varepsilon(v),\tau\rangle$ weiter ab. Da dies nicht in einem allgemeineren Fall der Form $\langle\mathbb{C}^{s_1}\varepsilon(v),\mathbb{C}^{s_2}\tau\rangle$ erreicht werden kann, wird auch die konkrete Wahl von $\|\mathbb{C}^{-1/2}\tau\|_{L^2}^2$ in den vorherigen Schritten klar.

Die L^2-orthogonale Zerlegung von τ in einen symmetrischen Anteil $\operatorname{sym}(\tau) := (\tau+\tau^\top)/2$ und einen asymmetrischen Anteil $\operatorname{asym}(\tau) := (\tau-\tau^\top)/2$ und die Tatsache, dass $\varepsilon(v)$ symmetrisch ist, erlauben mit einer partiellen Integration und der erneuten Anwedung der Cauchy-Schwarz-Ungleichung eine Abschätzung von $|\langle\varepsilon(v),\tau\rangle|$ durch

$$\begin{aligned}|\langle\varepsilon(v),\tau\rangle| &= |\langle(\tau+\tau^\top)/2,\nabla v\rangle|\\ &= |\langle\tau,\nabla v\rangle \;\; - \langle\operatorname{asym}(\tau),\nabla v\rangle|\\ &= |\langle\operatorname{div}\tau, v\rangle + \langle\operatorname{asym}(\tau),\nabla v\rangle|\\ &\leqslant \|\operatorname{div}\tau\|_{L^2}\|v\|_{L^2} + \|\operatorname{asym}(\tau)\|_{L^2}\|\nabla v\|_{L^2}\\ &\leqslant \|v\|_{H^1}(\mathrm{LS}_{\mathrm{div}}^{1/2} + \|\operatorname{asym}(\tau)\|_{L^2}).\end{aligned}$$

Der verbleibende asymetrische Teil $\operatorname{asym}(\tau)$ kann durch die Symmetrie von $\varepsilon(v)$ und nach Einfügen einer Null in Form von $\pm\mathbb{c}_2(s)\operatorname{tr}(\tau)I_{2\times 2}$ durch eine Dreiecksungleichung gegen das Least-Squares-Funktional $\mathrm{LS}_{\mathbb{C},s}^{1/2}$ abgeschätzt werden. Daher gilt

$$\begin{aligned}\|\operatorname{asym}(\tau)\|_{L^2} &= \mathbb{c}_1(s)^{-1}/2\|\mathbb{c}_1(s)\tau-\mathbb{c}_1(s)\tau^\top\|_{L^2}\\ &= \mathbb{c}_1(-s)/2\|\mathbb{c}_1(s)\tau-\mathbb{C}^{1+s}\varepsilon(v)-\mathbb{c}_1(s)\tau^\top+\mathbb{C}^{1+s}\varepsilon(v)^\top\|_{L^2}\\ &= \mathbb{c}_1(-s)/2\|\mathbb{C}^s\tau-\mathbb{C}^{1+s}\varepsilon(v)-\mathbb{C}^s\tau^\top+\mathbb{C}^{1+s}\varepsilon(v)^\top\|_{L^2}\\ &\leqslant \mathbb{c}_1(-s)\,\mathrm{LS}_{\mathbb{C},s}^{1/2}.\end{aligned}$$

Zusammengefasst ergibt sich mit der Konstate $c_4 := c_{21}(1/2 + \mathbb{c}_1(-s))$ die Abschätzung

$$\begin{aligned}
|\langle \varepsilon(v), \tau \rangle| &\leqslant \|v\|_{H^1} \left(\mathrm{LS}_{\mathrm{div}}^{1/2} + \|\operatorname{asym}(\tau)\|_{L^2} \right) \\
&\leqslant \left(c_{21} \mathrm{LS}_{\mathbb{C},s}^{1/2} + c_{22} \|\mathbb{C}^{-1/2}\tau\|_{L^2} \right) \left(\mathrm{LS}_{\mathrm{div}}^{1/2} + \mathbb{c}_1(-s) \mathrm{LS}_{\mathbb{C},s}^{1/2} \right) \\
&= c_{21} \mathrm{LS}_{\mathbb{C},s}^{1/2} \mathrm{LS}_{\mathrm{div}}^{1/2} + c_{21}\mathbb{c}_1(-s) \mathrm{LS}_{\mathbb{C},s} + c_{22} \|\mathbb{C}^{-1/2}\tau\|_{L^2} \left(\mathrm{LS}_{\mathrm{div}}^{1/2} + \mathbb{c}_1(-s) \mathrm{LS}_{\mathbb{C},s}^{1/2} \right) \\
&\leqslant c_{21}(1/2 + \mathbb{c}_1(-s)) \mathrm{LS}_{\mathbb{C},s} + c_{21}/2 \, \mathrm{LS}_{\mathrm{div}} + c_{22} \|\mathbb{C}^{-1/2}\tau\|_{L^2} \left(\mathrm{LS}_{\mathrm{div}}^{1/2} + \mathbb{c}_1(-s) \mathrm{LS}_{\mathbb{C},s}^{1/2} \right) \\
&\leqslant c_4 \, \mathrm{LS}_s + c_{22} \|\mathbb{C}^{-1/2}\tau\|_{L^2} \left(\mathrm{LS}_{\mathrm{div}}^{1/2} + \mathbb{c}_1(-s) \mathrm{LS}_{\mathbb{C},s}^{1/2} \right).
\end{aligned}$$

Schritt 5. Die in Schritt 3 begonnene Abschätzung von $\|\mathbb{C}^s\tau\|_{L^2}^2$ kann jetzt durch folgende Abschätzung von $\langle \varepsilon(v), \tau \rangle$ beendet werden

$$\begin{aligned}
\|\mathbb{C}^{-1/2}\tau\|_{L^2}^2 &\leqslant c_{31} \mathrm{LS}_{\mathbb{C},s}^{1/2} \|\mathbb{C}^{-1/2}\tau\|_{L^2} + c_{32} \, \mathrm{LS}_s + \langle \varepsilon(v), \tau \rangle \\
&\leqslant c_{31} \mathrm{LS}_{\mathbb{C},s}^{1/2} \|\mathbb{C}^{-1/2}\tau\|_{L^2} + c_{32} \, \mathrm{LS}_s \\
&\quad + c_4 \, \mathrm{LS}_s + c_{22} \|\mathbb{C}^{-1/2}\tau\|_{L^2} \left(\mathrm{LS}_{\mathrm{div}}^{1/2} + \mathbb{c}_1(-s) \mathrm{LS}_{\mathbb{C},s}^{1/2} \right) \\
&\leqslant (c_{32} + c_4) \, \mathrm{LS}_s + c_{22} \|\mathbb{C}^{-1/2}\tau\|_{L^2} \left(\mathrm{LS}_{\mathrm{div}}^{1/2} + (\mathbb{c}_1(-s) + c_{31}) \mathrm{LS}_{\mathbb{C},s}^{1/2} \right).
\end{aligned}$$

Da die linke Seite dieser Abschätzung quadratisch in $\|\mathbb{C}^{-1/2}\tau\|_{L^2}$ ist, die rechte Seite aber nur linear, kann $\|\mathbb{C}^{1-/2}\tau\|_{L^2}^2$ nicht unbeschränkt sein. Die beschränkende Konstante wird in Termen von $T := \|\mathbb{C}^{-1/2}\tau\|_{L^2}$, $A := (c_{32}+c_4) \, \mathrm{LS}_s$ und $B := c_{22}(\mathrm{LS}_{\mathrm{div}}^{1/2} + (\mathbb{c}_1(-s)+c_{31}) \mathrm{LS}_{\mathbb{C},s}^{1/2})$ angegeben. Mit diesen Substitutionen lautet die obere Abschätzung $T^2 \leqslant A + BT$. Durch Betrachtung der Ungleichung

$$0 \leqslant (T - B)^2 = T^2 - 2TB + B^2 \leqslant A - BT + B^2$$

folgt sofort $BT \leqslant A + B$, sowie $T^2 \leqslant 2A + B^2$ nach Einsetzten in die ursprüngliche Ungleichung. Eine Resubstitution ergibt schließlich

$$\begin{aligned}
\|\mathbb{C}^{-1/2}\tau\|_{L^2}^2 &\leqslant 2(c_{32} + c_4) \, \mathrm{LS}_s + c_{22}^2 \left(\mathrm{LS}_{\mathrm{div}}^{1/2} + (\mathbb{c}_1(-s) + c_{31}) \mathrm{LS}_{\mathbb{C},s}^{1/2} \right)^2 \\
&\leqslant 2(c_{32} + c_4) \, \mathrm{LS}_s + 2c_{22}^2 \left(\mathrm{LS}_{\mathrm{div}} + (\mathbb{c}_1(-s) + c_{31})^2 \, \mathrm{LS}_{\mathbb{C},s} \right) \\
&= \left(2(c_{32} + c_4) + 2c_{22}^2 \max\{1, (\mathbb{c}_1(-s) + c_{31})^2\} \right) \mathrm{LS}_s \\
&= c_5 \, \mathrm{LS}_s .
\end{aligned}$$

Schritt 6. Der letzte Schritt fasst nun die Ergebnisse der vorherigen zusammen. Zunächst kann die Abschätzung von $\|v\|_{H^1}$ durch die Abschätzung von $\|\mathbb{C}^s\tau\|_{L^2}$ aus Schritt 5 mit

der Konstante $c_6 := (c_{21} + c_{22}c_5)/2$ beendet werden

$$\begin{aligned}\|v\|_{H^1}^2 &\leqslant \left(c_{21}\,\mathrm{LS}_{\mathbb{C},s}^{1/2} + c_{22}\|\mathbb{C}^{-1/2}\tau\|_{L^2}\right)^2 \leqslant c_{21}/2\,\mathrm{LS}_{\mathbb{C},s} + c_{22}/2\|\mathbb{C}^{-1/2}\tau\|_{L^2}^2 \\ &\leqslant c_{21}/2\,\mathrm{LS}_{\mathbb{C},s} + c_{22}c_5/2\,\mathrm{LS}_s \leqslant c_6\,\mathrm{LS}_s\,.\end{aligned}$$

Schlussendlich folgt für die gesamte Norm

$$\|(\tau, v)\|_X^2 \leqslant c_{11}\,\mathrm{LS}_{\mathrm{div}} + c_{12}\|\mathbb{C}^{-1/2}\tau\|_{L^2}^2 + \|v\|_{H^1}^2 \leqslant c_{11}\,\mathrm{LS}_{\mathrm{div}} + c_{12}c_5\,\mathrm{LS}_s + c_6\,\mathrm{LS}_s \leqslant c_7\,\mathrm{LS}_s\,.\square$$

Zusammengefasst lässt sich aus der Betrachtung der Konstanten folgern, dass für $s \geqslant -1$ die Stetigkeitskonstante $\alpha(s, \lambda, \mu)$ vom Parameter λ abhängig ist. Genauer gilt für $\lambda \to \infty$, dass $\alpha(s, \lambda, \mu) \to \infty$ folgt. Umgekehrt für $s \leqslant -1$ und $\lambda \to \infty$ gilt für die Elliptizitätskonstante $\beta(s, \lambda, \mu) \to 0$.

Aus der Wohlgestelltheit folgt dann auch, dass $\mathcal{B} : X \to X$ eine symmetrische, stetige und elliptische Bilinearform, d.h. ein Skalarprodukt auf X ist. Der Rieszsche Darstellungssatz liefert also die Existenz und Eindeutigkeit von Lösungen des variationellen Problems.

2.3 Formulierung mit exakter Kontinuitätsgleichung

Wie in Abschnitt 1.2 bereits erwähnt, besteht ein großes Problem der Least-Squares-Methoden darin, nicht unabhängig von der Größe des Gebietes Ω zu sein. Das liegt daran, dass die einzelnen Beiträge des Least-Squares-Funktionals unterschiedlich Neben der beschriebenen Möglichkeit das Gebiet auf eine Referenzgröße zu skalieren, kann auch eine weitere Formulierung des Problems gefunden werden, welche keiner Skalierung bedarf. Die nachfolgend beschriebene Methode setzt die Ideen von Prof. Carstensen um. Dabei wird nur die zweite Gleichung des Systems erster Ordnung durch einen klassischen Least-Squares-Ansatz gelöst, die erste Gleichung aber als Nebenbedingung an den Ansatzraum gestellt. Dazu definiere den Raum $Q_N(f)$ durch

$$Q_N(f) := \{\tau \in H_N(\mathrm{div}, \Omega, \mathbb{R}^{2\times 2}) : f + \mathrm{div}\,\tau = 0 \text{ f.ü. in } \Omega\}.$$

Problem 4 (Least-Squares-Formulierung unter Nebenbedigung)**.** *Zu $f \in L^2(\Omega; \mathbb{R}^2)$ finde $(\sigma, u) \in Q_N(f) \times H_D^1(\Omega, \mathbb{R}^2)$, sodass*

$$(\sigma, u) = \operatorname*{argmin}_{(\tau,v)\in Q_N(f)\times H_D^1(\Omega,\mathbb{R}^2)} \|\mathbb{C}^s\tau - \mathbb{C}^{1+s}\varepsilon(v)\|_{L^2(\Omega)}^2.$$

Diese Formulierung als Minimierung unter Nebenbedingung kann für $\mu^L \in L^2(\Omega; \mathbb{R}^2)$ formuliert werden mit Hilfe des Lagrange-Funktionals

$$L(\tau, v, \mu^L) := \mathrm{LS}_{\mathbb{C},s}(\tau, v) + \int_\Omega \mu^L \cdot (f + \mathrm{div}\,\tau)\mathrm{d}x.$$

Problem 5 (Least-Squares-Formulierung mit Lagrange-Paramter)**.** *Zu* $f \in L^2(\Omega;\mathbb{R}^2)$ *finde* $(\sigma, u, \lambda^L) \in H_N(\mathrm{div}, \Omega; \mathbb{R}^{2\times 2}) \times H^1_D(\Omega;\mathbb{R}^2) \times L^2(\Omega;\mathbb{R}^2)$*, sodass*

$$(\sigma, u, \lambda^L) = \underset{\substack{(\tau,v,\mu^L)\in H_N(\mathrm{div},\Omega)\times \\ H^1_D(\Omega,\mathbb{R}^{2\times 2})\times L^2(\Omega,\mathbb{R}^2)}}{\operatorname{argmin}} L(\tau, v, \mu^L).$$

Es wird nun wie im vorherigen Abschnitt eine äquivalente variationelle Formulierung berechnet werden, welche die Grundlage der in dieser Arbeit umgesetzten Implementation liefert. Die Variation des Lagrange-Funktionals berechnet sich durch

$$\begin{aligned}
\mathrm{D}L(\sigma, u, \lambda^L)(\tau, v, \mu^L) &= \left[\frac{\mathrm{d}}{\mathrm{d}t} L(\sigma + t\tau, u + tv, \lambda^L + t\mu^L)\right]_{t=0} \\
&= \left[\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{LS}(\sigma + t\tau, u + tv,)\right]_{t=0} + \left[\frac{\mathrm{d}}{\mathrm{d}t}(\lambda^L + t\mu^L)\cdot\int_\Omega f + \mathrm{div}(\sigma + t\tau)\right]_{t=0} \\
&= \int_\Omega \big(\mathbb{C}^s\sigma - \mathbb{C}^{1+s}\varepsilon(u)\big) : \big(\mathbb{C}^s\tau - \mathbb{C}^{1+s}\varepsilon(v)\big)\mathrm{d}x \\
&\quad + \int_\Omega \lambda^L \cdot \mathrm{div}\,\tau + \mu^L \cdot \mathrm{div}\,\sigma \mathrm{d}x + \int_\Omega \mu^L \cdot f \mathrm{d}x.
\end{aligned}$$

Die Bedingungen an eine Minimalstelle kann also durch

$$-\int_\Omega \mu^L \cdot f\,\mathrm{d}x = \int_\Omega \big(\mathbb{C}^s\sigma - \mathbb{C}^{1+s}\varepsilon(u)\big) : \big(\mathbb{C}^s\tau - \mathbb{C}^{1+s}\varepsilon(v)\big)\mathrm{d}x + \int_\Omega \lambda^L \cdot \mathrm{div}\,\tau + \mu^L \cdot \mathrm{div}\,\sigma \mathrm{d}x$$

realisiert werden, und die variationelle Formulierung des Problems lautet somit wie folgt.

Problem 6 (Variationelle Formulierung des Lagrange-Funktionals)**.** *Zu* $f \in L^2(\Omega;\mathbb{R}^2)$ *finde* $(\sigma, u, \lambda^L) \in H_N(\mathrm{div}, \Omega; \mathbb{R}^{2\times 2}) \times H^1_D(\Omega;\mathbb{R}^2) \times L^2(\Omega;\mathbb{R}^2)$*, sodass für*

$$\begin{aligned}
\mathcal{B}^L\big((\sigma, u, \lambda^L), (\tau, v, \mu^L)\big) &:= \int_\Omega \big(\mathbb{C}^s\sigma - \mathbb{C}^{1+s}\varepsilon(u)\big) : \big(\mathbb{C}^s\tau - \mathbb{C}^{1+s}\varepsilon(v)\big)\,\mathrm{d}x \\
&\quad + \int_\Omega \lambda^L \cdot \mathrm{div}\,\tau + \mu^L \cdot \mathrm{div}\,\sigma\,\mathrm{d}x, \\
F^L\big((\tau, v, \mu^L)\big) &:= \int_\Omega \mu^L \cdot f\,\mathrm{d}x
\end{aligned}$$

und für alle $(\tau, v, \mu^L) \in H_N(\mathrm{div}, \Omega; \mathbb{R}^{2\times 2}) \times H^1_D(\Omega;\mathbb{R}^2) \times L^2(\Omega;\mathbb{R}^2)$ *gilt,*

$$\mathcal{B}^L\big((\sigma, u, \lambda^L), (\tau, v, \mu^L)\big) = -F^L(\tau, v, \mu^L).$$

Die diskrete Formulierung und die Berechnung entsprechender Steifigkeitsmatrizen zu dieser Methode kann im Abschnitt 3.5 gefunden werden. Experimente, die die Unabhängigkeit von der Größe des Gebietes bestätigen sind im Abschnitt 5.4 zu finden.

3 Diskretisierungen und Fehlerschätzer

Um die im vorangegangenen Abschnitt hergeleitete Variationsformulierung numerisch lösen zu können, ist es notwendig, diskrete Teilräume $X_h := \Sigma_h \times U_h$ zu dem Räumen $X = H_N(\mathrm{div}, \Omega; \mathbb{R}^{2\times 2}_{sym}) \times H^1_D(\Omega, \mathbb{R}^2)$ zu wählen.

Im einfachsten Fall einer konformen Approximation soll $X_h \subseteq X$ gelten. Wird X_h so konstruiert, sind die Lösungseigenschaften einer diskreten Variationsformulierung von denen der kontinuierlichen Variationsformulierung 3 übertragbar. Bei einer nicht-konformen Approximation, $X_h \nsubseteq X$, oder im Fall eines gemischten Problems wie in 2.3 müssen diese nochmals geprüft werden.

Um solche diskreten Räume zu konstruieren, ist es notwendig das Gebiet Ω in endlich viele Teilgebiete zu unterteilen. Sei also $\mathcal{T}$ eine reguläre Triangulierung von Ω in abgeschlossene Dreiecke. Mit $\mathcal{N}$ und $\mathcal{E}$ werden die Knoten bzw. die Kanten der Triangulierung bezeichnet. Weiter seien $\mathcal{N}(\Omega)$ und $\mathcal{N}(\partial\Omega)$ die Bezeichnungen für die Knoten im Inneren des Gebietes Ω, bzw. die Knoten auf dem Rand von Ω. In gleicher Weise sind $\mathcal{E}(\Omega)$ und $\mathcal{E}(\partial\Omega)$ definiert.

Im ersten Teil 3.1 dieses Abschnitts wird eine konforme Diskretisierung von niedrigster Ordnung vorgestellt. Allerdings sind die Approximationseigenschafeten von einer solchen Diskretisierung für divergenzfreie Funktionen auf bestimmten Triangulierungen schlecht. In der linearen Elastizitätstheorie treten solche Funktionen im Fall von fast-inkompressiblen Materialien und für spezielle Volumenkräfte auf. Diese Eigenschaft konformer Diskretisierungen niedrigster Ordnung wird als “Locking”-Verhalten bezeichnet und im Abschnitt 3.3 diskutiert. Im folgenden Abschnitt 3.4 wird daher eine nicht-konforme Approximation niedrigster Ordnung vorgestellt, welche das “Locking”-Verhalten vermeidet. Für die rein nicht-konforme Approximation mit Crouzeix-Raviart-Funktionen ist allerdings keine stabile Approximation zu erwarten, da für diesen Fall keine diskrete Kornsche Ungleichung gezeigt werden kann. Aus diesen Gründen wird die von Reijo Kouhia und Rolf Stenberg in [KS95] vorgeschlagene Methode, in der nur eine der beiden Komponenten der Lösung u nicht-konform, die andere aber konform approximiert wird, untersucht.

3.1 Konforme Diskretisierung

Als Grundlage der konformen Diskretisierung sollen der Raviart-Thomas- und der Courant-Finite-Elemente-Raum dienen.

Definition 3.1 (Courant-Finite-Elemente-Raum)**.** *Zu jedem Knoten $Z \in \mathcal{N}$ der Triangulierung $\mathcal{T}$ kann eine stückweise affine und global stetige Funktion $\varphi_Z : \Omega \to \mathbb{R}$ definiert*

werden, die genau im Punkt Z den Wert Eins und in allen anderen den Wert Null hat, d.h.

$$\varphi_Z(x) := \begin{cases} 1, & \textit{falls } x = z, \\ 0, & \textit{falls } x \in \mathcal{N}(\Omega)\backslash\{Z\}. \end{cases}$$

Die Linearkombinationen dieser Funktionen bilden den Splineraum $S^1(\mathcal{T})$. Die Funktionen dieses Raumes sind lokal affin und global stetig.

Durch $S^1(\mathcal{T};\mathbb{R}^2) := S^1(\mathcal{T}) \times S^1(\mathcal{T})$ lassen sich mehrkomponentige Funktionen mit denselben Eigenschaften konstruieren

$$\begin{aligned} \Phi_{Z,\kappa} &:= e_\kappa \varphi_Z, \\ S_0^1(\mathcal{T},\mathbb{R}^2) &= \operatorname{span}\left\{ \begin{pmatrix} \varphi_1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \varphi_1 \end{pmatrix}, \begin{pmatrix} \varphi_2 \\ 0 \end{pmatrix}, \dots \begin{pmatrix} 0 \\ \varphi_{|\mathcal{N}|} \end{pmatrix} \right\}. \end{aligned}$$

Offensichtlich sind die Funktionen aus $S^1(\mathcal{T};\mathbb{R}^2)$ fast überall stetig differenzierbar und global stetig, d.h. es gilt $S^1(\mathcal{T};\mathbb{R}^2) \subseteq H^1(\Omega,\mathbb{R}^2)$.

Definition 3.2 (Raviart-Thomas-Finite-Elemente-Raum)**.** *Wird für jede Kante $E \in \mathcal{E}$ in der Triangulierung $\mathcal{T}$ ein Normalenvektor ν_E ausgezeichnet und bezeichnet $\nu_\pm$ den Normalenvektor der Dreiecke $T_\pm$ mit der gemeinsamen Kante $E = T_+ \cap T_-$, so wird $\zeta_{E,T_\pm} := \nu_E \cdot \nu_{T_\pm} \in \{0,1\}$ als Vorzeichen der Kante definiert. Es kann eine stückweise, in zwei Komponenten affine und in Normalenrichtung global stetige Funktion $\psi_E : \Omega \to \mathbb{R}^2$ definiert werden, welche auf den der Kante gegenüberliegenden Punkten P_+ und P_- verschwindet und auf der Kante E selbst $\psi_E \cdot \nu_E = 1$ erfüllt, also*

$$\psi_E(x) := \begin{cases} \zeta_{E,T_\pm}|E|/(2|T_\pm|)(x - P_\pm), & \textit{falls } x \in T_\pm, \\ 0, & \textit{sonst.} \end{cases}$$

Die Linearkombinationen dieser Funktionen bilden den Raum $RT_0(\mathcal{T})$.

Lemma 3.3 (Eigenschaften der Raviart-Thomas-Basisfunktionen)**.** *Für ein Dreieck $T = \operatorname{conv}\{P,E\}$ und für $\kappa \in \{1,2\}$ gilt*

(i) $\psi_E = \zeta_E|E|/(2|T|)\sum_{\ell=1}^3 \varphi_\ell(P_\ell - P)$

(ii) $\operatorname{div}\psi_E = \zeta_E|E|/|T|$

(iii)

$$\int_T \psi_{E,\kappa}\,\mathrm{d}x = \zeta_E|E|/6\sum_{\ell=1}^3 \varphi_\ell(P_{\ell,\kappa} - P_\kappa)$$

(iv)

$$\int_T \psi_{E_\alpha,\kappa}\psi_{E_\beta,\kappa}\,\mathrm{d}x = \zeta_{E_\alpha}\zeta_{E_\beta}\frac{|E_\alpha||E_\beta|}{48|T|}\Bigg(\sum_{\ell=1}^{3}\sum_{m=1}^{3}(P_{\ell,\kappa}-P_{\alpha,\kappa})(P_{m,\kappa}-P_{\beta,\kappa}) + \sum_{d=1}^{3}(P_{d,\kappa}-P_{\alpha,\kappa})(P_{d,\kappa}-P_{\beta,\kappa})\Bigg).$$

Beweis. Die Beweise sind in [BC05, Lemma 4.1 und Lemma 4.2] zu finden, oder unmittelbar aus den dortigen abzuleiten. □

Durch $RT_0(\mathcal{T};\mathbb{R}^{2\times 2}) := RT_0(\mathcal{T})\times RT_0(\mathcal{T})$ lassen sich matrixwertige Funktionen darstellen mit

$$\Psi_{E,\kappa} := e_\kappa \otimes \psi_E,$$
$$RT_0(\mathcal{T},\mathbb{R}^{2\times 2}) = \operatorname{span}\left\{\begin{pmatrix}\psi_{1,1} & \psi_{1,2}\\ 0 & 0\end{pmatrix},\begin{pmatrix}0 & 0\\ \psi_{1,1} & \psi_{1,2}\end{pmatrix},\begin{pmatrix}\psi_{2,1} & \psi_{2,2}\\ 0 & 0\end{pmatrix},\dots,\begin{pmatrix}0 & 0\\ \psi_{|\mathcal{E}|,1} & \psi_{|\mathcal{E}|,2}\end{pmatrix}\right\}.$$

Dieser Raum von stellt den natürlichen konformen Unterraum von der Funktionen mit schwacher Divergenz dar, d.h.

$$RT_0(\mathcal{T},\mathbb{R}^{2\times 2}) \subseteq H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2}).$$

Damit ist $X_h := \Sigma_h \times U_h := RT_0(\mathcal{T},\mathbb{R}^{2\times 2}) \times S_0^1(\mathcal{T},\mathbb{R}^2)$ eine mögliche Wahl für einen konformen Unterraum von X, und es kann eine diskrete Variante der variationellen Formulierung aufgestellt werden, indem in der kontinuierlichen Variante einfach X durch X_h ersetzt wird.

Problem 7 (Diskrete Variationelle Formulierung)**.** *Gegeben $f \in L^2(\Omega;\mathbb{R}^2)$ finde $x_h := (\sigma_h, u_h) \in \Sigma_h \times U_h = X_h$, sodass für alle $(\tau_h, v_h) \in X_h$*

$$\mathcal{B}\big((\sigma_h,u_h),(\tau_h,v_h)\big) = F\big((\tau_h,v_h)\big).$$

Die Lösungen dieser Problemstellung haben Darstellungen entsprechend der oben eingeführten Basisfunktionen. Die Koeffizienten zu diesen Basisfunktionen sollen wie folgt bezeichnet sein

$$\sigma_h = \sum_{E=1}^{|\mathcal{E}|}\sum_{\kappa=1}^{2}\sigma_{2(E-1)+\kappa}\Psi_{E,\kappa} = \sum_{E=1}^{|\mathcal{E}|}\sum_{\kappa=1}^{2}x_{2(E-1)+\kappa}\Psi_{E,\kappa}$$
$$u_h = \sum_{Z=1}^{|\mathcal{N}|}\sum_{\kappa=1}^{2}u_{2(Z-1)+\kappa}\Phi_{Z,\kappa} = \sum_{Z=1}^{|\mathcal{N}|}\sum_{\kappa=1}^{2}x_{|2\mathcal{E}|+2(Z-1)+\kappa}\Phi_{Z,\kappa}.$$

Durch diese Darstellungen lässt sich Problem 7 auch als lineares Gleichungssystem schreiben, wenn wie gewöhnlich nur mit den Basisfunktionen getestet wird. Die darstellende Matrix dieses Gleichungssystems lässt sich dann in vier Blöcke unterteilen. In Block A sind die Beiträge der Basisfunktionen von Σ_h, in Block C die der Basisfunktionen von U_h zu finden. Block B bzw. $B^\top$ enthalten die gemischten Beiträge. Die Vektoren der Koeffizienten zu den Basisfunktionen werden mit σ bzw. u bezeichnet. Mit diesen Bezeichnungen liest sich das Gleichungssystem

$$\begin{pmatrix} A & B \\ B^\top & C \end{pmatrix} \begin{pmatrix} \sigma \\ u \end{pmatrix} = \begin{pmatrix} F \\ 0 \end{pmatrix}.$$

Die Einträge in den Blöcken, die zu den Kanten E_α und E_β, sowie Knoten Z_a und Z_b eines Dreiecks T in den Komponenten κ_1 und κ_2 gehören, sind gegeben durch

$$\begin{aligned}
A_{2(E_\alpha-1)+\kappa_1,2(E_\beta-1)+\kappa_2} &= \int_T \operatorname{div} \Psi_{E_\alpha,\kappa_1} \cdot \operatorname{div} \Psi_{E_\beta,\kappa_2} \,\mathrm{d}x + \int_T \mathbb{C}^s \Psi_{E_\alpha,\kappa_1} : \mathbb{C}^s \Psi_{E_\beta,\kappa_2} \,\mathrm{d}x \\
B_{2(E_\alpha-1)+\kappa_1,2(Z_b-1)+\kappa_2} &= -\int_T \mathbb{C}^s \Psi_{E_\alpha,\kappa_1} : \mathbb{C}^{1+s} \varepsilon(\Phi_{Z_b,\kappa_2}) \,\mathrm{d}x \\
C_{2(Z_a-1)+\kappa_1,2(Z_b-1)+\kappa_2} &= \int_T \mathbb{C}^{1+s} \varepsilon(\Phi_{Z_a,\kappa_1}) : \mathbb{C}^{1+s} \varepsilon(\Phi_{Z_b,\kappa_2}) \,\mathrm{d}x \\
F_{2(E_\alpha-1)+\kappa_1} &= -\int_T f \cdot \operatorname{div} \Psi_{E_\alpha,\kappa_1} \,\mathrm{d}x.
\end{aligned}$$

Die Blöcke A, B, C haben für feste κ_1, κ_2, alle höchstens 9 von Null verschiedene Einträge, sodass man sie in diesen Fällen auch als 3×3-Matrizen auffassen kann. Die folgende Bemerkung zeigt, wie diese 3×3-Matrizen in die 6×6-Blöcke eingeordnet werden können. Jeder der aufgeführten vier Fälle entspricht einer Wahl von κ_1 und κ_2.

Bemerkung 3.4. *Für eine 3×3-Matrix M mit den Einträgen $m_{j,k}$ mit $j, k = 1, 2, 3$ gilt*

$$S_1^\top \; M \; S_1 = \begin{bmatrix} m_{11} & 0 & m_{12} & 0 & m_{13} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ m_{21} & 0 & m_{22} & 0 & m_{23} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ m_{31} & 0 & m_{32} & 0 & m_{33} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad S_1^\top \; M \; S_2 = \begin{bmatrix} 0 & m_{11} & 0 & m_{12} & 0 & m_{13} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & m_{21} & 0 & m_{22} & 0 & m_{23} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & m_{31} & 0 & m_{32} & 0 & m_{33} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$S_2^\top \; M \; S_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ m_{11} & 0 & m_{12} & 0 & m_{13} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ m_{21} & 0 & m_{22} & 0 & m_{23} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ m_{31} & 0 & m_{32} & 0 & m_{33} & 0 \end{bmatrix}, \quad S_2^\top \; M \; S_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & m_{11} & 0 & m_{12} & 0 & m_{13} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & m_{21} & 0 & m_{22} & 0 & m_{23} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & m_{31} & 0 & m_{32} & 0 & m_{33} \end{bmatrix},$$

mit den Matrizen

$$S_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \; S_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Eine weitere Bemerkung hält die Wirkungen des Spuroperators auf die Basisfunktionen fest, wie sie bei der Berechnung der Steifigkeitsmatrix auftreten.

Bemerkung 3.5. *Es gilt*

$$\operatorname{tr}(\varepsilon(\Phi_{Z,\kappa})) = \frac{\partial \varphi_Z}{\partial x_\kappa} \quad \textit{und} \quad \operatorname{tr}(\Psi_{E,\kappa}) = \psi_{E,\kappa}.$$

3.1.1 Block A

Die Vorgehensweise bei der folgenden Berechnung aller Blöcke der Steifigkeitsmatrix ist im Grunde immer ähnlich. Zunächst wird der Block mit Lemma 2.3 in von $\mathbb{C}$ und s unabhängige Blöcke aufgeteilt. Dann wird ein beliebiger, aber fester Eintrag aus diesen Blöcken auf die bekannten Größen wie z.B. Längen von Kanten oder Koordinaten von Knoten zurückgeführt. Diese Einträge werden für feste Komponenten κ_1 und κ_2 in Matrizen zusammengefasst und zum Schluss mit Bemerkung 3.4 an die richtigen Positionen verschoben.

Die Berechnung des Blockes A kann mit Hilfe des Lemmas 2.3 auf die Berechnung von Blöcken $A_{\mathrm{div}}, A_{L^2}$ und A_{tr} zurückgeführt werden. Die zur Berechnung der neuen Blöcke notwendigen Schritte sind stark an die in [BC05] und [Bri13, Abschnitt 4] Rechnungen angelehnt. Es gilt

$$\begin{aligned} A_{2(E_\alpha-1)+\kappa_1,2(E_\beta-1)+\kappa_2} &= \underbrace{\int_T \operatorname{div} \Psi_{E_\alpha,\kappa_1} \cdot \operatorname{div} \Psi_{E_\beta,\kappa_2} \,\mathrm{d}x}_{A_{\mathrm{div}}} \\ &+ \mathbb{c}_1(2s) \underbrace{\int_T \Psi_{E_\alpha,\kappa_1} : \Psi_{E_\beta,\kappa_2} \,\mathrm{d}x}_{A_{L^2}} + \mathbb{c}_2(2s) \underbrace{\int_T \operatorname{tr}(\Psi_{E_\alpha,\kappa_1}) : \operatorname{tr}(\Psi_{E_\beta,\kappa_2}) \,\mathrm{d}x}_{A_{\mathrm{tr}}}. \end{aligned}$$

Berechnung von A_{div}. Mit Lemma 3.3 können die Divergenzen der Raviart-Thomas-Basisfunktionen berechnet und damit dann auch das Integral aufgelöst werden. Das heißt

,

$$A_{\mathrm{div},2(E_\alpha-1)+\kappa_1,2(E_\beta-1)+\kappa_2} = \int_T \operatorname{div}\Psi_{E_\alpha,\kappa_1}\cdot\operatorname{div}\Psi(E_\beta,\kappa_2)\,\mathrm{d}x$$
$$= \delta_{\kappa_1,\kappa_2}\int_T \zeta_{E_\alpha}\frac{|E_\alpha|}{|T|}\zeta_{E_\beta}\frac{|E_\beta|}{|T|}\,\mathrm{d}x = \delta_{\kappa_1,\kappa_2}\frac{1}{|T|}\zeta_{E_\alpha}\zeta_{E_\beta}|E_\alpha||E_\beta|.$$

Mit $L := \operatorname{diag}(\zeta_{E_1}|E_1|,\zeta_{E_2}|E_2|,\zeta_{E_3}|E_3|)$ lassen sich die 3×3-Blöcke, in denen $\kappa_1 = \kappa_2$ gilt, berechnen durch

$$\frac{1}{|T|}L\ I_{3\times 3}\ L = \begin{bmatrix} |E_1|^2 & \zeta_{E_1}\zeta_{E_2}|E_1||E_2| & \zeta_{E_1}\zeta_{E_3}|E_1||E_3| \\ \zeta_{E_2}\zeta_{E_1}|E_2||E_1| & |E_2|^2 & \zeta_{E_2}\zeta_{E_3}|E_2||E_3| \\ \zeta_{E_3}\zeta_{E_1}|E_3||E_1| & \zeta_{E_3}\zeta_{E_2}|E_3||E_2| & |E_3|^2 \end{bmatrix}.$$

Diese Blöcke können nun mit Bemerkung 3.4 an die entsprechenden Positionen der 6×6-Matrix verschoben werden

$$A_{\mathrm{div}} = \frac{1}{|T|}S_1^\top\ L\ I_{3\times 3}\ L\ S_1 + S_2^\top\ L\ I_{3\times 3}\ L\ S_2.$$

Berechnung von A_{L^2}. Zur Berechnung von A_{L^2} seien die Matrizen $N, M, M_{1,1}$ und $M_{2,2}$ wie folgt gegeben

$$N = \begin{bmatrix} Z_1-Z_3 & 0 & Z_1-Z_2 \\ Z_2-Z_3 & Z_2-Z_1 & 0 \\ 0 & Z_3-Z_1 & Z_3-Z_2 \end{bmatrix}, \qquad M_{1,1} = M\ S_1^\top\ S_1 = \begin{bmatrix} 2&0&1&0&1&0\\0&0&0&0&0&0\\1&0&2&0&1&0\\0&0&0&0&0&0\\1&0&1&0&2&0\\0&0&0&0&0&0 \end{bmatrix},$$

$$M = \begin{bmatrix} 2&0&1&0&1&0\\0&2&0&1&0&1\\1&0&2&0&1&0\\0&1&0&2&0&1\\1&0&1&0&2&0\\0&1&0&1&0&2 \end{bmatrix}, \qquad M_{2,2} = M\ S_2^\top\ S_2 = \begin{bmatrix} 0&0&0&0&0&0\\0&2&0&1&0&1\\0&0&0&0&0&0\\0&1&0&2&0&1\\0&0&0&0&0&0\\0&1&0&1&0&2 \end{bmatrix}.$$

Die Matrizen N, M entsprechen gerade denen aus [BC05, Lemma 4.2]. Nach Einsetzen der Definition der Basisfunktionen und Auswerten des Matrix- Skalarproduktes kann der vierte Teil des Lemmas 3.3 verwendet werden. Durch Nachrechnen verifiziert man dann, dass der erste Summand gerade den Einträgen von $LN^\top M_{1,1}NL/48|T|$ an den Stellen α,β entspricht. Ebenso entspricht der zweite Summand den Einträgen von $LN^\top M_{2,2}NL/48|T|$

an den Stellen α, β. Es gilt also

$$\begin{aligned} A_{L^2,2(E_\alpha-1)+\kappa_1,2(E_\beta-1)+\kappa_2} &= \int_T \Psi_{E_\alpha,\kappa_1} : \Psi_{E_\beta,\kappa_2} \,\mathrm{d}x \\ &= \delta_{\kappa_1,\kappa_2}\Big(\int_T \psi_{E_\alpha,1}\psi_{E_\beta,1} \,\mathrm{d}x + \int_T \psi_{E_\alpha,2}\psi_{E_\beta,2} \,\mathrm{d}x\Big) \\ &= \delta_{\kappa_1,\kappa_2}/48|T|\big((LN^\top M_{1,1}NL)_{\alpha,\beta} + (LN^\top M_{2,2}NL)_{\alpha,\beta}\big). \end{aligned}$$

Bemerkung 3.4 und $M_{1,1} + M_{2,2} = M$ liefern dann

$$\begin{aligned} A_{L^2} = \;& \frac{1}{48|T|}(S_1^\top LN^\top M_{1,1}NLS_1 + S_2^\top LN^\top M_{1,1}NLS_2) \\ &+ \frac{1}{48|T|}(S_1^\top LN^\top M_{2,2}NLS_1 + S_2^\top LN^\top M_{2,2}NLS_2) \\ = \;& \frac{1}{48|T|}(S_1^\top LN^\top MNLS_1 + S_2^\top LN^\top MNLS_2). \end{aligned}$$

Berechnung von A_{tr}. Mit der Bemerkung 3.5 folgt

$$A_{tr,2(E_\alpha-1)+\kappa_1,2(E_\beta-1)+\kappa_2} = \int_T tr\Psi(E_\alpha,\kappa_1)tr\Psi(E_\beta,\kappa_2) \,\mathrm{d}x = \int_T \psi_{E_\alpha,\kappa_1}\psi_{E_\beta,\kappa_2}$$

und man erkennt, dass für A_{tr} im wesentlichen dieselben Terme berechnet werden müssen wie für A_{L^2}. Dies gilt insbesondere für die Fälle $\kappa_1 = \kappa_2 = 1$ und $\kappa_1 = \kappa_2 = 2$. Die gemischten Fälle $\kappa_1 \neq \kappa_2$ können mit den Matrizen $M_{1,2}$ und $M_{2,1}$ durch

$$M_{1,2} = M\ S_1^\top\ S_2 = \begin{bmatrix} 0 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad M_{2,1} = M\ S_2^\top\ S_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 2 & 0 \end{bmatrix}$$

berechnet werden und A_{tr} ergibt sich als

$$\begin{aligned} A_{tr} = 1/48|T|(&S_1^\top LN^\top M_{1,1}NLS_1 + S_1^\top LN^\top M_{1,2}NLS_2 + \\ &S_2^\top LN^\top M_{2,1}NLS_1 + S_2^\top LN^\top M_{2,2}NLS_2). \end{aligned}$$

3.1.2 Block B

Für den Block B wird wiederum Lemma 2.3 verwendet, um die Berechnung auf einfachere Blöcke zu reduzieren. Damit ergibt sich

$$\begin{aligned} &B_{2(E_\alpha-1)+\kappa_1,2(Z_b-1)+\kappa_2} \\ &= -\mathbb{c}_1(2s+1)\underbrace{\int_T \Psi_{E_\alpha,\kappa_1} : \varepsilon(\Phi_{Z_b,\kappa_2})\,\mathrm{d}x}_{B_1} - \mathbb{c}_2(2s+1)\underbrace{\int_T \operatorname{tr}(\Psi_{E_\alpha,\kappa_1})\operatorname{tr}(\varepsilon(\Phi_{Z_b,\kappa_2}))\,\mathrm{d}x}_{B_{\mathrm{tr}}}. \end{aligned}$$

Ist durch κ eine Komponente einer Basisfunktion gegeben, so wird mit κ' im Folgenden die jeweils andere bezeichnet. Formal kann man κ' durch $\kappa' := 3 - \kappa$ definieren.

Berechnung von B_1. Nach Einsetzten der Definitionen der Basisfunktionen und Auswerten des Matrixskalarproduktes können die konstanten Ableitungen der nodalen Basisfunktionen aus dem Integral herausgezogen werden. Die verbleibenden Integrale können durch Lemma 3.3 wie folgt berechnet werden

$$B_{1,2(E_\alpha-1)+\kappa_1,2(Z_b-1)+\kappa_2} = \delta_{\kappa_1,\kappa_2}\frac{\partial\varphi_{Z_b}}{\partial x_{\kappa_2}}\int_T \psi_{E_\alpha,\kappa_1}\,\mathrm{d}x + \frac{1}{2}\frac{\partial\varphi_{Z_b}}{\partial x_{\kappa_2'}}\int_T \psi_{E_\alpha,\kappa_1'}\,\mathrm{d}x.$$

Die einzelnen Summanden entsprechen für eine Wahl von κ_1 und κ_2 den Einträgen der Matrix $LN^\top H_{\kappa_1,\kappa_2}G^\top/6$ bzw. $LN^\top H_{\kappa_1',\kappa_2'}G^\top/6$ an den Stellen α, b, wobei die Matrizen wie folgt gegeben sind

$$H_{1,1} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^\top, \qquad H_{1,2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}^\top,$$

$$H_{2,1} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^\top, \qquad H_{2,2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}^\top,$$

$$\text{und } G = \begin{bmatrix} \frac{\partial\varphi_{Z_1}}{\partial x_1} & \frac{\partial\varphi_{Z_2}}{\partial x_1} & \frac{\partial\varphi_{Z_3}}{\partial x_1} \\ \frac{\partial\varphi_{Z_1}}{\partial x_2} & \frac{\partial\varphi_{Z_2}}{\partial x_2} & \frac{\partial\varphi_{Z_3}}{\partial x_2} \end{bmatrix}^\top.$$

Nach Anwendung von Bemerkung 3.4 folgt

$$\begin{aligned} B_1 = \;& S_1^\top\left(LN^\top H_{1,1}G^\top/6\right)S_1 + S_2^\top\left(LN^\top H_{2,2}G^\top/6\right)S_2 \\ &+ S_1^\top\left(LN^\top H_{2,2}G^\top/12\right)S_1 + S_1^\top\left(LN^\top H_{2,1}G^\top/12\right)S_2 \\ &+ S_2^\top\left(LN^\top H_{1,2}G^\top/12\right)S_1 + S_2^\top\left(LN^\top H_{1,1}G^\top/12\right)S_2. \end{aligned}$$

Berechnung von B_{tr}. Wie schon bei A_{tr} lässt sich auch die Berechnung von B_{tr} mit Bemerkung 3.5 auf Terme zurückführen, die eben schon für B_1 berechnet wurden, also

$$B_{\mathrm{tr},2(E_\alpha-1)+\kappa_1,2(Z_b-1)+\kappa_2} = \int_T \operatorname{tr}(\Psi_{E_\alpha,\kappa_1}) \operatorname{tr}(\varepsilon(\Phi_{Z_b,\kappa_2}))\,\mathrm{d}x = \frac{\partial \varphi_{Z_b}}{\partial x_{\kappa_2}} \int_T \psi_{E_\alpha,\kappa_1}.$$

Nach Anwendung von Bemerkung 3.4 folgt

$$\begin{aligned} B_{\mathrm{tr}} = \; & S_1^\top (LN^\top H_{1,1} G^\top/6) S_1 + S_1^\top (LN^\top H_{1,2} G^\top/6) S_2 \\ & + S_2^\top (LN^\top H_{2,1} G^\top/6) S_1 + S_2^\top (LN^\top H_{2,2} G^\top/6) S_2. \end{aligned}$$

3.1.3 Block C

Durch eine erneute Anwendung des Lemmas 2.3, ergibt sich für C

$$\begin{aligned} C_{2(Z_a-1)+\kappa_1,2(Z_b-1)+\kappa_2} =& \mathbb{c}_1(2s+2) \underbrace{\int_T \varepsilon(\Phi(Z_a,\kappa_1)) : \varepsilon(\Phi(Z_b,\kappa_2))\,\mathrm{d}x}_{C_{L^2}} \\ & + \mathbb{c}_2(2s+2) \underbrace{\int_T \operatorname{tr}\varepsilon(\Phi(Z_a,\kappa_1)) \operatorname{tr}\varepsilon(\Phi(Z_b,\kappa_2))\,\mathrm{d}x}_{C_{\mathrm{tr}}}. \end{aligned}$$

Berechnung von C_{L^2}. Durch Einsetzen der Definition ergibt sich

$$C_{L^2,2(Z_a-1)+\kappa_1,2(Z_b-1)+\kappa_2} = \int_T \delta_{\kappa_1,\kappa_2} \frac{\partial \varphi_{Z_a}}{\partial x_{\kappa_1}} \frac{\partial \varphi_{Z_b}}{\partial x_{\kappa_2}} + \frac{1}{2} \frac{\partial \varphi_{Z_a}}{\partial x_{\kappa_1'}} \frac{\partial \varphi_{Z_b}}{\partial x_{\kappa_2'}}\,\mathrm{d}x$$

und mit den Matrizen

$$J_{1,1} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad J_{1,2} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad J_{2,1} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{und} \qquad J_{2,2} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

können die Summanden als Einträge an den Stellen a, b in den Matrizen $GJ_{\kappa_1,\kappa_2}G^\top$ bzw. $GJ_{\kappa_1',\kappa_2'}G^\top$ gefunden werden. Somit ergibt sich C_{L^2} mit Bemerkung 3.4 als

$$\begin{aligned} C_{L^2} = |T|(& S_1^\top G J_{1,1} G^\top S_1 + S_2^\top G J_{2,2} G^\top S_2 \\ & + S_1^\top G J_{2,2} G^\top S_1/2 + S_1^\top G J_{2,1} G^\top S_2/2 \\ & + S_2^\top G J_{1,2} G^\top S_1/2 + S_2^\top G J_{1,1} G^\top S_2/2). \end{aligned}$$

Berechnung von C_{tr}. Mit Bemerkung 3.5 ergibt sich

$$C_{\mathrm{tr},2(Z_a-1)+\kappa_1,2(Z_b-1)+\kappa_2} = \int_T \frac{\partial \varphi_{Z_a}}{\partial x_{\kappa_1}} \frac{\partial \varphi_{Z_b}}{\partial x_{\kappa_2}} \,\mathrm{d}x$$

und damit

$$C_{\mathrm{tr}} = |T|(S_1^\top G J_{1,1} G^\top S_1 + S_1^\top G J_{1,2} G^\top S_2 + S_2^\top G J_{2,1} G^\top S_1 + S_2^\top G J_{2,2} G^\top S_2).$$

3.1.4 Block F

Da im Block F kein $\mathbb{C}$ Term berechnet werden muss, wird Lemma 2.3 hier nicht angewendet. Stattdessen kann gleich Lemma 3.3 angewendet werden. Die Integrale über die Komponenten von f werden später über Quadraturverfahren approximiert. Damit ergibt sich

$$F_{2(E_\alpha-1+\kappa_1)} = -\int_T f \cdot \operatorname{div} \Psi(E_\alpha, \kappa) \,\mathrm{d}x = -\zeta_{E_\alpha} \frac{|E_\alpha|}{|T|} \int_T f_{\kappa_1} \,\mathrm{d}x.$$

Mit einer Abwandlung von Bemerkung 3.4 und der Matrix L ergibt sich

$$F = \big(- S_1^\top L \int_T f_1 \,\mathrm{d}x - S_2^\top L \int_T f_2 \,\mathrm{d}x\big)/|T|.$$

3.2 Least-Squares-Residuen und Berechnung exakter Fehler

3.2.1 Least-Squares-Residuum

Für einen adaptiven Algorithmus werden lokale Schätzungen der Fehlerbeiträge $\|x - x_h\|_X$ benötigt. Ein Vorteil der Least-Squares-Methoden ist die einfache Bestimmung von effizienten und zuverlässigen Schätzungen dieser Fehlerbeiträge durch die lokalen Least-Squares-Residuen. Zur Berechnung werden lediglich dieselben lokalen Matrizen A, B, C und F benötigt, die schon für die Berechnung der Lösung nötig waren.

Im Folgenden seien mit $x_h = (\sigma_h, u_h)$ die Koeffizienten der zu einem Dreieck T gehörenden Basisfunktion der Räume $X_h = \Sigma_h \times U_h$ bezeichnet. Dabei ist es nicht notwendig, zwischen der konformen und den nicht-konformen Diskretisierungen von X zu unterscheiden.

Die beiden Residuen sollen einzeln berechnet werden, um später verglichen werden zu

können. Es gilt

$$\begin{aligned}
\mathrm{LS}_{\mathrm{div}}(f;\sigma_{\mathrm{LS}}) =& \|f+\mathrm{div}(\sigma_{\mathrm{LS}})\|^2 = \|f\|^2 + 2\int_T f\cdot\mathrm{div}(\sigma_{\mathrm{LS}})\,\mathrm{d}x + \|\mathrm{div}(\sigma_{\mathrm{LS}})\|^2\\
=& \|f\|^2 + 2\sum_{\alpha=1}^{3}\sum_{\kappa_1=1}^{2} x_{h,2(\alpha-1)+\kappa_1}\int f\cdot\mathrm{div}(\Psi_{E_\alpha,\kappa_1})\,\mathrm{d}x\\
&+\sum_{\alpha,\beta=1}^{3}\sum_{\kappa_1,\kappa_2=1}^{2} x_{h,2(\alpha-1)+\kappa_1}\Big[\int_T \mathrm{div}(\Psi_{E_\alpha,\kappa_1})\cdot\mathrm{div}(\Psi_{E_\beta,\kappa_2})\,\mathrm{d}x\Big]x_{h,2(\beta-1)+\kappa_2}\\
=& \|f\|^2 - 2\sigma_h^\top F + \sigma_h^\top A_{\mathrm{div}}\sigma_h,
\end{aligned}$$

$$\begin{aligned}
\mathrm{LS}_{\mathbb{C},s}(\sigma_{\mathrm{LS}},u_{\mathrm{LS}}) =& \left\|\mathbb{C}^s\sigma_{\mathrm{LS}} - \mathbb{C}^{1+s}\varepsilon(u_{\mathrm{LS}})\right\|^2\\
=& \|\mathbb{C}^s\sigma_{\mathrm{LS}}\|^2 + 2\int_T \mathbb{C}^s\sigma_{\mathrm{LS}} : \mathbb{C}^{1+s}\varepsilon(u_{\mathrm{LS}})\,\mathrm{d}x + \left\|\mathbb{C}^{1+s}\varepsilon(u_{\mathrm{LS}})\right\|^2\\
=& \sigma_h^\top A_{\mathbb{C}}\sigma_h + \sigma_h^\top B u_h + u_h^\top B\sigma_h + u_h^\top C u_h.
\end{aligned}$$

Hierbei ist $A_{\mathbb{C}} := \mathbb{c}_1(2s)A_{L^2} + \mathbb{c}_2(2s)A_{\mathrm{tr}}$. Die L^2-Norm der Funktion f kann über ein Quadraturverfahren errechnet werden.

Lemma 3.6 (Zuverlässigkeit und Effizienz des LS-Schätzers)**.** *Für eine konforme Diskretisierung X_h des Raumes X, die Lösung x und ihre Approximation x_h gilt*

$$\begin{aligned}
\beta(s,\lambda,\mu)^2\|x-x_h\|_X^2 = \beta(s,\lambda,\mu)^2\|(\sigma-\sigma_h,u-u_h)\|_X^2 &\leqslant \mathrm{LS}_s(f;\sigma_h,u_h)\\
&\leqslant \alpha(s,\lambda,\mu)^2\|(\sigma-\sigma_h,u-u_h)\|_X^2 = \alpha(s,\lambda,\mu)^2\|x-x_h\|_X^2.
\end{aligned}$$

Beweis. Mit der Rechnung

$$\begin{aligned}
\mathrm{LS}_s(0;\sigma-\sigma_h,u-u_h) =& \|\,\mathrm{div}(\sigma-\sigma_h)\|_{L^2(\Omega,\mathbb{R}^2)}^2 + \|\mathbb{C}^s(\sigma-\sigma_h) - \mathbb{C}^{1+s}\varepsilon(u-u_h)\|_{L^2(\Omega,\mathbb{R}^{2\times 2})}^2\\
=& \|f-\mathrm{div}\,\sigma_h\|_{L^2(\Omega;\mathbb{R}^2)}^2 + \|\mathbb{C}^s\sigma_h - \mathbb{C}^{1+s}\varepsilon(u_h)\|_{L^2(\Omega;\mathbb{R}^{2\times 2})}^2\\
=& \mathrm{LS}_s(f;\sigma_h,u_h)
\end{aligned}$$

folgt die Behauptung sofort aus der Wohlgestelltheit aus Satz 2.15. □

Aus den Überlegungen zu den Konstanten α, und β aus Satz 2.15 lässt sich wiederum ableiten, dass die Zuverlässigkeits- und Effizienz-Eigenschaft für $s \neq -1$ kritisch vom Parameter λ abhängt. Für $s \geqslant -1$ und $\lambda \to \infty$ wird die Abschätzung $LS_s(f;\sigma_h,u_h) \leqslant \alpha(s\lambda,\mu)^2\|x-x_h\|_X$ wertlos, also geht die Effizienz-Eigenschaft verloren. Für $s \leqslant -1$ und $\lambda \to \infty$ wird die Abschätzung $\beta(s\lambda,\mu)^2\|x-x_h\|_X \leqslant \mathrm{LS}_s(f;\sigma_h,u_h)$ wertlos, also geht die Zuverlässigkeits-Eigenschaft verloren.

3.2.2 Exakte Fehler

Für einige wenige Beispiele ist die exakte Lösung des Systems erster Ordnung bekannt. An dieser Stelle soll noch kurz auf die Berechnung von exakten Fehlertermen eingegangen werden, welche zur Validierung von Implementationen der vorgestellten Methoden sehr hilfreich sind. Die Komponenten des Approximationsfehlers werden einzelnen berechnet um später genauere Aussagen über das Konvergenzverhalten zu erhalten. Die Fehlerkomponenten sind gegeben durch

$$\begin{aligned}\|x - x_{\mathrm{LS}}\|_X^2 &= \|u - u_{\mathrm{LS}}\|_{H^1(\Omega;\mathbb{R}^2)}^2 + \|\sigma - \sigma_{\mathrm{LS}}\|_{H(\mathrm{div},\Omega;\mathbb{R}^{2\times 2})} \\ &= \|u - u_{\mathrm{LS}}\|_{L^2(\Omega;\mathbb{R}^2)}^2 + \|D(u - u_{\mathrm{LS}})\|_{L^2(\Omega;\mathbb{R}^2)}^2 \\ &\quad + \|\sigma - \sigma_{\mathrm{LS}}\|_{L^2(\Omega;\mathbb{R}^{2\times 2})} + \|\operatorname{div}(\sigma - \sigma_{\mathrm{LS}})\|_{L^2(\Omega;\mathbb{R}^{2\times 2})}.\end{aligned}$$

Die Fehler zu der Lösungskomponente u_{LS} können durch Methoden des *AFEM*-Paketes berechnet werden. Diese sind in [Car+10, Kapitel 1.8.3] beschrieben. Da $\operatorname{div}\sigma = -f$ gilt, kann der Divergenz-Fehleranteil von σ_{LS} durch

$$\|\operatorname{div}(\sigma - \sigma_{\mathrm{LS}})\|_{L^2(\Omega;\mathbb{R}^{2\times 2})}^2 = \|f + \operatorname{div}\sigma_{\mathrm{LS}}\|_{L^2(\Omega;\mathbb{R}^{2\times 2})}^2 = \mathrm{LS}_{\mathrm{div}}(f;\sigma_{\mathrm{LS}})$$

wie in vorherigen Abschnitt berechnet werden. Der L^2-Anteil wird aufgespalten in,

$$\|\sigma - \sigma_{\mathrm{LS}}\|_{L^2(\Omega;\mathbb{R}^{2\times 2})}^2 = \|\sigma\|_{L^2(\Omega;\mathbb{R}^{2\times 2})}^2 - 2\int_\Omega \sigma : \sigma_{\mathrm{LS}}\mathrm{d}x + \|\sigma_{\mathrm{LS}}\|_{L^2(\Omega;\mathbb{R}^{2\times 2})}^2.$$

Die Norm der exakten Spannung σ kann dann wieder über die Quadraturverfahren aus dem *AFEM*-Paket berechnet werden und die Norm von σ_{LS} berechnet sich wie oben durch $\sigma_h^\top A_{L^2}\sigma_h$. Der verbleibende gemischte Term kann wie in in der Arbeit [Bri13, Abschnitt 4.3.3], welche wiederum Berechnungen aus [Gal12] nutzt.

3.2.3 Energiefehler

Für eine weitere Betrachtung soll noch der Fehler in der problemeingenen Energienorm $|||u - u_{LS}|||$ berechnet werden, d.h.

$$\begin{aligned}|||u - u_{LS}|||_T^2 &= |||u|||_T^2 - 2\int_T \mathbb{C}\varepsilon(u) : \varepsilon(u_{LS})\,\mathrm{d}x + |||u_{LS}|||_T^2 \\ &= |||u|||_T^2 - 2\varepsilon(u_{LS}) : \int_T \sigma\,\mathrm{d}x + |||u_{LS}|||_T^2.\end{aligned}$$

Dabei wird das Integral im letzten Schritt komponentenweise verstanden. Wiederum kann die Energienorm der exakten Funktion u über Quadraturmethoden errechnet werden. Da $\varepsilon(u_{\mathrm{LS}})$ einen stückweise konstante Funktion ist, können die restlichen Terme leicht berechnet werden.

3.3 Locking in konformen Diskretisierungen

Das Locking-Verhalten konformer Diskretisierungen niedriger Ordnung beschreibt ein schlechtes vorasymtotisches Konvergenzverhalten in Abhängigkeit des Materialparameters λ. Eine Intuition dafür kann durch die elliptischen regularitäts Abschätzungen für klassische Galerkinansätze gewonnen werden. In [BS02, Abschnitt 11.4] wird für das reine Verschiebugsproblem mit homogenen Dirichletranddaten

$$\begin{aligned} 2\mu\Delta u + 2\lambda\nabla \operatorname{div} u &= -f \qquad \text{in } \Omega \\ u &= 0 \qquad \text{auf } \Gamma_D \end{aligned}$$

die Abschätzung

$$\|u\|_{H^2(\Omega;\mathbb{R}^2)} + \lambda\| \operatorname{div} u\|_{H^1(\Omega)} \leqslant C\|f\|_{L^2(\Omega;\mathbb{R}^2)}$$

bewiesen. Für sehr große Werte von λ ist also zwangsläufig $\| \operatorname{div} u\|_{H^1(\Omega)}$ sehr klein. Für den Grenzwert $\lambda \to \infty$ folgt $\| \operatorname{div} u\|_{H^1(\Omega)} \to 0$, woraus dann weiter $\operatorname{div} u = 0$ folgt. Für den oben vorgestellten diskreten Funktionenraum $S_0^1(\mathcal{T};\mathbb{R}^2)$ und eine Approximation der Verschiebung $u_h \in S_0^1(\mathcal{T};\mathbb{R}^2)$ folgt aber aus $\operatorname{div} u_h = 0$ schon $u_h = 0$. Diese Eigenschaft des Approximationsraums $S_0^1(\mathcal{T};\mathbb{R}^2)$ kann am folgenden einfachen Beispiel nachvollzogen werden. Die Triangulierung $\mathcal{T}$ sei durch die Skizze 3.1 gegeben.

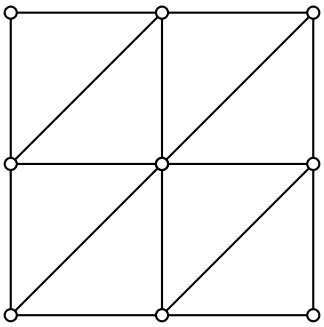

Abbildung 3.1: rotverfeinerte criss-Triangulierung des Einheitsquadrats

Für jedes Dreieck stehen $3 \cdot 2 = 6$ Freiheitsgrade zur Approximation der Verschiebung zur Verfügung. Da $|T| = 8$ ergeben sich daraus $8 \cdot 6 = 48$ potentielle Freiheitsgrade. Durch Stetigkeitsbedingungen entfallen $2 \cdot 15 = 30$, und durch Randbedingungnen $2 \cdot 8 = 16$, da $|\mathcal{N}(\partial\Omega)| = 8$ Randknoten gibt. Daher bleiben $48 - 30 - 16 = 2$ echte Freiheitsgrade für die Approximation.

Wenn nun aber $\operatorname{div} u_h = 0$ auf jedem Dreieck gelten soll, dann reduzieren sich die lokalen Freiheitsgrade. Wird die Funktion u_h lokal durch

$$u_h = \begin{pmatrix} a + bx + cy \\ d + ex + fy \end{pmatrix}$$

dargestellt, dann ergibt sich $0 = \operatorname{div} u_h = b + f$, und es verbleiben nur 5 lokale Freiheitsgrade. Die obere Überlegung liefert dann

$$40 - 30 - 16 = -6 < 0,$$

sodass es keine echten Freiheitsgrade zur Approximation gibt. Damit ist für eine solche Formulierung des Problems der linearen Elastizität u_h eine denkbar schlechte Approximation der tatsächlichen Verschiebung u.

In [BS02, Teorem 11.3.5] wird das Locking-Verhalten folgender Maßen formalisiert.

Satz 3.7 (Locking des reinen Verschiebungsproblems)**.** *Für $u \in H^2(\Omega;\mathbb{R}^2)$ als Lösung des reinen Verschiebungsporblems mit homogenen Dirichletranddaten und $u_C \in S^1_0(\mathcal{T};\mathbb{R}^2)$ als Lösung des diskreten Problems gilt*

$$\|u - u_C\|_{H^1(\Omega;\mathbb{R}^2)} \leqslant C_{(\mu,\lambda)} h \|u\|_{H^2(\Omega;\mathbb{R}^2)}.$$

Die Abhängigkeit der Konstante $C_{(\mu,\lambda)}$ von λ in der apiori-Abschätzung heißt, dass für jede Netzweite h ein λ existiert, sodass die Abschätzung und damit die Approximation noch beliebig schlecht ist.

Da für die Konstruktion der vorgestellten Least-Squares-Methoden auch der Raum $S^1(\mathcal{T};\mathbb{R}^2)$ verwendet wurde, stellt sich die Frage, ob und welche Fehleranteile ebenfalls eine schlechte vorasymtotische Konvergenz aufzeigen.

Sei u_C die $S^1(\mathcal{T};\mathbb{R}^2)$ Galerkinlösung zu u bzgl. des Energieskalarproduktes, sodass durch die Galerkinorthogonalität $|||u - u_C||| \leqslant |||u - v_C|||$ für alle $v_C \in S^1(\mathcal{T};\mathbb{R}^2)$ gilt. Dann folgt aus der Zuverlässigkeits-Eigenschaft aus Lemma 3.6, der Normäquivalenz zur Energienorm aus Bemerkung 2.14 und dem Locking Theorem 3.7, dass

$$\begin{aligned}
\mathrm{LS}(f;\sigma_{LS},u_{LS}) &\geqslant \beta(s,\lambda,\mu)^2(\|\sigma - \sigma_{LS}\|^2_{H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2})} + \|u - u_{LS}\|^2_{H^1(\Omega;\mathbb{R}^2)}) \\
&\geqslant \beta(s,\lambda,\mu)^2(\|\sigma - \sigma_{LS}\|^2_{H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2})} + (2\mu+2\lambda)^{-1}|||u - u_{LS}|||^2) \\
&\geqslant \beta(s,\lambda,\mu)^2(\|\sigma - \sigma_{LS}\|^2_{H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2})} + (2\mu+2\lambda)^{-1}|||u - u_C|||^2) \\
&\geqslant \beta(s,\lambda,\mu)^2(\|\sigma - \sigma_{LS}\|^2_{H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2})} \\
&\qquad + (2\mu+2\lambda)^{-1}(2\mu)c_{\mathrm{Korn}}^{-1/2}\|u - u_C\|^2_{H^1(\Omega;\mathbb{R}^2)}) \\
&\geqslant \beta(s,\lambda,\mu)^2(\|\sigma - \sigma_{LS}\|^2_{H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2})} \\
&\qquad + (2\mu+2\lambda)^{-1}(2\mu)c_{\mathrm{Korn}}^{-1/2}C^2_{(\mu,\lambda)}h^2\|u\|^2_{H^2(\Omega;\mathbb{R}^2)}).
\end{aligned}$$

Aus dieser Abschätzung kann man erkennen, dass selbst wenn $\beta(s,\lambda,\mu)$ linear von λ abhängig ist, sich nicht unbedingt ein Locking-Verhalten im Least-Squares-Funktional beobachten lässt.

3.4 Kouhia-Stenberg Diskretisierung der Verschiebung

Definition 3.8 (Crouzeix-Raviart-Finite-Elemente-Raum)**.** *Zu jedem Mittelpunkt einer Kante $E \in \mathcal{E}$ der Triangulierung $\mathcal{T}$ kann eine stückweise affine Funktion $\xi_E : \Omega \to \mathbb{R}$ definiert werden, die genau in diesem Mittelpunkt den Wert Eins und in allen anderen den Wert Null annimmt*

$$\xi_E(x) := \begin{cases} 1 & , \textit{falls } x = \mathrm{mid}(E) \\ 0 & , \textit{falls } x \in \mathrm{mid}(\mathcal{E}) \backslash \{\mathrm{mid}(E)\} \end{cases}.$$

Die Linearkombinationen dieser Funktionen bilden den Raum $CR^1(\mathcal{T})$.

Bemerkung 3.9. *Auf jedem Dreieck T der Triangulierung $\mathcal{T}$ sind sowohl alle ϕ_Z als auch alle ξ_E eindeutig bestimmt. Dies erlaubt die direkten Zusammenhänge*

$$\xi_{E_j} = 1 - 2\phi_{Z_{j-1}} \quad \textit{und} \quad \frac{\partial \xi_{E_j}}{\partial x_\kappa} = -2\frac{\partial \phi_{Z_{j-1}}}{\partial x_\kappa} \quad \textit{für alle } j = 1, 2, 3.$$

Dieser Zusammenhang kann für die Matrix G der partiellen Ableitungen der ϕ_{Z_j} in der Matrix D zusammengefasst werden.

Lemma 3.10. *Für die Matrizen*

$$G_{CR} := \begin{bmatrix} \frac{\partial \xi_{E_1}}{\partial x_1} & \frac{\partial \xi_{E_2}}{\partial x_1} & \frac{\partial \xi_{E_3}}{\partial x_1} \\ \frac{\partial \xi_{E_1}}{\partial x_2} & \frac{\partial \xi_{E_2}}{\partial x_2} & \frac{\partial \xi_{E_3}}{\partial x_2} \end{bmatrix}^\top \qquad \textit{und} \qquad D = \begin{bmatrix} 0 & 0 & -2 \\ -2 & 0 & 0 \\ 0 & -2 & 0 \end{bmatrix}$$

gilt $G_{CR} = DG$ bzw. $G_{CR}^\top = G^\top D^\top$.

Auf Basis des Crouiex-Raviart-Raumes, soll eine nicht-konforme Diskretisierung untersucht werden. Allerdings werden nicht beide Komponenten der Verschiebungsvariablen u nicht-konform approximiert, d.h. $U_h := CR^1(\mathcal{T}, \mathbb{R}^2)$, da für diese Approximation keine diskrete Kornsche Ungleichung existiert und somit das diskrete Problem nicht wohlgestellt ist. Stattdessen wird lediglich die zweite Komponente nicht-konform, die andere konform approximiert, d.h. $U_h := S^1(\mathcal{T}) \times CR^1(\mathcal{T})$. Basierend auf der diskreten variationellen Formulierung 7 können wieder Basisdarstellungen der Lösung notiert werden. Da sich im Vergleich zur ersten, konformen Diskretisierung nur die Approximation von u ändert, wird nur diese hier notiert und auch im folgenden betrachtet. Es gilt

$$u_h = \sum_{Z=1}^{|\mathcal{N}|} u_Z \varphi_Z + \sum_{E=1}^{|\mathcal{E}|} u_{|\mathcal{N}|+E} \psi_E = \sum_{Z=1}^{|\mathcal{N}|} x_{|2\mathcal{E}|+Z} \varphi_Z + \sum_{E=1}^{|\mathcal{E}|} x_{|2\mathcal{E}|+|\mathcal{N}|+E} \psi_E.$$

Zur Implementierung der Blöcke im Fall der Kouhia-Stenberg Diskretisierung werden die Vertauschungen mit der Matrix D in den Komponenten eingesetzt, in denen eine

Crouzeix-Raviart Diskretisierung zum Einsatz kommt. In den Blöcken B_1 und B_{tr} sind das die von $S_1^\top$ und S_2, sowie die von $S_2^\top$ und S_2 eingeschlossenen Blöcke. Es ergeben sich also mit den obigen Berechnungen für die konforme Diskretisierung

$$\begin{aligned} B_1 = \; & S_1^\top L N^\top H_{1,1} G^\top S_1/6 \quad + S_2^\top L N^\top H_{2,2} D G^\top S_2/6 \\ & + S_1^\top L N^\top H_{2,2} G^\top S_1/12 + S_1^\top L N^\top H_{2,1} D G^\top S_2/12 \\ & + S_2^\top L N^\top H_{1,2} G^\top S_1/12 + S_2^\top L N^\top H_{1,1} D G^\top S_2/12 \end{aligned}$$

$$\begin{aligned} B_{\text{tr}} = \; & S_1^\top L N^\top H_{1,1} G^\top S_1/6 \; + S_1^\top L N^\top H_{1,2} G^\top D^\top S_2/6 \\ & + S_2^\top L N^\top H_{2,1} G^\top S_1/6 + S_2^\top L N^\top H_{2,2} G^\top D^\top S_2/6. \end{aligned}$$

Für die Blöcke C_{L^2} und C_{tr} handelt es sich um alle Blöcke die nicht von $S_1^\top$ und S_1 eingeschlossen sind. Daher ergibt sich

$$\begin{aligned} C_{L^2} = & |T|(S_1^\top G J_{1,1} G^\top S_1 + S_2^\top D G J_{2,2} G^\top D^\top S_2 \\ & + S_1^\top G J_{2,2} G^\top S_1/2 + S_1^\top G J_{2,1} G^\top D^\top S_2/2 \\ & + S_2^\top D G J_{1,2} G^\top S_1/2 + S_2^\top D G J_{1,1} G^\top D^\top S_2/2) \end{aligned}$$

$$\begin{aligned} C_{\text{tr}} = & |T|(S_1^\top G J_{1,1} G^\top S_1 + S_1^\top G J_{1,2} G^\top D^\top S_2 \\ & + S_2^\top D G J_{2,1} G^\top S_1 + S_2^\top D G J_{2,2} G^\top D^\top S_2). \end{aligned}$$

3.5 Diskretisierung der Formulierung mit exakter Kontinuitätsgleichung

Die zum Ende des Abschnitts 2.3 hergeleitete variationelle Formulierung des Problems 6 kann erneut als diskretes Problem formuliert werden. Die benötigten konformen Unterräume $RT_0(\mathcal{T}; \mathbb{R}^{2\times 2})$ und $S^1(\mathcal{T}; \mathbb{R}^2)$ seien dabei genau so gegeben, wie in den vorherigen Abschnitten. Für die Diskretisierung des Lagrange-Paramters wird nun noch ein weiterer konformer Unterraum $L_h := P_0(\mathcal{T}; \mathbb{R}^2)$ von $L^2(\Omega; \mathbb{R}^2)$ gewählt. Der Raum $P_0(\mathcal{T}; \mathbb{R})$ sei dabei von den charakteristischen Funktionen χ_k der Dreiecke T_k aufgespannt. Der Raum der Funktionen mit Werten in $\mathbb{R}^2$ wird dann mit den Basisfunktionen $X_{k,\kappa} := \chi_k e_\kappa$ durch

$$P_0(\mathcal{T}; \mathbb{R}^2) = \operatorname{span}\left\{ \begin{pmatrix} \chi_1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \chi_1 \end{pmatrix}, \begin{pmatrix} \chi_2 \\ 0 \end{pmatrix}, \dots \begin{pmatrix} 0 \\ \chi_{|\mathcal{T}|} \end{pmatrix} \right\}$$

gebildet. Das diskrete Problem lautet mit diesen Bezeichnungen dann wie folgt.

Problem 8 (Diskrete Variationelle Formulierung des Lagrange-Funktionals)**.** *Zu* $f \in$

$L^2(\Omega;\mathbb{R}^2)$ *finde* $x_h := (\sigma_h, u_h, \lambda_h^L) \in \Sigma_h \times U_h \times L_h = X_h^L$, *sodass für alle* $(\tau_h, v_h, \mu^L) \in X_h^L$

$$\mathcal{B}^L\big((\sigma_h, u_h, \lambda^L), (\tau_h, v_h, \mu^L)\big) = F^L\big((\tau_h, v_h, \mu^L)\big).$$

Die Lösung des diskreten Problems besitzt damit die Basisdarstellung

$$\begin{aligned}
\sigma_h &= \sum_{E=1}^{|\mathcal{E}|}\sum_{\kappa=1}^{2} \sigma_{2(E-1)+\kappa}\Psi_{E,\kappa} = \sum_{E=1}^{|\mathcal{E}|}\sum_{\kappa=1}^{2} x_{2(E-1)+\kappa}\Psi_{E,\kappa} \\
u_h &= \sum_{Z=1}^{|\mathcal{N}|}\sum_{\kappa=1}^{2} u_{2(Z-1)+\kappa}\Phi_{Z,\kappa} = \sum_{Z=1}^{|\mathcal{N}|}\sum_{\kappa=1}^{2} x_{|\mathcal{E}|+2(Z-1)+\kappa}\Phi_{Z,\kappa} \\
\lambda_h^L &= \sum_{K=1}^{|\mathcal{N}|}\sum_{\kappa=1}^{2} \lambda^L_{2(K-1)+\kappa}X_{K,\kappa} = \sum_{K=1}^{|\mathcal{N}|}\sum_{\kappa=1}^{2} x_{|\mathcal{E}|+|\mathcal{N}|+2(K-1)+\kappa}X_{K,\kappa}
\end{aligned}$$

mit den Koeffizienten $\sigma_i, u_j, \lambda_k^L$ für $i = 1,\dots,2|\mathcal{E}|, j = 1,\dots,2|\mathcal{N}|, k = 1,\dots 2|\mathcal{T}|$ und ist gleichbedeutend mit der Lösung des linearen Gleichungssystems

$$\begin{bmatrix} A & B & D \\ B^\top & C & 0 \\ D^\top & 0 & 0 \end{bmatrix} \begin{pmatrix} \sigma \\ u \\ \lambda^L \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ F^L \end{pmatrix}$$

mit den lokalen Blöcken

$$\begin{aligned}
A_{\mathbb{C},2(E_\alpha-1)+\kappa_1,2(E_\beta-1)+\kappa_2} &= \int_T \mathbb{C}^s\Psi_{E_\alpha,\kappa_1} : \mathbb{C}^s\Psi_{E_\beta,\kappa_2}\mathrm{d}x \\
B_{2(E_\alpha-1)+\kappa_1,2(Z_b-1)+\kappa_2} &= -\int_T \mathbb{C}^s\Psi_{E_\alpha,\kappa_1} : \mathbb{C}^{1+s}\varepsilon(\Phi_{Z_b,\kappa_2})\mathrm{d}x \\
C_{2(Z_a-1)+\kappa_1,2(Z_b-1)+\kappa_2} &= \int_T \mathbb{C}^{1+s}\varepsilon(\Phi_{Z_a,\kappa_1}) : \mathbb{C}^{1+s}\varepsilon(\Phi_{Z_b,\kappa_2})\mathrm{d}x \\
D_{2(E_\alpha-1)+\kappa_1,2(T-1)+\kappa_2} &= \int_T X_{T,\kappa_2}\cdot \mathrm{div}(\Psi_{E_\alpha,\kappa_1})\mathrm{d}x \\
F^L_{2(T-1)+\kappa} &= -\int_T X_{T,\kappa_1}\cdot f\mathrm{d}x.
\end{aligned}$$

Dabei sind E_α, E_β, E_γ die Kanten und Z_a, Z_b und Z_c die Knoten des Dreiecks. Die Berechnung der Blöcke $A_\mathbb{C}$, B und C folgt dann, wie oben bereits beschrieben, wobei wieder $A_\mathbb{C} := \mathbb{c}_1(2s)A_{L^2} + \mathbb{c}_2(2s)A_{\mathrm{tr}}$ gilt.

3.5.1 Blöcke D und $\mathbf{F^L}$

Die Berechnung des Blockes D ergibt sich mit Lemma 3.3 durch

$$D_{2(E_\alpha-1)+\kappa_1,2(T-1)+\kappa_2} = \int_T X_{T,\kappa_2} \cdot \operatorname{div}(\Psi_{E_\alpha,\kappa_1})\mathrm{d}x = \delta_{\kappa_1,\kappa_2}\zeta_{E_\alpha}|E_\alpha|.$$

Mit Bemerkung 3.4 und den Matrizen $H_{j,k}$ kann D dann durch

$$D = S_1^\top L S_1 H_{1,1} + S_2^\top L S_2 H_{2,2}$$

dargestellt werden. Der Block F^L ergibt sich einfach als

$$F^L = -\int_T f\mathrm{d}x \begin{pmatrix}1\\1\end{pmatrix}.$$

4 Dokumentation der Software

Im Folgenden sollen die zu dieser Arbeit gehörenden Programme vorgestellt werden. Dabei wird zunächst die Steuerung der Hauptmethoden *mAfemEx* und *afemElasticityLS* erläutert und anschließend auf wird die wichtigsten Methoden des AFEM-Zyklus einzeln eingegangen. Im Vordergrund stehen dabei eine Beschreibung der Ein- und Ausgaben dieser Methoden sowie der Besonderheiten, die den vorangegangen Abschnitten nicht entnommen werden können.

Die Programme basieren auf dem Softwarepaket der Arbeitsgruppe "Numerische Analysis" von Prof. Carstensen, welches in [Car+10] dokumentiert ist, daher ist dieses für die Funktionsweise unabdingbar. Entwickelt wurde für die *MATLAB*-Version *R2013a*.

4.1 Hauptmethoden und Konfigurationsdateien

Dieser Abschnitt erklärt, wie die zur Arbeit gehörenden Programme gestartet und konfiguriert werden können. Es werden dafür die beiden Methoden erläutert, die zum Programmstart ausgeführt werden müssen, so wie die Syntax der Konfigurationsdateien und die Methoden, welche diese auslesen.

4.1.1 afemElasticityLS

Die Methode *afemElasticityLS(configFile)* steuert die AFEM-Schleife entsprechend der in der Konfigurationsdatei *configFile* angegebenen Parameter. Findet keine Übergabe einer Kofigurationsdatei an *afemElasticityLS* statt, so startet die Bearbeitung einer Default-Kofiguration. Die Parameter legen die Diskretisierung, die Abbruchbedingungen, die Ausgabe und vieles mehr fest. Die möglichen Werte und deren Einfluss auf den Programmablauf werden in Tabellen erläutert.

Die Konfigurationsdatei wird als String übergeben und muss sich entweder im Ordner *configs/* befinden, oder es muss ein relativer Pfad von diesem Ordner aus zur Datei angegeben werden.

Die Methode wird im Laufe der Ausführung Ausgaben zum aktuellen Stand der Berechnung machen. Dabei werden das aktuelle Verfeinerungslevel, die Anzahl der Freiheitsgrade, das errechnete Least-Squares-Residuum, die zur Berechnung benötigte Zeit sowie eventuell der exakte Fehler ausgegeben und in Relation zu den Abbruchkriterien gesetzt.

4.1.2 mAfemEx

Die Funktion *mAfemEx(mConfigFile)* erlaubt es, mehrere Durchläufe von *afemElasicityLS* automatisch zu starten. Der Funktion kann dafür ein *mConfigFile* übergeben werden, welches alle durchzuführenden Durchläufe spezifiziert. Wird keine Konfigurationsdatei übergeben, so wird eine default-Datei ausgewählt.

Die Konfigurationsdatei wird als String übergeben und muss sich entweder im Ordner *configs/* befinden, oder es muss ein relativer Pfad von diesem Ordner aus zur Datei angegeben werden.

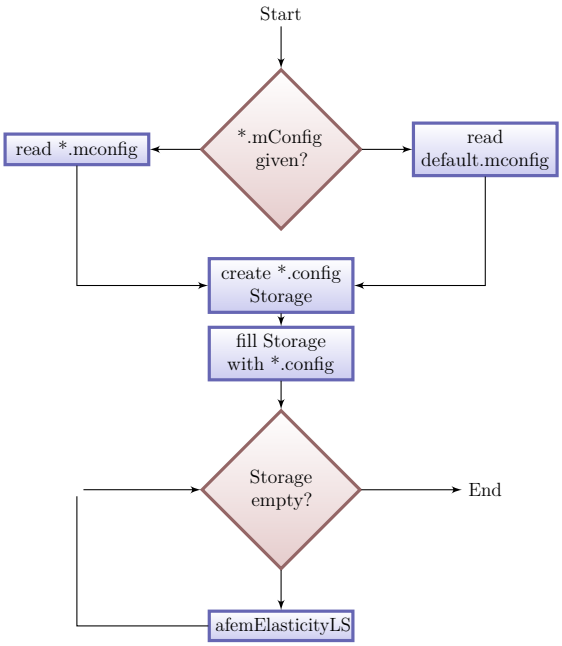

Abbildung 4.1: mAfemEx flowchart

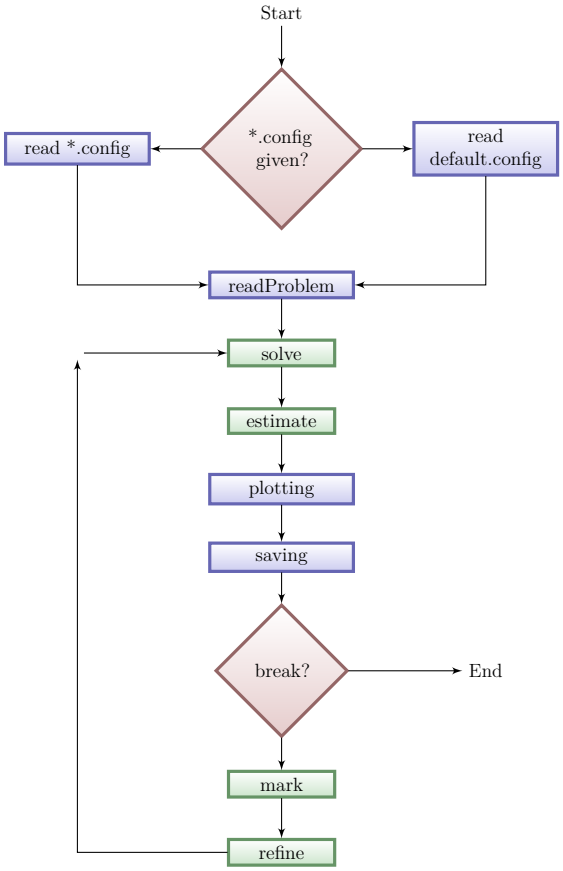

Abbildung 4.2: afemElasticityLS flowchart

4.1.3 Konfigurationsdateien

Die Syntax der Konfigurationsdateien soll im Folgenden beschrieben werden. Dabei unterscheidet sich die Syntax der Konfigurationsdateien für *afemElasticityLS* nicht wesentlich von der Syntax der Konfigurationsdateien für *mAfemEx*.

In der erst genannten sind die Parameter zeilenweise aufgezählt und durch ein Semikolon von ihrem Wert getrennt. Der Wert eines Parameters wird wieder durch ein Semikolon beendet. Auf diese folgt eine Beschreibung des Parameters und seiner Werte.

Für die **.mconfig*-Dateien können nach dem Parameternamen durch Kommata getrennt mehrere Werte für einen Parameter übergeben werden.

parameterName1; value11, value12, ... value1M; discription1
parameterName2; value21, value22, ... value2M; discription2
⋮
parameterNameN; valueN1, valueN2, ... valueNM; discriptionN

Abbildung 4.3: Schema einer Konfigurationsdatei für *mAfemExx*

4.1.4 Die Methoden readConfig und readMconfig

Die Methode *readMconfig* wird von *mAfemEx* aufgerufen, liest eine **.mconfig*-Datei und erstellt dann entsprechend aller Kombinationen an Werten der Parameter Konfigurationsdateien. Das heißt, werden zum Beispiel für den Vergleich von zwei unterschiedlichen Diskretisierungen die Werte *S1* und *KS* für den Parameter *method* angegeben, sowie für den Parameter *problemName* die Werte *LR* und *SQ* gesetzt, dann erstellt die Methode die Konfigurationsdateien

1. *c_S1_LR.config*, 2. *c_S1_SQ.config*, 3. *c_KS_LR.config* und 4. *c_KS_SQ.config*.

Die Methode *mAfemEx* ruft dann für jede Konfigurationsdatei einmal die Methode *afemElasicityLS* auf. Diese ruft wiederum die Methode *readConfig* auf, welche dann die Parameter für den aktuellen Durchlauf aus der Konfigurationsdatei ausliest.

Die nachfolgende Tabelle 4.1 dient als Legende für die Beschreibungen der Parameter in den folgenden Abschnitten.

Parametername	Beschreibung
mögliche Werte	

Tabelle 4.1: Beschreibung von Parametern

4.1.5 Parameter zur Problemspezifizierung - readProblem

Die Parameter in der Tabelle 4.2 dienen der Problemauswahl. Sie werden von der Methode *readProblem* bearbeitet, welche daraus weitere Datenstrukturen erstellt. Die Aufrufe erfolgen durch

[*parameter, geometryData, problemData, exSolutionData*] = *readProblem(parameter);*.

Insbesondere errechnet die Methode die Werte für λ und μ entsprechend den Zusammenhängen in Abschnitt 1.3, falls statt dieser die Werte für E und ν festgelegt wurden.

problemName	Wählt eine Geometrie, sowie die Randdaten- und Volumenkraftfunktion aus.
{SQ, LR, CI, KI, CW, LL}	
loadGeometry & loadPath	Wenn *loadGeometry* gesetzt ist, wird eine Geometrie verwendet, die sich an *loadPath* befindet.
$\{0, 1\}$	
geoScaling	Wenn *geoScaling* gesetzt ist, wird die in Abschnitt 1.2 beschriebene Skalierung verwendet.
$\{0, 1\}$	
lambda & lame_mu	Legt die Lamé-Parameter λ und μ fest. Wird nichts übergeben, werden die Werte aus E und ν errechnet.
$\mathbb{R}_+ \cup \{\ \}$	
E & nu	Legt den Elastizitätsmodul E und die Querkontraktionszahl ν fest, falls λ und μ nicht übergeben wurden.
$\mathbb{R}_+ \cup \{\ \}$	
muScaling	Wenn *muScaling* gesetzt ist, wird die in Abschnitt 1.3 beschriebene Skalierung verwendet.
$\{0, 1\}$	

Tabelle 4.2: Parameter für die Problemauswahl

4.1.6 globale Steuerungsparameter & Abbruchbedingungen

Die in Tabelle 4.3 aufgeführten Parameter legen Einstellungen fest welche das Verhalten des gesamten Programmablaufs beeinflussen. Dabei stehen die Einflüsse dieser Parameter nicht im theoretischen Fokus dieser Arbeit.

parallelComputing	Wenn *parallelComputing* gesetzt ist, werden die Rechenoperationen auf möglichst viele CPUs verteilt.
$\{0, 1\}$	
profileRun	Wenn *profileRun* gesetzt ist, werden Daten des *MATLAB*-Profilers gespeichert und angezeigt.
$\{0, 1\}$	
intDegree	Legt den Exaktheitsgrad der nötigen Numerischen Integrationen fest.
$\mathbb{N}$	

Tabelle 4.3: Steuerungsparameter

Die Tabelle 4.4 enthält die Abbruchkriterien, nach denen die AFEM-Zyklus abgebrochen werden kann. Der Zyklus wird abgebrochen, sobald mindestens eines der Kriterien erfüllt ist.

4.2 Datenstrukturen

Das Programm fasst einige Variablen in *struct*-Typen zusammen, um eine bessere Lesbarkeit des Programmcodes zu erzielen. Diese strukturierten Datentypen werden den

manualBreak	Wenn *manualBreak* gesetzt ist, wird nach jedem Durch-
$\{0, 1\}$	lauf erfragt, ob die Schleife abgebrochen werden soll.
minDof	Die Afem-Loop wird nach dem Durchlauf mit mindes-
$\mathbb{N} \cup \{\infty\}$	tens *minDof* Freiheitsgraden abgebrochen.
minLevel	Die Afem-Loop wird nach dem Durchlauf *minLevel*
$\mathbb{N} \cup \{\infty\}$	abgebrochen.
minTime	Die Afem-Loop wird abgebrochen, wenn insgesamt
$\mathbb{R}_+ \cup \{\infty\}$	mindestens *minTime* Sekunden vergangen sind.
minRes	Die Afem-Loop wird abgebrochen, wenn das Residuum
$\mathbb{R}_+ \cup \{\infty\}$	kleiner als *minRes* ist.
minErr	Die Afem-Loop wird abgebrochen, wenn der exakte
$\mathbb{R}_+ \cup \{\infty\}$	Fehler kleiner als *minRes* ist.

Tabelle 4.4: Abbruchbedingungen

nachfolgend aufgeführten Methoden übergeben. Sollten diese also unabhängig von *afemElasticityLS* aufgerufen werden, so müssen die Strukturen erst erstellt werden.

4.2.1 geometryData

Die Struktur *geometryData* enthält alle Informationen zur aktuellen Geometrie. Sie wird normalerweise von der Methode *readProblem* erstellt und enthält die acht Felder *c4n, n4e, n4sDb1C, n4sNb1C, n4sDb2C, n4sNb2C, componentBoundary* und *scale*. Dabei sind *c4n* und *n4e* die aus dem AFEM-Paket bekannten Datenstrukturen zur Beschreibung der Geometrie. Außerdem sind *n4sDb1C, n4sNb1C, n4sDb2C, n4sNb2C* an die Datenstruckturen *n4sDb, n4sNb* angelehnte Beschreibungen der Ränder Γ_D und Γ_N, wobei jedoch für jede Komponente der Lösung eine Randdaten-Information benötigt wird. In *componentBoundary* wird als Boolean die Information darüber, ob ein Unterschied der Randadaten in den einzelnen Komponenten existiert, repräsentiert. Die in Abschnitt 1.2 angegebene Skalierung $c \in \mathbb{R}$ wird in dem Eintrag *scale* gespeichert.

4.2.2 problemData

Die Struktur *problemData* enthält alle Informationen über die gegeben Funktionen f, g und t, als Function-Handle. Dabei dürfen alle nur von einer Variablen x der Dimension zwei abhängen und müssen Werte der Dimension zwei zurückgeben. Eventuelle Abhängigkeiten von λ und μ müssen implizit enthalten sein. Außerdem wird das Vorhandensein einer exakten Lösung in der Boolean Variable *exSolKnown* angegeben.

4.2.3 solutionData

Die Struktur *solutionData* wird von den Lösungsmethoden des nächsten Abschnittes erstellt und der Darstellungsfunktion *plotElasticityLS*, der Schätzmethode *computeElasticityLSResidual*, sowie der Fehlerberechnungsmethode *error4eElasticityLS*. übergeben. Neben den eigentlichen Lösungen *u* und *sigma*, als Vektoren der Koeffizienten mit der globalen Nummerierung, enthält die Struktur noch die Einträge *u4e* und *sigma4e*, welche die Koeffizienten spaltenweise in lokaler Nummerierung und zeilenweise für jedes Element beinhalten. Zusätzlich werden in den Lösungsmethoden auch die Ableitungen der Lösungen berechnet und in den Einträgen *Du1CDx4e, Du1CDy4e, Du2CDx4e, Du2CDy4e, gradU1C4e, gradU2C4e, divU4e, epsU4e* und *divSigma4e* elementweise gespeichert. Die Ableitungen sind als Vektoren bezüglich einer Basis aus charakteristischen Funktionen implementiert.

4.2.4 exSolutionData

Falls eine Lösung der Problemstellung existiert, so werden in der Struktur *exSolutionData* die exakten Funktionen *u1Cex, u2Cex, Du1CexDx, Du1CexDy, Du2CexDx, Du2CexDy, sigmaExact, gradUexact1C, gradUexact2C, divUexact* und *epsUex* als Function-Handle vorgehalten. Die Funktionen hängen wieder nur von einer Variablen x der Dimension 2 ab und Abhängigkeiten von den Materialparametern sind implizit.

4.2.5 localData

In der Struktur *localData* werden lokale Informationen wie zum Beispiel die lokalen Steifigkeitsmatrizen, die Flächeninhalte, die Koordinaten, u.ä. gesammelt. Eine komplette Liste enthält: *signs4e, G4e, L4e, N4e, A4e, B4e, C4e, F4e, Adiv4e, AC4e, AL24e, Atr4e, uEntries4e, sigmaEntries4e, area4e, coord4e, normal4e* und *normal4s*.

4.3 Lösungsmethoden

Zu den Programmen dieser Arbeit gehören drei Lösungsmethoden. Die erste implementiert den reinen Least-Squares-Ansatz aus Abschnitt 2.2, die zweite den Ansatz mit exakter Kontinuitätsgleichung aus Abschnitt 2.3, die dritte ist eine Wrapper-Methode für den in [ACFK02] vorgestellten Löser. Die Funktionen beinhalten die Errechnung der lokalen Steifigkeitsmatrizen, die Assemblierung dieser, die Verwaltung der Randdaten und Freiheitsgrade sowie das Lösen des linearen Gleichungssystems. Die Aufrufe erfolgen durch

[sigma, u, nrDof, STIMA, dof, localData, solutionData] ...
= solveElasticityLS(parameter, problemData, geometryData);

[sigma, u, nrDof, STIMA, dof, localData, solutionData] ...
= solveElasticityHybrid(parameter, problemData, geometryData);

und für die Wrapper-Methode durch

[sigma, u, nrDof, STIMA, dof, localData, solutionData] ...
= solveElasticityP1P1wrapper(parameter, problemData, geometryData);.

Für die Wrapper-Methode wird der Parameter *method* ignoriert, und es wird immer eine S^1-Diskretisierung verwendet. Außerdem können nicht alle Probleme mit der Wrapper-Methode gelöst werden, da das Behandeln unterschiedlicher Randbedigungen für die Komponenten im unterliegenden Löser nicht berücksichtigt ist. Um die Ausgabe des Lösers kompatibel zu den anderen beiden zu gestalten wird nach dem Lösen des Gleichungssystems noch eine Spannungsvariable errechnet. Dies geschieht auf zwei verschiedene Arten. Zum einen wird auf jedem Element $\mathbb{C}\varepsilon(u_h)$ berechnet und als Funktion $\sigma_0 \in P_0(\mathcal{T})$ gespeichert. Zum anderen wird eine Funktion $\sigma_h \in RT_0$ berechnet, indem die Koeffizienten zu den Basisfunktionen durch

$$\begin{pmatrix} \sigma_{2(E-1)+1} \\ \sigma_{2(E-1)+2} \end{pmatrix} = \frac{1}{2} \left(\mathbb{C}\varepsilon(u_h)_{|T_+} + \mathbb{C}\varepsilon(u_h)_{|T_-} \right) \nu_E$$

berechnet werden.

Die Parameter in der Tabelle 4.5 dienen zur Auswahl eines Lösers, einer Diskretisierung und einer Skalierung der Problemstellung entsprechend Abschnitt 2.2.

method	Wählt eine Diskretisierung für die Verschiebungsvariable aus.
{S1, KS, CR}	
solver	Legt die diskrete variationelle Formulierung fest, welche gelöst werden soll.
{LS, HYB, P1P1}	
s	Legt das Least-Squares Funktional fest, welches verwendet werden soll.
$\mathbb{R}$	
computeCond	Wenn *computeCond* gesetzt ist, wird die Konditionszahl der globalen Steifigkeitsmatrix errechnet.
{0, 1}	
computeDivU	Wenn *computeDivU* gesetzt ist, wird zusätzlich die Divergenz von u berechnet.
{0, 1}	

Tabelle 4.5: Parameter für die Löserkonfiguration

4.4 Fehlerschätzer und numerische Integration

Die Berechnung des Least-Squares-Residuums erfolgt mit der Methode *computeElasticityLSResidual* die durch

[residual4e, contrDiv4e, contrC4e] = computeElasticityLSResidual ...
(geometryData, problemData, localData, solutionData, parameter);

aufgerufen wird. Die Fehler zu den exakten Funktionen werden durch die folgenden Aufrufe berechnet. Die Fehler $\|u - u_h\|_{H^1(\Omega;\mathbb{R}^2)}$ und $\|\sigma - \sigma_h\|_{H(\mathrm{div},\Omega;\mathbb{R}^{2\times 2})}$ werden durch

[contrDu4e, contrSigma4e, error4e, contrDivSigma4e, contrL2Sigma4e, contrL2U4e, ...
contrU4e] = error4eElasticityLS(geometryData, problemData,...
exSolutionData, localData, solutionData, parameter);

errechnet. Da zur Untersuchen des Locking-Verhaltens ein besonderes Interesse an $\|\operatorname{div} u - \operatorname{div} u_h\|_{L^2(\Omega;\mathbb{R})}$ bestehen könnte kann dieser Fehler durch den Aufruf

[errorDivU4e, normDivUSq4e, normDivUexactSq4e] = error4eDivP1 ...
(geometryData,localData,solutionData,exSolutionData,parameter.intDegree);

gefunden werden. Aus gleichem Grund wurde eine Methode zur Errechnung von $|||u - u_h|||$ programmiert, welche durch

[elastErrorSq4e,elastEnergyUexSq4e,elastEnergyUsq4e] = errorP1elasticEnergy4e ...
(geometryData,exSolutionData,localData,solutionData,parameter);

aufgerufen wird. Für Probleme ohne bekannte Lösung kann ein Fehler zu einer deutlich besseren Approximation, welche als quasi exakte Lösung verstanden werden kann, berechnet werden. Für den Vergleich werden beide Lösungen auf ein Overlay der entsprechenden Gitter propagiert. Die Realisierung der Methode ist nicht optimiert und daher in vielen Situationen mit erheblichem Rechenaufwand verbunden. Der Aufruf folgt der Form

[quasiExactErr4e1C,quasiExactErr4e2C] = computeQeError
(paramter,solutionData,geometryData);.

Zur Errechnung der Datenapproximationsfehler in f, zum Beispiel für das separierte Markieren aus Abschnitt 4.5, steht die Methode

[dataApprox4e, fMean4e] = computeBAError
(geometryData.c4n,geometryData.n4e,problemData.f,parameter.intDegree);

der Arbeitsgruppe zur Verfügung. Ob und welche Fehler errechnet werden, wird mit den Parametern aus Tabelle 4.6 gesteuert.

useExactSolution	Wenn *useExactSolution* nicht gesetzt ist, werden keine exakten Fehler berechnet.
{0, 1}	
computeEnergyError	Wenn *computeEnergyError* gesetzt ist, wird der Energiefehler berechnet.
{0, 1}	
computeDataError	Wenn *computeDataError* gesetzt ist, wird der Datenapproximationsfehler berechnet.
{0, 1}	
computeQuasiExactError	Wenn *computeQuasiExactError* gesetzt ist, wird der Fehler zu einer anderen Approximation berechnet.
{0, 1}	
qeDataPath	Gibt den Pfad an, an dem die zweite Approximation gespeichert ist.
<PFAD>	

Tabelle 4.6: Parameter für die Fehlerschätzer und die Berechnung der exakten Fehler

4.5 Markieren und Verfeinern

Ein aktuelles Thema in der Forschung zu Least-Squares-Methoden liegt in der Untersuchung optimaler Konvergenzraten. Ein generelles Vorgehen bei der Untersuchung von optimalen Konvergenzraten liefert die Arbeit [CFPP14], in der vier Axiome aufgestellt werden, die ein Fehlerschätzer erfüllen muss, um bei adaptiver Verfeinerung optimale Konvergenzraten zu ermöglichen. Diese Axiome werden in den Schlüsselbegriffen Stabilität, Reduktion, diskrete Zuverlässigkeit und Quasiorthogonalität zusammengefasst. Obwohl für das Least-Squares-Residuum die Zuverlässigkeits- und Effizienz-Eigenschaft sofort nachgerechnet werden können, kann insbesondere eine Reduktionseingenschaft im Sinne der Axiome nicht bewiesen werden. Um dennoch ein Beweis für optimale Konvergenzraten führen zu können, werden alternative, zum Least-Squares-Funktional äquivalente, Fehlerschätzer gefunden, welche die Axiome erfüllen. Arbeiten mit diesem Vorgehen zum Poission-Modell-Problem [CP15], und den Differentialgleichungen des Stokes-Problems sind bereits erschienen [BC16]. Die in den Arbeiten bewiesenen Äquivalenzen der Least-Squares-Funktionals zu den Fehlerschätzern beinhalten zusätzlich die Datenapproximationsfehler der rechten Seite $f \in L^2(\Omega;\mathbb{R}^2)$ zu stückweise konstanten Approximationen $f_{0,\mathcal{T}}$. Die Resultate sind von der Form

$$\eta(\mathcal{T})^2 + \|f - f_{0,\mathcal{T}}\|^2_{L^2(\Omega;\mathbb{R}^2)} \approx \mathrm{LS}_s(f;\sigma_{\mathrm{LS}},u_{\mathrm{LS}}).$$

Eine weitere Arbeit mit diesem Vorgehen zur linearen Elastizität wird derzeit von P. Bringmann, C. Carstensen, und G. Starke angestrebt. Der vorgeschlagene alternative

Fehlerschätzer ist auf einem Dreieck T der Triangulierung $\mathcal{T}$ gegeben durch

$$\begin{aligned}\eta(\mathcal{T},T) = \quad & |T|\|\operatorname{div}(\operatorname{sym}\mathbb{C}^{-1}\sigma_{\mathrm{LS}} - \varepsilon(u_{\mathrm{LS}}))\|^2_{L^2(T)} \\ & + |T|\|\operatorname{curl}\mathbb{C}^{-1}(\mathbb{C}^{-1}\sigma_{\mathrm{LS}} - \varepsilon(u_{\mathrm{LS}}))\|^2_{L^2(T)} \\ & + |T|^{1/2} \sum_{E\in\mathcal{E}(T)\setminus\mathcal{E}(\Gamma_D)} \|[\operatorname{sym}\mathbb{C}^{-1}\sigma_{\mathrm{LS}} - \varepsilon(u_{\mathrm{LS}})]_E\nu_E\|^2_{L^2(E)} \\ & + |T|^{1/2} \sum_{E\in\mathcal{E}(T)\setminus\mathcal{E}(\Gamma_N)} \|[\mathbb{C}^{-1}(\mathbb{C}^{-1}\sigma_{\mathrm{LS}} - \varepsilon(u_{\mathrm{LS}}))]_E\tau_E\|^2_{L^2(E)} \\ & + |T|^{1/2} \sum_{E\in\mathcal{E}(T)\cup\mathcal{E}(\Gamma_N)} \|g - g_0\|^2_{L^2(E)}.\end{aligned}$$

Die Behandlung der Datenapproximationsfehler in einem adaptiven Algorithmus kann mit den separierten Markieren aus [Hel14a] erfolgen. Der Verfeinerungsalgorithmus wird dann von drei Parametern gesteuert. Der erste Parameter $0 < \kappa$ legt eine Gewichtung zwischen der Reduktion des Fehlerschätzers und der Reduktion des Datenfehlers fest. Im ersten Fall wird dann das Dörflermarkieren mit dem Bulk-Parameter $\Theta \in (0, 1]$ angewendet. Im zweiten Fall wird der Datenapproximationsfehler mindestens un den Faktor $\rho \in (0, 1]$ reduziert.

Input $0 < \Theta, \rho \leqslant 1$, $0 < \kappa < \infty$
for any level ℓ **do**
  Solve LS-FEM on $\mathcal{T}_\ell$.
  Estimate. Compute $\eta(\mathcal{T}_\ell)$ and $\|f - f_{0,\mathcal{T}_\ell}\|_{L^2(\Omega;\mathbb{R}^2)}$.
  if CASE A $\|f - f_{0,\mathcal{T}_\ell}\|^2_{L^2(\Omega;\mathbb{R}^2)} \leqslant \kappa\eta(\mathcal{T}_\ell)^2$ **then**
    Mark $\mathcal{M}_\ell$ of $\mathcal{T}_\ell$ with Dörflermarking s.t. $\Theta\eta(\mathcal{T}_\ell)^2 \leqslant \eta(\mathcal{M}_\ell)^2$
    Refine $\mathcal{T}_\ell$ s.t. $\mathcal{M} \subseteq \mathcal{T}_\ell\setminus\mathcal{T}_{\ell+1}$ using NVB.
  else CASE B $\kappa\eta_\ell^2 < \|f - f_{0,\mathcal{T}_\ell}\|_{L^2(\Omega;\mathbb{R}^2)}$
    Refine $\mathcal{T}_\ell$ s.t. $\|f - f_{0,\mathcal{T}_{\ell+1}}\|_{L^2(\Omega;\mathbb{R}^2)} \leqslant \rho\|f - f_{0,\mathcal{T}_\ell}\|_{L^2(\Omega;\mathbb{R}^2)}$

Da der alternative Fehlerschätzer noch nicht implementiert wurde, wird stattdessen die Differenz aus Least-Squares-Funktional und Datenapproximationsfehler verwendet. Die Schnittstelle für eine tatsächliche Implementation des alternativen Schätzers und die Verfeinerungsalgorithmen sind aber vorhanden.

Für die Markierung und Verfeinerung der Dreiecke stehen zwei Algorithmen zur Verfügung. Zum einen die Dörflermarkierung bezüglich der lokalen Beiträge des Least-Squares-Funktionals(*separateMarking*=0), oder bezüglich des alternativen Fehlerschätzers(*separateMarking*=2). Zum anderen kann das separierte Markieren bezüglich des alternativen Fehlerschätzers(*separateMarking*=1), verwendet werden. Die Tabelle 4.7 enthält Informationen zu den genannten Parametern.

theta $(0, 1]$	Gibt den Bulk-Parameter Θ für die Dörflermarkierung an.
separateMarking $\{0, 1, 2\}$	Wählt eine Markierungsstrategie.
kappa $\mathbb{R}_+$	Gibt den Parameter κ an, welcher die Fallunterscheidung beim separierten Markieren steuert.
rho $(0, 1]$	Gibt den Parameter ρ an, welcher die Reduktion des Datenfehlers festlegt.

Tabelle 4.7: Markierungsparameter

4.6 Graphische Darstellung und Speichern

Für eine direkte Auswertung der Ergebnisse stehen eine Reihe von Visualisierungsparametern zur Verfügung. Ist eine exakte Lösung bekannt, wird diese zum Vergleich immer mit ausgegeben. Die Parameter sind in Tabelle 4.8 zu finden. Ist die Methode entsprechend konfiguriert, so werden die errechneten Ergebnisse im Ordner *results/elasticity/<saveFolder>* abgelegt. Welche Daten gespeichert werden, wird über die Parameter in Tabelle 4.9 konfiguriert. Die Daten liegen dann sowohl als *.mat*- als auch als *.dat*-Datei vor.

generatePlots $\{0, 1, 2\}$	Wenn *generatePlots* 1 ist, werden die Ergebnisse des letzten Durchlaufs visualisiert. Bei 2 wird jeder Durchlauf visualisiert.
solutionPlotU $\{0, 1\}$	Wenn *solutionPlotU* gesetzt ist, wird die Lösung u komponentenweise geplottet.
solutionPlotSigma $\{0, 1\}$	Wenn *solutionPlotSigma* gesetzt ist, wird die Lösung *sigma* komponentenweise geplottet.
meshPlot $\{0, 1\}$	Wenn *meshPlot* gesetzt ist, wird die Triangulierung geplottet.
rhsPlot $\{0, 1\}$	Wenn *rhsPlot* gesetzt ist, wird die rechte Seite f und ihre Approximation geplottet.
STIMAPlot $\{0, 1\}$	Wenn *STIMAPlot* gesetzt ist, wird die Besetzung der Steifigkeitsmatrix geplottet.
divUPlot $\{0, 1\}$	Wenn *divUPlot* gesetzt ist, wird die Divergenz der Lösung u geplottet.
divSigmaPlot $\{0, 1\}$	Wenn *divSigmaPlot* gesetzt ist, wird die Divergenz der Lösung *sigma* geplottet.

Tabelle 4.8: Auswahl der Parameter zur Ausgabe von Plots

saveData $\{0, 1, 2\}$	Wenn *saveData* 1 ist, werden die Ergebnisse des letzten Durchlaufs gespeichert. Bei 2 wird nach jedem Durchlauf gespeichert.
saveFolder <PATH>	Gibt den Pfad an, an dem die Ergebnisse gespeichert werden sollen.
saveGeometry $\{0, 1\}$	Wenn *saveGeometry* gesetzt ist, werden die Geometrie Daten aus geometryData gespeichert.
saveSolution $\{0, 1\}$	Wenn *saveSolution* gesetzt ist, werden die Koeffizientenvektoren u und σ gespeichert.
savePlots $\{0, 1\}$	Wenn *savePlots* gesetzt ist, werden die erzeugten Plots als *.fig* Dateien gespeichert.
writeLog $\{0, 1\}$	Wenn *writeLog* gesetzt ist, werden die Ausgaben in der Konsole in der Datei *output.log* gespeichert.

Tabelle 4.9: Parameter zum Speichern von Ergebnissen

5 Numerische Ergebnisse und Schlussfolgerungen

Dieser Abschnitt befasst sich mit der Validierung der Implementation aus Abschnitt 3 und der Bestätigung der theoretischen Resultate aus Abschnitt 2 durch numerische Experimente. Die vorgestellten Resultate bilden dabei nur eine Auswahl der möglichen Berechnungen. Die erzeugten Daten liegen der Arbeit bei. Die Konfigurationsdateien zum Reproduzieren der Ergebnisse sind ebenfalls immer im jeweiligen Ordner zu finden. Die verwendeten Abbildungen 5.3, 5.8 und 5.9 zur Verdeutlichung der Gebiete Ω in den nachfolgenden Abschnitten sind aus [Hel14b] übernommen.

5.1 Akademisches Beispiel auf dem Einheitsquadrat

Das erste Experiment dieser Arbeit soll das generelle Konvergenzverhalten der implementierten Lösungsmethoden bestätigen. Dazu dient ein Problem auf dem Gebiet $\Omega = (0,1)^2$ mit homogenen Dirichletrandaten auf dem gesamten Rand $\Gamma_D = \partial\Omega$ und der Volumenkraft

$$f(x,y) = \begin{pmatrix} -2\mu\pi^3 \cos(\pi y)\sin(\pi y)(2\cos(2\pi x) - 1) \\ 2\mu\pi^3 \cos(\pi x)\sin(\pi x)(2\cos(2\pi y) - 1) \end{pmatrix}.$$

Als Lamé-Parameter wurden für diese erste Rechnung $\lambda = 1$ und $\mu = 1/2$ gesetzt, sodass keine Skalierung von μ entsprechend Abschnitt 1.3 nötig ist und keine negativen Auswirkungen auf die Konvergenzraten zu erwarten sind. Für dieses Problem ist die exakte Lösung

$$u(x,y) = \begin{pmatrix} \pi \cos(\pi y)\sin(\pi y)\sin^2(\pi x) \\ -\pi \cos(\pi x)\sin(\pi x)\sin^2(\pi y) \end{pmatrix}$$

bekannt, daher lassen sich die genauen Abweichungen der Approximationen berechnen. Die Abbildung 5.1 zeigt den gesamten Fehler bezüglich der Norm in X, die Fehleranteile in den Räumen $H(\mathrm{div}, \Omega; \mathbb{R}^{2\times 2})$ und $H^1(\Omega; \mathbb{R}^2)$, den Energiefehler und die Least-Squares-Residuen für unterschiedliche $s \in \mathbb{R}$. Die Approximationen wurden auf uniform verfeinerten Gittern berechnet, da aber Ω ein konvexes Gebiet ist und kein Sprung in den Randdaten vorliegt, sind optimale Konvergenzraten bezüglich der Freiheitsgrade zu erwarten. Dieser können in Abbildung 5.1 für alle Methoden und Fehler beobachtet werden. Die fast identische Abbildung 5.2 zeigt die gleichen Ergebnisse für den in 2.3 beschrieben Ansatz mit exakter Kontinuitätsgleichung, die zur Verkürzung im Rest dieses Abschnittest auch als hybride Methode bezeichnet werden soll. Zum Vergleich ist der Gesamtfehler im Raum X und das Least-Squares-Residuum für $s = -1$ jeder Methode in der Abbildung der jeweils anderen eingetragen. Es ist zu erkennen, dass die absolut Werte für die Fehler der Hybridmethode leicht über denen der reinen Least-Squares-Methode liegen. Die Ergebnisse befinden sich im Ordner *results/elasticity/ABGABE/exp11* und *results/elasticity/ABGABE/exp12*.

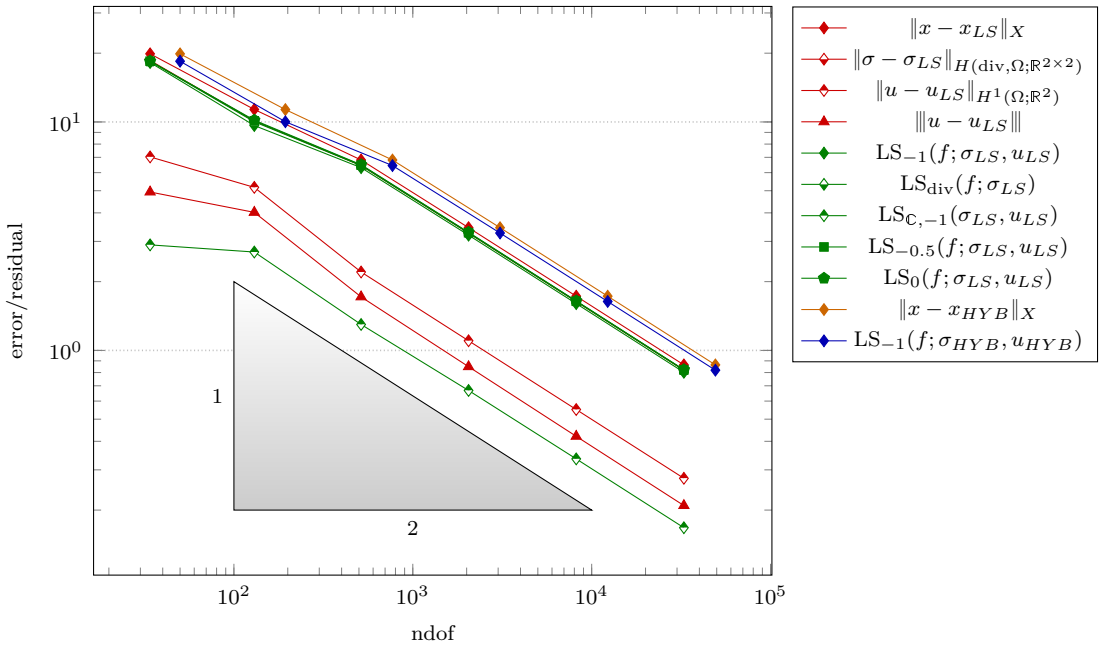

Abbildung 5.1: Exakte Fehler und Residuen der Least-Squares-Methoden für uniforme Verfeinerung; Experiment zu Abschnitt 5.1

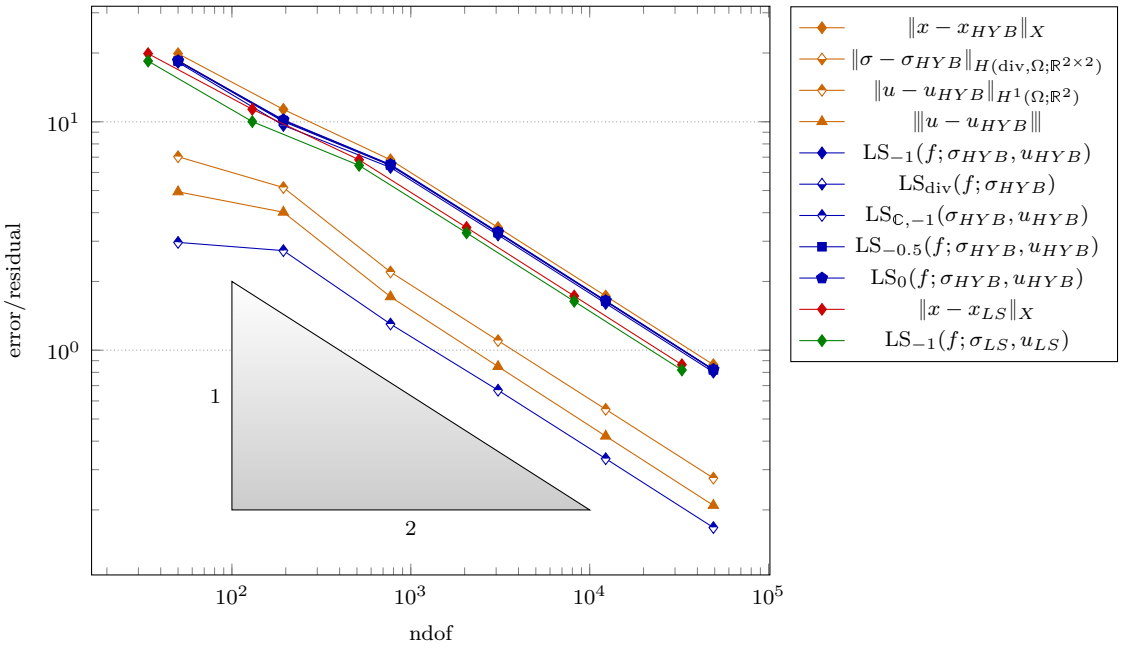

Abbildung 5.2: Exakte Fehler und Residuen der Hybridmethode für uniforme Verfeinerung; Experiment zu Abschnitt 5.1

5.2 Adaptive Verfeinerung und optimale Konvergenzraten

Ein zweites Experiment soll die Verbesserung der Konvergenzraten bei adaptivem Verfeinern mit Hilfe des Least-Squares-Residuums als Fehlerindikator nachweisen. Dazu wird ein weiteres Problem mit exakter Lösung, aber nicht-konvexem Gebiet herangezogen. Das Gebiet

$$\Omega = \operatorname{conv}\{(0,0),(-1,-1),(0,-2),(1,-1)\} \cup \operatorname{conv}\{(-1,1),(1,-1),(2,0),(0,2)\}$$

und die Ränder $\Gamma_N = \{(x,y) \in \Omega |\ |y| \leqslant 1, x \leqslant 0\}$ und $\Gamma_D = \partial\Omega \backslash \Gamma_N$ sind in der Abbildung 5.3 veranschaulicht.

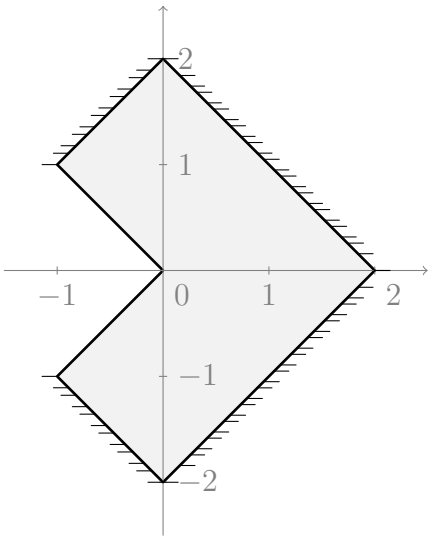

Abbildung 5.3: rotiertes L-Shape

Für das Problem sind die Volumenkraft $f \equiv 0$ und die Neumannranddaten $t \equiv 0$. Die Dirichletrandaten sind passend zur exakten Lösung gewählt. Diese kann in [CDFH00, Abschnitt 6] gefunden werden und lautet in Polarkoordinaten (r, φ)

$$\begin{aligned} u_r(r,\varphi) &= \frac{1}{2\mu} r^\alpha \left(-(\alpha+1)\cos((\alpha+1)\varphi) + (c_2 - \alpha - 1) c_1 \sin((\alpha-1)\varphi) \right) \\ u_\varphi(r,\varphi) &= \frac{1}{2\mu} r^\alpha \left(\ \ (\alpha+1)\sin((\alpha+1)\varphi) + (c_2 + \alpha - 1) c_1 \cos((\alpha-1)\varphi) \right), \end{aligned}$$

dabei ist α die positive Lösung der Gleichung $\alpha \sin(3\pi/2) + \sin(3\alpha\pi/2) = 0$ und die Konstanten gegeben durch

$$c_1 = -\frac{\cos(3\pi(\alpha+1)/4)}{\cos(3\pi(\alpha-1)/4)} \qquad \text{und} \qquad c_2 = \frac{2\lambda + 4\mu}{\lambda + \mu}.$$

Für die nachfolgenden Rechnungen sind die Lamé-Parameter wieder $\lambda = 1$ und $\mu = 0.5$. Die Abbildung 5.4 zeigt die exakten Fehler in der Norm des Raumes X und die Least-Squares-Residuen $\mathrm{LS}_{-1}(0; \sigma_{\mathrm{LS}}, u_{\mathrm{LS}})$ für uniforme und adaptive Triangulierungen bis rund 10^5 Freiheitsgrade. Für die uniformen Verfeinerungen sind suboptimale Konvergenzraten von rund 1/4 zu beobachten. Je restriktiver der Bulk-Parameter Θ gewählt wird, desto bessere Konvergenzraten lassen sich erzielen. Ab einem Wert von ungefähr $\Theta = 0.4$ ist die optimale Konvergenzrate erreicht. Hervorzuheben ist, dass $\Theta = 0.95$ schon deutlich bessere Konvergenzraten liefert, als die uniforme Verfeinerung mit $\Theta = 1$. Abbildung 5.5 zeigt ein adaptiv erzeugtes Gitter für $\Theta = 0.4$ nach acht Verfeinerungsschritten.

Auch für die in der Abbildung 5.4 nicht dargestellten Methoden lassen sich wörtlich dieselben Ergebnisse beobachten. Die Ergebnisse befinden sich im Ordner *results/elasticity/ABGABE/exp2*.

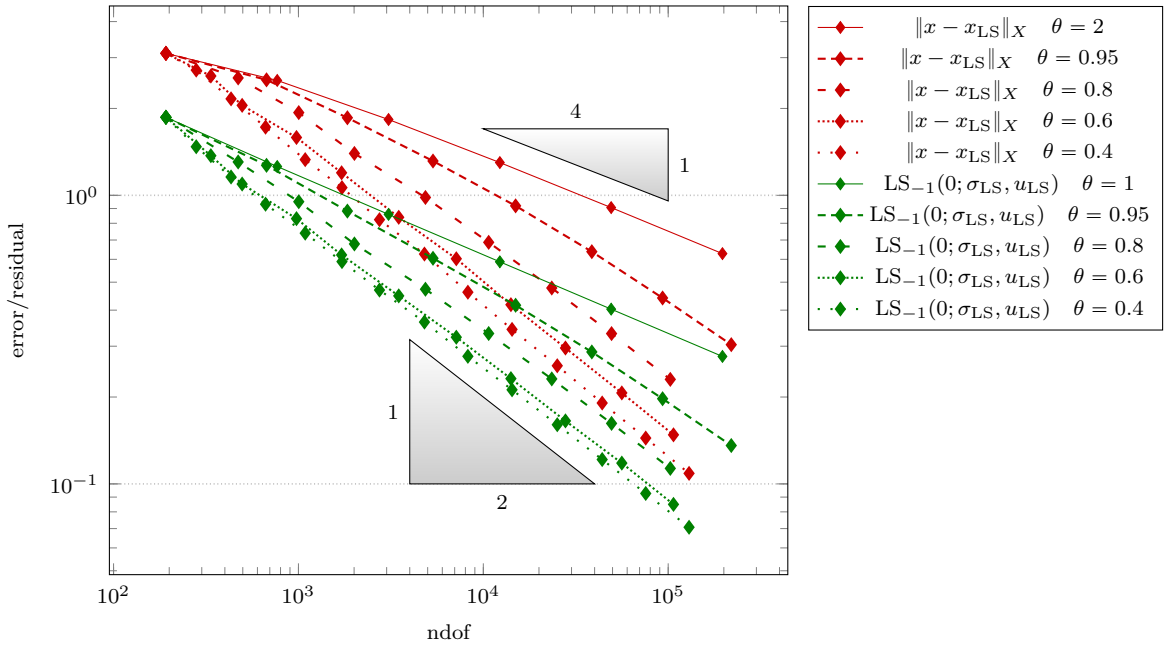

Abbildung 5.4: Konvergenzgraphen für unterschiedliche Bulk-Parameter; Experiment zu Abschnitt 5.2

5.3 Benchmarkproblem Cooks-Membran

Als nächstes soll ein bekanntes Benchmarkproblem betrachtet werden. Das Problem, von R.D.Cook entwickelt, betrachtet eine trapezförmige Platte, die an ihrer längeren Seite eingespannt und an ihrer kürzeren Seite durch eine Schubkraft belastet, wird. Gegeben ist $\Omega = \mathrm{conv}\{(0,0), (48,44), (48,60), (0,44)\}$ mit dem Dirichletrand $\Gamma_D = \{0\} \times [0, 44]$ und dem Neumannrand $\Gamma_N = \partial\Omega \backslash \Gamma_D$. Auf dem Dirichletrand werden Nullranddaten

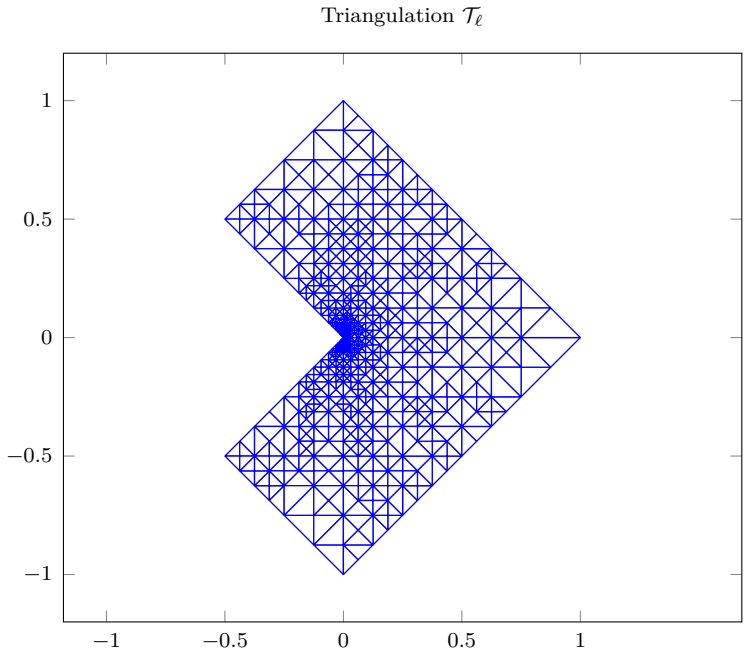

Abbildung 5.5: Gitterplot für adaptive Verfeinerung mit 1201 Elementen; Experiment zu Abschnitt 5.2

$g \equiv 0$ vorgegeben. Auf der kürzesten Seite des Neumannrandes $\{48\} \times [44, 60]$ sind die Randdaten durch $t = (0, 1)^\top$ gegeben. Auf dem restlichen Neumannrand ist $t = 0$. Die Situation wird in Abbildung 5.6 verdeutlicht. Die Lamé-Parameter werden aus dem Youngschen Modul $E = 2900$ und der Querkontraktionszahl $v = 0.4$ errechnet. Diese Konfiguration ist aus [ACFK02] entnommen und entspricht der Berechnung für einen Plexiglaswerkstück.

Für das Problem ist keine exakte Lösung bekannt, die Least-Squares-Residuen für $s = -1$ wurden in der Abbildung 5.10 abgetragen. Für die Hybridmethode ist nur das Residuum $\mathrm{LS}_{\mathbb{C}}$ eingetragen, da $\mathrm{LS}_{\mathrm{div}} \equiv 0$ ist. Auch in diesem Beispiel kann die optimale Konvergenzrate durch adaptive Verfeinerung erreicht werden. Nachfolgend werden die genauen Werte noch einmal in Tabellen 5.1 und 5.2 aufgelistet, um direkte Vergleiche zu ermöglichen.

Die Ergebnisse befinden sich im Ordner *results/elasticity/ABGABE/exp3*.

nrDof	$\mathrm{LS}(0;\sigma_{\mathrm{LS}},u_{\mathrm{LS}})$	$\mathrm{LS}_{Div}(0;\sigma_{\mathrm{LS}})$	$\mathrm{LS}_{\mathbb{C}}(\sigma_{\mathrm{LS}},u_{\mathrm{LS}})$	$\mathrm{LS}(\sigma_{HYB},u_{HYB})$
32	14.857120	3.845671	14.350778	15.404006
128	13.459280	3.128434	13.090650	13.847717
512	11.424967	2.201343	11.210886	11.645930
2048	8.390495	1.183753	8.306572	8.475757
8192	5.381060	0.504727	5.357337	5.404947
32768	3.269734	0.199942	3.263615	3.275871
131072	1.987856	0.080493	1.986225	1.989488
524288	1.231886	0.033487	1.231431	1.232341

Tabelle 5.1: Werte zu Abbildung 5.7 für $\Theta = 1$

nrDof	$\mathrm{LS}(0;\sigma_{\mathrm{LS}},u_{\mathrm{LS}})$	$\mathrm{LS}_{Div}(0;\sigma_{\mathrm{LS}})$	$\mathrm{LS}_{\mathbb{C}}(\sigma_{\mathrm{LS}},u_{\mathrm{LS}})$
32	14.857120	3.845671	14.350778
104	14.060353	3.672849	13.572166
192	12.973958	2.911958	12.642946
496	11.344856	2.085093	11.151598
592	11.085276	2.021981	10.899309
1860	8.179835	0.999221	8.118575
2296	7.608788	0.863683	7.559611
7148	4.891358	0.339363	4.879571
8688	4.517545	0.288898	4.508298
31008	2.517775	0.086436	2.516291
37112	2.329113	0.074291	2.327928
127692	1.273170	0.021808	1.272983
152296	1.171347	0.018634	1.171199
425752	0.698883	0.006640	0.698851
505636	0.643756	0.005631	0.643732

Tabelle 5.2: Ausgewählte Werte zu Abbildung 5.7 für $\Theta = 0.1$

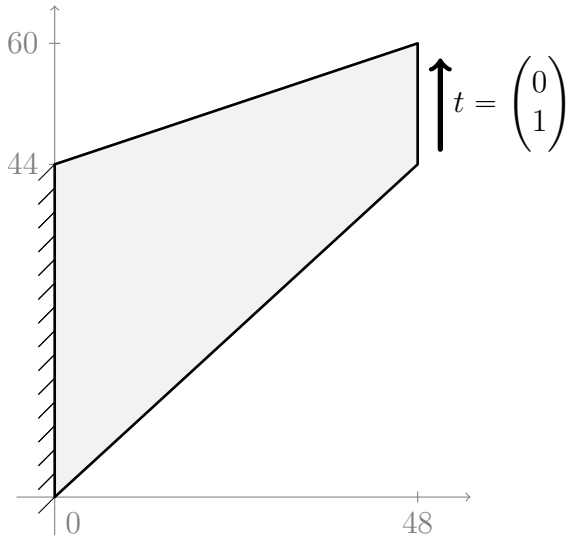

Abbildung 5.6: Cooks-Membran

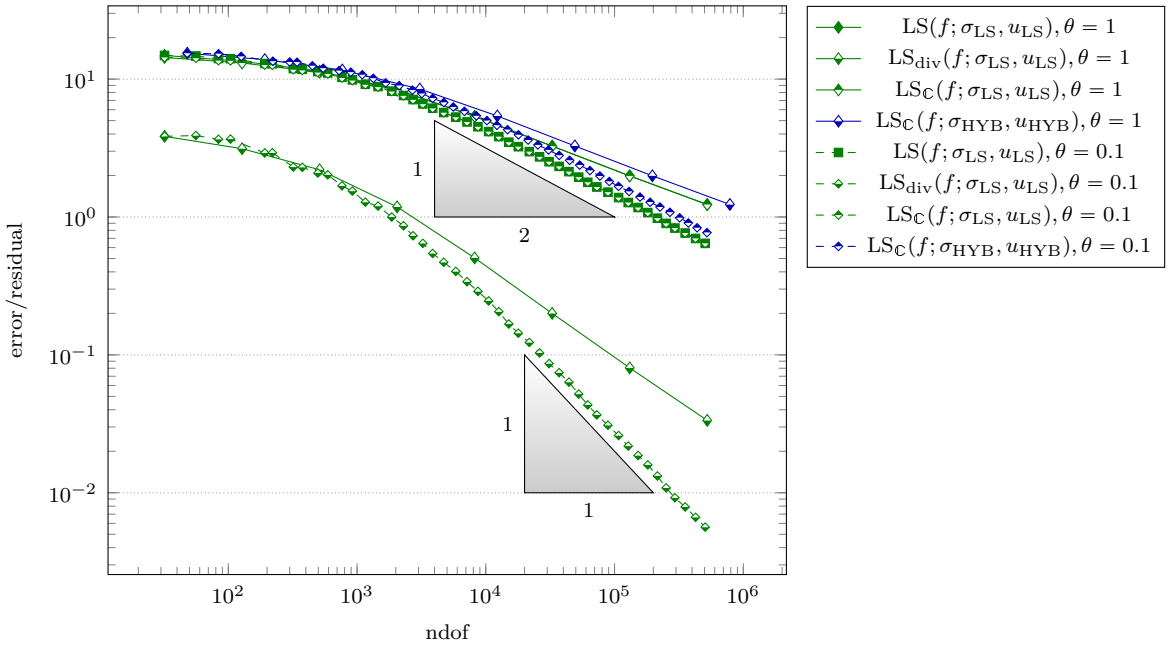

Abbildung 5.7: Konvergenzgraphen für das Benchmarkproblem aus Abschnitt 5.3

5.4 Einfluss und Skalierung der Gebietsgröße

Das nächste Experiment soll die Notwendigkeit der in Abschnitt 1.2 beschriebenen Skalierung des Gebietes für den reinen Least-Squares-Ansatz deutlich machen und die Überlegenheit der Formulierung aus Abschnitt 2.3 diesbezüglich aufzeigen. Dazu wird eine Abwandlung des nach Ernst Gustav Kirsch benannten Problems betrachtet. Das Gebiet ist dabei eine quadratische Platte mit Seitenlänge 100 in deren Mitte sich eine runde Aussparung mit Radius 25 befindet. D.h. $\hat{\Omega} = (-100, 100)^2 \backslash K_{25}(0,0)$. Es wird keine Volumenkraft aufgebracht, also $f \equiv 0$, aber auf die obere und untere Seite der Platte wirken normierte, orthogonale Zugkräfte, während der Rand des Loches fixiert wird. Die Abbildung 5.8 zeigt das Gebiet und die wirkenden Kräfte schematisch. Aus Symmetriegründen genügt es den in Abbildung 5.9 gezeigten Ausschnitt aus dem Gebiet $\hat{\Omega}$ zu betrachten. Das eigentlich berechnete Gebiet ist also $\Omega = (0, 100)^2 \backslash K_{25}(0,0)$. Die Randdaten an den mit Kreisen gekennzeichneten Kanten müssen dabei neu gesetzt werden. Da auf der Kante AB keine Verschiebung in vertikale Richtung zu erwarten ist, wird in der zweiten Lösungskomponente die Dirichletrandbedingung $g_2 \equiv 0$ vorgegeben. Aus dem gleichen Grund wird in der ersten Lösungskomponente $g_1 \equiv 0$ auf der Strecke DE vorgegeben. In den anderen Komponenten werden für den Übergang homogene Neumannranddaten gesetzt. Die Randdaten am Kreisrand sowie den Kanten DC und CB bleiben natürlich erhalten. Zu dem so gegeben Problem ist keine exakte Lösung bekannt, allerdings kann in [KS95] zu passenden Neumannranddaten die exakte Lösung

$$u(r,\varphi) = \begin{pmatrix} u_r(r,\varphi)\sin(\varphi) - u_\varphi(r,\varphi)\cos(\varphi) \\ u_r(r,\varphi)\sin(\varphi) + u_\varphi(r,\varphi)\cos(\varphi) \end{pmatrix}$$

in den Polarkoordinaten (r, φ) gefunden werden. Dabei sind die einzelnen Komponenten durch

$$\begin{aligned} u_r(r,\varphi) &= \frac{1}{8\mu r}((\kappa-1)r^2 + 2\gamma a^2 + (2r^2 - 2(\kappa+1)a^2/\kappa + 2a^4/(\kappa r^2))\cos(2\varphi)) \\ u_\varphi(r,\varphi) &= -\frac{1}{8\mu r}(2r^2 - 2(\kappa-1)a^2/\kappa - 2a^4/(\kappa r^2))\sin(2\varphi) \end{aligned}$$

mit den Konstanten $\kappa = 3 - 4\nu$, $\gamma = 2\nu - 1$ und $a = 25$ gegeben. Für die Berechnungen wird wieder $\lambda = 1$ und $\mu = 0.5$ gesetzt.

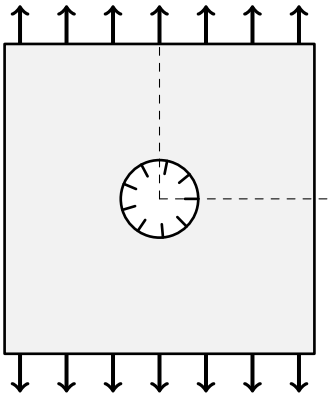

Abbildung 5.8: Lochplatte

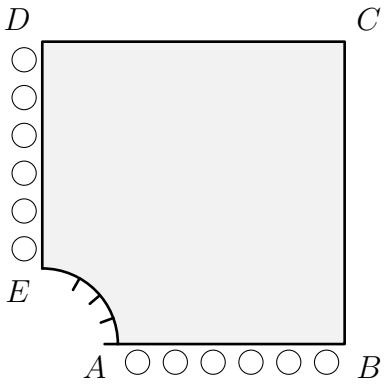

Abbildung 5.9: Ausschnitt aus der Lochplatte

Die Approximationsfehler und die Least-Squares-Residuen für uniforme Verfeinerungen und adaptive Verfeinerungen mit dem Bulk-Parameter $\Theta = 0.2$ sind in Abbildung 5.10 dargestellt. Das Gebiet Ω wird dabei in einem Durchlauf entsprechend Abschnitt 1.2 um den Faktor 100 verkleinert, in einem anderen nicht.

Die Ergebnisse befinden sich im Ordner *results/elasticity/ABGABE/exp5*.

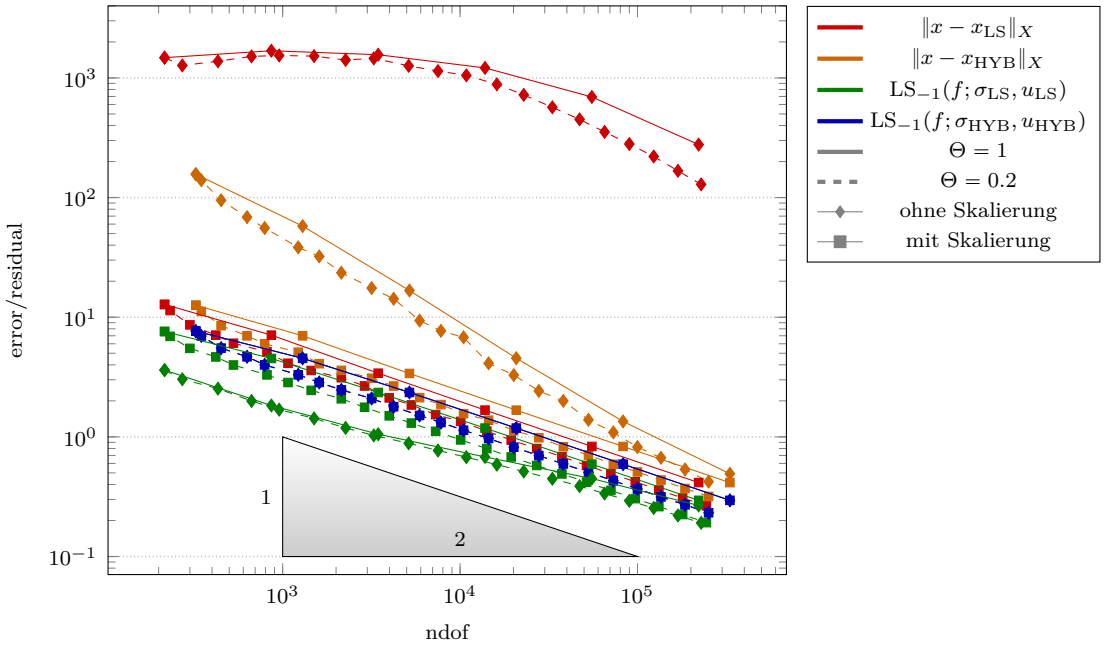

Abbildung 5.10: Konvergenzgraphen mit und ohne Skalierung des Gebiets; Experiment zu Abschnitt 5.4

Es lässt sich deutlich erkennen, dass ohne eine Skalierung für den reinen Least-Squares-Ansatz die Fehler unabhängig vom Bulk-Parameter Θ sehr viel größer sind als mit einer Skalierung. Die Konvergenzraten sind ebenso unabhängig von Θ noch bis 10^5 Freiheitsgrade deutlich schlechter als die optimal zu erwartende Rate. Die adaptive

Verfeinerung kann allerdings etwas schneller bessere Konvergenzraten herstellen. Ein wesentlichen Gewinn bringt aber nur die Skalierung des Gebietes. Mit dieser sind von Beginn der Iterationen an deutlich bessere Konvergenzraten zu beobachten. Die zweite Methode mit dem Lagrange-Ansatz für die Kontinuitätsgleichung zeigt deutlich bessere Konvergenzraten und Fehlerwerte, auch wenn das Gebiet nicht skaliert wird, aber auch hier kann noch eine Verbesserung durch eine Gebietsskalierung erreicht werden. Dieser Vorteil der Skalierung verliert sich aber ab ungefähr 10^5 Freiheitsgraden. Die Least-Squares-Residuen zeigen kaum einen Unterschied in den verschiedenen Konfigurationen.

5.5 Locking bei unterschiedlichen Querkontraktionszahlen

Für die Untersuchung der Konvergenzeigenschaften der Löser bei fast-inkompressiblen Materialen wird zunächst noch einmal das Problem aus Abschnitt 5.1 betrachtet. Zum Vergleich wird mit dem P1P1-Löser aus [ACFK02] das Problem für unterschiedliche Querkontraktionszahlen gelöst. Das Youngsche Modul wird hier immer auf $E = 1$ gesetzt. Aus den Zusammenhängen in Abschnitt 1.3 wird dann klar, dass Querkontraktionszahlen nahe $\nu \to 0.5$ gleichbedeutend mit Lamé-Parametern $\lambda \to \infty$ sind. Die Ergebnisse sind in der Abbildung 5.11 dargestellt. Es ist deutlich zu erkennen, dass die Fehler in der Approximation σ_0 der Stressvariable σ für Querkontraktionszahlen nahe 0.5 das Locking-Verhalten aufweisen. Dabei ist sogar zu beobachten, dass die Fehler nach dem ersten Verfeinerungsschritt nochmal zunehmen. Und dann erst mit sehr feinen Gittern wieder aufgelöst werden können. Je näher der Wert von ν an 0.5, desto länger dauert es, bis eine hinreichend kleine Gitterweite erreicht ist. Für $\nu = 0.49999$ ist bis 10^5 Freiheitsgrade keine Konvergenz beobachtbar. Der Energiefehler $|||u - u_{\mathrm{P1P1}}|||$ scheint aber nicht von den Querkontraktionszahlen beeinflusst zu werden. Dies entspricht nicht der in Abschnitt 3.3 vorgestellten Theorie.

Für $\nu = 0.49999$ wurde das Problem mit dem reinen Least-Squares-Ansatz und dem hybriden Ansatz für unterschiedliche Skalierungen $s \in \mathbb{R}$ berechnet. Die Ergebnisse werden in Abbildung 5.12 und 5.13 dargestellt. Im ersten Fall ist deutlich erkennbar, dass sowohl die exakten Fehler, als auch die Least-Squares-Residuen für die Parameter $s = -1/2$ und $s = -1$ aus [CS03], beziehungsweise [CS04] die Konvergenzrate $1/2$ erreichen. Aber auch die Residuen und Fehler der natürlichen Formulierung zeigen keine vorasymtotische Verschlechterung der Konvergenzraten. Lediglich extreme Skalierungen durch $s = 0.5$ und $s = -1.5$ zeigen ein abweichendes Verhalten. Dieses lässt sich durch die in Lemma 3.6 beschriebenen Auswirkungen des Parameters λ auf die Effizienz und Zuverlässigkeit begründen. Für $s = -1, 5 \leqslant -1$ ist der Fehler daher deutlich größer als das Least-Squares-Funktional. Für $s = 0.5 \geqslant -1$ ist das Least-Squares-Funktional größer als der exakte Fehler. Das für $s = 0 \geqslant -1$ und $s = -0.5 \geqslant -1$ dieses Verhalten nicht zu beobachten ist, kann durch das Lemma 3.6 nicht geklärt werden.

Im Fall der hybriden Methode sehen die Ergebnisse sehr ähnlich aus, allerdings ist hier auch für $s = 0.5$ keine Abweichung der Konvergenzrate von der optimalen Konvergenzrate zu

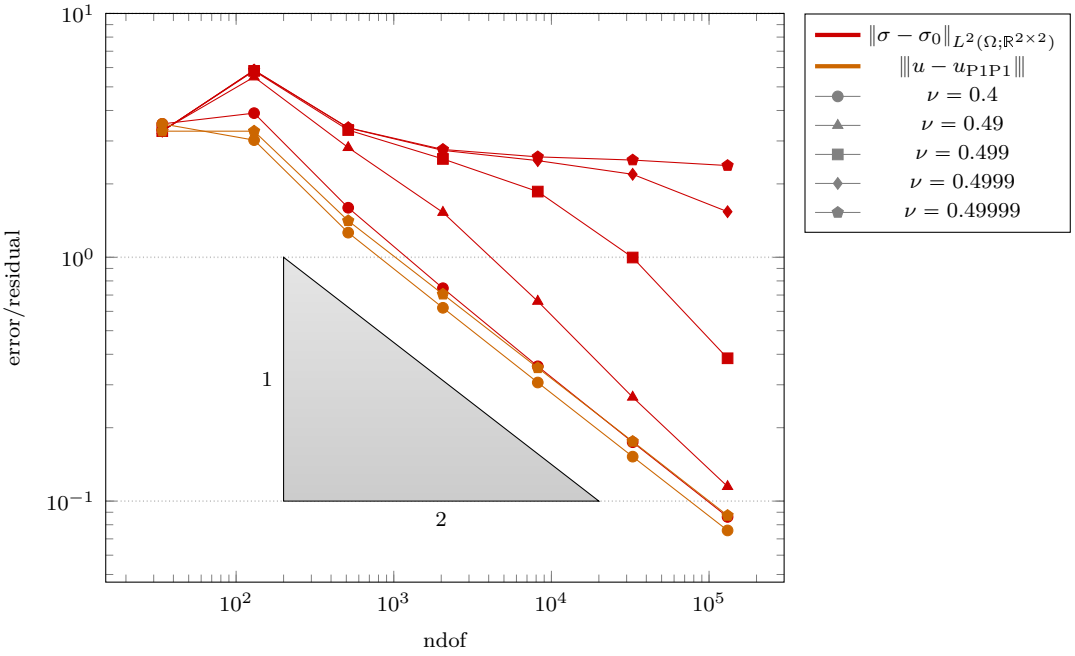

Abbildung 5.11: Locking der Stressvariable für den Vergleichslöser

beobachten. Alle Ergebnisse befinden sich im Ordner *results/elasticity/ABGABE/exp51*.

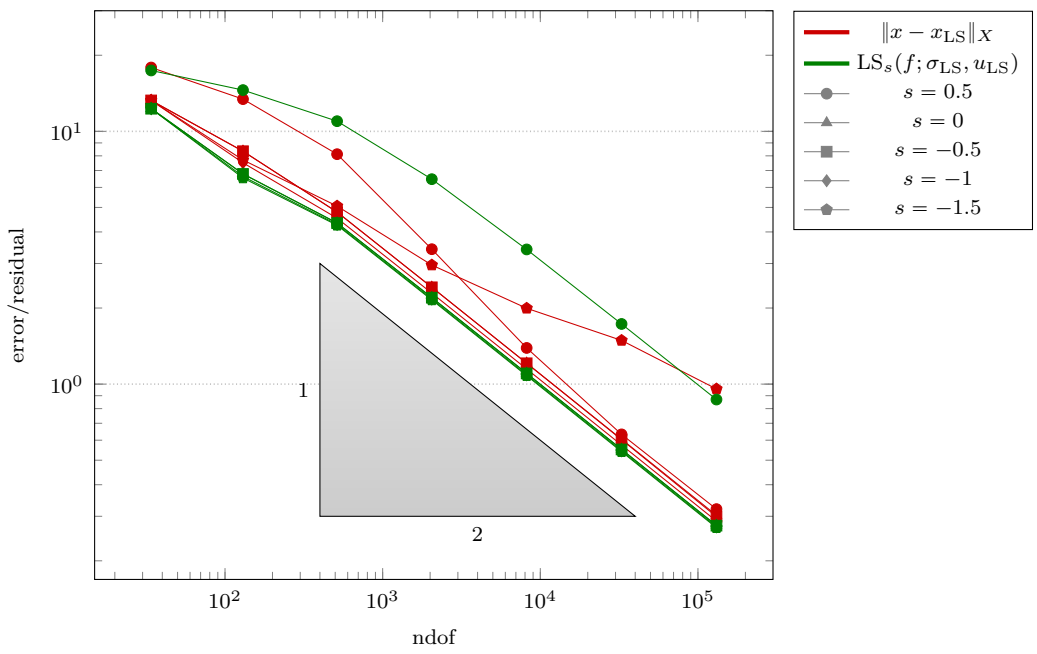

Abbildung 5.12: Konvergenzraten der LS-Methode für fast-inkompressibles Material ($\nu = 0.49999$); Experiment zu Abschnitt 5.5

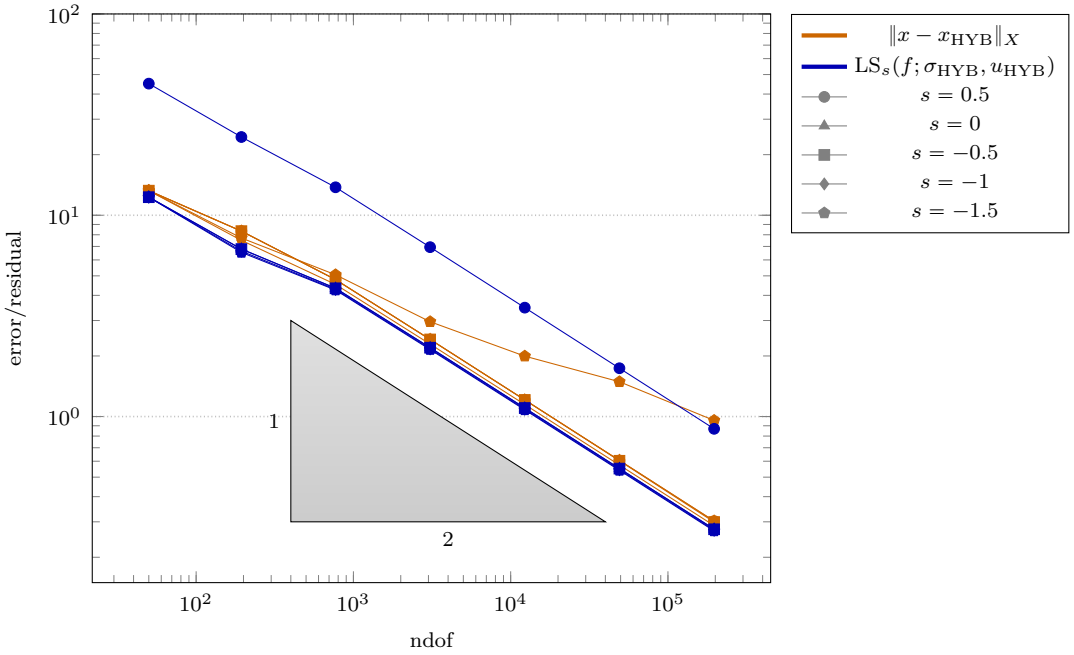

Abbildung 5.13: Konvergenzraten der HYB-Methode für fast-inkompressibles Material ($\nu = 0.49999$); Experiment zu Abschnitt 5.5

5.6 Einfluss und Skalierung des Schubmoduls

Das nächste Experiment untersucht den Einfluss des zweiten Lamé-Parameters μ und dessen Skalierung entsprechend dem Abschnitt 1.3 auf die Konvergenzraten der Methoden. Betrachtet wird dafür das gleiche Problem wie im zweiten Experiment 5.2. Der erste Lamé-Parameter wird für alle Testläufe auf $\lambda = 1$ gesetzt und mit $\Theta = 0.4$ verfeinert. Zunächst sollen die Konvergenzraten der Least-Squares-Methode ohne die Skalierung des Parameters betrachtet werden. In Abbildung 5.14 sind die exakten Fehler und die Least-Squares-Residuen für die Parameter $s = -1$ und $s = -0.5$ dargestellt. Für beide Methoden sind die absolut Werte der Fehler und der Residuen stark von der Wahl von μ abhängig. Je näher μ an Null, desto größer sind die absoluten Werte. Zusätzlich werden auch die Differenzen der Residuen zu den entsprechenden exakten Fehler ebenfalls größer, je näher μ an Null gewählt wird. Für die Wahl von $s = -0.5$ werden aber schon nach wenigen Verfeinerungen wieder optimale Konvergenzraten erreicht, was für die Wahl von $s = -1$ nicht der Fall ist. Genauer kann $\mu = 0.0005$ im Fall $s = -1$ bis 10^5 Freiheitsgraden keine Konvergenz mehr beobachten werden für $s = -0.5$ kann Konvergenz ab 10^3 Freiheitsgraden beobachtet werden. Dies kann wieder dadurch begründet werden, dass die Konstanten für die Zuverlässigkeit und Effizienz des Least-Squares-Schätzers natürlich von μ abhängig sind. Eine Auffälligkeit ist das Least-Squares-Residuum der Konfiguration $s = -0.5$ und $\mu = 0.0005$ bei rund 10^4 Freiheitsgraden zu beobachten. Diese könnte auf nicht exakte Berechnungen und numerische Fehler zurück zu führen sein.

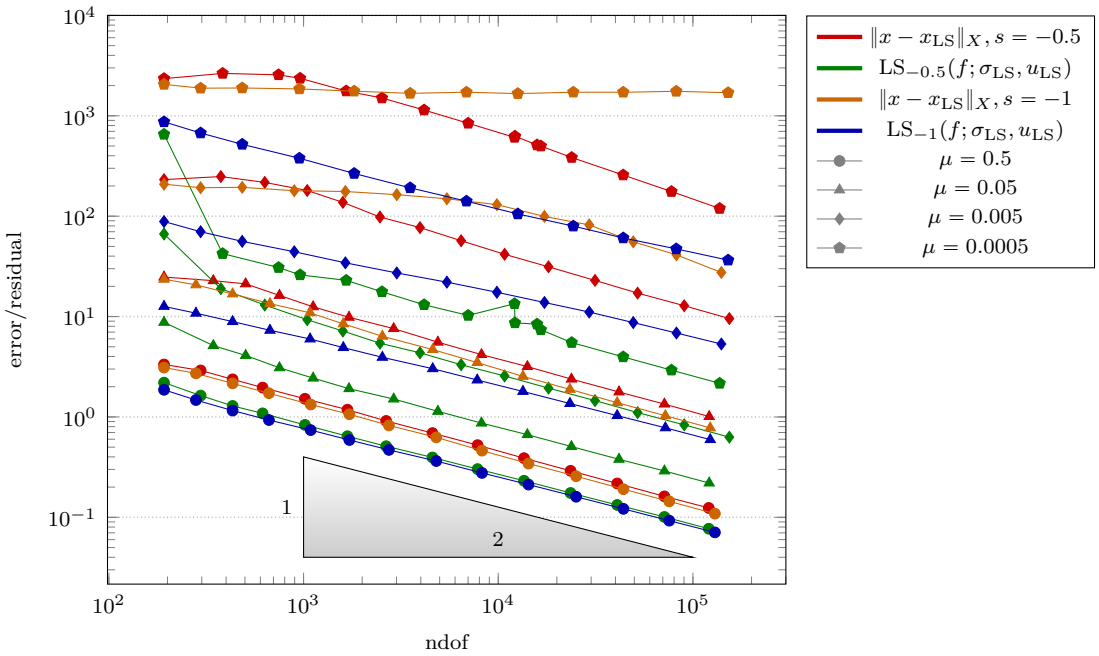

Abbildung 5.14: Konvergenzraten bei unterschiedlichen Schubmodulen; Experiment zu Abschnitt 5.6

5.7 Datenapproximation für hochoszillierende Volumenkräfte

Das nächste Experiment untersucht den Einfluss von schwer zu approximierenden Volumenkräften durch hohe Oszillationen. Dazu werden zwei unterschiedliche Probleme untersucht. In beiden ist das Gebiet Ω das Einheitsquadrat aus dem ersten Experiment 5.1. In beiden Problemen werden die Parameter $E = 1$ und $\nu = 0.4$ fixiert. Die Ergebnisse befinden sich im Ordner *results/elasticity/ABGABE/exp7*.

Im ersten Beispiel ist die Volumenkraft durch

$$f_1(x) = f_2(x) = \sin\left(\|x - x_m\|^{-1}\right)$$

gegeben, dabei ist $x_m := \operatorname{mid}\Omega = (1/2, 1/2)$ der Mittelpunkt des Einheitsquadrats. Diese Volumenkraft oszilliert beliebig schnell in der Nähe des Mittelpunkts x_m. Getestet wurden wieder die reine Least-Squares-Methode und die hybride Methode mit uniformer Verfeinerung und adaptiver Verfeinerung bezüglich der entsprechenden Least-Squares-Funktionale. Die Konvergenzgraphen sind in Abbildung 5.15 eingetragen. Es ist zu erkennen, dass für die uniforme Verfeinerung deutlich geringere Konvergenzraten erreicht werden als für die adaptive. Wie in allen Experimenten liegen die Funktionale für die hybride Methode dabei leicht über denen der reinen Least-Squares-Methode.

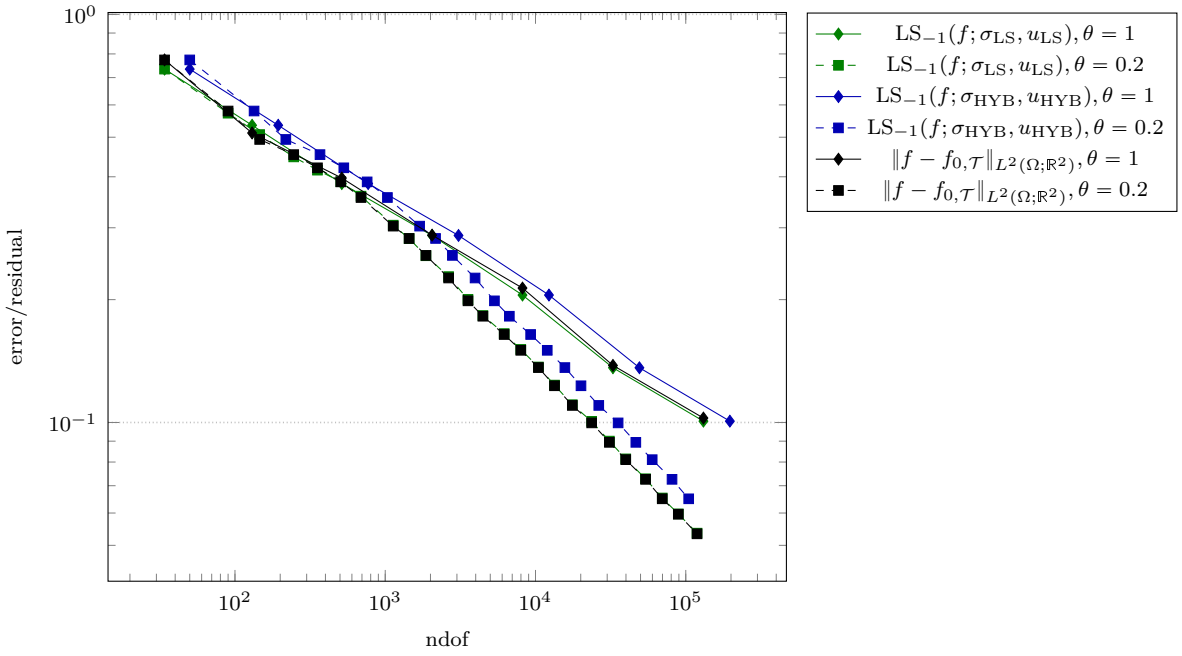

Abbildung 5.15: Konvergenzgraphen für hoch oszillierende f; Experiment zu Abschnitt 5.7

Ein zweites Beispiel betrachtet die Schachbrettfunktion als Volumenkraft. Die Schachbrettfunktion $f_{L,L_{\mathrm{fine}}}$ wird dabei abhängig von den zwei Parametern $L \in \mathbb{N}$ und $L_{\mathrm{fine}} \in \mathbb{N}$

rekursiv definiert. Die Anzahl der Iterationen wird durch die Parameter festgelegt und zu Beginn ist $f_{1,1}$ durch

$$f_{1,1}(x) := \begin{cases} 10 & \text{wenn } x_1, x_2 < 0.5 \text{ oder } x_1, x_2 > 0.5 \\ -10 & \text{sonst.} \end{cases}$$

gegeben. Diese Ausgangssituation wird in Abbildung 5.16 dargestellt. Nachfolgend werden mit Q_ℓ Quadrate der Seitenlänge $2^{-\ell}$ bezeichnet, auf denen $f_{\ell,1}$ einen konstanten Wert annimmt.

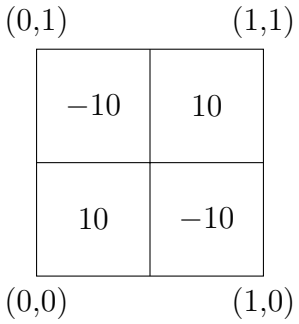

Abbildung 5.16: Initialisierung der Schachbrettfunktion

Nun folgen noch $L-1$ viele Iterationen. In jeder Iteration $\ell = 2, \ldots, L$ wird $f_{\ell,1}$ durch die Werte von $f_{\ell-1,1}$ folgender Maßen definiert. Alle Quadrate $Q_{\ell-1}$ werden in vier kleinere Quadrate $Q_{\ell,1}, \ldots, Q_{\ell,4}$ mit der Seitenlänge $2^{-\ell}$ aufgeteilt, ähnlich der Aufteilung von $Q_0 = \Omega$ in der Initialisierung. Dabei liegt $Q_{\ell,1}$ im unteren rechten Viertel von $Q_{\ell-1}$ und $Q_{\ell,4}$ im oberen linken. Wenn der konstante Wert von $f_{\ell-1,1}$ auf $Q_{\ell-1}$ mit $v_{\ell-1}$ bezeichnet wird, ist $f_{\ell,1}$ gegeben durch

$$f_{\ell,1}(x) := \begin{cases} 2v_{\ell-1} & \text{für } x \in Q_{\ell,1} \cup Q_{\ell,4}, \\ 0 & \text{für } x \in Q_{\ell,2} \cup Q_{\ell,3}. \end{cases}$$

Die Iterationen werden in Abbildung 5.17 verbildlicht. Danach wird für Quadrate Q_L, in denen $v_L = 0$ gilt, $f_{L,1}$ ersetzt durch

$$f_{L,1}(x) := \begin{cases} 1 & \text{für } x \in Q_{L+1,1} \cup Q_{L+1,4}, \\ -1 & \text{für } x \in Q_{L+1,2} \cup Q_{L+1,3}. \end{cases}$$

Nun folgen nochmal $L_{\text{fine}} - 1$ viele Iterationen $\ell = L+1, \ldots, L+L_{\text{fine}}$, in denen die Iteration von oben wiederholt wird. Dabei werden die Funktionen $f_{L,\ell-L+1}$ über die Funktionen $f_{L.\ell-L}$ definiert.

Die Konstruktion ist der Art, dass das Integralmittel $\fint_\Omega f_{L,L_{\text{fine}}} \,\mathrm{d}x$ in jeder Iteration verschwindet, aber die Funktion für grobe Gitter große Oszillationen aufweist. Ab einem von

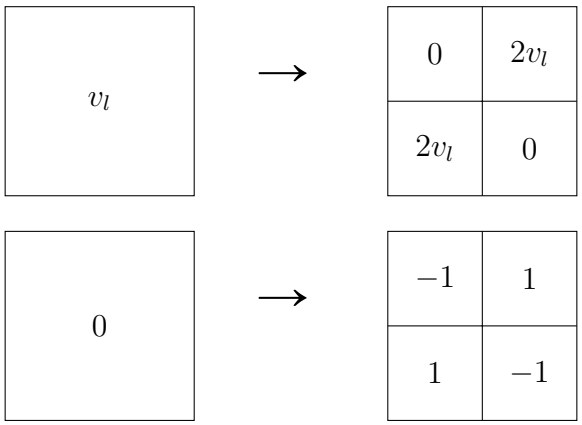

Abbildung 5.17: Iterationen bei der Schachbrettfunktion

L und L_{fine} abhängenden Freiheitsgrade sind aber alle Oszillationen vollständig aufgelöst. Damit verschwindet der Datenapproximationsfehler $\|f - f_l\|$ für jede Triangulierung, die fein genug ist. Im Experiment werden die Verschiedenen Verfeinerungsalgorithmen verglichen, es ist dabei zu beachten, dass wie in Abschnitt 4.5 nicht tatsächlich der alternative Schätzer berechnet wird. Die Markierungsparameter sind dabei $\Theta = 0.2$, $\kappa = 1$ und $\rho = 0.8$. Abbildung 5.18 enthält die Visualisierung der Ergebnisse. Es ist zu beobachten, dass im Falle der Dörflermarkierung bezüglich des Least-Squares-Funktionals (*separateMarking = 0*) die Datenfehler in sehr ähnlicher Rate aufgelöst werden, wie im Fall des separierten Markieren bezüglich der Differenz aus Least-Squares-Funktional und Datenfeher (*separateMarking = 1*). Zum Vergleich ist auch die Dörflermarkierung bezüglich der Differenz eingetragen (*separateMarking = 2*), die Rechnung wurde hier allerdings nach 100 Verfeinerungen abgebrochen. In allen Fällen ist zu beobachten, wie nach dem Auflösen der groben Oszillationen, bei ungefähr 100 Freiheitsgraden, der Fehlerwert signifikant fällt. Die feineren Oszillationen werden bei den ersten beiden Algorithmen bei ungefähr 10^4 Freiheitsgraden aufgelöst. Das Dörflermarkieren bezüglich der Differenz aus Least-Squares-Funktional und Datenfeher erreicht diese Anzahl an Freiheitsgraden nicht in den ersten 100 Iterationen, während die Oszillationen beim separierten Markieren im wesentlichen durch eine *CASE B* Verfeinerung aufgelöst werden.

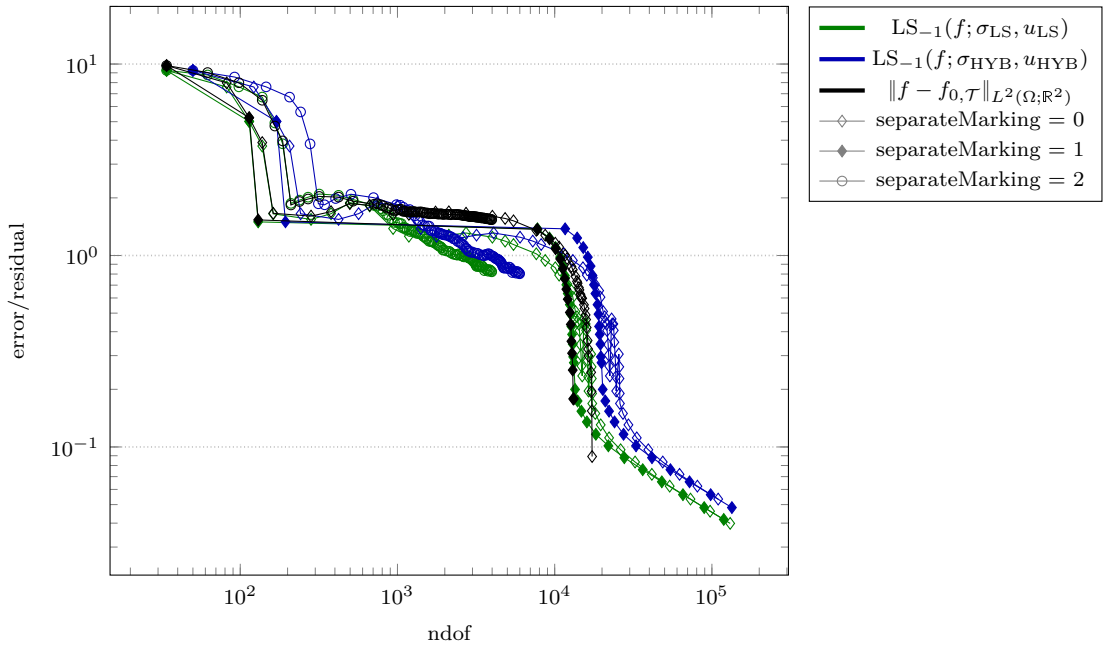

Abbildung 5.18: Konvergenzgraphen für die Schachbrettfunktion; Experiment zu Abschnitt 5.7

Danksagung

Ich möchte mich sehr bei Prof. Carstensen und Phillip Bringmann für die hilfreiche und umfangreiche Betreuung dieser Arbeit und die erhellenden und anregenden Gespräche bedanken.

Literatur

[ACFK02] J. Alberty, C. Carstensen, S. A. Funken und R. Klose. *Matlab implementation of the finite element method in elasticity. Computing* **69**.3 (2002), 239–263.

[BC05] C. Bahriawati und C. Carstensen. *Three MATLAB implementations of the lowest-order Raviart-Thomas MFEM with a posteriori error control. Comput. Methods Appl. Math.* **5**.4 (2005), 333–361 (electronic).

[BG09] Pavel B. Bochev und Max D. Gunzburger. *Least-squares finite element methods*. Bd. 166. Applied Mathematical Sciences. Springer, New York, 2009, S. xxii+660.

[BBF13] Daniele Boffi, Franco Brezzi und Michel Fortin. *Mixed finite element methods and applications*. Bd. 44. Springer Series in Computational Mathematics. Springer, Heidelberg, 2013, S. xiv+685.

[BS02] Susanne C. Brenner und L. Ridgway Scott. *The mathematical theory of finite element methods*. Second. Bd. 15. Texts in Applied Mathematics. Springer-Verlag, New York, 2002, S. xvi+361.

[Bri13] P. Bringmann. *Least-Squares Finite-Elemente-Methoden für die Stokes-Gleichungen* (2013). Bachelorarbeit, Humboldt-Universität zu Berlin, lokal verfügbar.

[BC16] Philipp Bringmann und Carsten Carstensen. *An adaptive least-squares FEM for the Stokes equations with optimal convergence rates. Numerische Mathematik* (2016 (published online)).

[CKS05] Zhiqiang Cai, Johannes Korsawe und Gerhard Starke. *An adaptive least squares mixed finite element method for the stress-displacement formulation of linear elasticity. Numer. Methods Partial Differential Equations* **21**.1 (2005), 132–148.

[CS03] Zhiqiang Cai und Gerhard Starke. *First-order system least squares for the stress-displacement formulation: linear elasticity. SIAM J. Numer. Anal.* **41**.2 (2003), 715–730 (electronic).

[CS04] Zhiqiang Cai und Gerhard Starke. *Least-squares methods for linear elasticity. SIAM J. Numer. Anal.* **42**.2 (2004), 826–842 (electronic).

[CH16] C. Carsten und F. Hellwig. *Low-order discontinuous Petrov-Galerkin finite element methods for linear elasticity* (2016). Eingereicht.

[Car+10] C. Carsten, J. Gedicke, L. Kern, J. Neumann, H. Rabus und M. Rozova. *AFEM-Dokumentation* (2010). lokal verfügbar.

[CDFH00] C. Carstensen, G. Dolzmann, S. A. Funken und D. S. Helm. *Locking-free adaptive mixed finite element methods in linear elasticity. Comput. Methods Appl. Mech. Engrg.* **190**.13-14 (2000), 1701–1718.

[CFPP14] C. Carstensen, M. Feischl, M. Page und D. Praetorius. *Axioms of adaptivity. Comput. Math. Appl.* **67**.6 (2014), 1195–1253.

[CP15] Carsten Carstensen und Eun-Jae Park. *Convergence and optimality of adaptive least squares finite element methods. SIAM J. Numer. Anal.* **53**.1 (2015), 43–62.

[EGK02] Christof Eck, Harald Garcke und Peter Knabner. *Mathematische Modellierung*. Second. Bd. 2. Springer-Verlag, Berlin heidelberg, 2002, S. xiv+513.

[Gal12] D. Gallistl. *Über nichtkonforme Finite-Elemente-Diskretisierungen der biharmonischen Gleichung* (2012). Diplomarbeit, Humboldt-Universität zu Berlin.

[Hel14a] R. Hella. *On the quasi-optimal convergence of adaptive nonconforming finite element methods in three examples* (2014). Diplomarbeit, Humboldt-Universität zu Berlin.

[Hel14b] F. Hellwig. *Drei dPG-Methoden niedriger Ordnung für Lineare Elastizität* (2014). Masterarbeit, Humboldt-Universität zu Berlin, lokal verfügbar.

[KS95] Reijo Kouhia und Rolf Stenberg. *A linear nonconforming finite element method for nearly incompressible elasticity and Stokes flow. Comput. Methods Appl. Mech. Engrg.* **124**.3 (1995), 195–212.

Anhang - Verzeichnisstrucktur

```
afem-lsfem/
├── afemElasicityLS.m
├── mAfemEx.m
├── bethke/
│   ├── FigureAdvanced.m
│   ├── c.m
│   ├── collectMdata.m
│   └── readProblem.m
├── configs/
│   ├── readConfig.m
│   ├── readMconfig.m
│   ├── default.config
│   └── default.mconfig
├── common/
│   ├── loadGeometry_elasticity.m
│   └── shiftMembraneWithHole.m
├── data/
│   ├── Checkerboard.m
│   ├── CircularInclusion.m
│   ├── CookMembrane.m
│   ├── Kirsch.m
│   ├── LShapeLocking.m
│   ├── LshapeRotExact.m
│   ├── QuickOscillation.m
│   └── SquareExact.m
├── estimate/
│   ├── computeElasticityLSResidual.m
│   └── computeQeError.m
├── integrate/
│   ├── parIntegrate.m
│   ├── error4eElasticityLS.m
│   └── errorP1elasticEnergy.m
├── plot/
│   ├── plotTriangulationDebug.m
│   └── plotElasticityLS.m
├── refine/refineBi3GB_elasticity.m
├── results/elasticity/ABGABE/
└── solve/
    ├── solveElasticityLS.m
    ├── solveElasticityHybrid.m
    └── solveElasticityP1P1wrapper.m
```

Selbstständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe und ich zum ersten Mal eine Bachelorarbeit in diesem Studiengang einreiche.

Berlin, den 8. Juli 2016