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

Eine neue dPG-Finite-Elemente-Methode niedriger Ordnung für Stokes Gleichungen

Masterarbeit

zur Erlangung des akademischen Grades
Master of Science (M. Sc.)

eingereicht von:	Sophie Louise Puttkammer
geboren am:	29.04.1991
geboren in:	Berlin
Betreuer:	Prof. Dr. Carsten Carstensen

Berlin, den 29. September 2015

Inhaltsverzeichnis

1 Einleitung

Seien es Bauteile von Windrädern oder Flugzeugen, Wärmeverteilungen in Reinigungsanlagen, elektromagnetische Effekte in Mikrochips, in zahlreichen Anwendungen müssen partielle Differentialgleichungen gelöst werden, um Ergebnisse zu optimieren oder vorherzusagen. Das liegt daran, dass die mathematischen Modelle vieler physikalischer Prozesse genau solche Gleichungen sind. So gibt es die Maxwell Gleichungen für den Elektromagnetismus, die Navier-Lamé Gleichungen in der linearen Elastizität oder die Wärmeleitungsgleichungen, um die Temperaturverteilung in einem Medium zu charakterisieren. Diese Gleichungen haben gemein, dass sie nur in Spezialfällen analytisch gelöst werden können. Daher wurden eine Vielzahl numerischer Methoden entwickelt, um diese Probleme anzugehen.

In dieser Arbeit wird eine neue diskontinuierliche Petrov-Galerkin (dPG) Finite-Elemente-Methode zur Lösung von Stokes Gleichungen vorgestellt.

In der ersten Hälfte des 19. Jahrhunderts formulierten Claude Navier und Georg Stokes unabhängig voneinander ein System partieller Differentialgleichungen zur Beschreibung des Strömungsverhaltens Newtonscher Flüssigkeiten. Ein Spezialfall und grundlegendes Modellproblem der Strömungsmechanik sind die Stokes Gleichungen für inkompressible Newtonsche Flüssigkeiten. In Abschnitt 2.6 werden einige der zahlreichen Formulierungen dieses Systems erläutert und hergeleitet. Das Modell in dieser Arbeit basiert auf der Pseudospannungs-Geschwindigkeits Formulierung. Gegeben ein Gebiet Ω, eine äußere Kraftdichte $f \in L^2(\Omega;\mathbb{R}^n)$ und Randdaten $g \in L^2(\partial\Omega;\mathbb{R}^n)$, werden dabei ein Geschwindigkeitsfeld $u : \Omega \to \mathbb{R}^n$ und eine Pseudospannung $\boldsymbol{\sigma} : \Omega \to \mathbb{R}^{n\times n}$ gesucht, für die gilt

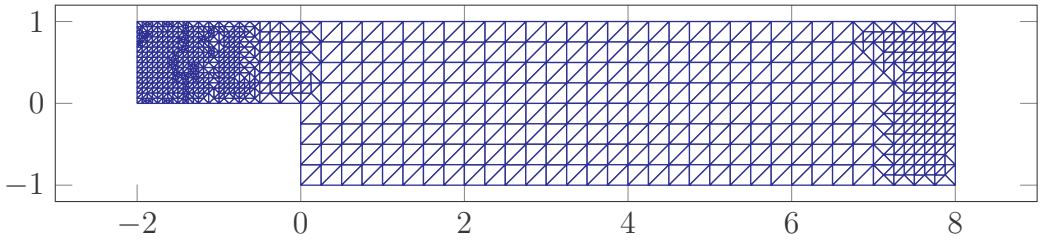
$$\begin{aligned} \operatorname{div}\boldsymbol{\sigma} + f &= 0 \quad && \text{in } \Omega, \\ \operatorname{dev}\boldsymbol{\sigma} - \mathrm{D}\,u &= 0 && \text{in } \Omega, \\ u &= g && \text{entlang } \partial\Omega. \end{aligned}$$

Diese Gleichungen spielen beispielsweise eine Rolle, wenn Strömungen an einer rückwärts gewandten Stufe betrachtet werden, wie im "backward facing step" Beispiel. Für diese Anwendung kann im Allgemeinen analytisch keine Lösung bestimmt werden. Gegeben sind ein leicht verzerrtes L-Gebiet, $\Omega = ((-2,8\times(-1,1)) \setminus ((-2,0)\times(-1,0))$, sowie die konstante rechte Seite $f \equiv 0$ und die Dirichlet Randdaten für $(x_1,x_2)^\top \in \partial\Omega$

$$g\left((x_1,x_2)^\top\right) = \begin{cases} (0,0)^\top & \text{für } -2 < x_1 < 8, \\ 1/10\,(-x_2(x_2-1),0)^\top & \text{für } x_1 = -2, \\ 1/80\,(-(x_2-1)(x_2+1),0)^\top & \text{für } x_1 = 8. \end{cases}$$

Mit dem im Zuge dieser Arbeit implementierten Programm kann für dieses Problem das in Abbildung 1.1 dargestellte Geschwindigkeitsfeld ermittelt werden. In Kapitel 5

(a) Triangulierung $\mathcal{T}_\ell$



(b) Geschwindigkeitsfeld u_ℓ

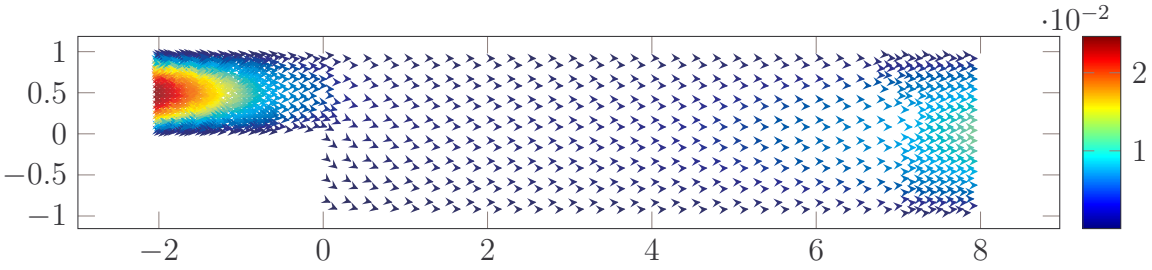


Abbildung 1.1: Triangulierung mit 1551 Elementen (15511 Freiheitsgraden) und entsprechendes Geschwindigkeitsfeld im "backward facing Step" Beispiel bei adaptiver Verfeinerung mit $\theta = 0.3$

werden dieses und weitere mit der neuen dPG-Methode untersuchte Beispiele genauer diskutiert.

Dem Namen nach ist die dPG-Methode eine Verbindung aus einer diskontinuierlichen Galerkin-Methode, bei welcher gebrochene Test- und Ansatzfunktionen möglich sind, und einer Petrov-Galerkin-Methode, bei der sich Test- und Ansatzraum unterscheiden dürfen. Der Begriff diskontinuierliche Petrov-Galerkin Methode wurde zuerst in [BMS02] verwendet. Dort wurde bereits eine ultra-schwache Formulierung betrachtet, es wurden also alle Ableitungen auf die Testfunktionen verschoben. Außerdem wurden die numerischen Flüsse der diskontinuierlichen Galerkin-Methode durch unabhängige, nur auf Kanten definierte Variablen ersetzt. Die Idee, die optimalen Testräume, welche die Methode ursprünglich motivierten, zu bestimmen, wurde von Leszek Demkowicz und Jay Gopalakrishnan aufgegriffen. In [DG10, DG11a, DG11b, DGN12] untersuchten sie diese Klasse von Methoden und ihre Anwendung beispielsweise für Wellentransportprobleme und Konvektions-Diffusions-Gleichungen. In [RBTD14] betrachteten Demkowicz, Roberts und Bui-Than die Anwendungsmöglichkeiten der dPG-Methode für Stokes Probleme. Sie verwenden dabei stückweise Polynome vom Grad $k \geq 1$ zur Approximation der vorkommenden L^2-Funktionen, also des Spannungstensors und des Geschwindigkeitsfeldes, und für die Normalenspuren der Spannung in $H^{-1/2}(\partial\mathcal{T};\mathbb{R}^n)$. Die Spuren der Geschwindigkeit in $H^{1/2}(\partial\mathcal{T};\mathbb{R}^n)$ werden mit stückweisen Polynomen des Grades $k+1$ diskretisiert. Der Testraum wird bei dieser Methode durch sukzessives Anreichern und Vergleichen der Ergebnisse ermittelt. Ziel ist es, die minimale Anreicherung zu finden, bei der die Ergebnisse ungefähr so gut sind wie bei größeren Anreicherungen. In ihren Experimenten wird meist der Anreicherungsgrad 1 verwendet, d.h. stückweise Polynome und Raviart-Thomas Funktionen vom Grad $k+2$.
Die in dieser Arbeit vorgestellte neue dPG-Methode verringert den Rechenaufwand enorm, denn sie kommt mit dem diskreten Ansatzraum

$$X_h := P_0\left(\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R} \times P_0\left(\mathcal{T};\mathbb{R}^n\right) \times S_0^1\left(\mathcal{E};\mathbb{R}^n\right) \times P_0\left(\mathcal{E};\mathbb{R}^n\right)$$

und dem diskreten Testraum

$$Y_h := RT_0^{\mathrm{pw}}\left(\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R} \times P_1\left(\mathcal{T};\mathbb{R}^n\right)$$

aus. Dass hier sämtliche theoretischen Betrachtungen für eine beliebige Dimension $n \in \mathbb{N}$ durchgeführt werden, erlaubt dabei eine problemlose Verallgemeinerung der vorgestellten zweidimensionalen Implementierung.

Im Allgemeinen kann eine dPG-Methode als Least-Squares Finite-Elemente-Methode, die das Residuum der ultra-schwachen Formulierung in einer Nicht-Standardnorm minimiert, oder als Finite-Elemente-Methode mit Nicht-Standardteträumen aufgefasst werden. Den Betrachtungen liegt dabei eine wohlgestellte Variationsformulierung einer partiellen Differentialgleichung zu Grunde. Seien dazu zwei Banachräume X und Y, eine stetige Bilinearform $b : X \times Y \to \mathbb{R}$ und ein Funktional $F \in Y^*$ gegeben. Gesucht wird dann

eine Lösung $x \in X$, welche für jedes $y \in Y$ die Gleichung

$$b(x, y) = F(y)$$

erfüllt. Wie bereits erwähnt ist eine Besonderheit der dPG-Methoden, dass schon auf diesem kontinuierlichen Level eine Triangulierung zu Grunde gelegt wird, also gebrochene Testfunktionen betrachtet werden. In der praktischen dPG-Methode wird für diskrete Unterräume $X_h \subseteq X$ und $Y_h \subseteq Y$ ein $x_h \in X_h$ gesucht, für das gilt

$$x_h \in \operatorname*{argmin}_{\xi_h \in X_h} \|F - b(\xi_h, \bullet)\|_{Y_h^*} .$$

Aus [CDG14] und [CH15] geht hervor, dass die folgenden Bedingungen genügen, um gewisse Fehlerabschätzungen zu garantieren. Im Hilbertraum Fall ist zu zeigen, dass die Bilinearform b beschränkt ist und einen trivialen Kern besitzt und dass die folgende kontinuierliche Inf-Sup-Bedingung erfüllt ist

$$0 < \beta := \inf_{x \in X \setminus \{0\}} \sup_{y \in Y \setminus \{0\}} \frac{b(x, y)}{\|x\|_X \|y\|_Y}.$$

Für die konforme Diskretisierung ist dann lediglich die diskrete Inf-Sup-Bedingung

$$0 < \beta_h := \inf_{x_h \in X_h \setminus \{0\}} \sup_{y_h \in Y_h \setminus \{0\}} \frac{b(x_h, y_h)}{\|x_h\|_X \|y_h\|_Y}$$

zu überprüfen. Dieser letzte Punkt ist natürlich leichter zu garantieren, je größer der diskrete Testraum $Y_h \subseteq Y$ gewählt wird. Sind die genannten Bedingungen gleichzeitig erfüllt, so sind optimale Konvergenz und diskrete Stabilität im Sinn der folgenden a-priori-Abschätzung gewährleistet

$$\|x - x_h\|_X \leq C \min_{\xi_h \in X_h} \|x - \xi_h\|_X ,$$

wobei $C \leq \|b\|^2 / \beta\beta_h$ eine generische nur von Ω abhängige Konstante ist. Außerdem gilt die a-posteriori-Abschätzung

$$\begin{aligned} \beta\|x - \xi_h\|_X \leq \|b\|/\beta_h \, \|F - b(\xi_h, \bullet)\|_{Y_h^*} + \|F \circ (1 - \Pi)\|_{Y^*} \\ \leq 2\|b\|^2/\beta_h \|x - \xi_h\|_X \end{aligned}$$

für jedes $\xi_h \in X_h$, die Methode ist also auch bei inexaktem Lösen stabil. In dieser Abschätzung kommt neben dem Fehlerschätzer $\|F - b(\xi_h, \bullet)\|_{Y_h^*}$, der für adaptives Verfeinern verwendet wird, auch einen Datenapproximationsfehler $\|F \circ (1 - \Pi)\|_{Y^*}$ vor. Dabei wird mit $\Pi : Y \to Y_h$ ein Fortin-Interpolator bezeichnet, also ein beschränkter, linearer Operator, der die Eigenschaft $b(\bullet, (1 - \Pi)y) = 0$ in X^* für alle $y \in Y$ erfüllt. Die Existenz dieses Operators ist im Grunde zu der diskreten Inf-Sup-Bedingung äquivalent. In den meisten Arbeiten wird dieser Operator daher direkt konstruiert und auf den Beweis der

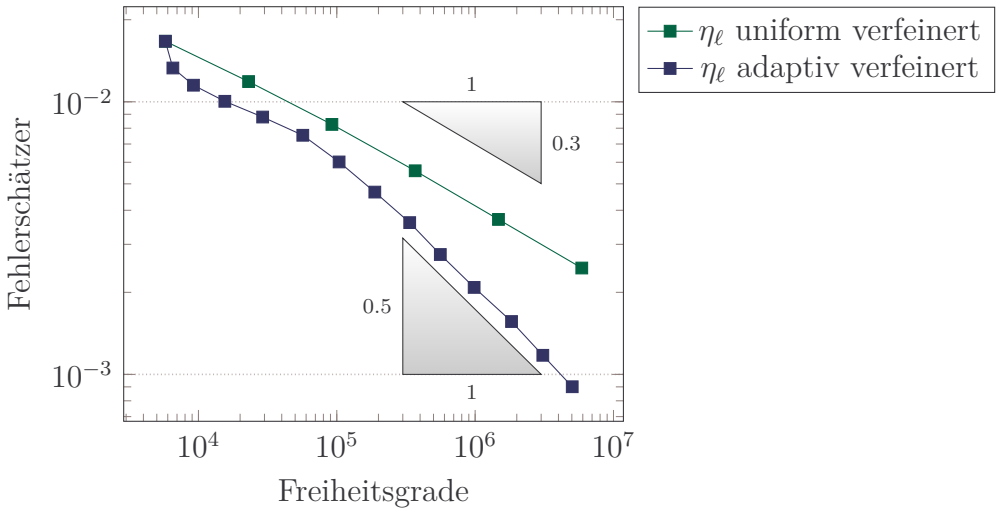


Abbildung 1.2: Konvergenzplot des Fehlerschätzers für das "backward facing step" Beispiel bei uniformem Verfeinern und adaptivem Verfeinern mit Bulkparameter $\theta = 0.3$

letzten Bedingung verzichtet. Die hier gewählte Herangehensweise, die diskrete Inf-Sup-Bedingung direkt nachzuweisen, hat den Vorteil, dass ein deutlich kleinerer Testraum gewählt werden kann. Dafür ist der Datenapproximationsfehler schwerer zu bestimmen. Im zweidimensionalen Fall kann allerdings ein solcher Operator Π angegeben werden. Damit schließt diese Arbeit direkt an die aktuelle Forschung an. Die numerischen Ergebnisse zeigen, dass bereits mit Hilfe von $\|F - b(\xi_h, \bullet)\|_{Y_h^*}$ optimale Konvergenzraten des adaptiven Verfahrens erzielt werden, wie z.B. in Abbildung 1.2 zu erkennen.

In diesem allgemeinen Rahmen wird die neue dPG-Methode hergeleitet. Dabei wird das Residuum der ultra-schwachen Formulierung in der Dualnorm des gebrochenen Testraumes minimiert. In der ultra-schwachen Form der Pseudostress-Geschwindigkeits Formulierung der Stokes Gleichungen werden die Spuren der Geschwindigkeit und Normalspuren der Pseudospannung auf dem Rand der einzelnen Elemente der Triangulierung durch neue Variablen substituiert. Dies führt zu folgender Bilinearform $b : X \times Y \to \mathbb{R}$, für alle

$$\begin{aligned} x = (\boldsymbol{\sigma}, u, s, t) \in X :=& L^2\left(\Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R} \times L^2(\Omega; \mathbb{R}^n) \times H_0^{1/2}(\partial\mathcal{T}; \mathbb{R}^n), \\ & \times H^{-1/2}(\partial\mathcal{T}; \mathbb{R}^n) \\ y = (\boldsymbol{\tau}, v) \in Y :=& H\left(\operatorname{div}, \mathcal{T}; \mathbb{R}^{n\times n}\right)/\mathbb{R} \times H^1(\mathcal{T}; \mathbb{R}^n) \end{aligned}$$

sei

$$\begin{aligned} b(x, y) :=& -\langle t, \gamma_0 v\rangle_{\partial\mathcal{T}} - \langle \gamma_\nu \boldsymbol{\tau}, s\rangle_{\partial\mathcal{T}} \\ & + \int_\Omega \boldsymbol{\sigma} : \mathrm{D}_{\mathrm{NC}}\, v \,\mathrm{d}x + \int_\Omega \operatorname{dev}\boldsymbol{\sigma} : \boldsymbol{\tau} \,\mathrm{d}x + \int_\Omega u \cdot \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau} \,\mathrm{d}x. \end{aligned}$$

Für diese Bilinearform die zuvor erwähnten Bedingungen nachzuweisen, wird einen Großteil dieser Arbeit ausmachen. Der Beweis der kontinuierlichen Inf-Sup-Bedingung wird durch das Splitting Lemma und die Wohlgestelltheit der dualen gemischten Formulierung ermöglicht. Die diskrete Inf-Sup-Bedingung wird nachgewiesen, in dem zu jedem $x_h \in X_h$ ein $\tilde{y}_h \in Y_h$ so konstruiert wird, dass die Ungleichungen $\|\tilde{y}_h\|_Y \leq C_{\tilde{y}_h} b(x_h, \tilde{y}_h)$ und $\|x_h\|_X \leq C_{x_h} b(x_h, \tilde{y}_h)$ für zwei direkt angegebene generische Konstanten $C_{\tilde{y}_h}$, C_{x_h} gelten. Dabei erweisen sich die Spurterme als besonders kompliziert, weshalb Fortsetzungen $w_c \in P_1(\mathcal{T}; \mathbb{R}^n)$ von $s_1 \in S_0^1(\mathcal{E}; \mathbb{R}^n)$ bzw. $\boldsymbol{q}_{\mathrm{RT}} \in RT_0(\mathcal{T}; \mathbb{R}^{n\times n})$ von $t \in P_0(\mathcal{E}; \mathbb{R}^n)$ verwendet werden.

Diese Arbeit ist wie folgt aufgebaut. Der nächste Abschnitt beschäftigt sich mit der Einführung bzw. Wiederholung grundlegender Notationen und Definitionen. Des Weiteren werden die kontinuierlichen und diskreten Funktionsräume vorgestellt und einige bekannte bzw. häufiger benötigte Lemmata zitiert bzw. nachgewiesen. Zuletzt erfolgt eine kurze Einführung zu den Stokes Gleichungen und einigen numerischen Verfahren zu deren Lösung.
Im dritten Kapitel wird die neue dPG-Methode analysiert. Dazu werden zunächst die allgemeinen Grundlagen zu dieser Klasse numerischer Methoden erläutert. Es werden die verschiedenen Problemformulierungen eingeführt und die Bedingungen für die bereits erwähnten a-priori- und a-posteriori-Fehlerabschätzungen vorgestellt. Diese Bedingungen werden anschließend nachgewiesen – zunächst für das auf kontinuierlichem Level formulierte Problem und danach für die gewählte Diskretisierung. Als Letztes wird in diesem Abschnitt noch der Fortin-Interpolator dieser Methode untersucht, der Gegenstand aktueller Forschung ist.
Der vierte Abschnitt widmet sich der Implementierung. Dazu werden zunächst die grundlegenden Berechnungen durchgeführt und anschließend die im Zuge dieser Arbeit entstandenen Programme kurz vorgestellt.
Vor der Zusammenfassung werden im fünften Kapitel einige numerischen Experimente ausgewertet, die mit Hilfe der entstandenen Programme durchgeführt wurden. Diese Benchmark Probleme für die Stokes Gleichungen erlauben, die Ergebnisse mit denen anderer Methoden zu vergleichen.

2 Vorbereitung

In diesem Kapitel werden grundlegende Notationen eingeführt bzw. wiederholt und gewisse nützliche Eigenschaften der verwendeten Räume und Operatoren festgehalten. In dieser Arbeit wird die Schreibweise $a \lesssim b$ verwendet, wenn eine generische Konstante C existiert, so dass $a \leq Cb$ gilt. Des Weiteren wird $a \lesssim b \lesssim a$ durch $a \approx b$ abgekürzt.

2.1 Triangulierungen

In dieser Arbeit sei $\Omega \subseteq \mathbb{R}^n$ stets ein beschränktes, polygonal berandetes Lipschitz-Gebiet mit Dirichlet Rand Γ. Dieses Gebiet wird in abgeschlossene n-Simplexe unterteilt, so dass sich eine reguläre Triangulierung ergibt.

Definition 2.1 (ℓ-Simplex). Sei $\ell \in \mathbb{N}$ und $\ell \leq n$. Eine Menge $T \subset \mathbb{R}^n$ heißt *ℓ-Simplex*, falls T die konvexe Hülle von $\ell + 1$ paarweise linear unabhängigen Punkten im $\mathbb{R}^n$ ist.

Bei einen n-Simplex T bezeichne $\mathcal{N}(T) := \{P_1, \ldots, P_{n+1}\}$ die Menge aller Eckpunkte bzw. Knoten von T, $\mathcal{E}(T) = \{\operatorname{conv}(\mathcal{P}) : \mathcal{P} \subset \mathcal{N}(T), |\mathcal{P}| = n\}$ die Menge aller Seiten von T und $\operatorname{int}(T)$ alle inneren Punkte.

Definition 2.2 (Reguläre Triangulierung). Eine *reguläre Triangulierung* von Ω ist eine endlich Menge $\mathcal{T}$ aus abgeschlossenen n-Simplexen mit folgenden Eigenschaften

(i) $\operatorname{int}(T) \neq \emptyset$ für alle $T \in \mathcal{T}$,

(ii) $\overline{\Omega} = \bigcup_{T \in \mathcal{T}} T$,

(iii) $\operatorname{int}(T) \cap \operatorname{int}(\hat{T}) = \emptyset$ für alle $T, \hat{T} \in \mathcal{T}$ mit $T \neq \hat{T}$,

(iv) wenn für $T, \hat{T} \in \mathcal{T}$ mit $T \neq \hat{T}$ gilt $T \cap \hat{T} \neq \emptyset$, so ist der Schnitt von T und $\hat{T}$ die konvexe Hülle der $\ell \leq n$ gemeinsamen Knoten und $T \cap \hat{T} = \operatorname{conv}\left(\mathcal{N}(T) \cap \mathcal{N}\left(\hat{T}\right)\right)$ ist ein $\ell - 1$ Simplex, d.h. es gibt in der Triangulierung keine hängenden Knoten.

Für eine reguläre Triangulierung $\mathcal{T}$ von Ω sei $\mathcal{N} := \bigcup_{T \in \mathcal{T}} \mathcal{N}(T)$ die Menge aller Knoten der Triangulierung und $\mathcal{N}(\Omega)$ bzw. $\mathcal{N}(\partial\Omega)$ die Mengen der inneren Knoten bzw. Randknoten. Mit $\mathcal{E}$, $\mathcal{E}(\Omega)$ bzw. $\mathcal{E}(\partial\Omega)$ werden die Menge aller Seiten, der inneren Seiten bzw. der Randseiten der Triangulierung beschrieben.

Definition 2.3 (Skelett). Das *Skelett* einer regulären Triangulierung $\mathcal{T}$ ist durch

$$\partial\mathcal{T} := \bigcup_{T \in \mathcal{T}} \bigcup_{E \in \mathcal{E}(T)} E$$

definiert.

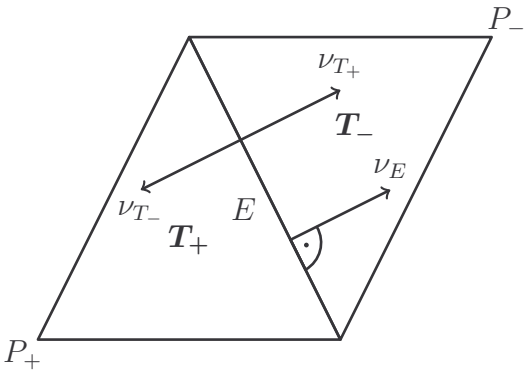


Abbildung 2.1: Seitenpatch ω_E für $E \in \mathcal{E}(\Omega)$ für $\Omega \subseteq \mathbb{R}^2$

Auf jedem n-Simplex $T \in \mathcal{T}$ bezeichne ν_T das äußere Einheitsnormalenvektorfeld entlang des Randes ∂T. Außerdem ist jede Seite $E \in \mathcal{E}$ mit einer Orientierung der Einheitsnormale ν_E, wie in Abbildung 2.1 dargestellt, ausgestattet. Entlang einer Außenseite, $E \in \mathcal{E}(\partial\Omega)$, zeige ν_E stets nach außen. Entlang einer inneren Seite, $E = \partial T_+ \cap \partial T_- \in \mathcal{E}(\Omega)$, werde eine der beiden möglichen Orientierungen der Normale ν_E fixiert. Die Nachbarelemente werden stets so benannt, dass ν_E von T_+ nach T_- zeigt, also $\nu_E = \nu_{T_+}$.
Damit wird die folgende Vorzeichenfunktion definiert

$$\operatorname{sgn}(T, E) := \nu_E \cdot \nu_T \in \{\pm 1\}, \text{ für alle } T \in \mathcal{T},\ E \in \mathcal{E}(T). \tag{2.1}$$

Die Konvention der Elementbenennung wird auch bei der Definition von Sprüngen verwendet, welche bei nur stückweise stetigen Funktionen auf der Triangulierung $\mathcal{T}$ interessant werden.

Definition 2.4 (Sprung). Der *Sprung* einer gegebenen Lebesgue integrierbaren Funktion $v \in L^2(\mathcal{T}; \mathbb{M})$ (siehe Abschnitt 2.3) mit $\mathbb{M} \subset \mathbb{R}^{m\times n}$, n, $m \in \mathbb{N}_{>0}$ entlang einer inneren Seite $E \in \mathcal{E}(\Omega)$ ist definiert als $[v]_E := (v|_{T_+} - v|_{T_-})\big|_E \in L^2(E; \mathbb{M})$ und entlang einer Randseite $E \in \mathcal{E}(\partial\Omega)$ durch $[v]_E := v|_E \in L^2(E; \mathbb{M})$.

Die folgenden Größen der Triangulierung bzw. der Elemente werden in dieser Arbeit ebenfalls eine Rolle spielen.

Es bezeichne $P_0(\mathcal{T}; \mathbb{M})$ mit $\mathbb{M} \subset \mathbb{R}^{m\times n}$, n, $m \in \mathbb{N}_{>0}$, die stückweise konstanten Funktionen, die auf jedem Element $T \in \mathcal{T}$ einen konstanten Wert in $\mathbb{M}$ haben, also in $P_0(T; \mathbb{M})$ liegen (siehe auch Abschnitt 2.4).

Für jedes Element $T \in \mathcal{T}$ sei

$$h_T := \operatorname{diam}(T) = \max_{x,\, y \in T} |x - y| \in P_0(T)$$

der Durchmesser und

$$h_{\max} := \max_{T \in \mathcal{T}} h_T$$

bezeichne die maximale Netzweite der Triangulierung $\mathcal{T}$. Des Weiteren sei die Funktion $h_{\mathcal{T}} \in P_0(\mathcal{T};\mathbb{R})$ für jedes $T \in \mathcal{T}$ definiert durch $h_{\mathcal{T}}|_T := h_T$.

Der Schwerpunkt eines n-Simplexes $T \in \mathcal{T}$ wird mit

$$\operatorname{mid}(T) := \fint_T x \,\mathrm{d}x = |T|^{-1} \int_T x \,\mathrm{d}x = \frac{1}{n+1} \sum_{z \in \mathcal{N}(T)} z$$

bezeichnet. Die Funktion $\operatorname{mid}(\mathcal{T}) \in P_0(\mathcal{T};\mathbb{R}^n)$ habe den Wert $\operatorname{mid}(T)$ auf jedem $T \in \mathcal{T}$. Wenn außerdem $\bullet$ die Identitätsabbildung in $\mathbb{R}^n$ und $\perp$ die Orthogonalität in $L^2(T)$ bezeichnen, so gilt

$$\bullet - \operatorname{mid}(T) \perp P_0(\mathcal{T}) \quad \text{und} \quad \|\bullet - \operatorname{mid}(T)\|_{L^\infty(\Omega)} \leq h_{\max}\, n/(n+1). \tag{2.2}$$

Die erste Tatsache ergibt sich sofort aus der Integraldefinition des Schwerpunkts, die zweite ergibt sich auf $T = \operatorname{conv}\{P_1, \ldots, P_{n+1}\} \in \mathcal{T}$, da das Maximum von $|x - \operatorname{mid}(T)|$ für $x \in T$ in einem Eckpunkt angenommen wird. O.B.d.A werde das Maximum bei P_1 erreicht, dann gilt für alle $x \in T$

$$\begin{aligned} |x - \operatorname{mid}(T)| &\leq \left| P_1 - \frac{1}{n+1} \sum_{j=1}^{n+1} P_j \right| = \frac{1}{n+1} \left| n P_1 - \sum_{j=2}^{n+1} P_j \right| \\ &\leq \frac{1}{n+1} \sum_{j=2}^{n+1} \left| P_1 - P_j \right| \leq h_T \frac{n}{n+1}. \end{aligned}$$

2.2 Vektor und Matrix Notation

Dieser Paragraph stellt Details der in dieser Arbeit verwendeten Matrix und Vektor Notation klar.

Für zwei Vektoren a, $b \in \mathbb{R}^n$ wird das Skalarprodukt wie folgt notiert,

$$a \cdot b = a^\top b = \sum_{j=1}^{n} a_j b_j \in \mathbb{R},$$

wohingegen $A : B$ das Skalarprodukt von zwei $n \times n$ Matrizen A, $B \in \mathbb{R}^{n \times n}$ symbolisiert,

$$A : B = \sum_{j,\,k=1}^{n} A_{jk} B_{jk} \in \mathbb{R}.$$

Für einen Vektor $a \in \mathbb{R}^n$ und eine Matrix $B \in \mathbb{R}^{n \times m}$ sei $a \cdot B = a^\top B \in R^{1 \times m}$.

Das dyadische Produkt zweier Vektoren a, $b \in \mathbb{R}^n$ ist definiert als

$$a \otimes b := a b^\top \in \mathbb{R}^{n \times n}.$$

Es gilt $|a \otimes b| = |a||b|$.

Mit $\bullet$ wird jeweils die Identitätsabbildung bezeichnet. Die Notation $|\bullet|$ ist kontextabhängig, sie steht für die durch $\cdot$ bzw. : induzierte Norm auf $\mathbb{R}^n$ bzw. $\mathbb{R}^{n\times n}$, die Kardinalität einer endlichen Menge oder das n- bzw. $(n-1)$-dimensionale Lebesque-Maß einer Teilmenge von $\mathbb{R}^n$.

In einem normierten Vektorraum $\left(X, \|\bullet\|_X\right)$ wird die Sphäre wie folgt definiert

$$S(X) := \left\{x \in X : \|x\|_X = 1\right\}.$$

Der Deviator einer Matrix $A \in \mathbb{R}^{n\times n}$ ist ein linearer Operator definiert durch $\operatorname{dev} A := A - 1/n \ \operatorname{tr}(A)\ \mathrm{I}_{n\times n}$, wobei der lineare Operator $\operatorname{tr} A = A_{11} + \cdots + A_{nn} = \sum_{j=1}^{n} A_{jj}$ die Spur der Matrix bezeichnet. Der Raum $\mathbb{R}^{n\times n}_{dev}$ enthält alle deviatorischen, also spurfreien, $n \times n$ Matrizen, d.h.

$$\mathbb{R}^{n\times n}_{dev} := \left\{A \in \mathbb{R}^{n\times n} : \operatorname{dev} A = A\right\} = \left\{A \in \mathbb{R}^{n\times n} : \operatorname{tr} A = 0\right\}. \tag{2.3}$$

Lemma 2.5 (Eigenschaften des Deviators). *Es gilt*

(i) $\operatorname{tr}\operatorname{dev} A = 0$ *für alle* $A \in \mathbb{R}^{n\times n}$ *und damit* $\operatorname{dev}\operatorname{dev} A = \operatorname{dev} A$,

(ii) $\operatorname{dev} A : \operatorname{dev} B = \operatorname{dev} A : B = A : \operatorname{dev} B$ *für alle* $A, B \in \mathbb{R}^{n\times n}$,

(iii) $\|\boldsymbol{\tau}\|^2_{L^2(\Omega)} = 1/n \ \|\operatorname{tr}\boldsymbol{\tau}\|^2_{L^2(\Omega)} + \|\operatorname{dev}\boldsymbol{\tau}\|^2_{L^2(\Omega)}$ *für alle* $\boldsymbol{\tau} \in L^2\left(\Omega; \mathbb{R}^{n\times n}\right)$, *insbesondere* $\|\operatorname{dev}\boldsymbol{\tau}\|_{L^2(\Omega)} \leq \|\boldsymbol{\tau}\|_{L^2(\Omega)}$.

Beweis. Punkt (i) gilt, da $\operatorname{tr}(\mathrm{I}_{n\times n}) = n$. Punkt (ii) ergibt sich, weil $\mathrm{I}_{n\times n} : A = \operatorname{tr}(A)$ für alle $A \in \mathbb{R}^{n\times n}$ und damit $\operatorname{dev} A : \operatorname{dev} B = A : B - 1/n \operatorname{tr}(A)\operatorname{tr}(B) = \operatorname{dev} A : B = A : \operatorname{dev} B$. Punkt (iii) folgt aus der Orthogonalität von deviatorischem und isochorischem Anteil, denn

$$\operatorname{dev}\boldsymbol{\tau} : \operatorname{tr}(\boldsymbol{\tau})\ \mathrm{I}_{n\times n} = (\boldsymbol{\tau} - 1/n \ \operatorname{tr}(\boldsymbol{\tau})\ \mathrm{I}_{n\times n}) : \operatorname{tr}(\boldsymbol{\tau})\ \mathrm{I}_{n\times n} = \operatorname{tr}(\boldsymbol{\tau})^2 - \operatorname{tr}(\boldsymbol{\tau})^2 = 0$$

und $\boldsymbol{\tau} = \operatorname{dev}\boldsymbol{\tau} + 1/n \operatorname{tr}(\boldsymbol{\tau})\mathrm{I}_{n\times n}$. Dies ist lediglich in die Definition der L_2-Norm aus Abschnitt 2.3 einzusetzen. □

2.3 Funktionenräume

In diesem Paragraphen werden die benötigten Funktionenräume im Überblick dargestellt, um die Notationen zu klären.

Der Lebesgue-Raum $L^2(\Omega; \mathbb{R}^n)$ wird mit der von dem L^2-Skalarprodukt $(u, v)_{L^2(\Omega)} := \int_\Omega u \cdot v \, \mathrm{d}x$ induzierten Norm $\|\bullet\|_{L^2(\Omega)}$ ausgestattet, wobei u, $v \in L^2(\Omega; \mathbb{R}^n)$, d.h. messbar

mit $\|v\|_{L^2(\Omega)}, \|u\|_{L^2(\Omega)} < \infty$. Bei dem matrixwertigen Äquivalent $L^2\left(\Omega, \mathbb{R}^{n\times n}\right)$ lautet das entsprechende L^2-Skalarprodukt $(\boldsymbol{u}, \boldsymbol{v})_{L^2(\Omega)} := \int_\Omega \boldsymbol{u} : \boldsymbol{v} \, \mathrm{d}x$ für $\boldsymbol{u}, \boldsymbol{v} \in L^2\left(\Omega, \mathbb{R}^{n\times n}\right)$. In dieser Arbeit spielen auch folgende Quotientenräume eine Rolle,

$$Q/\mathbb{R} := \left\{ \boldsymbol{\tau} \in Q : \int_\Omega \operatorname{tr} \boldsymbol{\tau} \, \mathrm{d}x = 0 \right\}, \tag{2.4}$$

wobei ein Unterraum $Q \subseteq L^2\left(\Omega; \mathbb{R}^{n\times n}\right)$ betrachtet wird. Es wird also zusätzlich eine lineare Nebenbedingung verlangt.

Für $k \in \mathbb{N}$ sei $C^k(\Omega)$ der Raum der k-fach stetig differenzierbaren Funktionen auf Ω, $C^\infty(\Omega) := \bigcap_{k\in\mathbb{N}} C^k(\Omega)$ und $C_0^\infty := \{f \in C^\infty(\Omega) : \operatorname{supp}(f) \subset \Omega \text{ kompakt}\}$ der Raum der Testfunktionen.

Eine Zusammenfassung der Differentialoperatoren scheint ebenfalls angebracht. Für $\partial/\partial x_k$ für $1 \le k \le n$ wird die Kurzschreibweise ∂_k verwendet. Für $u \in C^1(\Omega)$ sei $\nabla u := (\partial_1 u, \ldots, \partial_n u)^\top \in \mathbb{R}^n$ der Gradient und für $u \in C^1(\Omega; \mathbb{R}^n)$ sei $\mathrm{D}\, u \in \mathbb{R}^{n\times n}$ mit $(\mathrm{D}\, u)_{jk} := \partial_k u_j$ für $1 \le j, k \le n$ die Jakobi-Matrix. Die Divergenz sei $\operatorname{div} u := \sum_{j=1}^n \partial_j u_j = \operatorname{tr}(\mathrm{D}\, u) \in \mathbb{R}$ für $u \in C^1(\Omega; \mathbb{R}^n)$. Für matrix-wertige Funktionen $\boldsymbol{u} \in C^1\left(\Omega; \mathbb{R}^{n\times n}\right)$ wird die Divergenz zeilenweise angewendet, d.h. $\operatorname{div} \boldsymbol{u} := (\operatorname{div} u_1, \ldots, \operatorname{div} u_n)^\top \in \mathbb{R}^n$, wobei u_j für $1 \le j \le n$ die j-te Zeile der Matrix $\boldsymbol{u} \in C^1\left(\Omega; \mathbb{R}^{n\times n}\right)$ beschreibt.

Um den Sobolevraum $H^1(\Omega; \mathbb{R}^n)$ zu definieren wird der Begriff der schwachen Ableitung benötigt [Wer11, S. 210]. Für $1 \le j \le n$ sei die j-te schwache Ableitung von $u \in L^2(\Omega)$ eine Funktion $\tilde{u} \in L^2(\Omega)$, so dass für alle Testfunktionen $\varphi \in C_0^\infty(\Omega)$ gilt $\int_\Omega u \, \partial_j \varphi \, \mathrm{d}x = -\int_\Omega \tilde{u}\varphi \, \mathrm{d}x$. Diese schwache Ableitung $\tilde{u}$ wird mit $\partial_j u$ bezeichnet. Im mehrdimensionalen Fall $u \in L^2(\Omega; \mathbb{R}^n)$ ist die Existenz der j-ten schwachen Ableitung von u_k für alle $1 \le j, k \le n$ äquivalent zu der Gleichheit $\int_\Omega u \cdot \operatorname{div} \boldsymbol{\varphi} \, \mathrm{d}x = -\int_\Omega \mathrm{D}\, u : \boldsymbol{\varphi} \, \mathrm{d}x$ für alle Testfunktionen $\boldsymbol{\varphi} \in C_0^\infty\left(\Omega; \mathbb{R}^{n\times n}\right)$, $\mathrm{D}\, u$ bezeichnet dann die schwache Ableitung.
Es sei

$$H^1(\Omega; \mathbb{R}^n) := \left\{ u \in L^2(\Omega; \mathbb{R}^n) : \text{ im schwachen Sinne existiert } \mathrm{D}\, u \in L^2\left(\Omega; \mathbb{R}^{n\times n}\right) \right\}.$$

Dieser Sobolevraum wird mit der H^1-Norm

$$\|u\|_{H^1(\Omega)}^2 := \|u\|_{L^2(\Omega)}^2 + \|\mathrm{D}\, u\|_{L^2(\Omega)}^2$$

versehen. Wie die Friedrichs Ungleichung, Lemma 2.16, zeigt, ist die Energienorm

$$|\!|\!| u |\!|\!| := \|\mathrm{D}\, u\|_{L^2(\Omega)}$$

dazu äquivalent. Dirichlet Randdaten werden in diesen Raum mit Hilfe des Spuroperators $\gamma_0 : H^1(\Omega; \mathbb{R}^n) \to L^2(\partial\Omega; \mathbb{R}^n)$ mit $\gamma_0 u = u|_{\partial\Omega}$ (siehe Definition 2.7) einbezogen

$$H_0^1(\Omega; \mathbb{R}^n) := \left\{ u \in H^1(\Omega; \mathbb{R}^n) : \gamma_0(u) = 0 \text{ fast überall auf } \Gamma \right\}.$$

Es ist bekannt, dass $H_0^1(\Omega;\mathbb{R}^n)$ mit $\|\bullet\|_{H^1(\Omega)}$ bzw. $|||\bullet|||$ ein Banachraum ist [Alt06, S.68].

Um die Räume $H(\mathrm{div},\Omega)$ und $H\left(\mathrm{div},\Omega;\mathbb{R}^{n\times n}\right)$ zu definieren, wird der Begriff der schwachen Divergenz benötigt. Die schwache Divergenz der Funktion $u \in L^2(\Omega;\mathbb{R}^n)$ sei $\hat{u}:\Omega\to\mathbb{R}$ bezeichnet als $\mathrm{div}\,u$, wenn $\int_\Omega \varphi\hat{u}\,\mathrm{d}x = -\int_\Omega \nabla\varphi\cdot u\,\mathrm{d}x$ für alle Testfunktionen $\varphi\in C_0^\infty(\Omega)$. Definiere damit

$$H(\mathrm{div},\Omega) := \left\{u\in L^2(\Omega;\mathbb{R}^n):\ \text{im schwachen Sinne existiert }\ \mathrm{div}\,u\in L^2(\Omega)\right\},$$
$$H\left(\mathrm{div},\Omega;\mathbb{R}^{n\times n}\right) := \left\{\boldsymbol{u}\in L^2\left(\Omega;\mathbb{R}^{n\times n}\right):\ u_j\in H(\mathrm{div},\Omega)\ \forall\ 1\le j\le n\right\},$$

wobei $u_j\in L^2(\Omega;\mathbb{R}^n)$ wieder die j-te Zeile der Matrix $\boldsymbol{u}\in L^2\left(\Omega;\mathbb{R}^{n\times n}\right)$ sei. Auch im schwachen Sinne wird also der Divergenzoperator auf Matrizen zeilenweise angewendet. Diese Räume werden ausgestattet mit der Norm

$$\|u\|^2_{H(\mathrm{div})} := \|u\|^2_{L^2(\Omega)} + \|\mathrm{div}\,u\|^2_{L^2(\Omega)}$$

für $u\in H(\mathrm{div},\Omega)$ bzw. $u\in H\left(\mathrm{div},\Omega;\mathbb{R}^{n\times n}\right)$ zu Banachräumen [Bra13, S.139].

Die in dieser Arbeit verwendeten Testräume erfüllen die Regularitätsbedingungen nur stückweise auf den einzelnen Elementen $T\in\mathcal{T}$ der regulären Triangulierung $\mathcal{T}$. Es seien

$$H\left(\mathrm{div},\mathcal{T};\mathbb{R}^{n\times n}\right) := \left\{\boldsymbol{\tau}\in L^2\left(\Omega;\mathbb{R}^{n\times n}\right):\forall T\in\mathcal{T}\ \boldsymbol{\tau}|_T\in H(\mathrm{div},T;\mathbb{R}^{n\times n})\right\},$$
$$H^1(\mathcal{T};\mathbb{R}^n) := \left\{v\in L^2(\Omega;\mathbb{R}^n):\ \forall T\in\mathcal{T}\ v|_T\in H^1(T;\mathbb{R}^n)\right\}.$$

Für $\boldsymbol{\tau}\in H\left(\mathrm{div},\mathcal{T};\mathbb{R}^{n\times n}\right)$ bzw. $v\in H^1(\mathcal{T};\mathbb{R}^n)$ werden die stückweise Divergenz mit $\mathrm{div}_{\mathrm{NC}}$ bzw. die schwachen Ableitungen mit D_{NC} bezeichnet, d.h.

$$\mathrm{div}_{\mathrm{NC}}\,\boldsymbol{\tau}|_T := \mathrm{div}\,(\boldsymbol{\tau}|_T)\ \text{ für alle } T\in\mathcal{T},$$
$$\mathrm{D}_{\mathrm{NC}}\,v|_T := \mathrm{D}\,(v|_T)\ \text{ für alle } T\in\mathcal{T}.$$

Damit ergeben sich die Normen für $H\left(\mathrm{div},\mathcal{T};\mathbb{R}^{n\times n}\right)$ und $H^1(\mathcal{T};\mathbb{R}^n)$, als

$$\|\boldsymbol{\tau}\|^2_{H(\mathrm{div},\mathcal{T})} := \|\boldsymbol{\tau}\|^2_{H(\mathrm{div},\mathcal{T};\mathbb{R}^{n\times n})} := \|\boldsymbol{\tau}\|^2_{L^2(\Omega)} + \|\mathrm{div}_{\mathrm{NC}}\,\boldsymbol{\tau}\|^2_{L^2(\Omega)},$$
$$\|v\|^2_{H^1(\mathcal{T})} := \|v\|^2_{H^1(\mathcal{T};\mathbb{R}^n)} := \|v\|^2_{L^2(\Omega)} + \|\mathrm{D}_{\mathrm{NC}}\,v\|^2_{L^2(\Omega)} := \|v\|^2_{L^2(\Omega)} + |||v|||^2_{\mathrm{NC}}.$$

Zusätzlich werden noch Spurräume benötigt. In [Alt06, S.265] und [GR86, S.27 f.] werden die Spursätze bewiesen, die die Definitionen der Spuroperatoren ermöglichen.

Satz 2.6. *Es sei $U\subseteq\mathbb{R}^n$ ein offenes, beschränktes Lipschitz-Gebiet. Dann existiert*

genau eine stetige, lineare Abbildung $\gamma_0 : H^1(U) \to L^2(\partial U)$, *so dass*

$$\gamma_0 w = w|_{\partial U}, \quad \text{für alle } w \in H^1(U) \cap C^0(\overline{U}).$$

Definition 2.7. Der Bildraum dieses Spuroperators wird mit $H^{1/2}(\partial U) := \gamma_0\left(H^1(U)\right)$, sein Dualraum mit $H^{-1/2}(\partial U) = \left(H^{1/2}(\partial U)\right)^*$ bezeichnet.

Satz 2.8. *Es sei* $U \subseteq \mathbb{R}^n$ *ein offenes, beschränktes Lipschitz-Gebiet. Dann existiert genau eine stetige, lineare Abbildung* $\gamma_\nu : H(\mathrm{div}, U) \to H^{-1/2}(\partial U)$, *so dass*

$$\gamma_\nu q = (q|_{\partial U}) \cdot \nu, \quad \text{für alle } q \in H(\mathrm{div}, U).$$

Für alle $q \in H(\mathrm{div}, U)$ *und* $w \in H^1(U)$ *gilt*

$$\langle \gamma_\nu q, \gamma_0 w \rangle_{\partial U} = \int_U q \nabla w \, \mathrm{d}x + \int_U \mathrm{div}(q)\, w \, \mathrm{d}x.$$

Eine komponentenweise Anwendung dieser Abbildung erlaubt die Definition der mehrdimensionalen Spuroperatoren

$$\gamma_0 : H^1(U; \mathbb{R}^n) \to H^{1/2}(\partial U; \mathbb{R}^n),$$
$$\gamma_\nu : H\left(\mathrm{div}, U; \mathbb{R}^{n\times n}\right) \to H^{-1/2}(\partial U; \mathbb{R}^n).$$

Sei $\mathcal{T}$ eine regulären Triangulierung von Ω, so können Spurabbildung auf das Skelett $\partial\mathcal{T}$ definiert werden. Die Funktion $\gamma_0^{\mathcal{T}} \colon H^1(\mathcal{T}; \mathbb{R}^n) \to \prod_{T\in\mathcal{T}} H^{1/2}(\partial T; \mathbb{R}^n)$ bildet $w \in H^1(\mathcal{T}; \mathbb{R}^n)$ auf $\gamma_0^{\mathcal{T}} w := (s_T)_{T\in\mathcal{T}}$ mit $s_T := \gamma_0(w|_T)$ für alle $T \in \mathcal{T}$ ab. Die Funktion $\gamma_\nu^{\mathcal{T}} \colon H\left(\mathrm{div}, \mathcal{T}; \mathbb{R}^{n\times n}\right) \to \prod_{T\in\mathcal{T}} H^{-1/2}(\partial T; \mathbb{R}^n)$ bildet $\boldsymbol{q} \in H\left(\mathrm{div}, \mathcal{T}; \mathbb{R}^{n\times n}\right)$ auf $\gamma_\nu^{\mathcal{T}} \boldsymbol{q} := (t_T)_{T\in\mathcal{T}}$ mit $t_T := \gamma_\nu(\boldsymbol{q}|_T)$ für alle $T \in \mathcal{T}$ ab. Die assoziierten Spurräume werden wie folgt definiert.

$$H_0^{1/2}(\partial\mathcal{T}; \mathbb{R}^n) := \gamma_0^{\mathcal{T}}\left(H_0^1(\Omega; \mathbb{R}^n)\right) \tag{2.5}$$
$$H^{-1/2}(\partial\mathcal{T}; \mathbb{R}^n) := \gamma_\nu^{\mathcal{T}}\left(H\left(\mathrm{div}, \Omega; \mathbb{R}^{n\times n}\right)\right) \tag{2.6}$$

Diese Räume werden mit den folgenden Normen ausgestattet

$$\|s\|_{H_0^{1/2}(\partial\mathcal{T})} := \|s\|_{H_0^{1/2}(\partial\mathcal{T};\mathbb{R}^n)} := \inf_{\substack{w \in H_0^1(\Omega;\mathbb{R}^n) \\ \gamma_0^{\mathcal{T}} w = s}} \|w\|_{H^1(\Omega)},$$
$$\|t\|_{H^{-1/2}(\partial\mathcal{T})} := \|t\|_{H^{-1/2}(\partial\mathcal{T};\mathbb{R}^n)} := \inf_{\substack{\boldsymbol{q} \in H(\mathrm{div},\Omega;\mathbb{R}^{n\times n}) \\ \gamma_\nu^{\mathcal{T}} \boldsymbol{q} = t}} \|\boldsymbol{q}\|_{H(\mathrm{div},\Omega)}.$$

Definition 2.9. Die Dualitätsklammer auf dem Skelett sei für alle $t = (t_T)_{T\in\mathcal{T}} \in \prod_{T\in\mathcal{T}} H^{-1/2}(T;\mathbb{R}^n)$ und $s = (s_T)_{T\in\mathcal{T}} \in \prod_{T\in\mathcal{T}} H^{1/2}(\partial T;\mathbb{R}^n)$ definiert durch

$$\langle t, s\rangle_{\partial\mathcal{T}} := \sum_{T\in\mathcal{T}} \langle t_T, s_T\rangle_{\partial T}\,.$$

Bemerkung 2.10 Sei $v \in H^1_0(\mathcal{T};\mathbb{R}^n)$ mit $\gamma_0^{\mathcal{T}} v \in H_0^{1/2}(\partial\mathcal{T};\mathbb{R}^n)$ so gilt bereits $v \in H^1_0(\Omega;\mathbb{R}^n)$. Für den Nachweis sei $w \in H^1_0(\Omega;\mathbb{R}^n)$ mit $\gamma_0^{\mathcal{T}} w = \gamma_0^{\mathcal{T}} v$ und $\boldsymbol{\varphi} \in C_0^\infty\left(\Omega;\mathbb{R}^{n\times n}\right)$ eine beliebige Testfunktion, dann gilt die Formel aus Satz 2.8 für jedes $T \in \mathcal{T}$ und damit

$$\begin{aligned}
\int_\Omega \mathrm{D_{NC}}\, v : \boldsymbol{\varphi}\,\mathrm{d}x &= \left\langle \gamma_\nu^{\mathcal{T}}\boldsymbol{\varphi}, \gamma_0^{\mathcal{T}} v\right\rangle_{\partial\mathcal{T}} - \int_\Omega \operatorname{div}\boldsymbol{\varphi}\cdot v\,\mathrm{d}x\\
&= \left\langle \gamma_\nu^{\mathcal{T}}\boldsymbol{\varphi}, \gamma_0^{\mathcal{T}} w\right\rangle_{\partial\mathcal{T}} - \int_\Omega \operatorname{div}\boldsymbol{\varphi}\cdot v\,\mathrm{d}x\\
&= \int_\Omega \mathrm{D}\, w : \boldsymbol{\varphi}\,\mathrm{d}x + \int_\Omega \operatorname{div}\boldsymbol{\varphi}\cdot w\,\mathrm{d}x - \int_\Omega \operatorname{div}\boldsymbol{\varphi}\cdot v\,\mathrm{d}x\\
&= \int_{\partial\Omega} w\boldsymbol{\varphi}\cdot\nu\,\mathrm{d}x - \int_\Omega \operatorname{div}\boldsymbol{\varphi}\cdot v\,\mathrm{d}x = -\int_\Omega \operatorname{div}\boldsymbol{\varphi}\cdot v\,\mathrm{d}x.
\end{aligned}$$

Somit gilt nach Definition der schwachen Ableitung, dass $\mathrm{D_{NC}}\, v = \mathrm{D}\, v$. Analog folgt für $\boldsymbol{\tau} \in H\left(\operatorname{div},\mathcal{T};\mathbb{R}^{n\times n}\right)$ mit $\gamma_\nu^{\mathcal{T}}\boldsymbol{\tau} \in H^{-1/2}(\partial\mathcal{T};\mathbb{R}^n)$, dass bereits $\boldsymbol{\tau} \in H\left(\operatorname{div},\Omega;\mathbb{R}^{n\times n}\right)$ gilt. Sei dazu $\boldsymbol{q} \in H(\operatorname{div},\Omega;\mathbb{R}^{n\times n})$ mit $\gamma_\nu^{\mathcal{T}}\boldsymbol{q} = \gamma_\nu^{\mathcal{T}}\boldsymbol{\tau}$ und $\varphi \in C_0^\infty(\Omega;\mathbb{R}^n)$ eine beliebige Testfunktion, so gilt

$$\begin{aligned}
\int_\Omega \operatorname{div_{NC}}\boldsymbol{\tau}\cdot\varphi\,\mathrm{d}x &= \left\langle \gamma_\nu^{\mathcal{T}}\boldsymbol{\tau}, \gamma_0^{\mathcal{T}}\varphi\right\rangle_{\partial\mathcal{T}} - \int_\Omega \mathrm{D}\,\varphi : \boldsymbol{\tau}\,\mathrm{d}x\\
&= \left\langle \gamma_\nu^{\mathcal{T}}\boldsymbol{q}, \gamma_0^{\mathcal{T}}\varphi\right\rangle_{\partial\mathcal{T}} - \int_\Omega \mathrm{D}\,\varphi : \boldsymbol{\tau}\,\mathrm{d}x\\
&= \int_\Omega \mathrm{D}\,\varphi : \boldsymbol{q}\,\mathrm{d}x + \int_\Omega \operatorname{div}\boldsymbol{q}\cdot\varphi\,\mathrm{d}x - \int_\Omega \mathrm{D}\,\varphi : \boldsymbol{\tau}\,\mathrm{d}x\\
&= \int_{\partial\Omega} \varphi\boldsymbol{q}\cdot\nu\,\mathrm{d}x - \int_\Omega \mathrm{D}\,\varphi : \boldsymbol{\tau}\,\mathrm{d}x = -\int_\Omega \mathrm{D}\,\varphi : \boldsymbol{\tau}\,\mathrm{d}x.
\end{aligned}$$

Also gilt nach Definition der schwachen Ableitung, dass $\operatorname{div_{NC}}\boldsymbol{\tau} = \operatorname{div}\boldsymbol{\tau}$.

Bemerkung 2.11 Mit Hilfe der vorangehenden Bemerkung 2.10 wird umgehend deutlich, dass die Normen $\|\bullet\|_{H_0^{1/2}(\partial\mathcal{T})}$ bzw. $\|\bullet\|_{H^{-1/2}(\partial\mathcal{T})}$ Minimumsnormen sind.
Sei $s \in H_0^{1/2}(\Omega;\mathbb{R}^n)$, so ist die Funktion $w \in H_0^1(\Omega;\mathbb{R}^n)$ mit $\gamma_0^{\mathcal{T}} w = s$ und $\|s\|_{H_0^{1/2}(\partial\mathcal{T})} = \|w\|_{H^1(\Omega)}$ nach den Euler-Lagrange Gleichungen für jedes $T \in \mathcal{T}$ Lösung der folgenden Differentialgleichung in $H^1(T;\mathbb{R}^n)$

$$\begin{aligned}
w &= s \quad \text{entlang } \partial T,\\
-\operatorname{div}(\mathrm{D}\, w) + w &= 0 \quad \text{in } T.
\end{aligned}$$

Für $t \in H^{-1/2}(\partial\mathcal{T};\mathbb{R}^n)$ ist die Funktion $\boldsymbol{q} \in H(\operatorname{div},\Omega;\mathbb{R}^{n\times n})$ mit $\gamma_\nu^{\mathcal{T}}\boldsymbol{q} = t$ und $\|t\|_{H^{-1/2}(\partial\mathcal{T})} =$

$\|\boldsymbol{q}\|_{H(\mathrm{div},\Omega)}$ auf jedem $T \in \mathcal{T}$ Lösung der folgenden Differentialgleichung in $H(\mathrm{div}, T; \mathbb{R}^{n\times n})$

$$\begin{aligned} \boldsymbol{q}\cdot\nu &= t \quad \text{entlang } \partial T,\\ -\mathrm{D}(\mathrm{div}\,\boldsymbol{q}) + \boldsymbol{q} &= 0 \quad \text{in } T. \end{aligned}$$

Somit ist es stets möglich eine Fortsetzung $w \in H_0^1(\Omega;\mathbb{R}^n)$ bzw. $\boldsymbol{q} \in H(\mathrm{div},\Omega;\mathbb{R}^{n\times n})$ mit derselben Norm zu finden.

Elementweise sind $H^{1/2}(\partial T)$ und $H^{-1/2}(\partial T)$ nach Definition 2.7 dual. Dies ist nicht auf die Räume $H^{1/2}(\partial\mathcal{T})$ und $H^{-1/2}(\partial\mathcal{T})$ übertragbar. Es können allerdings die folgenden nützlichen Dualitätsaussagen wie in [CDG15b, Lemma 3.1] bzw. [CDG15a, Theorem 2.3] bewiesen werden.

Lemma 2.12 (Dualitätslemma I). *Für alle $s \in H_0^{1/2}(\partial\mathcal{T};\mathbb{R}^n)$ gilt*

$$\|s\|_{H^{1/2}(\partial\mathcal{T})} = \sup_{\substack{\boldsymbol{\tau}\in H\left(\mathrm{div},\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R}\\ \boldsymbol{\tau}\neq 0}} \frac{\left\langle \gamma_\nu^{\mathcal{T}}\boldsymbol{\tau}, s\right\rangle_{\partial\mathcal{T}}}{\|\boldsymbol{\tau}\|_{H(\mathrm{div},\mathcal{T})}}.$$

Beweis. Sei zunächst $w \in H_0^1(\Omega;\mathbb{R}^n)$ die Fortsetzung von s mit derselben Norm aus Bemerkung 2.11. Setze $\boldsymbol{q} := \mathrm{D}\,w$, dann gilt nach obiger Bemerkung $\mathrm{div}\,\boldsymbol{q}|_T = \mathrm{div}(\mathrm{D}\,w|_T) = w|_T$ auf jedem Element $T \in \mathcal{T}$ und

$$\int_\Omega \mathrm{tr}\,\boldsymbol{q}\;\mathrm{d}x = \int_\Omega \mathrm{div}\,w\;\mathrm{d}x = \int_{\partial\Omega} w\cdot\nu\;\mathrm{d}s = 0,$$

also $\boldsymbol{q} \in H(\mathrm{div},\mathcal{T};\mathbb{R}^n)/\mathbb{R}$. Weiter ergibt sich nach der Wahl von $\boldsymbol{q}$

$$\|\boldsymbol{q}\|^2_{H(\mathrm{div},\mathcal{T})} = \|\boldsymbol{q}\|^2_{L^2(\Omega)} + \|\mathrm{div}_{\mathrm{NC}}\,\boldsymbol{q}\|^2_{L^2(\Omega)} = \|w\|^2_{H^1(\Omega)}.$$

Nun bedeutet Satz 2.8

$$\left\langle \gamma_\nu^{\mathcal{T}}\boldsymbol{q}, s\right\rangle_{\partial\mathcal{T}} = \left\langle \gamma_\nu^{\mathcal{T}}\boldsymbol{q}, \gamma_0^{\mathcal{T}} w\right\rangle_{\partial\mathcal{T}} = \int_\Omega \boldsymbol{q} : \mathrm{D}\,w\;\mathrm{d}x + \int_\Omega \mathrm{div}_{\mathrm{NC}}\,\boldsymbol{q}\cdot w\;\mathrm{d}x = \|w\|^2_{H^1(\Omega)}.$$

Dies ist die eine Richtung der behaupteten Gleichheit, denn

$$\|s\|_{H^{1/2}(\partial\mathcal{T})} = \|w\|_{H^1(\Omega)} = \frac{\left\langle \gamma_\nu^{\mathcal{T}}\boldsymbol{q}, s\right\rangle_{\partial\mathcal{T}}}{\|\boldsymbol{q}\|_{H(\mathrm{div},\mathcal{T})}} \le \sup_{\substack{\boldsymbol{\tau}\in H\left(\mathrm{div},\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R}\\ \boldsymbol{\tau}\neq 0}} \frac{\left\langle \gamma_\nu^{\mathcal{T}}\boldsymbol{\tau}, s\right\rangle_{\partial\mathcal{T}}}{\|\boldsymbol{\tau}\|_{H(\mathrm{div},\mathcal{T})}}.$$

Für die andere Richtung betrachte außer der minimalen Fortsetzung $w \in H_0^1(\Omega;\mathbb{R}^n)$ ein beliebiges $\boldsymbol{q} \in S(H(\mathrm{div},\mathcal{T};\mathbb{R}^{n\times n})/\mathbb{R})$. Dann zeigen partielle Integration und Cauchy-Schwarz-Ungleichung auf jedem Element $T \in \mathcal{T}$

$$\langle \gamma_0 w, \gamma_\nu \boldsymbol{q}|_T\rangle_{\partial T} = \int_T \boldsymbol{q} : \mathrm{D}\,v\;\mathrm{d}x + \int_T v\cdot\mathrm{div}\,\boldsymbol{q}\;\mathrm{d}x \le \|v\|_{H^1(T)}\|\boldsymbol{q}\|_{H(\mathrm{div},T)} \le \|v\|_{H^1(T)},$$

da $\boldsymbol{q}$ normiert ist. In der Summe ergibt sich damit

$$\left\langle \gamma_0^{\mathcal{T}} w, \gamma_\nu^{\mathcal{T}} \boldsymbol{q} \right\rangle_{\partial\mathcal{T}} \leq \sum_{T\in\mathcal{T}} \|w\|_{H^1(T)} = \|w\|_{H^1(\Omega)}.$$

Daraus folgt die Behauptung, da $\boldsymbol{q} \in S(H(\operatorname{div}, \mathcal{T}; \mathbb{R}^{n\times n})/\mathbb{R})$ beliebig war. □

Lemma 2.13 (Dualitätslemma II). *Für alle $t \in H^{-1/2}(\partial\mathcal{T}; \mathbb{R}^n)$ gilt*

$$\|t\|_{H^{-1/2}(\partial\mathcal{T})} = \sup_{v\in H^1(\mathcal{T};\mathbb{R}^n)\setminus\{0\}} \frac{\left\langle t, \gamma_0^{\mathcal{T}} v \right\rangle_{\partial\mathcal{T}}}{\|v\|_{H^1(\mathcal{T})}}.$$

Beweis. Sei zunächst $w \in H^1(\mathcal{T}; \mathbb{R}^n)$ auf jedem Element $T \in \mathcal{T}$ die Lösung des folgenden Neumann Problems

$$\operatorname{div} \mathrm{D}\, w - w = 0 \text{ in } T \quad \text{und} \quad \mathrm{D}\, w \cdot \nu_T = t_T \text{ auf } \partial T.$$

Setze $\boldsymbol{q} := \mathrm{D}_{\mathrm{NC}}\, w \in H\left(\operatorname{div}, \mathcal{T}; \mathbb{R}^{n\times n}\right)$, dann gilt $\gamma_\nu^{\mathcal{T}} \boldsymbol{q} = t \in H^{-1/2}(\partial\mathcal{T}; \mathbb{R}^n)$ also nach Bemerkung 2.10 $\boldsymbol{q} \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)$. Außerdem ist $\operatorname{div} \boldsymbol{q} = w$ und daher ergibt sich

$$\|\boldsymbol{q}\|^2_{H(\operatorname{div})} = \|\boldsymbol{q}\|^2_{L^2(\Omega)} + \|\operatorname{div} \boldsymbol{q}\|^2_{L^2(\Omega)} = \|\mathrm{D}_{\mathrm{NC}}\, w\|^2_{L^2(\Omega)} + \|w\|^2_{L^2(\Omega)} = \|w\|^2_{H^1(\mathcal{T})}.$$

Nun bedeutet Satz 2.8

$$\left\langle t, \gamma_0^{\mathcal{T}} w \right\rangle_{\partial\mathcal{T}} = \left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{q}, \gamma_0^{\mathcal{T}} w \right\rangle_{\partial\mathcal{T}} = \int_\Omega \boldsymbol{q} : \mathrm{D}_{\mathrm{NC}}\, w \,\mathrm{d}x + \int_\Omega \operatorname{div} \boldsymbol{q} \cdot w \,\mathrm{d}x = \|\boldsymbol{q}\|^2_{H(\operatorname{div})}.$$

Damit ergibt sich die erste Richtung

$$\begin{aligned}\|t\|_{H^{-1/2}(\partial\mathcal{T})} &= \inf_{\substack{\boldsymbol{\tau}\in H(\operatorname{div},\Omega;\mathbb{R}^n)\\ \gamma_\nu^{\mathcal{T}}\boldsymbol{\tau}=t}} \|\boldsymbol{\tau}\|_{H(\operatorname{div})} \leq \|\boldsymbol{q}\|_{H(\operatorname{div})} = \frac{\left\langle t, \gamma_0^{\mathcal{T}} w \right\rangle_{\partial\mathcal{T}}}{\|w\|_{H^1(\mathcal{T})}} \\ &\leq \sup_{v\in H^1(\mathcal{T};\mathbb{R}^n)\setminus\{0\}} \frac{\left\langle t, \gamma_0^{\mathcal{T}} v \right\rangle_{\partial\mathcal{T}}}{\|v\|_{H^1(\mathcal{T})}}.\end{aligned}$$

Der Beweis der anderen Ungleichung funktioniert analog zu Lemma 2.12. Diesmal wird allerdings zu dem $\boldsymbol{q} \in H(\operatorname{div}, \Omega; \mathbb{R}^{n\times n})/\mathbb{R}$ mit $\gamma_\nu^{\mathcal{T}} \boldsymbol{q} = t$ und minimaler Norm aus Bemerkung 2.11 ein beliebiges $w \in S(H_0^1(\mathcal{T}; \mathbb{R}^n))$ gewählt. Damit ergibt sich mit elementweise partieller Integration und Cauchy-Schwarz Ungleichung diesmal

$$\left\langle \gamma_0^{\mathcal{T}} w, \gamma_\nu^{\mathcal{T}} \boldsymbol{q} \right\rangle_{\partial\mathcal{T}} = \sum_{T\in\mathcal{T}} \langle \gamma_0 w, \gamma_\nu \boldsymbol{q} \rangle_{\partial T} \leq \sum_{T\in\mathcal{T}} \|v\|_{H^1(T)} \|\boldsymbol{q}\|_{H(\operatorname{div},T)} \leq \|\boldsymbol{q}\|_{H(\operatorname{div},\Omega)}.$$

Da dies für jedes $w \in S(H_0^1(\mathcal{T}; \mathbb{R}^n))$ gilt, folgt der zweite Teil der Behauptung. □

Mit den Definitionen aus diesem Paragraphen können nun die in der dPG-Formulierung des Stokes Problems auftretenden Produkträume X und Y eingeführt werden. Der

Ansatzraum X, in dem die Lösung der Differentialgleichung gesucht wird, lautet

$$X := L^2\left(\Omega;\mathbb{R}^{n\times n}\right)/\mathbb{R} \times L^2\left(\Omega;\mathbb{R}^n\right) \times H_0^{1/2}\left(\partial\mathcal{T};\mathbb{R}^n\right) \times H^{-1/2}\left(\partial\mathcal{T};\mathbb{R}^n\right), \tag{2.7}$$

und sei mit folgender Norm ausgestattet

$$\|x\|_X^2 := \|\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega)}^2 + \|s\|_{H_0^{1/2}(\partial\mathcal{T})}^2 + \|t\|_{H^{-1/2}(\partial\mathcal{T})}^2 \qquad \text{für jedes } x = (\boldsymbol{\sigma}, u, s, t) \in X.$$

Der *Testraum* Y, mit dessen Elementen die Variationsformulierung der Bilinearform ausgewertet wird, liest sich

$$Y := H\left(\operatorname{div},\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R} \times H^1\left(\mathcal{T};\mathbb{R}^n\right) \tag{2.8}$$

und sei mit folgender Norm ausgestattet

$$\|y\|_Y^2 := \|\boldsymbol{\tau}\|_{H(\operatorname{div},\mathcal{T})}^2 + \|v\|_{H^1(\mathcal{T})}^2 \qquad \text{für jedes } y = (\boldsymbol{\tau}, v) \in Y.$$

Damit liegt den Räumen X und Y bereits auf kontinuierlichem Level eine Triangulierung zu Grunde.

Bemerkung 2.14 Sowohl der Ansatzraum als auch der Testraum sind Hilberträume. Die Norm $\|\bullet\|_Y$ erfüllt die Parallelogrammgleichung

$$\|y+\hat{y}\|_Y^2 + \|y-\hat{y}\|_Y^2 = 2\left(\|y\|_Y^2 + \|\hat{y}\|_Y^2\right).$$

Damit lässt sich mittels

$$\begin{aligned}\langle y, \hat{y}\rangle_Y :=& \frac{1}{4}\left(\|y+\hat{y}\|_Y^2 - \|y-\hat{y}\|_Y^2\right)\\ =& \int_\Omega \boldsymbol{\tau} : \hat{\boldsymbol{\tau}}\,\mathrm{d}x + \int_\Omega \operatorname{div}_{\mathrm{NC}}\boldsymbol{\tau}\cdot\operatorname{div}_{\mathrm{NC}}\hat{\boldsymbol{\tau}}\,\mathrm{d}x + \int_\Omega v\cdot\hat{v}\,\mathrm{d}x + \int_\Omega \mathrm{D}_{\mathrm{NC}}\, v : \mathrm{D}_{\mathrm{NC}}\,\hat{v}\,\mathrm{d}x\end{aligned}$$

ein Skalarprodukt für alle $y = (\boldsymbol{\tau}, v)$, $\hat{y} = (\hat{\boldsymbol{\tau}}, \hat{v}) \in Y$ definieren [Wer11, S.205].

2.4 Diskrete Funktionenräume

In diesem Paragraphen werden die in der Diskretisierung verwendeten Räume vorgestellt. Die endlich-dimensionalen Unterräume des Ansatzraumes $X_h \subset X$ und Testraumes $Y_h \subset Y$ sind Räume stückweise polynomialer Funktionen.

Zunächst seien für $\mathbb{M} \subseteq \mathbb{R}^m$ oder $\mathbb{M} \subseteq \mathbb{R}^{m\times n}$ und k, m, $n \in \mathbb{N}_{>0}$,

$$\begin{aligned}P_k(T) &:= \{p_k \in L^\infty(T) : p_k \text{ ist auf } T \text{ ein Polynom vom Grad } \leq k\},\\ P_k(\mathcal{T}) &:= \{p_k \in L^\infty(\Omega) : \ \forall T \in \mathcal{T}, p_k|_T \in P_k(T)\},\\ P_k(T;\mathbb{M}) &:= \{\boldsymbol{q}_k \in L^\infty(T;\mathbb{M}) : \text{ jede Komponente von } \boldsymbol{q}_k \text{ gehört zu } P_k(T)\},\\ P_k(\mathcal{T},\mathbb{M}) &:= \{\boldsymbol{q}_k \in L^\infty(\Omega;\mathbb{M}) : \ \forall T \in \mathcal{T}, \boldsymbol{q}_k|_T \in P_k(T;\mathbb{M})\}.\end{aligned}$$

Auf dem Skelett seien analog definiert

$$\begin{aligned}
P_k(E) &:= \{p_k \in L^\infty(E) : p_k|_E \text{ ist ein Polynom vom Grad } \leq k\}, \\
P_k(\mathcal{E}) &:= \left\{p_k \in L^\infty(\bigcup \mathcal{E}) : \ \forall E \in \mathcal{E},\, p_k|_E \in P_k(E)\right\}, \\
P_k(E;\mathbb{M}) &:= \{\boldsymbol{q}_k \in L^\infty(E;\mathbb{M}) : \text{ jede Komponente von } \boldsymbol{q}_k \text{ gehört zu } P_k(E)\}, \\
P_k(\mathcal{E};\mathbb{M}) &:= \left\{\boldsymbol{q}_k \in L^\infty(\bigcup \mathcal{E};\mathbb{M}) : \ \forall E \in \mathcal{E},\, \boldsymbol{q}_k|_E \in P_k(E;\mathbb{M})\right\}.
\end{aligned}$$

Die stetigen und stückweise P_k-Finite-Elemente-Funktionen auf $\mathcal{T}$ werden mit

$$S^k(\mathcal{T};\mathbb{M}) := P_k(\mathcal{T}) \cap C(\overline{\Omega}) \text{ und } S_0^k(\mathcal{T};\mathbb{M}) := S^k(\mathcal{T};\mathbb{M}) \cap C_0(\Omega)$$

und auf dem Skelett mit

$$S_0^k(\mathcal{E};\mathbb{M}) := P_k(\mathcal{E};\mathbb{M}) \cap C_0(\bigcup \mathcal{E};\mathbb{M})$$

bezeichnet.

Die (stückweise) Raviart-Thomas Funktionen niedrigster Ordnung sind definiert als

$$\begin{aligned}
RT_0^{\mathrm{pw}}\left(\mathcal{T};\mathbb{R}^{n\times n}\right) &:= \Big\{\boldsymbol{q}_{\mathrm{RT}} \in L^\infty\left(\Omega;\mathbb{R}^{n\times n}\right) : \ \exists A \in P_0\left(\mathcal{T};\mathbb{R}^{n\times n}\right),\ \exists b \in P_0\left(\mathcal{T};\mathbb{R}^n\right), \\
&\qquad\qquad \boldsymbol{q}_{\mathrm{RT}} = A + b \otimes (\bullet - \mathrm{mid}(\mathcal{T}))\Big\}, \\
RT_0\left(\mathcal{T};\mathbb{R}^{n\times n}\right) &:= RT_0^{\mathrm{pw}}\left(\mathcal{T};\mathbb{R}^{n\times n}\right) \cap H\left(\mathrm{div},\Omega;\mathbb{R}^{n\times n}\right).
\end{aligned}$$

Es kann jede Raviart-Thomas Funktion $\boldsymbol{q}_{\mathrm{RT}} \in RT_0^{\mathrm{pw}}\left(\mathcal{T};\mathbb{R}^{n\times n}\right)$ auf jedem Simplex $T \in \mathcal{T}$ geschrieben werden als $\boldsymbol{q}_{\mathrm{RT}}|_T = A + 1/n \ \mathrm{div}\, \boldsymbol{q}_{\mathrm{RT}} \otimes (\bullet - \mathrm{mid}(T))$ für ein $A \in \mathbb{R}^{n\times n}$. Wenn also Π_0 die P_0-Projektion von $f \in L^2(\Omega;\mathbb{M})$ auf $P_0(\mathcal{T})$,

$$\Pi_0 f|_T := |T|^{-1} \int_T f \ \mathrm{d}x = \fint_T f \ \mathrm{d}x,$$

bezeichnet, dann gilt nach (2.2)

$$(1 - \Pi_0)\, \boldsymbol{q}_{\mathrm{RT}} = 1/n \ \mathrm{div}\, \boldsymbol{q}_{\mathrm{RT}} \otimes (\bullet - \mathrm{mid}(\mathcal{T})) \perp P_0(\mathcal{T}). \tag{2.9}$$

Außerdem gilt für alle $f,\, g \in L^2(\Omega)$

$$\int_\Omega f \Pi_0 g \ \mathrm{d}x = \int_\Omega \Pi_0 f \Pi_0 g \ \mathrm{d}x, \tag{2.10}$$

da auf jedem Element $\int_T f\Pi_0 g \ \mathrm{d}x = \Pi_0 g \int_T f \ \mathrm{d}x = \Pi_0 g\, |T| \fint_T f \ \mathrm{d}x = \int_T \Pi_0 f \Pi_0 g \ \mathrm{d}x$.

Die Raviart-Thomas Funktionen ermöglichen außerdem eine Einbettung von $P_0(\mathcal{E};\mathbb{R}^n)$ in $H^{-1/2}(\partial T;\mathbb{R}^n)$.

Bemerkung 2.15 Es gilt für jedes $\boldsymbol{q}_{\mathrm{RT}} \in RT_0\left(\mathcal{T};\mathbb{R}^{n\times n}\right)$ und für alle $E \in \mathcal{E}$, $\boldsymbol{q}_{\mathrm{RT}}|_E \nu_E \in$

$P_0(E;\mathbb{R}^n)$. Umgekehrt, existiert zu jedem $t_0 \in P_0(\mathcal{E};\mathbb{R}^n)$ genau ein $\boldsymbol{q}_{\mathrm{RT}} \in RT_0\left(\mathcal{T};\mathbb{R}^{n\times n}\right)$, so dass $\boldsymbol{q}_{\mathrm{RT}}|_E \nu_E = t_0|_E$. Beides folgt direkt aus [BC05, Lemma 4.1].

Nun ist $P_0(\mathcal{E};\mathbb{R}^n)$ ein Unterraum von $H^{-1/2}(\partial T;\mathbb{R}^n)$ vermöge der Einbettung

$$P_0(\mathcal{E};\mathbb{R}^n) \hookrightarrow H^{-1/2}(\partial\mathcal{T};\mathbb{R}^n), \quad t_0 \mapsto t = (t_T)_{T\in\mathcal{T}} \text{ mit } t_T = \boldsymbol{q}_{\mathrm{RT}}\nu_T|_{\partial T},$$

wobei $\boldsymbol{q}_{\mathrm{RT}} \in RT_0\left(\mathcal{T};\mathbb{R}^{n\times n}\right)$ die laut Bemerkung 2.15 eindeutige Funktion mit $\boldsymbol{q}_{\mathrm{RT}}|_E\nu_E = t_0|_E$ auf allen Seiten $E \in \mathcal{E}$ ist. Offensichtlich gilt auf jedem Element $T \in \mathcal{T}$ und jeder Seite $E \in \mathcal{E}(T)$ die Gleichheit $(\boldsymbol{q}_{\mathrm{RT}}\nu_T)|_E = \mathrm{sgn}(T,E)t_0$.

Damit lassen sich nun für eine gegebene reguläre Triangulierung $\mathcal{T}$ von Ω die endlich dimensionalen Unterräume der konformen Diskretisierung angeben

$$X_h := P_0\left(\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R} \times P_0(\mathcal{T};\mathbb{R}^n) \times S_0^1(\mathcal{E};\mathbb{R}^n) \times P_0(\mathcal{E};\mathbb{R}^n), \tag{2.11}$$

$$Y_h := RT_0^{\mathrm{pw}}\left(\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R} \times P_1(\mathcal{T};\mathbb{R}^n). \tag{2.12}$$

2.5 Weitere Hilfsmittel

In diesem Paragraphen werden einige nützliche Tatsachen vorgestellt.

Die folgenden Lemmata werden für verschiedene Normabschätzung in dieser Arbeit benötigt.

Lemma 2.16 (Friedrichs Ungleichung). *Für jedes beschränkte, polygonal berandete Lipschitzgebiet Ω existiert eine nur von Ω abhängige Konstante C_F, so dass für jedes $v \in H_0^1(\Omega;\mathbb{R}^n)$ gilt, dass*

$$\|v\|_{L^2(\Omega)} \le C_F \|\mathrm{D}\, v\|_{L^2(\Omega)}.$$

Die nur von Ω abhängige Friedrichskonstante kann auf verschiedenen Wegen abgeschätzt werden. Laut [Bra13, S.29] gilt, falls Ω in einem n-dimensionalen Würfel der Kantenlänge q enthalten ist, so gilt $C_F \le q$. In [Car09b, Theorem 0.49] ist bewiesen, dass $C_F \le \mathrm{width}(\Omega)/\pi$, wobei $\mathrm{width}(\Omega) := L := \beta - \alpha$ die kürzeste Länge ist, so dass Ω zwischen den parallelen Hyperebenen $\{x\cdot\nu = \alpha\}$ und $\{x\cdot\nu = \beta\}$ mit Abstand L für einen Einheitsvektor $\nu \in \mathbb{R}^n$ liegt.

Lemma 2.17 (Poincaré Ungleichung). *Für jedes beschränkte, konvexe Gebiet Ω und jedes $v \in H^1(\Omega;\mathbb{R}^n)$ mit $\bar{v} := \fint_\Omega v\,\mathrm{d}x := |\Omega|^{-1}\int_\Omega v\,\mathrm{d}x$ existiert eine nur von Ω abhängige Konstante C_P, so dass*

$$\|v - \bar{v}\|_{L^2(\Omega)} \le C_P \|\mathrm{D}\, v\|_{L^2(\Omega)}.$$

Auch diese Konstante kann genauer angegeben werden. Für beliebige Dimensionen ist in [PW60] bewiesen, dass $C_P \leq \operatorname{diam}(\Omega)/\pi$, wobei $\operatorname{diam}(\Omega) := \sup_{x,y\in\Omega}|x-y|$. Für den zweidimensionalen Fall ist in [LS10] eine optimale Konstante hergeleitet. Sei $\mathcal{T}$ eine reguläre Triangulierung des beschränkten, konvexen Gebietes $\Omega \subseteq \mathbb{R}^2$. So gilt auf jeden Dreieck $T \in \mathcal{T}$ mit $0 < C_P(T) \leq h_T/j_{1,1}$, wobei $j_{1,1} > \pi$ die erste Wurzel der ersten Besselfunktion ist, dass

$$\left\| v - \fint_T v \, \mathrm{d}x \right\|_{L^2(T)} \leq C_P(T) \|\mathrm{D}\, v\|_{L^2(T)} .$$

So ergibt sich auf ganz Ω folgende Abschätzung für die Poincaré Konstante $0 < C_P \leq h_{\max}/j_{1,1}$.

Das folgende häufig verwendete Lemma ist aus [CGS14] entnommen.

Lemma 2.18 (Tr-Div-Dev). *Sei $\Sigma_0 \subseteq H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)$ ein abgeschlossener Unterraum, der den konstanten Tensor $I_{n\times n}$ nicht enthält, dann existiert eine nur von Ω abhängige Konstante $C_{tdd} < \infty$, so dass für alle $\boldsymbol{\tau} \in \Sigma_0$ gilt*

$$\|\operatorname{tr}\boldsymbol{\tau}\|_{L^2(\Omega)} \leq C_{tdd}\left(\|\operatorname{dev}\boldsymbol{\tau}\|_{L^2(\Omega)} + \|\operatorname{div}\boldsymbol{\tau}\|_{L^2(\Omega)}\right).$$

Bemerkung 2.19 In dieser Arbeit wird Lemma 2.18 lediglich auf den Raum $\Sigma_0 := \left\{\boldsymbol{\tau} \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right) : \int_\Omega \operatorname{tr}\boldsymbol{\tau} \, \mathrm{d}x = 0\right\}$ angewendet. Es ist wie in [Hel14, S.27] leicht zu zeigen, dass dieses Σ_0 die in Lemma 2.18 geforderten Bedingungen erfüllt. Σ_0 enthält $\mathrm{I}_{n\times n}$ offensichtlich nicht. Um außerdem die Abgeschlossenheit in $H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)$ zu prüfen, sei $\left(\boldsymbol{\tau}^j\right)_{j\in\mathbb{N}} \in \Sigma_0$ eine konvergente Folge mit Grenzwert $\boldsymbol{\tau} \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)$. Dann gilt wegen der Linearität der Spur, $\boldsymbol{\tau}^j \in \Sigma_0$ und Cauchy-Schwarz für $L^2(\Omega)$ für alle $j \in \mathbb{N}$

$$\begin{aligned}
\left|\int_\Omega \operatorname{tr}\boldsymbol{\tau} \, \mathrm{d}x\right| &= \left|\int_\Omega \operatorname{tr}\left(\boldsymbol{\tau} - \boldsymbol{\tau}^j\right) \mathrm{d}x\right| \leq \int_\Omega \left|\operatorname{tr}\left(\boldsymbol{\tau} - \boldsymbol{\tau}^j\right)\right| \mathrm{d}x \\
&= \left\|\operatorname{tr}\left(\boldsymbol{\tau} - \boldsymbol{\tau}^j\right)\right\|_{L^1(\Omega)} = \left\|\sum_{k=1}^n \left(\tau_{kk} - \tau_{kk}^j\right)\right\|_{L^1(\Omega)} \\
&\leq \sum_{k=1}^n \left\|\left(\tau_{kk} - \tau_{kk}^j\right)\right\|_{L^1(\Omega)} \leq |\Omega|^{1/2} \sum_{k=1}^n \left\|\left(\tau_{kk} - \tau_{kk}^j\right)\right\|_{L^2(\Omega)} .
\end{aligned}$$

Die Cauchy-Schwarz-Ungleichung in $\mathbb{R}^n$ erlaubt weiter

$$\begin{aligned}
\sum_{k=1}^n \left\|\left(\tau_{kk} - \tau_{kk}^j\right)\right\|_{L^2(\Omega)} &= (1,\ldots,1) \cdot \left(\left\|\left(\tau_{11} - \tau_{11}^j\right)\right\|_{L^2(\Omega)}, \ldots, \left\|\left(\tau_{nn} - \tau_{nn}^j\right)\right\|_{L^2(\Omega)}\right)^\top \\
&\leq \sqrt{n}\sqrt{\sum_{k=1}^n \left\|\left(\tau_{kk} - \tau_{kk}^j\right)\right\|_{L^2(\Omega)}^2} .
\end{aligned}$$

Außerdem gilt nach der Definition des L_2-Skalarprodukts für Funktionen nach $\mathbb{R}^{n\times n}$,

$$\sum_{k=1}^{n}\left\|\left(\tau_{kk}-\tau_{kk}^{j}\right)\right\|_{L^2(\Omega)}^2 \leq \sum_{k,l=1}^{n}\left\|\left(\tau_{kl}-\tau_{kl}^{j}\right)\right\|_{L^2(\Omega)}^2 = \left\|\boldsymbol{\tau}-\boldsymbol{\tau}^{j}\right\|_{L^2(\Omega)}^2 \to 0, \quad \text{für } j\to\infty.$$

Daher gilt $\int_\Omega \operatorname{tr}\boldsymbol{\tau}\ \mathrm{d}x = 0$, also $\boldsymbol{\tau}\in\Sigma_0$.

In den Beweisen der folgenden Kapitel ist auch die im Folgende zu einer Raviart-Thomas Funktion $\boldsymbol{q}_{\mathrm{RT}}$ definierte Hilfsfunktion $\tilde{\boldsymbol{q}}_{\mathrm{RT}}$ nützlich. Einige ihrer Eigenschaften werden hier aufgeführt.

Bemerkung 2.20 Sei $\boldsymbol{q}_{\mathrm{RT}}\in RT_0\left(\mathcal{T};\mathbb{R}^{n\times n}\right)$ und definiere dazu

$$\tilde{\boldsymbol{q}}_{\mathrm{RT}} := \boldsymbol{q}_{\mathrm{RT}} - 1/n\left(\fint_\Omega \operatorname{tr}\boldsymbol{q}_{\mathrm{RT}}\ \mathrm{d}x\right)\mathrm{I}_{n\times n}\in H\left(\operatorname{div},\Omega;\mathbb{R}^{n\times n}\right).$$

Dann gilt offensichtlich $\int_\Omega \operatorname{tr}\tilde{\boldsymbol{q}}_{\mathrm{RT}}\ \mathrm{d}x = 0$, d.h. Lemma 2.18 ist auf $\tilde{\boldsymbol{q}}_{\mathrm{RT}}$ anwendbar. Außerdem gilt

(i) $\|\operatorname{dev}\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} = \|\operatorname{dev}\tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)}$ und $\|\operatorname{div}\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} = \|\operatorname{div}\tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)}$, da $\operatorname{dev}\mathrm{I}_{n\times n} = \operatorname{div}\mathrm{I}_{n\times n} = 0$,

(ii) Für $\boldsymbol{\sigma}_0\in L^2(\Omega;\mathbb{R}^{n\times n})/\mathbb{R}$ gilt

$$\|\boldsymbol{\sigma}_0-\tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} \leq \|\boldsymbol{\sigma}_0-\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)},$$

denn $A:\mathrm{I}_{n\times n} = \operatorname{tr}A$ für alle $A\in\mathbb{R}^{n\times n}$ und $\int_\Omega \operatorname{tr}\boldsymbol{\sigma}_0\ \mathrm{d}x = 0$ und damit

$$\begin{aligned}
\|\boldsymbol{\sigma}_0-\tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 =& \left\|\boldsymbol{\sigma}_0-\boldsymbol{q}_{\mathrm{RT}}+1/n\left(\fint_\Omega \operatorname{tr}\boldsymbol{q}_{\mathrm{RT}}\ \mathrm{d}x\right)\mathrm{I}_{n\times n}\right\|_{L^2(\Omega)}^2\\
=&\|\boldsymbol{\sigma}_0-\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2+\left\|1/n\left(\fint_\Omega \operatorname{tr}\boldsymbol{q}_{\mathrm{RT}}\ \mathrm{d}x\right)\mathrm{I}_{n\times n}\right\|_{L^2(\Omega)}^2\\
&+2/n\left(\fint_\Omega \operatorname{tr}\boldsymbol{q}_{\mathrm{RT}}\ \mathrm{d}x\right)\left(\int_\Omega(\boldsymbol{\sigma}_0-\boldsymbol{q}_{\mathrm{RT}}):\mathrm{I}_{n\times n}\ \mathrm{d}x\right)\\
=&\|\boldsymbol{\sigma}_0-\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2+1/n^2\left(\fint_\Omega \operatorname{tr}\boldsymbol{q}_{\mathrm{RT}}\ \mathrm{d}x\right)^2\left(\int_\Omega \mathrm{I}_{n\times n}:\mathrm{I}_{n\times n}\ \mathrm{d}x\right)\\
&+2/n\left(\fint_\Omega \operatorname{tr}\boldsymbol{q}_{\mathrm{RT}}\ \mathrm{d}x\right)\left(\int_\Omega \operatorname{tr}\boldsymbol{\sigma}_0-\operatorname{tr}\boldsymbol{q}_{\mathrm{RT}}\ \mathrm{d}x\right)\\
=&\|\boldsymbol{\sigma}_0-\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2-1/n\left(\fint_\Omega \operatorname{tr}\boldsymbol{q}_{\mathrm{RT}}\ \mathrm{d}x\right)\left(\int_\Omega \operatorname{tr}\boldsymbol{q}_{\mathrm{RT}}\ \mathrm{d}x\right)\\
=&\|\boldsymbol{\sigma}_0-\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2-|\Omega|/n\left(\int_\Omega \operatorname{tr}\boldsymbol{q}_{\mathrm{RT}}\ \mathrm{d}x\right)^2\\
\leq&\|\boldsymbol{\sigma}_0-\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2.
\end{aligned}$$

(iii) Wenn $f \in L^2(\Omega;\mathbb{R})$ und $\int_\Omega f \,\mathrm{d}x = 0$, dann ist $\int_\Omega f \operatorname{tr} \boldsymbol{q}_{\mathrm{RT}} \,\mathrm{d}x = \int_\Omega f \operatorname{tr} \tilde{\boldsymbol{q}}_{\mathrm{RT}} \,\mathrm{d}x$, da $\operatorname{tr} \boldsymbol{q}_{\mathrm{RT}} - \operatorname{tr} \tilde{\boldsymbol{q}}_{\mathrm{RT}} = \fint_\Omega \operatorname{tr} \boldsymbol{q}_{\mathrm{RT}} \,\mathrm{d}x$ konstant ist.

Desweiteren ist folgende Zerlegung sehr hilfreich.

Lemma 2.21. *Für alle deviatorischen Funktionen $u \in P_0\left(\mathcal{T};\mathbb{R}^{n\times n}\right)$ existieren $z \in H_0^1\left(\Omega;\mathbb{R}^n\right)$ mit $\operatorname{div} z = 0$ und $b \in L^2\left(\Omega;\mathbb{R}^{n\times n}\right)$, so dass*

$$u = \mathrm{D}\, z + \operatorname{dev} b$$

eine orthogonal Zerlegung in $L^2\left(\Omega;\mathbb{R}^{n\times n}\right)$ ist.

Beweis. Die Lösbarkeit des Stokes Problems in n Dimensionen (siehe beispielsweise [Bra13, S.153 ff.]) garantiert , dass es $\zeta \in L_0^2(\Omega) := \left\{v \in L^2(\Omega) : \int_\Omega v \,\mathrm{d}x = 0\right\}$ und $z \in H_0^1\left(\Omega;\mathbb{R}^n\right)$ gibt, so dass für alle Testfunktionen $\varphi \in H_0^1\left(\Omega;\mathbb{R}^n\right)$ gilt

$$\int_\Omega u : \mathrm{D}\,\varphi \,\mathrm{d}x = \int_\Omega \mathrm{D}\, z : \mathrm{D}\,\varphi \,\mathrm{d}x - \int_\Omega \zeta \cdot \operatorname{div}\varphi \,\mathrm{d}x \tag{2.13}$$

$$\operatorname{div} z = 0. \tag{2.14}$$

Sei

$$b := u - \mathrm{D}\, z \in L^2\left(\Omega;\mathbb{R}^{n\times n}\right),$$

so gilt wegen Lemma 2.5 Punkt (i) und (2.14)

$$\operatorname{tr} b = \operatorname{tr}\left(u - \mathrm{D}\, z\right) = 0 - \operatorname{div} z = 0.$$

Da u deviatorisch und z divergenzfrei sind, ist also auch b deviatorisch und $b = \operatorname{dev} b$. Weil außerdem $z \in H_0^1\left(\Omega;\mathbb{R}^n\right)$ liegt, gilt nach (2.13)–(2.14)

$$\begin{aligned}\int_\Omega b : \mathrm{D}\, z \,\mathrm{d}x &= \int_\Omega u : \mathrm{D}\, z \,\mathrm{d}x - \int_\Omega \mathrm{D}\, z : \mathrm{D}\, z \,\mathrm{d}x \\ &= \int_\Omega \mathrm{D}\, z : \mathrm{D}\, z \,\mathrm{d}x - \int_\Omega \zeta \cdot \operatorname{div} z \,\mathrm{d}x - \int_\Omega \mathrm{D}\, z : \mathrm{D}\, z \,\mathrm{d}x = 0.\end{aligned}$$

Insbesondere ergibt sich

$$\|\mathrm{D}\, z\|_{L^2(\Omega)}^2 + \|b\|_{L^2(\Omega)}^2 = \|u\|_{L^2(\Omega)}^2. \qquad \square$$

Dieser Paragraph schließt mit einer nützlichen Abschätzung, die später Schreibarbeit erspart.

Bemerkung 2.22 Seien $a, b \in \mathbb{R}^+$, wobei $\mathbb{R}^+ := \{x \in \mathbb{R} : 0 < x\}$, und $C_1, C_2, K_1, K_2 \in \mathbb{R}^+$ Konstanten. In den Beweisen werden häufiger Abschätzungen der folgenden Form benötigt

$$C_1 a + C_2 b \leq C\sqrt{K_1 a^2 + K_2 b^2}.$$

Folgender Trick führt zu einer kleinen Konstante C

$$\begin{aligned} C_1a + C_2b =& \sqrt{(C_1a + C_2b)^2} \leq \min_{\lambda\in\mathbb{R}^+} \sqrt{(1+\lambda)C_1^2a^2 + (1+1/\lambda)C_2^2b^2} \\ &\leq \min_{\lambda\in\mathbb{R}^+} \max\left\{\sqrt{(1+\lambda)C_1^2/K_1},\ \sqrt{(1+1/\lambda)C_2^2/K_2}\right\}\sqrt{K_1a^2+K_2b^2}. \end{aligned}$$

C wird offenbar minimal, wenn $(1+\lambda)C_1^2/K_1 = (1+1/\lambda)C_2^2/K_2$, also $\lambda = \left(C_2^2K_1\right) / \left(C_1^2K_2\right)$ und $C = \sqrt{C_1^2/K_1 + C_2^2/K_2}$.

2.6 Modellierung

In diesem Abschnitt sollen die Stokes bzw. Navier-Stokes Gleichungen genauer betrachtet werden, die es numerisch mit der dPG-Methode zu lösen gilt. Die Ausführungen richten sich nach [CTVW10], [CW07] und [Bra13, III.6].

Mit Hilfe der Navier-Stokes Gleichungen wird die Bewegung einer inkompressiblen, zähen Flüssigkeit in einem n-dimensionalen Körper Ω (mit $n = 2$ oder $n = 3$) beschrieben.

Sei $f:\ \Omega \to \mathbb{R}^n$ die gegebene äußere Kraftdichte und $\eta/\rho > 0$ die kinematische Viskosität, der Quotient aus dynamischer Viskosität η und Dichte ρ. Die Zähigkeit kann durch geeignete Skalierung später als $\eta/\rho = 1$ angesetzt werden, da die Dichte der Flüssigkeit als praktisch konstant betrachtet wird. Zu bestimmen sind das Geschwindigkeitsfeld $u:\ \Omega \to \mathbb{R}^n$, die Druckverteilung $p:\ \Omega \to \mathbb{R}$ und der Spannungstensor $\boldsymbol{\sigma}_{\mathrm{ph}}:\ \Omega \to \mathbb{R}^{n\times n}$. Die Standardgleichungen für inkompressible Newtonsche Flüssigkeiten, die diese Größen verbinden sind, das Materialgesetz

$$\boldsymbol{\sigma}_{\mathrm{ph}} + p\mathrm{I}_{n\times n} - 2\frac{\eta}{\rho}\epsilon(u) = 0, \tag{2.15}$$

mit dem symmetrischen Gradienten $\epsilon(u) := \left(\mathrm{D}\,u + (\mathrm{D}\,u)^\top\right)/2$, die Impulserhaltung

$$\frac{\partial u}{\partial t} + u\cdot \mathrm{D}\,u - \operatorname{div}\boldsymbol{\sigma}_{\mathrm{ph}} = f \tag{2.16}$$

und das Gesetz über die Erhaltung der Masse, das in einem System ohne Quellen und Senken besagt, dass

$$\operatorname{div} u = 0. \tag{2.17}$$

Zu diesem System gehören noch eine Anfangsbedingung $u|_{t=0} = u_0$ und Randbedingungen. In den meisten Anwendungen inkompressibler Newtonscher Flüsse handelt es sich um Dirichlet Randbedingungen an die Geschwindigkeit. Davon soll auch hier ausgegangen

werden. Es ist also

$$u = g \quad \text{entlang } \partial\Omega \tag{2.18}$$

gefordert. Aus dem Gaußschen Integralsatz ergibt sich eine Verträglichkeitsbedingung an die Randwerte $g : \Omega \to \mathbb{R}^n$, denn

$$\int_{\partial\Omega} g \cdot \nu \, \mathrm{d}s = \int_{\partial\Omega} u \cdot \nu \, \mathrm{d}s = \int_{\Omega} \operatorname{div} u \, \mathrm{d}x = 0. \tag{2.19}$$

Da der Druck durch (2.15)–(2.17) lediglich bis auf eine additive Konstante bestimmt wird, ist weiterhin eine Normierung der Form

$$\int_{\Omega} p \, \mathrm{d}x = 0$$

üblich. Die Gleichungen (2.15)–(2.17) werden als Spannungs-Geschwindigkeits-Druck Formulierung der inkompressiblen Navier-Stokes Gleichungen bezeichnet.

Aus diesem System kann die Spannung $\boldsymbol{\sigma}_{\mathrm{ph}}$ eliminiert werden, in dem (2.15) in (2.16) eingesetzt wird. Da bei inkompressiblen Flüssigkeiten die Divergenz von $(\mathrm{D}\, u)^\top$ verschwindet, lautet das entstehende System

$$\begin{aligned} \frac{\partial u}{\partial t} + u \cdot \mathrm{D}\, u + \nabla p - \frac{\eta}{\rho}\Delta(u) &= f \quad \text{in } \Omega, \\ \operatorname{div} u &= 0 \quad \text{in } \Omega \end{aligned} \tag{2.20}$$

und wird als Geschwindigkeits-Druck Formulierung der Navier-Stokes Gleichung bezeichnet. Hierbei sei auch der Laplace-Operator zeilenweise definiert $\Delta u := (\Delta u_1, \ldots, \Delta u_n)^\top = \sum_{k=1}^n (\partial_{kk} u_1, \ldots, \partial_{kk} u_n)^\top = \operatorname{div} \mathrm{D}\, u$. Ohne den nichtlinearen Term $u \cdot \mathrm{D}\, u$ werden diese Gleichungen als Stokes Gleichungen bezeichnet. Das klassische stationäre Stokes Problem entsteht, wenn die Zähigkeit zu 1 skaliert wird. Es werden also ein Geschwindigkeitsfeld $u : \Omega \to \mathbb{R}^n$ und eine Druckverteilung $p : \Omega \to \mathbb{R}$ gesucht, die folgenden Gleichungen erfüllen

$$\begin{aligned} \Delta u + \nabla p &= -f \quad \text{in } \Omega, \\ \operatorname{div} u &= 0 \quad \text{in } \Omega, \\ u &= g \quad \text{entlang } \partial\Omega. \end{aligned} \tag{2.21}$$

Dabei genüge g der Verträglichkeitsbedingung (2.19). Als klassische Lösung des Stokes Problem werden $u \in C^2(\Omega, \mathbb{R}^n) \cap C(\overline{\Omega}; \mathbb{R}^n)$ und $p \in C^0(\Omega)$ bezeichnet, die das System (2.21) erfüllen, sollten diese existieren.

Unter Verwendung der homogenen Randbedingungen $g \equiv 0$ mit den Räumen $\tilde{X} := H_0^1(\Omega; \mathbb{R}^n)$, $\tilde{Y} := L_0^2 := \{g \in L^2(\Omega) : \int_\Omega g \, \mathrm{d}x = 0\}$ ergibt sich das folgende Sattelpunkt-

problem. Zu gegebenem $f \in L^2(\Omega;\mathbb{R}^n)$ sind $u \in \tilde{X}$ und $p \in \tilde{Y}$ gesucht mit

$$\begin{aligned} \int_\Omega \mathrm{D}\, u : \mathrm{D}\, v \,\mathrm{d}x + \int_\Omega \operatorname{div} v \cdot p \,\mathrm{d}x &= \int_\Omega f \cdot v \,\mathrm{d}x \quad && \text{für alle } v \in \tilde{X}, \\ \int_\Omega \operatorname{div} u \cdot q \,\mathrm{d}x &= 0 && \text{für alle } q \in \tilde{Y}. \end{aligned}$$

Die eindeutige Lösbarkeit dieser gemischten Formulierung wird mit Hilfe des Brezzi-Splitting Theorems bewiesen, in dem das Ladyshenskaya Lemma und Eigenschaften von D ausgenutzt werden, um die Inf-Sup-Bedingung zu beweisen [Bra13, S.154ff]. Zu dieser variationellen Formulierung sind zahlreiche Diskretisierungen bekannt. Beispielsweise das Taylor-Hood-Element mit $\tilde{X}_h := C_0(\Omega;\mathbb{R}^2) \cap P_2\left(\mathcal{T};\mathbb{R}^2\right)$ und $\tilde{Y}_h := C(\overline{\Omega}) \cap L_0^2(\Omega) \cap P_1(\mathcal{T})$ [Bra13, S.163] oder das MINI-Element welches in $\tilde{X}_h$ bubble Funktionen, die sich als Produkt der Baryzentrischen Koordinaten ergeben, einsetzt [Bra13, S.164].

Die Beschäftigung mit nicht Newtonschen Flüssigkeiten lenkt mehr Interesse auf die Spannungs-Geschwindigkeits-Druck Formulierung, denn in dieser Situation ist (2.15) nicht linear, die Spannung kann damit nicht einfach eliminiert werden. Ein anderer Vorteil dieses Systems ist, dass die Spannung direkt berechnet wird. Dies verhindert, dass durch die numerische Ableitung der berechneten Geschwindigkeit, die Approximationsordnung von $\boldsymbol{\sigma}_{\mathrm{ph}}$ vermindert wird. Die Spannungs-Geschwindigkeits-Druck Formulierung weist allerdings auch Nachteile auf. Zum einen ist die Zahl der Unbekannten deutlich höher, zum anderen ist die Symmetrieanforderung an die Spannung kompliziert. Daher wird die Pseudospannungs-Geschwindigkeits Formulierung betrachtet, die akkurate Approximation der physikalischen Größen Spannung und Wirbelstärke ermöglicht und keine Zusatzbedingungen an die Approximationsräume stellt. Dazu wird die Pseudospannung

$$\boldsymbol{\sigma} := \frac{\eta}{\rho}\,\mathrm{D}\, u - p\mathrm{I}_{n\times n}$$

definiert. Auf Grund der Divergenzfreiheit von u, (2.17), gilt

$$p = -\frac{1}{n}\operatorname{tr}\sigma.$$

Der Druck ist also leicht zu bestimmen. Für den spurfreien Anteil $\operatorname{dev}\boldsymbol{\sigma} = \boldsymbol{\sigma} - 1/n \operatorname{tr}\boldsymbol{\sigma}\mathrm{I}_{n\times n}$, gilt also

$$\operatorname{dev}\boldsymbol{\sigma} = \frac{\eta}{\rho}\,\mathrm{D}\, u.$$

Da zusätzlich nach (2.15) und wegen $\operatorname{div}\left((\mathrm{D}\, u)^\top\right) = 0$ gilt

$$\operatorname{div}\boldsymbol{\sigma}_{\mathrm{ph}} = \operatorname{div}\boldsymbol{\sigma},$$

ergeben sich die folgenden Gleichungen als Pseudo-Spannung-Geschwindigkeits Formulie-

rung

$$\frac{\partial u}{\partial t} + u \cdot \mathrm{D}\, u - \operatorname{div} \boldsymbol{\sigma} = f \quad \text{in } \Omega, \tag{2.22}$$

$$\operatorname{dev} \boldsymbol{\sigma} - \frac{\eta}{\rho} \mathrm{D}\, u = 0 \quad \text{in } \Omega, \tag{2.23}$$

$$u = g \quad \text{entlang } \partial\Omega.$$

wobei g wie zuvor (2.19) erfülle. Die Normierung $\int_\Omega p \, \mathrm{d}x = 0$, wird zu der linearen Nebenbedingung

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

die in $H(\operatorname{div}, \Omega; \mathbb{R}^{n\times n})/\mathbb{R}$ gefordert wird. Aus der Pseudospannung $\boldsymbol{\sigma}$ lassen sich neben dem Druck, p, auch ohne Verringerung der Approximationsgüte die Spannung, $\boldsymbol{\sigma}_{\mathrm{ph}}$, und die Wirbelstärke, $\omega = \nabla \times u = 1/2 \left(\mathrm{D}\, u - (\mathrm{D}\, u)^\top\right)$, wie folgt berechnen

$$\boldsymbol{\sigma}_{\mathrm{ph}} = -p\mathrm{I}_{n\times n} + \eta/\rho \left(\mathrm{D}\, u + (\mathrm{D}\, u)^\top\right) = \boldsymbol{\sigma} + \eta/\rho \, (\mathrm{D}\, u)^\top = \boldsymbol{\sigma} + (\operatorname{dev} \boldsymbol{\sigma})^\top,$$
$$\omega = \rho/(2\eta) \left(\operatorname{dev} \boldsymbol{\sigma} - (\operatorname{dev} \boldsymbol{\sigma})^\top\right).$$

In dieser Arbeit wird das stationäre Stokes Problem in der Pseudospannungs-Geschwindigkeits Formulierung betrachtet, bei dem die als konstant angenommene Viskosität η/ρ durch Skalierung von $\boldsymbol{\sigma}$ und f mit dieser Größe nicht mehr auftritt.

Problem 1 (Pseudospannungs-Geschwindigkeits Formulierung)
Zu gegebenem $f \in L^2(\Omega; \mathbb{R}^n)$ und $g \in H^{1/2}(\partial\Omega; \mathbb{R}^n)$ sind $u \in H^1(\Omega; \mathbb{R}^n)$ und $\boldsymbol{\sigma} \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R}$ gesucht mit

$$\operatorname{div} \boldsymbol{\sigma} + f = 0 \quad \text{in } \Omega,$$
$$\operatorname{dev} \boldsymbol{\sigma} - \mathrm{D}\, u = 0 \quad \text{in } \Omega,$$
$$u = g \quad \text{entlang } \partial\Omega,$$

wobei g entlang jeder der M Zusammenhangskomponenten Γ_m, $1 \le m \le M$, des Randes $\Gamma = \bigcup_{m=1}^M \Gamma_m$ die Kompatibilitätsbedingung

$$\int_{\Gamma_m} \nu \cdot g \, \mathrm{d}s = 0$$

erfüllt.

Eine gemischte Formulierung zu diesem Problem verwendet $\tilde{X} := H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R}$ und $\tilde{Y} := L^2(\Omega)$ und sucht $\boldsymbol{\sigma} \in \tilde{X}$ und $u \in \tilde{Y}$ mit (2.18) zu gegebenem $f \in L^2(\Omega)$, so

dass

$$\begin{aligned}\int_\Omega \operatorname{dev}\boldsymbol{\sigma} : \boldsymbol{\tau}\ \mathrm{d}x + \int_\Omega u \cdot \operatorname{div}\boldsymbol{\tau}\ \mathrm{d}x &= \int_{\partial\Omega} g \cdot (\boldsymbol{\tau}\nu)\ \mathrm{d}s \qquad \text{für alle } \boldsymbol{\tau} \in \tilde{X},\\ \int_\Omega \operatorname{div}\boldsymbol{\sigma} \cdot v\ \mathrm{d}x &= -\int_\Omega f \cdot v\ \mathrm{d}x \qquad \text{für alle } v \in \tilde{Y}.\end{aligned}$$

Dieses Variationsproblem hat nach [CTVW10, Theorem 2.3] eine eindeutige Lösung.

Bemerkung 2.23 Sei $Z := H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R} \times L^2(\Omega)$ für jedes $(\boldsymbol{\rho}, w) \in Z$ versehen mit der Norm $\|(\boldsymbol{\rho}, w)\|_Z^2 := \|\boldsymbol{\rho}\|_{H(\operatorname{div},\Omega)}^2 + \|w\|_{L^2(\Omega)}^2$. In [CTVW10, Theorem 2.3] wird mit Hilfe des Brezzi-Splitting Theorems [Bra13, Satz 4.3] bewiesen, dass die lineare Abbildung $L : Z \to Z^*$, die durch

$$\langle L(z), \hat{z}\rangle = \tilde{b}(z, \hat{z}) \quad \text{für } \hat{z} \in Z$$

definiert ist, ein Isomorphismus ist. Also erfüllt die Bilinearform $\tilde{b} : Z \times Z \to \mathbb{R}$

$$\tilde{b}((\boldsymbol{\tau}, v), (\boldsymbol{\rho}, w)) := \int_\Omega \operatorname{dev}\boldsymbol{\tau} : \boldsymbol{\rho}\ \mathrm{d}x + \int_\Omega v \cdot \operatorname{div}\boldsymbol{\rho}\ \mathrm{d}x + \int_\Omega \operatorname{div}\boldsymbol{\tau} \cdot w\ \mathrm{d}x,$$

die folgende Inf-Sup-Bedingung

$$\gamma := \inf_{z\in Z\setminus\{0\}} \sup_{\hat{z}\in Z\setminus\{0\}} \frac{\tilde{b}(z,\hat{z})}{\|z\|_Z \|\hat{z}\|_Z} > 0.$$

Dies wird später für den Beweis einer ähnlichen Inf-Sup-Bedingung in der ultra-schwachen Formulierung verwendet, siehe Theorem 3.8.

In der dPG-Version der Pseudospannungs-Geschwindigkeits Formulierung wird schon auf dem kontinuierlichen Level eine reguläre Triangulierung $\mathcal{T}$ zu Grunde gelegt. Wie auch bei der gemischten Formulierung werden die Gleichungen aus Problem 1 mit jeweils einer Testfunktion multipliziert und über Ω integriert. Anschließend werden beide Gleichungen partiell integriert, da die Testfunktionen im vorliegenden Setting gebrochen sind, entstehen dabei Randintegrale auf dem Skelett. Es ergeben sich folgende Gleichungen

$$\int_\Omega \boldsymbol{\sigma} : \mathrm{D}_{\mathrm{NC}} v\ \mathrm{d}x - \left\langle \gamma_\nu^{\mathcal{T}}\boldsymbol{\sigma}, \gamma_0^{\mathcal{T}} v\right\rangle_{\partial\mathcal{T}} = \int_\Omega f \cdot v\ \mathrm{d}x \quad \text{für alle } v \in H^1(\mathcal{T}; \mathbb{R}^n), \tag{2.24}$$

$$\int_\Omega \operatorname{dev}\boldsymbol{\sigma} : \boldsymbol{\tau} + \int_\Omega u \cdot \operatorname{div}_{\mathrm{NC}}\boldsymbol{\tau}\ \mathrm{d}x - \left\langle \gamma_\nu^{\mathcal{T}}\boldsymbol{\tau}, \gamma_0^{\mathcal{T}} u\right\rangle = 0 \quad \text{für alle } \boldsymbol{\tau} \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R}. \tag{2.25}$$

Es werden neue Variablen für die Spuren der Geschwindigkeit u auf dem Skelett $\gamma_0^{\mathcal{T}} u = s$ und die Normalenspuren des Pseudostress $\boldsymbol{\sigma}$ auf dem Skelett $\gamma_\nu^{\mathcal{T}}\boldsymbol{\sigma} = t$ eingeführt. Dadurch ist es möglich weder an u noch an $\boldsymbol{\sigma}$ Differenzierbarkeitsbedingungen zu stellen. Auch die Geschwindigkeitsrandbedingungen werden auf s übertragen, also wird $s|_\Gamma = g$ gefordert. Die Komponenten einer Lösung zu (2.24)–(2.25) werden also in folgenden Räumen gesucht, $\boldsymbol{\sigma} \in L^2\left(\Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R}$, $u \in L^2(\Omega; \mathbb{R}^n)$, $s \in H^{1/2}(\partial\mathcal{T}; \mathbb{R}^n)$ und $t \in H^{-1/2}(\partial\mathcal{T}; \mathbb{R}^n)$.

Die dPG Methode löst dieses System, in dem $x \in X$ wie in (2.7) gesucht wird, so dass für alle $y \in Y$ wie in (2.8)

$$b(x, y) = F(y). \tag{2.26}$$

Dabei ist die Bilinearform b, wie in Definition 3.2 definiert, die Summe der linken Seiten von (2.24) und (2.25). Das Funktional F aus Definition 3.2 ist leicht modifiziert, es enthält neben den Summen der rechten Seiten noch den Term $\left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}, \hat{s} \right\rangle_{\partial\mathcal{T}}$ wobei $\hat{s} \in H^{1/2}(\partial\mathcal{T}; \mathbb{R}^n)$ im Wesentlichen eine Funktion ist, die die Randdaten $\hat{s}|_\Gamma = g$ realisiert. Dies ermöglicht ein diskretes Problem mit homogenen Randbedingungen, also $s \in H_0^{1/2}(\partial\mathcal{T}; \mathbb{R}^n)$, zu lösen. Aus einer Lösung u und $\boldsymbol{\sigma}$ von Problem 1 kann die Lösung dieses modifizierten Problems, leicht gewonnen werden. Es gilt nach Satz 3.3, dass $x = \left(u, \boldsymbol{\sigma}, \gamma_0^{\mathcal{T}} u - \hat{s}, \gamma_\nu^{\mathcal{T}} \boldsymbol{\sigma}\right)$ (2.26) erfüllt.

3 Analysis der dPG-Methode

Zunächst einmal werden in Abschnitt 3.1 einige allgemeine Informationen zu der Klasse der dPG-Methoden gegeben. In den folgenden Paragraphen wird dann die neue Methode detaillierter vorgestellt und analysiert.

3.1 Allgemeine Informationen zur dPG-Methode

Dieser Paragraph gibt einen Überblick über die allgemeine Theorie der dPG-Methoden inklusive der a-posteriori- Fehlerabschätzung aus [CDG14].

Die diskontinuierliche Petrov-Galerkin-Methode (dPG-Methode) kann als Finite-Elemente-Methode mit Nicht-Standardtesträumen oder als Least-Squares Finite-Elemente-Methode, die das Residuum der ultra-schwachen Formulierung in einer Nicht-Standardnorm minimiert, aufgefasst werden [DG10],[DG11b]. Diese Dualität wird auch in den verschiedenen Problemformulierungen deutlich. Dazu werden zunächst die Banachräume $(X,\|\bullet\|_X)$ und $(Y,\|\bullet\|_Y)$ und zwei abgeschlossene, endlich dimensionale Unterräume $X_h \subseteq X$ und $Y_h \subseteq Y$ betrachtet. Außerdem sei $b\colon\ X \times Y \to \mathbb{R}$ eine reelle, beschränkte Bilinearform. In diesem Setting können folgende Formulierungen betrachtet werden.

Problem 2 (Variationsformulierung des exakten Problems)
Zu gegebenem $F \in Y^*$ ist $x \in X$ gesucht mit $b(x, \bullet) = F$ in Y^*.

Dieses Problem ist wegen $\|F - b(x, \bullet)\|_{Y^*} = 0 \leq \|F - b(\xi, \bullet)\|_{Y^*}$ für alle $\xi \in X$ äquivalent zu folgender Minimierungsaufgabe.

Problem 3 (Exaktes Problem als Minimierung des Residuums)
Zu gegebenem $F \in Y^*$ ist $x \in X$ gesucht mit $x \in \operatorname{argmin}_{\xi\in X}\|F - b(\xi, \bullet)\|_{Y^*}$.

Diese Zielfunktion wird unter Umständen auch quadriert betrachtet, weshalb die dPG-Methode als Least-Square-Methode verstanden werden kann. In der idealisierten dPG-Methode wird das Problem im endlich dimensionalen Unterraum $X_h \subseteq X$ gelöst.

Problem 4 (Idealisierte dPG-Methode als Minimierung des Residuums)
Zu gegebenem $F \in Y^*$ ist $x_h \in X_h$ gesucht mit $x_h \in \operatorname{argmin}_{\xi_h\in X_h}\|F - b(\xi_h, \bullet)\|_{Y^*}$.

Ist Y insbesondere ein Hilbertraum mit Skalaprprodukt $\langle \bullet, \bullet \rangle_Y$, so erlaubt der sogenannte *trial-to-test*-Operator $T : X \to Y$ mit $\langle Tx, y\rangle_Y = b(x, y)$ für alle $y \in Y$ eine äquivalente Formulierung, [RBTD14]. Wenn $R\colon Y \to Y^*$ der Riesz-Isomorphismus ist [Alt06, S.163] und $B : X \to Y^*$ die Abbildung $Bx(y) := b(x, y)$, so gilt $T = R^{-1} \circ B$.

Problem 5 (Idealisierte dPG-Methode als Variationsformulierung)
Zu gegebenem $F \in Y^*$ ist $x_h \in X_h$ gesucht mit $b(x_h, \bullet) = F$ in $(T(X_h))^*$.

Im Allgemeinen ist Y ein unendlich dimensionaler Vektorraum und der trial-to-test-Operator T bzw. $\|F - b(\xi_h, \bullet)\|_{Y^*}$ sind in der idealisierten dPG-Methode kaum zu berechnen. Darum wird in der praktischen dPG-Methode statt Y ein ebenfalls endlich dimensionaler Unterraum $Y_h \subseteq Y$ betrachtet.

Problem 6 (Praktische dPG-Methode als Minimierung des Residuums)
Zu gegebenem $F \in Y^*$ ist $x_h \in X_h$ gesucht mit $x_h \in \operatorname{argmin}_{\xi_h \in X_h} \|F - b(\xi_h, \bullet)\|_{Y_h^*}$.

Dazu existiert ebenfalls eine äquivalente Variationsformulierung. Dafür wird der approximierte trial-to-test-Operator $T_h : X \to Y_h$ mit $\langle Tx, y_h \rangle_Y = b(x, y_h)$ für alle $y_h \in Y_h$ definiert.

Problem 7 (Praktische dPG-Methode als Variationsformulierung)
Zu gegebenem $F \in Y^*$ ist $x_h \in X_h$ gesucht mit $b(x_h, \bullet) = F$ in $(T_h(X_h))^*$.

Für die praktische dPG Methode wurde in [CDG14] eine a-posteriori-Fehlerkontrolle vorgestellt. Dabei wird vorausgesetzt, dass $(X, \|\bullet\|_X)$ ein reflexiver reeller Banachraum ist und $(Y, \langle \bullet, \bullet \rangle_Y)$ ein reeller Hilbertraum. Dies ist für die in dieser Arbeit verwendeten Räume aus (2.7) bzw. (2.8) gewährleistet. Um die Stabilität, die quasi-optimale Konvergenz sowie die Zuverlässigkeit der dPG-Methode und die Effizienz des Residuumsschätzers zu folgern, werden außerdem die folgenden Bedingung gefordert, deren Nachweis sich der Rest des Kapitels widmet.

Bedingung (B1) „Beschränktheits- und Eindeutigkeitsbedingung "
Es ist $b\colon X \times Y \to \mathbb{R}$ eine beschränkte Bilinearform mit $\|b\| := \sup_{x \in S(X)} \sup_{y \in S(Y)} b(x, y) < \infty$ und trivialem Kern $N := \{y \in Y : b(\bullet, y) = 0 \text{ in } X^*\} = \{0\}$.

Bedingung (B2) „Kontinuierliche Inf-Sup-Bedingung"
Es gilt $0 < \beta := \inf_{x \in S(X)} \sup_{y \in S(Y)} b(x, y)$.

Bedingung (B3) „Existenz eines Fortin-Operators"
Die abgeschlossenen Unterräume $X_h \subseteq X$, $Y_h \subseteq Y$ erlauben die Definition einer beschränkten, linearen Abbildung $\Pi\colon Y \to Y_h$ mit Operatornorm $\|\Pi\|$ und der Eigenschaft $b(\xi_h, (1 - \Pi)y) = 0$ für alle $\xi_h \in X_h$, $y \in Y$.

Aus (B1) und (B2) folgt nach [Bra13, Satz 3.6], dass Problem 2 bzw. Problem 3 eine eindeutige Lösung $x \in X$ besitzen. Außerdem ergibt sich, dass die Zielfunktion der Minimierungsaufgabe, $\|F - b(\xi_h, \bullet)\|_{Y^*}$, in Problem 4 bzw. Problem 5 ein zuverlässiger und effizienter Fehlerschätzer ist.

Lemma 3.1 (Zuverlässigkeit und Effizienz der idealisierten dPG-Methode). *Unter den Voraussetzungen (B1) und (B2) gilt für jedes $\xi_h \in X_h$*

$$\beta\|x-\xi_h\|_X \leq \|F-b(\xi_h,\bullet)\|_{Y^*} \leq \|b\|\,\|x-\xi_h\|_X\,.$$

Beweis. Es gilt

$$\beta = \inf_{\zeta\in X\setminus\{0\}} \frac{\|b(\zeta,\bullet)\|_{Y^*}}{\|\zeta\|_X} \leq \frac{\|b(x-\xi_h,\bullet)\|_{Y^*}}{\|x-\xi_h\|_X} \leq \frac{\|b\|\,\|x-\xi_h\|_X}{\|x-\xi_h\|_X}.$$

Dies ist die Behauptung, da $\|b(x-\xi_h,\bullet)\|_{Y^*} = \|F-b(\xi_h,\bullet)\|_{Y^*}$. □

Für die praktische dPG Methode wurde in [GQ14, Theorem 2.1] unter der Zusatzbedingung, dass X ein Hilbertraum ist, bereits gezeigt, dass aus den Bedingungen (B1), (B2) und (B3) die diskrete Stabilität und quasi-optimale Konvergenz in dem Sinn folgen, dass

$$\|x-x_h\|_X \leq \frac{\|\Pi\|\,\|b\|}{\beta} \min_{\xi_h\in X_h} \|x-\xi_h\|_X$$

gilt. Wie bereits erwähnt, wurde die a-posteriori-Fehlerabschätzung für die praktische dPG-Methode unter den genannten Bedingungen in [CDG14] vervollständigt. Die Einbeziehung eines Datenapproximationsfehlers neben dem Residuumsfehler ergibt einen zuverlässigen und effizienten Schätzer für die praktische dPG-Methode. Genauer gilt für alle $\xi_h \in X_h$

$$\begin{aligned}\beta\|x-\xi_h\|_X \leq \sqrt{1+\|\Pi\|^2}\|F-b\,(\xi_h,\bullet)\|_{Y_h^*} + &\|F\circ(1-\Pi)\|_{Y^*}\\ &\leq \|b\|\left(\sqrt{1+\|\Pi\|^2}+\|1-\Pi\|\right)\|x-\xi_h\|_X\end{aligned}$$

Da diese Abschätzungen für alle Elemente aus X_h und nicht nur für die diskrete Lösung x_h gelten, ist Stabilität beim inexakten Lösen gewährleistet.

In [CH15, Lemma 2.5] wird bewiesen, dass die letzte Bedingung durch die folgende diskrete Inf-Sup-Bedingung ersetzt werden kann.

Bedingung (B4) „Diskrete Inf-Sup-Bedingung“
Es gilt $0 < \beta_h := \inf_{x_h\in S(X_h)} \sup_{y_h\in S(Y_h)} b(x_h,y_h)$ für die abgeschlossenen Unterräume $X_h \subseteq X$, $Y_h \subseteq Y$.

Zusätzlich wird lediglich gefordert, dass auch X wie im vorliegenden Fall ein Hilbertraum ist. Ist (B2) erfüllt, so sind (B3) und (B4) sogar äquivalent und es gilt $\beta \leq \beta_h\|\Pi\|$ bzw. $\|\Pi\| \leq \|b\|/\beta_h$. Nach [CH15, Lemma 2.5] gilt dann die folgende Fehlerabschätzung für alle $\xi_h \in X_h$

$$\begin{aligned}\beta\|x-\xi_h\|_X \leq \|b\|/\beta_h\,\|F-b\,(\xi_h,\bullet)\|_{Y_h^*} + &\|F\circ(1-\Pi)\|_{Y^*}\\ &\leq 2\|b\|^2/\beta_h\|x-\xi_h\|_X\,. \qquad (3.1)\end{aligned}$$

In dieser Arbeit wird der Ansatz verfolgt, die Bedingungen (B1), (B2) und (B4) für die vorgestellte neue Methode zu überprüfen. In dem gegebenen Setting folgt aus (B4) die eindeutige Lösbarkeit von Problem 6 bzw. Problem 7. In [Hel14, Kapitel 3] sind auch Bedingungen und Zuverlässig- bzw. Effizienzabschätzungen für den Fall, dass Y kein Hilbertraum ist, zu finden.

3.2 Problemformulierung und kontinuierliche Bedingungen

In diesem Paragraphen werden die Bilinearform b und das Funktional F definiert. Anschließend werden die Annahmen (B1) und (B2) aus dem vorangegangenen Abschnitt überprüft.

Gegeben seien dazu der Ansatzraum X und der Testraum Y

$$\begin{aligned} X :=& L^2\left(\Omega;\mathbb{R}^{n\times n}\right)/\mathbb{R} \times L^2\left(\Omega;\mathbb{R}^n\right) \times H_0^{1/2}\left(\partial\mathcal{T};\mathbb{R}^n\right) \times H^{-1/2}\left(\partial\mathcal{T};\mathbb{R}^n\right), \\ Y :=& H\left(\operatorname{div},\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R} \times H^1\left(\mathcal{T};\mathbb{R}^n\right) \end{aligned}$$

ausgestattet mit den Normen

$$\begin{aligned} \|(\boldsymbol{\sigma},u,s,t)\|_X^2 :=& \|\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega)}^2 + \|s\|_{H_0^{1/2}(\partial\mathcal{T})}^2 + \|t\|_{H^{-1/2}(\partial\mathcal{T})}^2, \\ \|(\boldsymbol{\tau},v)\|_Y^2 :=& \|\boldsymbol{\tau}\|_{H(\operatorname{div},\mathcal{T})}^2 + \|v\|_{H^1(\mathcal{T})}^2. \end{aligned}$$

X und Y sind nach Bemerkung 2.14 insbesondere Hilberträume.

Zunächst werden aus den gegebenen Stokes Gleichungen die Bilinearform b und das Funktional F zur Formulierung der dPG-Methode hergeleitet, dabei wird die bereits erwähnte Pseudospannungs-Geschwindigkeits Formulierung verwendet.

Problem 8 (Pseudospannungs-Geschwindigkeits Formulierung)
Zu gegebenem $f \in L^2\left(\Omega;\mathbb{R}^n\right)$ und $g \in H^{1/2}\left(\partial\Omega;\mathbb{R}^n\right)$ sind $u \in H^1\left(\Omega;\mathbb{R}^n\right)$ und $\boldsymbol{\sigma} \in H\left(\operatorname{div},\Omega;\mathbb{R}^{n\times n}\right)/\mathbb{R}$ gesucht mit

$$\operatorname{div}\boldsymbol{\sigma} + f = 0 \quad \text{in } \Omega, \tag{3.2}$$

$$\operatorname{dev}\boldsymbol{\sigma} - \mathrm{D}\,u = 0 \quad \text{in } \Omega, \tag{3.3}$$

$$u = g \quad \text{entlang } \Gamma, \tag{3.4}$$

wobei g entlang jeder der M Zusammenhangskomponenten Γ_m, $1 \le m \le M$, des Randes $\Gamma = \bigcup_{m=1}^{M} \Gamma_m$ die Kompatibilitätsbedingung

$$\int_{\Gamma_m} \nu \cdot g \,\mathrm{d}s = 0 \tag{3.5}$$

erfüllt.

Definition 3.2. Die Bilinearform $b : X \times Y \to \mathbb{R}$ sei für alle $x = (\boldsymbol{\sigma}, u, s, t) \in X$ und $y = (\boldsymbol{\tau}, v) \in Y$ definiert durch

$$b(x,y) := \int_\Omega \boldsymbol{\sigma} : \mathrm{D}_{\mathrm{NC}}\, v \,\mathrm{d}x + \int_\Omega \operatorname{dev} \boldsymbol{\sigma} : \boldsymbol{\tau} \,\mathrm{d}x + \int_\Omega u \cdot \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau} \,\mathrm{d}x - \left\langle t, \gamma_0^{\mathcal{T}} v \right\rangle_{\partial\mathcal{T}} - \left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}, s \right\rangle_{\partial\mathcal{T}}. \tag{3.6}$$

Zu den gegebenen Daten $f \in L^2(\Omega;\mathbb{R}^n)$ und $\hat{s} \in H^{1/2}(\partial\mathcal{T};\mathbb{R}^n)$ wird das Funktional $F \in Y^*$ für alle $y = (\boldsymbol{\tau}, v) \in Y$ definiert durch

$$F(y) := \int_\Omega f \cdot v \,\mathrm{d}x + \left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}, \hat{s} \right\rangle_{\partial\mathcal{T}}. \tag{3.7}$$

Mit (2.7)–(2.8) und (3.6)–(3.7) lässt sich die ultra-schwache Formulierung analog Problem 2 aufstellen.

Problem 9 (Ultra-schwache Formulierung)
Zu gegebenem $f \in L^2(\Omega;\mathbb{R}^n)$ und $\hat{s} \in H^{1/2}(\partial\mathcal{T};\mathbb{R}^n)$ ist $x = (\boldsymbol{\sigma}, u, t, s) \in X$ gesucht mit

$$b(x,y) = F(y) \quad \text{für alle } y = (\boldsymbol{\tau}, v) \in Y. \tag{3.8}$$

Im Folgenden wird der Zusammenhang zwischen der ultra-schwachen und der Pseudospannungs-Geschwindigkeits Formulierung untersucht und gezeigt, dass für die theoretischen Betrachtungen homogene Nullranddaten genügen, wenn $\hat{s} = g$ entlang Γ gilt.

Satz 3.3. *Es sei $(u, \boldsymbol{\sigma}) \in H^1(\Omega;\mathbb{R}^n) \times H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R}$ eine Lösung von Problem 8. Außerdem sei $\hat{u} \in H^1(\Omega;\mathbb{R}^n)$ mit $\hat{u} = g$ auf Γ und $\hat{s} := \gamma_0^{\mathcal{T}} \hat{u} \in H^{1/2}(\partial\mathcal{T};\mathbb{R}^n)$. Setze nun $s := \gamma_0^{\mathcal{T}} u - \hat{s} \in H_0^{1/2}(\partial\mathcal{T};\mathbb{R}^n)$ und $t := \gamma_\nu^{\mathcal{T}} \boldsymbol{\sigma} \in H^{-1/2}(\partial\mathcal{T};\mathbb{R}^n)$.*

Dann löst $(\boldsymbol{\sigma}, u, s, t) \in X$ Problem 9 zu den gewünschten Daten $f \in L^2(\Omega;\mathbb{R}^n)$ und $\hat{s} \in H^{1/2}(\partial\mathcal{T};\mathbb{R}^n)$.

Beweis. Zunächst einmal gilt nun $(\boldsymbol{\sigma}, u, s, t) \in X$, da $u - \hat{u} = 0$ auf Γ und somit $s \in H_0^{1/2}(\partial\mathcal{T};\mathbb{R}^n)$. Es seien nun $(\boldsymbol{\tau}, v) \in Y$ und $T \in \mathcal{T}$ beliebig. Mit Hilfe der partiellen Integration aus Satz 2.8 und Gleichung (3.2) aus Problem 8 ergibt sich

$$\langle \gamma_\nu \boldsymbol{\sigma}, \gamma_0 v \rangle_{\partial T} = \int_T \boldsymbol{\sigma} : \mathrm{D}\, v \,\mathrm{d}x + \int_T \operatorname{div} \boldsymbol{\sigma} \cdot v \,\mathrm{d}x = \int_T \boldsymbol{\sigma} : \mathrm{D}\, v \,\mathrm{d}x - \int_T f \cdot v \,\mathrm{d}x$$

und damit in Summe über alle Simplices $T \in \mathcal{T}$, da $t = \gamma_\nu^{\mathcal{T}} \boldsymbol{\sigma}$,

$$\left\langle t, \gamma_0^{\mathcal{T}} v \right\rangle_{\partial\mathcal{T}} = \int_\Omega \boldsymbol{\sigma} : \mathrm{D}_{\mathrm{NC}}\, v \,\mathrm{d}x - \int_\Omega f \cdot v \,\mathrm{d}x. \tag{3.9}$$

Ebenso ergibt sich aus Satz 2.8 und Gleichung (3.3) aus Problem 8

$$\langle \gamma_\nu \boldsymbol{\tau}, \gamma_0 u \rangle_{\partial T} = \int_T \boldsymbol{\tau} : \mathrm{D}\, u \,\mathrm{d}x + \int_T \operatorname{div} \boldsymbol{\tau} \cdot u \,\mathrm{d}x = \int_T \boldsymbol{\tau} : \operatorname{dev} \boldsymbol{\sigma} \,\mathrm{d}x + \int_T \operatorname{div} \boldsymbol{\tau} \cdot u \,\mathrm{d}x$$

und nach Summation über alle Elemente $T \in \mathcal{T}$, da $\gamma_0^{\mathcal{T}} u = \hat{s} + s$,

$$\left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}, \hat{s} + s \right\rangle_{\partial \mathcal{T}} = \int_\Omega \boldsymbol{\tau} : \operatorname{dev} \boldsymbol{\sigma} \,\mathrm{d}x + \int_\Omega \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau} \cdot u \,\mathrm{d}x. \tag{3.10}$$

Die Addition von Gleichung (3.9) und Gleichung (3.10) ergibt

$$\begin{aligned} \int_\Omega f \cdot v \,\mathrm{d}x + \left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}, \hat{s} \right\rangle_{\partial \mathcal{T}} = &- \left\langle t, \gamma_0^{\mathcal{T}} v \right\rangle_{\partial \mathcal{T}} + \int_\Omega \boldsymbol{\sigma} : \mathrm{D}_{\mathrm{NC}}\, v \,\mathrm{d}x \\ &- \left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}, s \right\rangle_{\partial \mathcal{T}} + \int_\Omega \boldsymbol{\tau} : \operatorname{dev} \boldsymbol{\sigma} \,\mathrm{d}x + \int_\Omega \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau} \cdot u \,\mathrm{d}x, \end{aligned}$$

also ist nach Definition 3.2 Gleichung (3.8) erfüllt und die Behauptung bewiesen. □

Zunächst wird nun der erste Teil von (B1), die Beschränktheit von b, bewiesen.

Satz 3.4 (Beschränktheit von b). *Die in (3.6) definierten Bilinearform b ist beschränkt, denn es gilt*

$$|b(x,y)| \leq M \|x\|_X \|y\|_Y \quad \textit{für alle } x \in X,\, y \in Y,$$

wobei $M = \|b\| \leq \sqrt{3}$.

Beweis. Es seien $x = (\boldsymbol{\sigma}, u, t, s) \in X$ und $y = (\boldsymbol{\tau}, v) \in Y$ gegeben. Nach Bemerkung 2.11 existieren $w \in H_0^1(\Omega; \mathbb{R}^n)$ mit $s = \gamma_0^{\mathcal{T}} w$ und $\|w\|_{H^1(\Omega)} = \|s\|_{H^{1/2}(\partial \mathcal{T})}$ sowie $\boldsymbol{q} \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{n \times n}\right)$ mit $t = \gamma_\nu^{\mathcal{T}} \boldsymbol{q}$ und $\|\boldsymbol{q}\|_{H(\operatorname{div})} = \|t\|_{H^{-1/2}(\partial \mathcal{T})}$. Die partielle Integrationsformel (2.8) bedeutet nach Summation über alle $T \in \mathcal{T}$

$$\left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}, s \right\rangle_{\partial \mathcal{T}} = \int_\Omega \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau} \cdot w \,\mathrm{d}x + \int_\Omega \boldsymbol{\tau} : \mathrm{D}\, w \,\mathrm{d}x$$

und

$$\left\langle t, \gamma_0^{\mathcal{T}} v \right\rangle_{\partial \mathcal{T}} = \int_\Omega \operatorname{div} \boldsymbol{q} \cdot v \,\mathrm{d}x + \int_\Omega \boldsymbol{q} : \mathrm{D}_{\mathrm{NC}}\, v \,\mathrm{d}x.$$

Damit kann b wie folgt geschrieben werden

$$\begin{aligned} b(x,y) = &- \int_\Omega \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau} \cdot w \,\mathrm{d}x - \int_\Omega \boldsymbol{\tau} : \mathrm{D}\, w \,\mathrm{d}x - \int_\Omega \operatorname{div} \boldsymbol{q} \cdot v \,\mathrm{d}x - \int_\Omega \boldsymbol{q} : \mathrm{D}_{\mathrm{NC}}\, v \,\mathrm{d}x \\ &+ \int_\Omega \boldsymbol{\sigma} : \mathrm{D}_{\mathrm{NC}}\, v \,\mathrm{d}x + \int_\Omega \operatorname{dev} \boldsymbol{\sigma} : \boldsymbol{\tau} \,\mathrm{d}x + \int_\Omega u \cdot \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau} \,\mathrm{d}x. \end{aligned}$$

Die Anwendung von Dreiecks- und Cauchy-Schwarz-Ungleichung sowie Lemma 2.5 Punkt (iii) ergibt

$$|b(x,y)| \leq \|\boldsymbol{\tau}\|_{L^2(\Omega)} \left(\|\mathrm{D}\, w\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}\|_{L^2(\Omega)} \right) + \|\operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\|_{L^2(\Omega)} \left(\|w\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)} \right)$$

$$\|v\|_{L^2(\Omega)}\|\operatorname{div} \boldsymbol{q}\|_{L^2(\Omega)} + \|\mathrm{D}_{\mathrm{NC}}\, v\|_{L^2(\Omega)} \left(\|\boldsymbol{q}\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}\|_{L^2(\Omega)}\right).$$

Dies kann als $\mathbb{R}^4$-Skalarprodukt der Vektoren

$$\left(\|\mathrm{D}\, w\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}\|_{L^2(\Omega)},\ \ \|w\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)},\ \ \|\operatorname{div} \boldsymbol{q}\|_{L^2(\Omega)},\ \ \|\boldsymbol{q}\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}\|_{L^2(\Omega)}\right)^{\top}.$$

und

$$\left(\|\boldsymbol{\tau}\|_{L^2(\Omega)},\ \ \|\operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\|_{L^2(\Omega)},\ \ \|v\|_{L^2(\Omega)},\ \ \|\mathrm{D}_{\mathrm{NC}}\, v\|_{L^2(\Omega)}\right)^{\top}.$$

betrachtet werden. Dafür ergibt die Cauchy-Schwarz-Ungleichung folgende Abschätzung

$$\begin{aligned}|b(x,y)| \leq & \|y\|_Y \Bigg(\left(\|\mathrm{D}\, w\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}\|_{L^2(\Omega)}\right)^2 + \left(\|w\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}\right)^2 + \|\operatorname{div} \boldsymbol{q}\|^2_{L^2(\Omega)} \\ & + \left(\|\boldsymbol{q}\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}\|_{L^2(\Omega)}\right)^2 \Bigg)^{1/2}.\end{aligned}$$

Der zweite Faktor kann wie folgt gegen $\|x\|_X$ abgeschätzt werden, in dem die Forderung an die Norm von w und $\boldsymbol{q}$ genutzt werden

$$\begin{aligned}&\left(\|\mathrm{D}\, w\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}\|_{L^2(\Omega)}\right)^2 + \left(\|w\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}\right)^2 + \|\operatorname{div} \boldsymbol{q}\|^2_{L^2(\Omega)} + \left(\|\boldsymbol{q}\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}\|_{L^2(\Omega)}\right)^2 \\ &\quad \leq \min_{\lambda,\mu,\kappa\in\mathbb{R}^+} (1+\lambda)\|\mathrm{D}\, w\|^2_{L^2(\Omega)} + (1+1/\lambda)\|\boldsymbol{\sigma}\|^2_{L^2(\Omega)} + (1+\mu)\|w\|^2_{L^2(\Omega)} + \|\operatorname{div} \boldsymbol{q}\|^2_{L^2(\Omega)} \\ &\qquad + (1+1/\mu)\|u\|^2_{L^2(\Omega)} + (1+\kappa)\|\boldsymbol{q}\|^2_{L^2(\Omega)} + (1+1/\kappa)\|\boldsymbol{\sigma}\|^2_{L^2(\Omega)} \\ &\quad \leq \max\left\{ \min_{\lambda,\mu,\kappa\in\mathbb{R}^+} \{2+1/\lambda+1/\kappa, 1+1/\mu, 1+\mu, 1+\lambda, 1+\kappa, 1\} \right\} \\ &\qquad \left(\|\boldsymbol{\sigma}\|^2_{L^2(\Omega)} + \|u\|^2_{L^2(\Omega)} + \|w\|^2_{H^1(\Omega)} + \|\boldsymbol{q}\|^2_{H(\operatorname{div})}\right) \\ &\quad \leq 3\left(\|\boldsymbol{\sigma}\|^2_{L^2(\Omega)} + \|u\|^2_{L^2(\Omega)} + \|s\|^2_{H^{1/2}(\partial\mathcal{T})} + \|t\|^2_{H^{-1/2}(\partial\mathcal{T})}\right) \\ &\quad = 3\ \|x\|^2_X.\end{aligned}$$

Die Konstante 3 ergibt sich dabei mit $\lambda = \kappa = 2$ und $\mu = 1$. Es gilt also

$$|b(x,y)| \leq \sqrt{3}\,\|x\|_X \|y\|_Y. \qquad \square$$

Des Weiteren ist der Kern der Bilinearform b, $N := \{y \in Y :\ b(\,\bullet\,, y) = 0 \in X^*\}$, zu untersuchen, um (B1) zu zeigen.

Lemma 3.5. *Für Bilinearform b aus Definition 3.2 gilt*

$$N = \{y \in Y :\ b(\,\bullet\,, y) = 0 \in X^*\} = \{0\}.$$

Beweis. Sei $y = (\boldsymbol{\tau}, v) \in Y$, sodass $b(x,y) = 0$ für jedes $x = (\boldsymbol{\sigma}, u, s, t) \in X$.

Die Wahl $x = (0, u, 0, 0) \in X$ führt für jedes $u \in C_0^\infty(\Omega;\mathbb{R}^n) \subseteq L^2(\Omega;\mathbb{R}^n)$ zu der Gleichung $0 = \int_\Omega \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau} \cdot u \,\mathrm{d}x$. Nach dem Fundamentallemma der Variationsrechnung

gilt daher $\operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau} \equiv 0$, d.h. insbesondere $\boldsymbol{\tau} \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R}$ mit $\operatorname{div} \boldsymbol{\tau} = 0$. Also gilt nach (2.6) $\gamma_\nu^{\mathcal{T}} \boldsymbol{\tau} \in H^{-1/2}\left(\partial\mathcal{T}; \mathbb{R}^n\right)$ und damit $x = \left(0, 0, 0, \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}\right) \in X$. Daraus ergibt sich

$$0 = \left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}, \gamma_0^{\mathcal{T}} v \right\rangle_{\partial\mathcal{T}} = \int_\Omega v \cdot \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau} \, \mathrm{d}x + \int_\Omega \mathrm{D}_{\mathrm{NC}} v : \boldsymbol{\tau} \, \mathrm{d}x = \int_\Omega \mathrm{D}_{\mathrm{NC}} v : \boldsymbol{\tau} \, \mathrm{d}x.$$

Mit $x = (\boldsymbol{\tau}, 0, 0, 0) \in X$, da $H\left(\operatorname{div}, \mathcal{T}; \mathbb{R}^{n\times n}\right)/\mathbb{R} \subseteq L^2\left(\Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R}$, folgt

$$0 = \int_\Omega \boldsymbol{\tau} : \mathrm{D}_{\mathrm{NC}} v \, \mathrm{d}x + \int_\Omega \operatorname{dev} \boldsymbol{\tau} : \boldsymbol{\tau} \, \mathrm{d}x = \int_\Omega \operatorname{dev} \boldsymbol{\tau} : \boldsymbol{\tau} \, \mathrm{d}x = \|\operatorname{dev} \boldsymbol{\tau}\|_{L^2(\Omega)}^2,$$

also $\operatorname{dev} \boldsymbol{\tau} = 0$. Nach dem Tr-Div-Dev Lemma, Lemma 2.18, für $\boldsymbol{\tau} \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R}$, gilt außerdem

$$\|\operatorname{tr} \boldsymbol{\tau}\|_{L^2(\Omega)} \leq C_{\mathrm{tdd}} \left(\|\operatorname{dev} \boldsymbol{\tau}\|_{L^2(\Omega)} + \|\operatorname{div} \boldsymbol{\tau}\|_{L^2(\Omega)} \right) = 0,$$

also auch $\operatorname{tr} \boldsymbol{\tau} = 0$ und somit $\boldsymbol{\tau} \equiv 0$.

Damit ergibt sich für jedes $x = (\boldsymbol{\sigma}, 0, 0, 0) \in X$,

$$0 = \int_\Omega \boldsymbol{\sigma} : \mathrm{D}_{\mathrm{NC}} v \, \mathrm{d}x. \tag{3.11}$$

Sei $w \in H_0^1\left(\Omega; \mathbb{R}^n\right)$ die eindeutig bestimmt schwache Lösung des Laplace Problems

$$\operatorname{div} \mathrm{D}\, w = \Delta w = v \text{ in } \Omega \quad \text{ and } \quad w = 0 \text{ auf } \Gamma.$$

Dann gilt $\boldsymbol{q} := \mathrm{D}\, w \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)$ mit $\operatorname{div} \boldsymbol{q} = v$ und $\gamma_\nu^{\mathcal{T}} \boldsymbol{q} \in H^{-1/2}\left(\partial\mathcal{T}; \mathbb{R}^n\right)$. Nach dem Gaußschen Integralsatz ist insbesondere

$$\int_\Omega \operatorname{tr} \boldsymbol{q} \, \mathrm{d}x = \int_\Omega \operatorname{div} w \, \mathrm{d}x = \int_{\partial\Omega} w \cdot \nu \, \mathrm{d}x = 0,$$

also $\boldsymbol{q} \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R} \subseteq L^2\left(\Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R}$. Mit $x = \left(0, 0, 0, \gamma_\nu^{\mathcal{T}} \boldsymbol{q}\right) \in X$ ergibt sich aus Satz 2.8 und (3.11)

$$0 = \left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{q}, \gamma_0^{\mathcal{T}} v \right\rangle_{\partial\mathcal{T}} = \int_\Omega v \cdot \operatorname{div} \boldsymbol{q} \, \mathrm{d}x + \int_\Omega \boldsymbol{q} : \mathrm{D}_{\mathrm{NC}} v \, \mathrm{d}x = \|v\|_{L^2(\Omega)},$$

also $v = 0$.

Damit ist $N = \{0\}$ bewiesen. □

Es gilt also (B1) für die gegebene Bilinearform und es bleibt (B2) zu beweisen. Dazu wird folgendes Lemma, [CDG15b, Lemma 2.1], verwendet. Dieses Lemma ist in einer speziell an das vorliegende Setting angepassten Form auch in [CDG15a, Theorem 3.1] zu finden.

Lemma 3.6 (Splitting Lemma). *Es seien X und Y Banachräume mit Unterräumen $X_1, X_2 \subseteq X$ und $Y_1 \subseteq Y$, so dass X (nicht notwendigerweise direkt) zerlegbar ist in $X = X_1 + X_2$. Des Weiteren gelte*

$$b(x_2, y_1) = 0 \quad \text{für alle } x_2 \in X_2,\ y_1 \in Y_1$$

sowie die Inf-Sup-Bedingungen

$$0 < \beta_1 := \inf_{x_1 \in S(X_1)} \sup_{y_1 \in S(Y_1)} b(x_1, y_1) \quad \text{und} \quad 0 < \beta_2 := \inf_{x_2 \in S(X_2)} \sup_{y \in S(Y)} b(x_2, y).$$

In diesem Fall gilt (B2) mit

$$\beta \geq \frac{\beta_1 \beta_2}{\beta_1 + \beta_2 + \|b\|}.$$

Dieses Lemma wird mit der Zerlegung $X_1 := L^2\left(\Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R} \times L^2(\Omega; \mathbb{R}^n) \times \{0\} \times \{0\}$ und $X_2 := \{0\} \times \{0\} \times H_0^{1/2}(\partial\mathcal{T}; \mathbb{R}^n) \times H^{-1/2}(\partial\mathcal{T}; \mathbb{R}^n)$ und dem Unterraum $Y_1 := H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R} \times H_0^1(\Omega; \mathbb{R}^n)$ verwendet um (B2) zu zeigen. Dafür wird jedoch eine Hilfsbetrachtung benötigt.

Lemma 3.7. *Es gilt die folgende Inf-Sup-Bedingung*

$$0 < \hat{\beta} := \inf_{(\boldsymbol{\sigma}, u) \in S(\hat{X})} \sup_{(\boldsymbol{\tau}, v) \in S(Y_1)} \left(\int_\Omega \boldsymbol{\sigma} : \mathrm{D}\, v \, \mathrm{d}x + \int_\Omega \operatorname{dev} \boldsymbol{\sigma} : \boldsymbol{\tau} \, \mathrm{d}x + \int_\Omega u \cdot \operatorname{div} \boldsymbol{\tau} \, \mathrm{d}x \right) \quad (3.12)$$

mit $\hat{X} = L^2\left(\Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R} \times L^2(\Omega; \mathbb{R}^n)$ ausgestattet mit der Norm $\|(\boldsymbol{\sigma}, u)\|_{\hat{X}}^2 := \|\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega)}$ und $Y_1 = H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R} \times H_0^1(\Omega; \mathbb{R}^n)$ mit der Norm von Y.

Beweis. Es sei $0 \neq (\boldsymbol{\sigma}, u) \in \hat{X}$ vorgegeben. Dazu ist $(\boldsymbol{\tau}, v) \in Y$ zu finden, so dass die Bedingung (3.12) erfüllt ist. Zu diesem Zweck sei $\tilde{F} \in Z^*$ definiert durch

$$\tilde{F}(\boldsymbol{\rho}, w) := \int_\Omega \boldsymbol{\sigma} : \boldsymbol{\rho} + u \cdot w \, \mathrm{d}x \quad \text{für alle } (\boldsymbol{\rho}, w) \in Z := H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)/\mathbb{R} \times L^2(\Omega; \mathbb{R}^n),$$

also mit Z wie in Bemerkung 2.23. Die Cauchy-Schwarz-Ungleichung zeigt

$$\left\|\tilde{F}\right\|_{Z^*} \leq \|\boldsymbol{\sigma}\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)} \leq \sqrt{2}\|(\boldsymbol{\sigma}, u)\|_{\hat{X}}.$$

Sei für alle $(\boldsymbol{\tau}, v), (\boldsymbol{\rho}, w) \in Z$

$$\tilde{b}((\boldsymbol{\tau}, v), (\boldsymbol{\rho}, w)) := \int_\Omega \operatorname{dev} \boldsymbol{\tau} : \boldsymbol{\rho} \, \mathrm{d}x + \int_\Omega v \cdot \operatorname{div} \boldsymbol{\rho} \, \mathrm{d}x + \int_\Omega \operatorname{div} \boldsymbol{\tau} \cdot w \, \mathrm{d}x,$$

ebenfalls wie in Bemerkung 2.23. Nach ebendieser Bemerkung existiert $(\boldsymbol{\tau}, -v) \in Z$ mit $\tilde{b}((\boldsymbol{\tau}, -v), \bullet) = \tilde{F}$ in Z^*. Es gilt also für jedes $(\boldsymbol{\rho}, w) \in Z$

$$0 = \int_\Omega (\boldsymbol{\sigma} - \operatorname{dev} \boldsymbol{\tau}) : \boldsymbol{\rho} \, \mathrm{d}x + \int_\Omega \operatorname{div} \boldsymbol{\rho} \cdot v \, \mathrm{d}x + \int_\Omega (u - \operatorname{div} \boldsymbol{\tau}) \cdot w \, \mathrm{d}x. \quad (3.13)$$

Das Fundamentallemma der Variationsrechnung liefert $\operatorname{div}\boldsymbol{\tau} = u$. Nach der folgenden Argumentation gilt sogar $v \in H^1_0(\Omega;\mathbb{R}^n)$ mit $\mathrm{D}\, v = \boldsymbol{\sigma} - \operatorname{dev}\boldsymbol{\tau}$.

Um dies zu sehen, ist zunächst für alle $\boldsymbol{\varphi} \in C^\infty_0(\Omega;\mathbb{R}^{n\times n})$ zu zeigen, dass

$$\int_\Omega \operatorname{div}\boldsymbol{\varphi}\cdot v \,\mathrm{d}x = -\int_\Omega (\boldsymbol{\sigma} - \operatorname{dev}\boldsymbol{\tau}) : \boldsymbol{\varphi} \,\mathrm{d}x. \tag{3.14}$$

Betrachte dazu statt $\boldsymbol{\varphi} \in C^\infty_0(\Omega;\mathbb{R}^{n\times n})$ wie in Bemerkung 2.20

$$\tilde{\boldsymbol{\varphi}} := \boldsymbol{\varphi} - 1/n \left(\fint_\Omega \operatorname{tr}\boldsymbol{\varphi}\,\mathrm{d}x\right) \mathrm{I}_{n\times n} \in C^\infty_0(\Omega;\mathbb{R}^{n\times n})/\mathbb{R}.$$

Es gilt $\operatorname{div}\boldsymbol{\varphi} = \operatorname{div}\tilde{\boldsymbol{\varphi}}$. Außerdem gilt nach Lemma 2.5 und da $\boldsymbol{\sigma} \in L^2\left(\Omega;\mathbb{R}^{n\times n}\right)/\mathbb{R}$

$$\begin{aligned}
\int_\Omega (\boldsymbol{\sigma} - \operatorname{dev}\boldsymbol{\tau}) : \tilde{\boldsymbol{\varphi}} \,\mathrm{d}x &= \int_\Omega (\boldsymbol{\sigma} - \operatorname{dev}\boldsymbol{\tau}) : \boldsymbol{\varphi} \,\mathrm{d}x - 1/n \fint_\Omega \operatorname{tr}\boldsymbol{\varphi}\,\mathrm{d}x \int_\Omega (\boldsymbol{\sigma} - \operatorname{dev}\boldsymbol{\tau}) : \mathrm{I}_{n\times n}\,\mathrm{d}x \\
&= \int_\Omega (\boldsymbol{\sigma} - \operatorname{dev}\boldsymbol{\tau}) : \boldsymbol{\varphi} \,\mathrm{d}x - 1/n \fint_\Omega \operatorname{tr}\boldsymbol{\varphi}\,\mathrm{d}x \int_\Omega \operatorname{tr}\boldsymbol{\sigma} - 0 \,\mathrm{d}x \\
&= \int_\Omega (\boldsymbol{\sigma} - \operatorname{dev}\boldsymbol{\tau}) : \boldsymbol{\varphi} \,\mathrm{d}x.
\end{aligned}$$

Da $C^\infty_0(\Omega;\mathbb{R}^{n\times n})/\mathbb{R} \subseteq H\left(\operatorname{div},\Omega;\mathbb{R}^{n\times n}\right)/\mathbb{R}$, gilt nach (3.13) auch (3.14), d.h. es liegt $v \in H^1(\Omega;\mathbb{R}^n)$ mit schwacher Ableitung $\mathrm{D}\, v = \boldsymbol{\sigma} - \operatorname{dev}\boldsymbol{\tau}$. Für alle $\boldsymbol{\varphi} \in C^\infty_0(\Omega;\mathbb{R}^{n\times n})$ ergibt sich mit partieller Integration und Anwendung von (3.13), nach analoger Argumentation,

$$\begin{aligned}
\int_{\partial\Omega} v\cdot\boldsymbol{\varphi}\nu \,\mathrm{d}s &= \int_\Omega \mathrm{D}\, v : \boldsymbol{\varphi}\,\mathrm{d}x + \int_\Omega v\cdot\operatorname{div}\boldsymbol{\varphi}\,\mathrm{d}x \\
&= \int_\Omega (\boldsymbol{\sigma} - \operatorname{dev}\boldsymbol{\tau}) : \boldsymbol{\varphi}\,\mathrm{d}x + \int_\Omega \operatorname{div}\boldsymbol{\varphi}\cdot v\,\mathrm{d}x = 0,
\end{aligned}$$

also hat v homogene Randdaten. Insbesondere gilt also $(\boldsymbol{\tau}, v) \in Y_1$.

Wird $(\boldsymbol{\tau}, v) \in Y_1$ in die zu untersuchende Bilinearform aus (3.12) eingesetzt, so ergibt sich nach Lemma 2.5 Punkt (ii)

$$\begin{aligned}
&\int_\Omega \boldsymbol{\sigma} : \mathrm{D}\, v \,\mathrm{d}x + \int_\Omega \operatorname{dev}\boldsymbol{\sigma} : \boldsymbol{\tau}\,\mathrm{d}x + \int_\Omega u\cdot\operatorname{div}\boldsymbol{\tau}\,\mathrm{d}x \\
&\qquad = \int_\Omega \boldsymbol{\sigma} : (\boldsymbol{\sigma} - \operatorname{dev}\boldsymbol{\tau})\,\mathrm{d}x + \int_\Omega \boldsymbol{\sigma} : \operatorname{dev}\boldsymbol{\tau}\,\mathrm{d}x + \int_\Omega u\cdot u\,\mathrm{d}x \\
&\qquad = \|\boldsymbol{\sigma}\|^2_{L^2(\Omega)} + \|u\|^2_{L^2(\Omega)} = \|(\boldsymbol{\sigma}, u)\|^2_{\hat{X}}\,.
\end{aligned}$$

Die Norm $\|(\boldsymbol{\tau}, v)\|_Z$ lässt sich mit der Inf-Sup-Bedingung aus Bemerkung 2.23 abschätzen

$$\gamma\|(\boldsymbol{\tau}, v)\|_Z = \gamma\|(\boldsymbol{\tau}, -v)\|_Z \le \left\|\tilde{b}\left((\boldsymbol{\tau}, -v)\,,\,\bullet\right)\right\|_{Z^*} = \|F\|_{Z^*} \le \sqrt{2}\|(\boldsymbol{\sigma}, u)\|_{\hat{X}}\,.$$

Desweiteren gilt mit der Dreiecksungleichung

$$\|\mathrm{D}\, v\|_{L^2(\Omega)}^2 = \|\boldsymbol{\sigma} - \operatorname{dev} \boldsymbol{\tau}\|_{L^2(\Omega)}^2 \leq 2\|\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 + 2\|\operatorname{dev} \boldsymbol{\tau}\|_{L^2(\Omega)}^2 \leq 2\|\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 + 2\|\boldsymbol{\tau}\|_{L^2(\Omega)}^2 .$$

Diese letzten beiden Ungleichungen erlauben die Abschätzung

$$\begin{aligned}
\|(\boldsymbol{\tau}, v)\|_Y^2 &= \|\boldsymbol{\tau}\|_{L^2(\Omega)}^2 + \|\operatorname{div} \boldsymbol{\tau}\|_{L^2(\Omega)}^2 + \|v\|_{L^2(\Omega)}^2 + \|\mathrm{D}\, v\|_{L^2(\Omega)}^2 \\
&\leq 3\|\boldsymbol{\tau}\|_{L^2(\Omega)}^2 + \|\operatorname{div} \boldsymbol{\tau}\|_{L^2(\Omega)}^2 + \|v\|_{L^2(\Omega)}^2 + 2\|\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 \\
&\leq 3\|(\boldsymbol{\tau}, v)\|_Z^2 + 2\|\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 \\
&\leq 6/\gamma^2 \left(\|\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega)}^2\right) + 2\|\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 \\
&\leq \left(6/\gamma^2 + 2\right)\|(\boldsymbol{\sigma}, u)\|_{\hat{X}}^2 .
\end{aligned}$$

Nun ergibt sich

$$\begin{aligned}
\sup_{(\boldsymbol{\rho}, w) \in Y_1 \setminus \{0\}} & \frac{\int_\Omega \boldsymbol{\sigma} : \mathrm{D}\, w \, \mathrm{d}x + \int_\Omega \operatorname{dev} \boldsymbol{\sigma} : \boldsymbol{\rho} \, \mathrm{d}x + \int_\Omega u \cdot \operatorname{div} \boldsymbol{\rho} \, \mathrm{d}x}{\|(\boldsymbol{\rho}, w)\|_Y} \\
&\geq \frac{\int_\Omega \boldsymbol{\sigma} : \mathrm{D}\, v \, \mathrm{d}x + \int_\Omega \operatorname{dev} \boldsymbol{\sigma} : \boldsymbol{\tau} \, \mathrm{d}x + \int_\Omega u \cdot \operatorname{div} \boldsymbol{\tau} \, \mathrm{d}x}{\|(\boldsymbol{\tau}, v)\|_Y} \\
&\geq \frac{\|(\boldsymbol{\sigma}, u)\|_{\hat{X}}^2}{\sqrt{6/\gamma^2 + 2}\, \|(\boldsymbol{\sigma}, u)\|_{\hat{X}}} \\
&= \left(6/\gamma^2 + 2\right)^{-1/2} \|(\boldsymbol{\sigma}, u)\|_{\hat{X}} .
\end{aligned}$$

Da $(\boldsymbol{\sigma}, u) \in \hat{X}$ beliebig war, beweist dies die behauptete Inf-Sup-Bedingung mit

$$\hat{\beta} \geq \left(6/\gamma^2 + 2\right)^{-1/2} > 0. \qquad \square$$

Mit dieser Vorarbeit ist es möglich die kontinuierliche Inf-Sup-Bedingung für b zu beweisen.

Theorem 3.8 (Kontinuierliche Inf-Sup-Bedingung für b). *Es gilt für die Bilinearform b definiert in Gleichung* (3.6)

$$0 < \beta = \inf_{x \in S(X)} \sup_{y \in S(Y)} b(x, y).$$

Beweis. Wie bereits angedeutet wird für diesen Beweis das Splitting Lemma, Lemma 3.6, verwendet. Dabei seien

$$\begin{aligned}
X_1 &:= L^2\left(\Omega; \mathbb{R}^{n \times n}\right) / \mathbb{R} \times L^2\left(\Omega; \mathbb{R}^n\right) \times \{0\} \times \{0\} \subseteq X, \\
X_2 &:= \{0\} \times \{0\} \times H_0^{1/2}\left(\partial \mathcal{T}; \mathbb{R}^n\right) \times H^{-1/2}\left(\partial \mathcal{T}; \mathbb{R}^n\right) \subseteq X, \\
Y_1 &:= Y_1 = H\left(\operatorname{div}, \Omega; \mathbb{R}^{n \times n}\right) / \mathbb{R} \times H_0^1\left(\Omega; \mathbb{R}^n\right) \subseteq Y.
\end{aligned}$$

Für diese Räume sind nun die Voraussetzungen von Lemma 3.6 zu prüfen. Es gilt $X = X_1 + X_2$. Für jedes $y_1 := (\boldsymbol{\tau}, v) \in Y_1$ und alle $\boldsymbol{q} \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}\right)$ und $w \in H_0^1(\Omega; \mathbb{R}^n)$, also jedes $x_2 := \left(0, 0, \gamma_0^{\mathcal{T}} w, \gamma_\nu^{\mathcal{T}} \boldsymbol{q}\right) \in X_2$, zeigt sich außerdem mit partieller Integration

$$\begin{aligned} b(x_2, y_1) &= -\left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}, \gamma_0^{\mathcal{T}} w \right\rangle_{\partial\mathcal{T}} - \left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{q}, \gamma_0^{\mathcal{T}} v \right\rangle_{\partial\mathcal{T}} \\ &= -\int_\Omega \boldsymbol{q} : \mathrm{D}\, v \,\mathrm{d}x - \int_\Omega \operatorname{div} \boldsymbol{q} \cdot v \,\mathrm{d}x - \int_\Omega \boldsymbol{\tau} : \mathrm{D}\, w \,\mathrm{d}x - \int_\Omega \operatorname{div} \boldsymbol{\tau} \cdot w \,\mathrm{d}x \\ &= -\int_{\partial\Omega} v \cdot \boldsymbol{q}\nu \,\mathrm{d}s - \int_{\partial\Omega} w \cdot \boldsymbol{\tau}\nu \,\mathrm{d}s \\ &= 0. \end{aligned}$$

Die Bedingung

$$0 < \beta_2 := \inf_{x_2 \in S(X_2)} \sup_{y \in S(Y)} b(x_2, y)$$

wird mit Hilfe von Lemma 2.12 und Lemma 2.13 gezeigt. Für jedes $x_2 = (0, 0, s, t) \in X_2$ gilt

$$\begin{aligned} \|x_2\|_X &= \sqrt{\|s\|^2_{H^{1/2}(\partial\mathcal{T})} + \|t\|^2_{H^{-1/2}(\partial\mathcal{T})}} \\ &\leq \|s\|_{H^{1/2}(\partial\mathcal{T})} + \|t\|_{H^{-1/2}(\partial\mathcal{T})} \\ &= \sup_{\boldsymbol{q} \in S\left(H\left(\operatorname{div},\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R}\right)} \left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{q}, s \right\rangle_{\partial\mathcal{T}} + \sup_{w \in S\left(H^1(\mathcal{T};\mathbb{R}^n)\right)} \left\langle t, \gamma_0^{\mathcal{T}} w \right\rangle_{\partial\mathcal{T}} \\ &\leq \sup_{\substack{\boldsymbol{q} \in S\left(H\left(\operatorname{div},\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R}\right) \\ w \in S\left(H^1(\mathcal{T};\mathbb{R}^n)\right)}} -b(x_2, (\boldsymbol{q}, w)) \\ &= \sup_{\substack{\boldsymbol{q} \in S\left(H\left(\operatorname{div},\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R}\right) \\ w \in S\left(H^1(\mathcal{T};\mathbb{R}^n)\right)}} \frac{-b(x_2, (\boldsymbol{q}, w))\|(\boldsymbol{q}, w)\|_Y}{\|(\boldsymbol{q}, w)\|_Y} \\ &\leq \sqrt{2} \sup_{y \in S(Y)} b(x_2, y), \end{aligned}$$

also gilt die Inf-Sup-Bedingung mit $\beta_2 \geq \sqrt{2}$. Die Bedingung

$$0 < \beta_1 := \inf_{x_1 \in S(X_1)} \sup_{y_1 \in S(Y_1)} b(x_1, y_1)$$

gilt nach Lemma 3.7 mit $\beta_1 = \hat{\beta} \geq \left(6/\gamma^2 + 2\right)^{-1/2}$, wobei γ die Inf-Sup-Konstante aus Bemerkung 2.23 ist. Damit lässt sich die kontinuierliche Inf-Sup-Konstante wie in Lemma 3.6 durch

$$\beta \geq \frac{\beta_1 \beta_2}{\beta_1 + \beta_2 + \|b\|}$$

abschätzen. □

Der folgende Abschnitt widmet sich nach einigen Vorbereitungen dem Nachweis von (B4).

3.3 Diskrete Inf-Sup-Bedingung

Die vorgeschlagenen diskreten Räume lauten

$$X_h := P_0\left(\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R} \times P_0\left(\mathcal{T};\mathbb{R}^n\right) \times S_0^1\left(\mathcal{E};\mathbb{R}^n\right) \times P_0\left(\mathcal{E};\mathbb{R}^n\right),$$
$$Y_h := RT_0^{\mathrm{pw}}\left(\mathcal{T};\mathbb{R}^{n\times n}\right)/\mathbb{R} \times P_1\left(\mathcal{T};\mathbb{R}^n\right).$$

Damit lautet das diskrete zu lösende Problem analog zu Problem 6 wie folgt.

Problem 10 (Praktische dPG-Methode im gegebenen Setting)
Zu gegebenem $f \in L^2\left(\Omega;\mathbb{R}^n\right)$ und $\hat{s} \in H^{1/2}\left(\partial\mathcal{T};\mathbb{R}^n\right)$ ist $x_h \in X_h$ gesucht mit

$$x_h \in \operatorname*{argmin}_{\xi_h \in X_h} \|F - b(\xi_h, \bullet)\|_{Y_h^*},$$

wobei b und F wie in Definition 3.2 definiert sind.

Die diskreten Spuren in $S_0^1\left(\mathcal{E};\mathbb{R}^n\right)$ bzw. $P_0\left(\mathcal{E};\mathbb{R}^n\right)$ können eindeutig zu globalen Funktionen in $S_0^1(\mathcal{T};\mathbb{R}^n)$ bzw. $RT_0(\mathcal{T};\mathbb{R}^{n\times n})$ fortgesetzt werden. Mit Hilfe dieser Fortsetzungen wird in Bemerkung 3.11 eine äquivalenten Bilinearform definiert, die den Beweis der diskreten Inf-Sup-Bedingung vereinfacht. Der Vollständigkeit halber wird in den folgenden Lemmata die Normäquivalenz zwischen den diskreten Spuren und ihren Fortsetzungen bewiesen.

Lemma 3.9. *Die Spur $s_1 := \gamma_0 w_c \in S_0^1\left(\mathcal{E};\mathbb{R}^n\right)$ von einer Funktion $w_c \in S_0^1\left(\mathcal{T};\mathbb{R}^n\right)$ gehört zu $H_0^{1/2}\left(\partial\mathcal{T};\mathbb{R}^n\right)$ und genügt folgender Normäquivalenz*

$$\|s_1\|_{H_0^{1/2}(\partial\mathcal{T})} \approx \|w_c\|_{H_0^1(\Omega)}.$$

Beweis. Es gilt, die nicht triviale Richtung

$$\|w_{\mathrm{c}}\|_{H_0^1(\Omega)} \lesssim \|s_1\|_{H_0^{1/2}(\partial\mathcal{T})} = \inf_{\substack{v\in H_0^1(\Omega,\mathbb{R}^n)\\ \gamma_0 v = s_1}} \|v\|_{H_0^1(\Omega)},$$

zu beweisen. Für alle $v \in H_0^1\left(\Omega;\mathbb{R}^n\right)$ mit Spur $\gamma_0 v = s_1$ und für die elementweise konstante Ableitung $\mathrm{D}\, w_{\mathrm{c}}$ gilt auf jedem n-Simplex $T \in \mathcal{T}$

$$\mathrm{D}\, w_{\mathrm{c}}|_T = |T|^{-1} \int_T \mathrm{D}\, w_{\mathrm{c}}\, \mathrm{d}x = |T|^{-1} \int_{\partial T} \nu_T s_1\, \mathrm{d}s = |T|^{-1} \int_T \mathrm{D}\, v\, \mathrm{d}x.$$

Daher gilt $\mathrm{D}\, w_\mathrm{c} = \Pi_0 \mathrm{D}\, v$ und für die Norm

$$\|\mathrm{D}\, w_\mathrm{c}\|_{L^2(\Omega)} = \|\Pi_0 \mathrm{D}\, v\|_{L^2(\Omega)} \leq \|\mathrm{D}\, v\|_{L^2(\Omega)} \leq \|v\|_{H^1_0(\Omega)}\,.$$

Zusammen mit der Friedrichs Ungleichung aus Lemma 2.16, $\|w_\mathrm{c}\|_{L^2(\Omega)} \leq C_F \|\mathrm{D}\, w_\mathrm{c}\|_{L^2(\Omega)}$, ergibt sich

$$\|w_\mathrm{c}\|_{H^1_0(\Omega)} \leq \sqrt{1 + C_F^2}\ \|v\|_{H^1_0(\Omega)}\,.$$

Da $v \in H^1_0(\Omega; \mathbb{R}^n)$ eine beliebige Fortsetzung ist, beweist dies das Lemma. Genauer gilt mit $C_s := \sqrt{1 + C_F^2}$

$$\|s_1\|_{H^{1/2}_0(\partial\mathcal{T})} \leq \|w_\mathrm{c}\|_{H^1_0(\Omega)} \leq C_s \|s_1\|_{H^{1/2}_0(\partial\mathcal{T})}\,. \tag{3.15}$$

□

Lemma 3.10. *Die Normalenspur $t_0 = \gamma_\nu \boldsymbol{q}_{RT} \in P_0(\mathcal{E}; \mathbb{R}^n)$ der Raviart-Thomas Funktion $\boldsymbol{q}_{RT} \in RT_0\left(\mathcal{T}; \mathbb{R}^{n\times n}\right)$ gehört zu $H^{-1/2}(\partial\mathcal{T}; \mathbb{R}^n)$ und genügt folgender Normäquivalenz*

$$\|t_0\|_{H^{-1/2}(\partial\mathcal{T})} \approx \|\boldsymbol{q}_{RT}\|_{H(\mathrm{div},\mathcal{T})}\,.$$

Beweis. Es gilt, die nicht triviale Richtung

$$\|\boldsymbol{q}_\mathrm{RT}\|_{H(\mathrm{div},\mathcal{T})} \lesssim \|t_0\|_{H^{-1/2}(\partial\mathcal{T})} = \inf_{\substack{\boldsymbol{\rho}\in H(\mathrm{div},\mathcal{T};\mathbb{R}^{n\times n}) \\ \gamma_\nu\boldsymbol{\rho}=t_0}} \|\boldsymbol{\rho}\|_{H(\mathrm{div},\mathcal{T})}\,,$$

zu beweisen. Für alle $\boldsymbol{\rho} \in H(\mathrm{div}, \mathcal{T}; \mathbb{R}^n)$ mit Normalenspur $\gamma_\nu\boldsymbol{\rho} = t_0$ und die elementweise konstante Divergenz $\mathrm{div}\, \boldsymbol{q}_\mathrm{RT}$ gilt auf jedem n-Simplex $T \in \mathcal{T}$

$$\mathrm{div}\, \boldsymbol{q}_\mathrm{RT}|_T = |T|^{-1} \int_T \mathrm{div}\, \boldsymbol{q}_\mathrm{RT}\ \mathrm{d}x = |T|^{-1} \int_{\partial T} t\ \mathrm{d}s = |T|^{-1} \int_T \mathrm{div}\, \boldsymbol{\rho}\ \mathrm{d}x.$$

Also ist $\mathrm{div}\, \boldsymbol{q}_\mathrm{RT} = \Pi_0 \mathrm{div}\, \boldsymbol{\rho}$ und für die Norm gilt

$$\|\mathrm{div}\, \boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)} = \|\Pi_0 \mathrm{div}\, \boldsymbol{\rho}\|_{L^2(\Omega)} \leq \|\mathrm{div}\, \boldsymbol{\rho}\|_{L^2(\Omega)} \leq \|\boldsymbol{\rho}\|_{H(\mathrm{div},\mathcal{T})}\,. \tag{3.16}$$

Außerdem gilt für jedes $x \in T \in \mathcal{T}$

$$\begin{aligned}\boldsymbol{q}_\mathrm{RT}|_T(x) &= \Pi_0 \boldsymbol{q}_\mathrm{RT} + 1/n\ \mathrm{div}\, \boldsymbol{q}_\mathrm{RT} \otimes (x - \mathrm{mid}(T)) \\ &= \mathrm{D}\left(\Pi_0 \boldsymbol{q}_\mathrm{RT} x + 1/(2n)\ \mathrm{div}\, \boldsymbol{q}_\mathrm{RT}\ |x - \mathrm{mid}(T)|^2\right).\end{aligned}$$

Daher existiert für alle $T \in \mathcal{T}$ ein $g_T \in H^1(T; \mathbb{R}^n)$, so dass $\boldsymbol{q}_\mathrm{RT} = \mathrm{D}\, g_T$ und per Normierung $\int_T g_T\ \mathrm{d}x = 0$. Mit Hilfe zweier partieller Integrationen und der Poincaré Ungleichung aus Lemma 2.17 mit Konstante $C_P \leq \mathrm{diam}(\Omega)/\pi$, zeigt sich

$$\begin{aligned}\|\mathrm{D}\, g_T\|^2_{L^2(T)} &= \int_T \mathrm{D}\, g_T : \boldsymbol{q}_\mathrm{RT}\ \mathrm{d}x \\ &= -\int_T g_T \cdot \mathrm{div}\, \boldsymbol{q}_\mathrm{RT}\ \mathrm{d}x + \int_{\partial T} g_T \cdot t_0\ \mathrm{d}s\end{aligned}$$

$$
\begin{aligned}
&= -\int_T g_T \cdot \Pi_0 \operatorname{div} \boldsymbol{\rho} \, \mathrm{d}x + \int_{\partial T} g_T \cdot \gamma_\nu \boldsymbol{\rho} \, \mathrm{d}s \\
&= \int_T g_T \cdot (1 - \Pi_0) \operatorname{div} \boldsymbol{\rho} \, \mathrm{d}x + \int_T \boldsymbol{\rho} : \mathrm{D}\, g_T \, \mathrm{d}x \\
&\leq C_P \|\mathrm{D}\, g_T\|_{L^2(T)} \|(1 - \Pi_0) \operatorname{div} \boldsymbol{\rho}\|_{L^2(T)} + \|\mathrm{D}\, g_T\|_{L^2(T)} \|\boldsymbol{\rho}\|_{L^2(T)} .
\end{aligned}
$$

Daher gilt

$$
\|\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(T)} = \|\mathrm{D}\, g_T\|_{L^2(T)} \leq C_P \|(1 - \Pi_0) \operatorname{div} \boldsymbol{\rho}\|_{L^2(T)} + \|\boldsymbol{\rho}\|_{L^2(T)} .
$$

Das Aufsummieren über alle Dreiecke und (3.16) zeigen

$$
\begin{aligned}
\|\boldsymbol{q}_{\mathrm{RT}}\|^2_{H(\operatorname{div},\mathcal{T})} &= \|\boldsymbol{q}_{\mathrm{RT}}\|^2_{L^2(\Omega)} + \|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|^2_{L^2(\Omega)} \\
&\leq \|\operatorname{div} \boldsymbol{\rho}\|^2_{L^2(\Omega)} + \left(C_P \|\operatorname{div} \boldsymbol{\rho}\|_{L^2(\Omega)} + \|\boldsymbol{\rho}\|_{L^2(\Omega)}\right)^2 \\
&\leq \min_{\lambda \in \mathbb{R}^+} \left\{ (1 + \lambda)\, C_P^2 \|\operatorname{div} \boldsymbol{\rho}\|^2_{L^2(\Omega)} + (1 + 1/\lambda) \|\boldsymbol{\rho}\|^2_{L^2(\Omega)} \right\} \\
&\leq \min_{\lambda \in \mathbb{R}^+} \max \left\{ 1 + (1 + \lambda)\, C_P^2,\, 1 + 1/\lambda \right\} \|\boldsymbol{\rho}\|^2_{H(\operatorname{div},\mathcal{T})} \\
&\leq C_t^2 \|\boldsymbol{\rho}\|^2_{H(\operatorname{div},\mathcal{T})} ,
\end{aligned}
$$

wobei die Wahl $\lambda = 1/2 \left(\sqrt{C_P^2 + 4}/C_P - 1 \right)$, so dass $1 + (1 + \lambda)\, C_P^2 = (1 + 1/\lambda)$, die Konstante

$$
C_t = \sqrt{1 + \left(C_P + \sqrt{C_P^2 + 4} \right) C_P / 2}
$$

ergibt. Da $\boldsymbol{\rho} \in H\left(\operatorname{div}, \mathcal{T}; \mathbb{R}^{n \times n}\right)$ eine beliebige Fortsetzung war beweist dies das Lemma und insbesondere gilt

$$
\|t_0\|_{H^{-1/2}(\partial\mathcal{T})} \leq \|\boldsymbol{q}_{\mathrm{RT}}\|_{H(\operatorname{div},\mathcal{T})} \leq C_t \|t_0\|_{H^{-1/2}(\partial\mathcal{T})} . \tag{3.17}
$$

□

Nun folgt wie angekündigt die Definition einer alternativen Bilinearform.

Bemerkung 3.11 (Alternative Bilinearform) Die Fortsetzung der Spurterme $s_1 \in S^1_0(\mathcal{E}; \mathbb{R}^n)$ bzw. $t_0 \in P_0(\mathcal{E}; \mathbb{R}^n)$ durch $w_{\mathrm{c}} \in S^1_0(\mathcal{T}; \mathbb{R}^n)$ bzw. $\boldsymbol{q}_{\mathrm{RT}} \in RT_0\left(\mathcal{T}; \mathbb{R}^{n \times n}\right)$ erlaubt eine weitere partielle Integration, die zu folgender alternativer Bilinearform führt

$$
\begin{aligned}
\hat{b}(\hat{x}_h, y_h) = &\int_\Omega \boldsymbol{\sigma}_0 : \mathrm{D}_{\mathrm{NC}}\, v_1 \, \mathrm{d}x + \int_\Omega \operatorname{dev} \boldsymbol{\sigma}_0 : \boldsymbol{\tau}_{\mathrm{RT}} \, \mathrm{d}x + \int_\Omega u_0 \cdot \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}_{\mathrm{RT}} \, \mathrm{d}x \\
&- \int_\Omega \left(v_1 \cdot \operatorname{div} \boldsymbol{q}_{\mathrm{RT}} + \boldsymbol{q}_{\mathrm{RT}} : \mathrm{D}_{\mathrm{NC}}\, v_1 \right) \mathrm{d}x - \int_\Omega \left(\boldsymbol{\tau}_{\mathrm{RT}} : \mathrm{D}\, w_{\mathrm{c}} + w_{\mathrm{c}} \cdot \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}_{\mathrm{RT}} \right) \mathrm{d}x.
\end{aligned}
$$

Hier ist der Testraum unverändert, d.h. $y_h \in Y_h$ wie in (2.12), und die Ansatzfunktion $\hat{x}_h = (\boldsymbol{\sigma}_0, u_0, w_{\mathrm{c}}, \boldsymbol{q}_{\mathrm{RT}}) \in \hat{X}_h$ stammt aus dem folgenden Raum

$$
\hat{X}_h := P_0\left(\mathcal{T}; \mathbb{R}^{n \times n}\right) / \mathbb{R} \times P_0(\mathcal{T}; \mathbb{R}^n) \times S^1_0(\mathcal{T}, \mathbb{R}^n) \times RT_0\left(\mathcal{T}; \mathbb{R}^{n \times n}\right) . \tag{3.18}
$$

Dieser Raum sei mit der Norm des kontinuierlichen Äquivalents $\hat{X} := L^2\left(\Omega : \mathbb{R}^{n\times n}\right)/\mathbb{R} \times L^2(\Omega;\mathbb{R}^n) \times H_0^1(\Omega;\mathbb{R}^n) \times H\left(\operatorname{div},\Omega;\mathbb{R}^{n\times n}\right)$ versehen,

$$\|x_h\|_{\hat{X}}^2 = \|\boldsymbol{\sigma}_0\|_{L^2(\Omega)}^2 + \|u_0\|_{L^2(\Omega)}^2 + \|w_{\mathrm{c}}\|_{H_0^1(\Omega)}^2 + \|\boldsymbol{q}_{\mathrm{RT}}\|_{H(\operatorname{div},\Omega)}^2 . \tag{3.19}$$

Die diskrete Inf-Sup-Bedingung wird nun mit Hilfe des Tr-dev-div Lemmas, Lemma 2.18, und der Zerlegung aus Lemma 2.21, aus Paragraph 2.5, zunächst für die alternative Bilinearform aus Bemerkung 3.11 bewiesen. Dabei kann die Inf-Sup-Konstante direkt berechnet werden.

Theorem 3.12. *In den diskreten Räumen aus (2.11)–(2.12), erfüllt die Bilinearform b (3.6) die Inf-Sup-Bedingung*

$$1 \lesssim \beta_h := \inf_{x_h \in X_h \setminus \{0\}} \sup_{y_h \in Y_h} \frac{b(x_h, y_h)}{\|y_h\|_Y \|x_h\|_X}.$$

Die Konstante β_h wird im Beweis berechnet.

Beweis. Zunächst wird wie angekündigt mit der alternativen Bilinearform aus Bemerkung 3.11 gearbeitet.

Schritt 1. Diskrete Testfunktionen. Zu einem gegebenen $\hat{x}_h = (\boldsymbol{\sigma}_0, u_0, w_{\mathrm{c}}, \boldsymbol{q}_{\mathrm{RT}}) \in \hat{X}_h \setminus \{0\}$ wie in (3.18) wird folgendes $\tilde{y}_h = (\tilde{\boldsymbol{\tau}}_{\mathrm{RT}}, \tilde{v}_1)$ gewählt

$$\begin{aligned} \tilde{\boldsymbol{\tau}}_{\mathrm{RT}} &:= \operatorname{dev} \boldsymbol{\sigma}_0 - \mathrm{D}\, w_{\mathrm{c}} + 1/n \;\; (u_0 - \Pi_0 w_{\mathrm{c}}) \otimes (\bullet - \operatorname{mid}(T)) , \\ \tilde{v}_1 &:= -\operatorname{div} \boldsymbol{q}_{\mathrm{RT}} + (\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}) (\bullet - \operatorname{mid}(T)) . \end{aligned}$$

Dies ist eine zulässige Testfunktion, denn offensichtlich liegen $\tilde{\boldsymbol{\tau}}_{\mathrm{RT}} \in RT_0^{\mathrm{pw}}(\mathcal{T};\mathbb{R}^{n\times n})$ und $\tilde{v}_1 \in P_1(\mathcal{T};\mathbb{R}^n)$ und die globale Nebenbedingung $\int_\Omega \operatorname{tr} \tilde{\boldsymbol{\tau}}_{\mathrm{RT}} \,\mathrm{d}x = 0$ ist erfüllt. Letzteres gilt, da $\operatorname{dev} \boldsymbol{\sigma}_0$ per Definition spurfrei ist, $u_0 - \Pi_0 w_{\mathrm{c}}$ elementweise konstant und nach (2.2) $\int_T \bullet - \operatorname{mid}(T) \,\mathrm{d}x = 0$ sowie $\int_\Omega \operatorname{div} w_{\mathrm{c}} \,\mathrm{d}x = \int_{\partial\Omega} w_{\mathrm{c}} \cdot \nu \,\mathrm{d}s = 0$. Außerdem gilt, $\operatorname{div}_{\mathrm{NC}} \tilde{\boldsymbol{\tau}}_{\mathrm{RT}} = u_0 - \Pi_0 w_{\mathrm{c}}$, $\mathrm{D}_{\mathrm{NC}}\, \tilde{v}_1 = \boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}$ und $\Pi_0 \tilde{v}_1 = -\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}$.

Schritt 2. Es gilt

$$\begin{aligned} \hat{b}(\hat{x}_h, \tilde{y}_h) = &\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 + \|\operatorname{dev} \boldsymbol{\sigma}_0 - \mathrm{D}\, w_{\mathrm{c}}\|_{L^2(\Omega)}^2 \\ &+ \|u_0 - \Pi_0 w_{\mathrm{c}}\|_{L^2(\Omega)}^2 + \|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 . \end{aligned}$$

Beweis von Schritt 2. Die Bilinearform $\hat{b}$ mit $\hat{x}_h = (\boldsymbol{\sigma}_0, u_0, w_{\mathrm{c}}, \boldsymbol{q}_{\mathrm{RT}})$ und $y_h = (\boldsymbol{\tau}_{\mathrm{RT}}, v_1)$ lautet

$$\begin{aligned} \hat{b}(\hat{x}_h, y_h) = &\int_\Omega (u_0 - w_{\mathrm{c}}) \cdot \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}_{\mathrm{RT}} \,\mathrm{d}x + \int_\Omega (\operatorname{dev} \boldsymbol{\sigma}_0 - \mathrm{D}\, w_{\mathrm{c}}) : \boldsymbol{\tau}_{\mathrm{RT}} \,\mathrm{d}x \\ &+ \int_\Omega (\boldsymbol{\sigma}_0 - \boldsymbol{q}_{\mathrm{RT}}) : \mathrm{D}_{\mathrm{NC}}\, v_1 \,\mathrm{d}x - \int_\Omega v_1 \cdot \operatorname{div} \boldsymbol{q}_{\mathrm{RT}} \,\mathrm{d}x. \end{aligned}$$

Das Einsetzen der in *Schritt 1.* zu $\hat{x}_h \in \hat{X}_h$ gewählten Testfunktion $\tilde{y}_h \in Y_h$ ergibt nach obigen Überlegungen zu $\operatorname{div}_{\mathrm{NC}} \tilde{\boldsymbol{\tau}}_{\mathrm{RT}}$, $\mathrm{D}_{\mathrm{NC}} \tilde{v}_1$ und $\Pi_0 \tilde{v}_1$

$$\begin{aligned}\hat{b}\left(\hat{x}_h, \tilde{y}_h\right) = &\int_\Omega \left(u_0 - w_{\mathrm{c}}\right) \cdot \left(u_0 - \Pi_0 w_{\mathrm{c}}\right) \,\mathrm{d}x + \int_\Omega \left(\boldsymbol{\sigma}_0 - \boldsymbol{q}_{\mathrm{RT}}\right) : \left(\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\right) \,\mathrm{d}x \\ &+ \int_\Omega (\operatorname{dev} \boldsymbol{\sigma}_0 - \mathrm{D}\, w_{\mathrm{c}}) : (\operatorname{dev} \boldsymbol{\sigma}_0 - \mathrm{D}\, w_{\mathrm{c}} + (u_0 - \Pi_0 w_{\mathrm{c}}) \otimes (\bullet - \operatorname{mid}(T))/n) \,\mathrm{d}x \\ &- \int_\Omega \Big(- \operatorname{div} \boldsymbol{q}_{\mathrm{RT}} + \left(\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\right) (\bullet - \operatorname{mid}(T)) \Big) \cdot \operatorname{div} \boldsymbol{q}_{\mathrm{RT}} \,\mathrm{d}x.\end{aligned}$$

Daraus folgt mit (2.10) und (2.2), genauer weil $(\bullet - \operatorname{mid}(T))$ senkrecht zu den elementweise konstanten Funktionen $P_0(\mathcal{T})$, wie $\operatorname{dev} \boldsymbol{\sigma}_0$, $\mathrm{D}\, w_{\mathrm{c}}$ und $\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}$ steht, sofort die behauptete Darstellung von $\hat{b}$. □

Um die diskrete Inf-Sup-Bedingung zu beweisen werden nun $\|\hat{x}_h\|_{\hat{X}}$ und $\|\tilde{y}_h\|_Y$ gegen die Bilinearform $\hat{b}\left(\hat{x}_h, \tilde{y}_h\right)$ in der soeben entwickelten Form abgeschätzt.

Schritt 3. Es gilt $\|\tilde{y}_h\|_Y^2 \leq \left(1 + h_{\max}^2 n^2/(n+1)^2\right) \hat{b}\left(\hat{x}_h, \tilde{y}_h\right)$.

Beweis von Schritt 3. Die Gesetzmäßigkeiten aus (2.2) sowie die Gleichungen für $\operatorname{div}_{\mathrm{NC}} \tilde{\boldsymbol{\tau}}_{\mathrm{RT}}$ und $\mathrm{D}_{\mathrm{NC}} \tilde{v}_1$ ergeben

$$\begin{aligned}\|\tilde{y}_h\|_Y^2 =& \|\tilde{\boldsymbol{\tau}}_{\mathrm{RT}}\|_{H(\operatorname{div},\mathcal{T})}^2 + \|\tilde{v}_1\|_{H^1(\mathcal{T})}^2 \\ =& \left\|\operatorname{dev} \boldsymbol{\sigma}_0 - \mathrm{D}\, w_{\mathrm{c}} + 1/n \;\; (u_0 - \Pi_0 w_{\mathrm{c}}) \otimes (\bullet - \operatorname{mid}(T))\right\|_{L^2(\Omega)}^2 + \|u_0 - \Pi_0 w_{\mathrm{c}}\|_{L^2(\Omega)}^2 \\ &+ \left\| - \operatorname{div} \boldsymbol{q}_{\mathrm{RT}} + \left(\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\right) (\bullet - \operatorname{mid}(T))\right\|_{L^2(\Omega)}^2 + \|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 \\ \leq& \|\operatorname{dev} \boldsymbol{\sigma}_0 - \mathrm{D}\, w_{\mathrm{c}}\|_{L^2(\Omega)}^2 + \left(1/n^2 \|\bullet - \operatorname{mid}(T)\|_{L^\infty(\Omega)}^2 + 1\right) \|u_0 - \Pi_0 w_{\mathrm{c}}\|_{L^2(\Omega)}^2 \\ &+ \|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 + \left(\|\bullet - \operatorname{mid}(T)\|_{L^\infty(\Omega)}^2 + 1\right) \|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 \\ \leq& \|\operatorname{dev} \boldsymbol{\sigma}_0 - \mathrm{D}\, w_{\mathrm{c}}\|_{L^2(\Omega)}^2 + \left(h_{\max}^2/(n+1)^2 + 1\right) \|u_0 - \Pi_0 w_{\mathrm{c}}\|_{L^2(\Omega)}^2 \\ &+ \|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 + \left(h_{\max}^2 n^2/(n+1)^2 + 1\right) \|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 .\end{aligned}$$

Zusammen mit der Gleichheit aus *Schritt 2.* folgt die Behauptung. □

Der zentrale Schritt, um eine ähnliche Abschätzung für $\|\hat{x}_h\|_{\hat{X}}$ herzuleiten, ist die ein wenig technische Abschätzung von $|\!|\!| w_{\mathrm{c}} |\!|\!| := \|\mathrm{D}\, w_{\mathrm{c}}\|_{L^2(\Omega)}$ gegen $\hat{b}\left(\hat{x}_h, \tilde{y}_h\right)$. Dabei sind die Zerlegung aus Lemma 2.21, und das Tr-Div-Dev Lemma, Lemma 2.18, entscheidend.

Schritt 4. Es gilt $|\!|\!| w_{\mathrm{c}} |\!|\!|^2 \leq C_{w_{\mathrm{c}}} \hat{b}\left(\hat{x}_h, \tilde{y}_h\right)$.

Beweis von Schritt 4. Es wird die Zerlegung

$$|\!|\!| w_{\mathrm{c}} |\!|\!|^2 \leq \min_{\lambda \in \mathbb{R}^+} \left\{ (1+\lambda) |\!|\!| w_{\mathrm{c}} + z |\!|\!|^2 + (1 + 1/\lambda) |\!|\!| z |\!|\!|^2 \right\} \tag{3.20}$$

betrachtet. Die Kontrolle beider Terme auf der rechten Seite durch $B^2 := \hat{b}\left(\hat{x}_h, \tilde{y}_h\right)$ gelingt, wenn $z \in H_0^1\left(\Omega; \mathbb{R}^n\right)$ divergenzfrei und $b \in L^2\left(\Omega; \mathbb{R}^{n\times n}\right)$ so gewählt werden, dass

$$\operatorname{dev}\left(\boldsymbol{\sigma}_0 - \mathrm{D}\, w_\mathrm{c}\right) = \mathrm{D}\, z + \operatorname{dev} b \tag{3.21}$$

eine orthogonale Zerlegung ist. Eine solche Wahl ist nach Lemma 2.21 möglich. Die Normgleichheit aus Lemma 2.5 Punkt (iii) ermöglicht mit dieser orthogonalen Zerlegung folgende Abschätzung

$$\begin{aligned}\|\operatorname{dev}\boldsymbol{\sigma}_0 - \mathrm{D}\, w_\mathrm{c}\|_{L^2(\Omega)}^2 =& 1/n\|\operatorname{div} w_\mathrm{c}\|_{L^2(\Omega)}^2 + \|\operatorname{dev}\left(\mathrm{D}\, w_\mathrm{c} - \operatorname{dev}\boldsymbol{\sigma}_0\right)\|_{L^2(\Omega)}^2\\ =& 1/n\|\operatorname{div} w_\mathrm{c}\|_{L^2(\Omega)}^2 + |\!|\!| z |\!|\!|^2 + \|\operatorname{dev} b\|_{L^2(\Omega)}^2 .\end{aligned}$$

Aus *Schritt 2.* folgt die Beschränkung der Summe und damit aller Summanden nach oben durch $\|\operatorname{dev}\boldsymbol{\sigma}_0 - \mathrm{D}\, w_\mathrm{c}\|_{L^2(\Omega)}^2 \leq B^2$. Der eine Summand in Gleichung (3.20) ist also bereits kontrolliert. Lemma 2.5 Punkt (iii) und die Divergenzfreiheit von $z \in H_0^1\left(\Omega; \mathbb{R}^n\right)$ ermöglichen folgende Zerlegung des zweiten Summanden

$$\begin{aligned}A^2 := |\!|\!| w_\mathrm{c} + z |\!|\!|^2 =& \|\operatorname{dev}\mathrm{D}\left(w_\mathrm{c} + z\right)\|_{L^2(\Omega)}^2 + 1/n\|\operatorname{div} w_\mathrm{c}\|_{L^2(\Omega)}^2\\ \leq& \|\operatorname{dev}\mathrm{D}\left(w_\mathrm{c} + z\right)\|_{L^2(\Omega)}^2 + \|\operatorname{dev}\boldsymbol{\sigma}_0 - \mathrm{D}\, w_\mathrm{c}\|_{L^2(\Omega)}^2\\ \leq& \|\operatorname{dev}\mathrm{D}\left(w_\mathrm{c} + z\right)\|_{L^2(\Omega)}^2 + B^2\end{aligned} \tag{3.22}$$

Die Rechenregel für den deviatorischen Anteil aus Lemma 2.5 sowie Gleichung (3.21) erlauben

$$\begin{aligned}\|\operatorname{dev}\mathrm{D}\left(w_\mathrm{c} + z\right)\|_{L^2(\Omega)}^2 &= \int_\Omega \left(\operatorname{dev}\mathrm{D}\left(w_\mathrm{c} + z\right)\right) : \mathrm{D}\left(w_\mathrm{c} + z\right)\,\mathrm{d}x\\ &= \int_\Omega \left(\operatorname{dev}\boldsymbol{\sigma}_0 - \operatorname{dev} b\right) : \mathrm{D}\left(w_\mathrm{c} + z\right)\,\mathrm{d}x.\end{aligned}$$

Die Cauchy-Schwarz-Ungleichung, die Definition und Rechenregeln für den Deviator und die Divergenzfreiheit von z ermöglichen weiter

$$\begin{aligned}\|\operatorname{dev}\mathrm{D}\left(w_\mathrm{c} + z\right)\|_{L^2(\Omega)}^2 \leq& \|\operatorname{dev} b\|_{L^2(\Omega)}\, |\!|\!| w_\mathrm{c} + z |\!|\!| + \int_\Omega \operatorname{dev}\left(\boldsymbol{\sigma}_0 - \boldsymbol{q}_\mathrm{RT}\right) : \mathrm{D}\left(w_\mathrm{c} + z\right)\,\mathrm{d}x\\ &+ \int_\Omega \boldsymbol{q}_\mathrm{RT} : \left(\mathrm{D}\left(w_\mathrm{c} + z\right) - \operatorname{div} w_\mathrm{c}/n\ \mathrm{I}_{n\times n}\right)\,\mathrm{d}x\\ \leq& \left(\|\operatorname{dev} b\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}_0 - \boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}\right) |\!|\!| w_\mathrm{c} + z |\!|\!|\\ &+ \int_\Omega \boldsymbol{q}_\mathrm{RT} : \mathrm{D}\left(w_\mathrm{c} + z\right)\,\mathrm{d}x - \frac{1}{n}\int_\Omega \operatorname{tr}\boldsymbol{q}_\mathrm{RT}\operatorname{div} w_\mathrm{c}\,\mathrm{d}x.\end{aligned} \tag{3.23}$$

Eine stückweise partielle Integration ergibt

$$\int_\Omega \boldsymbol{q}_\mathrm{RT} : \mathrm{D}\left(w_\mathrm{c} + z\right)\,\mathrm{d}x = -\int_\Omega \left(w_\mathrm{c} + z\right)\cdot\operatorname{div}\boldsymbol{q}_\mathrm{RT}\,\mathrm{d}x,$$

da die normalen Komponenten der Raviart-Thomas Funktionen stückweise konstant

und stetig sind, während die Sprünge der stetigen Funktionen $z \in H_0^1(\Omega;\mathbb{R}^n)$ und $w_c \in S_0^1(\mathcal{T};\mathbb{R}^n)$ verschwinden und somit keine Randintegrale auftreten. Da außerdem $\int_\Omega \operatorname{div} w_c \,\mathrm{d}x = 0$ gilt, kann nach Bemerkung 2.20 (iii) $\boldsymbol{q}_{\mathrm{RT}}$ im letzten Term von (3.23) durch

$$\tilde{\boldsymbol{q}}_{\mathrm{RT}} := \boldsymbol{q}_{\mathrm{RT}} - 1/n \left(\fint_\Omega \operatorname{tr} \boldsymbol{q}_{\mathrm{RT}} \,\mathrm{d}x\right) \mathrm{I}_{n\times n} \in H\left(\operatorname{div},\Omega;\mathbb{R}^{n\times n}\right)/\mathbb{R} \tag{3.24}$$

ersetzt werden. Durch Einsetzen dieser Erkenntnisse in die Formel für $\|\operatorname{dev} \mathrm{D}(w_c+z)\|_{L^2(\Omega)}^2$, ergibt sich unter Verwendung der Cauchy-Schwarz und der Friedrichsungleichung mit Konstante $C_F \le \operatorname{width}(\Omega)/\pi$ aus Lemma 2.16

$$\begin{aligned}
\|\operatorname{dev} \mathrm{D}(w_c+z)\|_{L^2(\Omega)}^2 \le& 1/n \;\|\operatorname{tr} \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} \|\operatorname{div} w_c\|_{L^2(\Omega)} + |\!|\!| w_c + z |\!|\!| \\
&\times \left(\|\operatorname{dev} b\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}_0 - \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} + C_F \|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}\right) \\
\le& 1/n \;\|\operatorname{tr} \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} \|\operatorname{div} w_c\|_{L^2(\Omega)} + |\!|\!| w_c + z |\!|\!| \\
&\times \left(\|\operatorname{dev} \boldsymbol{\sigma}_0 - \mathrm{D}\, w_c\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}_0 - \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} + C_F \|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}\right) \\
\le& 1/\sqrt{n} \;\|\operatorname{tr} \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} B + A\left(\sqrt{1+C_F^2}\; B + \|\boldsymbol{\sigma}_0 - \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}\right).
\end{aligned}$$

Der Faktor $1/\sqrt{n} = \sqrt{n}/n$ vor $\|\operatorname{tr} \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} B$ ergibt sich, da $1/n\|\operatorname{div} w_c\|_{L^2(\Omega)}^2 \le B^2$. Die Konstante $\sqrt{1+C_F^2}$ ergibt sich aus Bemerkung 2.22 mit $a = \|\operatorname{dev} \boldsymbol{\sigma}_0 - \mathrm{D}\, w_c\|_{L^2(\Omega)}$, $b = \|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}$, $C_1 = K_1 = K_2 = 1$ und $C_2 = C_F$. Die Eigenschaft (2.9) für Raviart-Thomas Funktionen und (2.2) erlauben weiter die Abschätzung

$$\begin{aligned}
\|\boldsymbol{\sigma}_0 - \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 &= \|(1-\Pi_0)\,\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 + \|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 \\
&= 1/n^2 \left\|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\,(\bullet - \operatorname{mid}(T))\right\|_{L^2(\Omega)}^2 + \|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 \\
&\le h_{\max}^2/(n+1)^2 \;\|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 + \|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 \\
&\le \max\left\{1, h_{\max}^2/(n+1)^2\right\} B^2
\end{aligned} \tag{3.25}$$

Das in (2.20) definierte $\tilde{\boldsymbol{q}}_{\mathrm{RT}}$ erfüllt die Voraussetzungen des Tr-Div-Dev Lemmas, Lemma 2.18. Zusammen mit der Dreiecksungleichung und Bemerkung 2.20 (i) ergibt sich

$$\begin{aligned}
\frac{1}{C_{\mathrm{tdd}}} \|\operatorname{tr} \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} &\le \|\operatorname{dev} \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} + \|\operatorname{div} \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} \\
&= \|\operatorname{dev} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} + \|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} \\
&\le \|\operatorname{dev}(\boldsymbol{\sigma}_0 - \boldsymbol{q}_{\mathrm{RT}})\|_{L^2(\Omega)} + \|\operatorname{dev} \boldsymbol{\sigma}_0\|_{L^2(\Omega)} + \|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} \\
&\le \max\{1, h_{\max}/(n+1)\}\, B + \|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} + \|\operatorname{dev} \boldsymbol{\sigma}_0\|_{L^2(\Omega)} \\
&\le \left(1 + \max\{1, h_{\max}/(n+1)\}\right) B + \|\operatorname{dev} \boldsymbol{\sigma}_0\|_{L^2(\Omega)}.
\end{aligned}$$

Eine erneute Anwendung der Zerlegung (3.21) erlaubt gemeinsam mit der Dreiecksun-

gleichung und der Definition des Deviators, die Abschätzung des letzten Terms

$$\begin{aligned}\|\operatorname{dev}\boldsymbol{\sigma}_0\|_{L^2(\Omega)} &\leq \|\mathrm{D}\,(z+w_\mathrm{c}) - 1/n\,\operatorname{div}(w_\mathrm{c})\mathrm{I}_{n\times n} + \operatorname{dev} b\|_{L^2(\Omega)}\\ &\leq \|\!|\!| z+w_\mathrm{c} |\!|\!| + \sqrt{1/n}\ \|\operatorname{div} w_\mathrm{c}\|_{L^2(\Omega)} + \|\operatorname{dev} b\|_{L^2(\Omega)}\\ &\leq A + \sqrt{2}\|\operatorname{dev}\boldsymbol{\sigma}_0 - \mathrm{D}\,w_\mathrm{c}\|_{L^2(\Omega)}\\ &\leq A + \sqrt{2}B.\end{aligned}$$

Der Faktor $\sqrt{2}$ ergibt sich wie in Bemerkung 2.22 mit $a = \|\operatorname{div} w_\mathrm{c}\|_{L^2(\Omega)}$, $b = \|\operatorname{dev} b\|_{L^2(\Omega)}$, $C_1 = \sqrt{1/n}$, $K_1 = 1/n$ und $C_2 = K_2 = 1$. Insgesamt hat sich damit aus der Ungleichung (3.22) Folgendes ergeben

$$\begin{aligned}A^2 \leq& B^2 + A\left(B\sqrt{1+C_F^2} + \|\boldsymbol{\sigma}_0 - \boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}\right) + B/\sqrt{n}\ \|\operatorname{tr}\tilde{\boldsymbol{q}}_\mathrm{RT}\|_{L^2(\Omega)}\\ \leq& B^2 + AB\left(\sqrt{1+C_F^2} + \max\{1, h_\mathrm{max}/(n+1)\}\right)\\ &+ BC_\mathrm{tdd}/\sqrt{n}\left(B\left(1+\max\{1,h_\mathrm{max}/(n+1)\}\right) + \|\operatorname{dev}\boldsymbol{\sigma}_0\|_{L^2(\Omega)}\right)\\ \leq& B^2\left(1 + C_\mathrm{tdd}/\sqrt{n}\ \left(1+\max\{1,h_\mathrm{max}/(n+1)\}\right)\right) + BC_\mathrm{tdd}/\sqrt{n}\left(A + B\sqrt{2}\right)\\ &+ AB\left(\sqrt{1+C_F^2} + \max\{1, h_\mathrm{max}/(n+1)\}\right)\\ \leq& B^2\left(1 + C_\mathrm{tdd}/\sqrt{n}\ \left(1+\sqrt{2}+\max\{1,h_\mathrm{max}/(n+1)\}\right)\right)\\ &+ AB\left(\sqrt{1+C_F^2} + C_\mathrm{tdd}/\sqrt{n} + \max\{1, h_\mathrm{max}/(n+1)\}\right).\end{aligned}$$

Mit $C_1 := 1 + C_\mathrm{tdd}/\sqrt{n}\ \left(1+\sqrt{2}+\max\{1,h_\mathrm{max}/(n+1)\}\right)$ und $C_2 := \sqrt{1+C_F^2} + C_\mathrm{tdd}/\sqrt{n} + \max\{1, h_\mathrm{max}/(n+1)\}$ liest sich dies

$$A^2 \leq C_1 B^2 + C_2 AB.$$

Aus dieser Ungleichung für A^2 ergibt sich

$$0 \leq (A - C_2 B)^2 = A^2 - 2C_2 AB + C_2^2 B^2 \leq C_1 B^2 - C_2 AB + C_2^2 B^2$$

und damit $A^2 \leq C_1 B^2 + C_2 AB \leq 2C_1 B^2 + C_2^2 B^2 \leq (2C_1 + C_2^2)B^2$. Also garantiert die Wahl $C_{w_\mathrm{c}} = \min_{\lambda\in\mathbb{R}^+}\left\{(1+\lambda)(2C_1+C_2^2) + (1+1/\lambda)\right\}$

$$\|\!|\!| w_\mathrm{c} |\!|\!|^2 \leq C_{w_\mathrm{c}}\hat{b}\,(\hat{x}_h, \tilde{y}_h)\,.$$

$(1+\lambda)(2C_1+C_2^2) + (1+1/\lambda)$ wird durch $\lambda = 1/\sqrt{2C_1+C_2^2}$ minimiert, dies ergibt $C_{w_\mathrm{c}} = \left(1+\sqrt{2C_1+C_2^2}\right)^2$. □

Schritt 5. Es gilt $\|\hat{x}_h\|_{\hat{X}}^2 \lesssim \hat{b}\,(\hat{x}_h, \tilde{y}_h)$.

Beweis von Schritt 5. Hier werden die einzelnen Teile von $\|\hat{x}_h\|_{\hat{X}}^2$ gegen die Terme aus $\hat{b}\,(\hat{x}_h, y_h)$ in der Form von *Schritt 2.* abgeschätzt. Zunächst erlauben die Friedrichsungleichung 2.16 mit Konstant $C_F \leq \operatorname{width}(\Omega)/\pi$ und die Dreiecksungleichung folgende Abschätzung

$$\begin{aligned}
\|\hat{x}_h\|_{\hat{X}}^2 =& \|\boldsymbol{\sigma}_0\|_{L^2(\Omega)}^2 + \|u_0\|_{L^2(\Omega)}^2 + \|\boldsymbol{q}_{\mathrm{RT}}\|_{H(\mathrm{div},\Omega)}^2 + \|w_{\mathrm{c}}\|_{H^1(\Omega)}^2 \\
\leq& \|\boldsymbol{\sigma}_0\|_{L^2(\Omega)}^2 + \left(\|u_0 - \Pi_0 w_{\mathrm{c}}\|_{L^2(\Omega)} + \|\Pi_0 w_{\mathrm{c}}\|_{L^2(\Omega)}\right)^2 + \|\boldsymbol{q}_{\mathrm{RT}}\|_{H(\mathrm{div},\Omega)}^2 + (1+C_F^2)\|\!|w_{\mathrm{c}}|\!\|^2 \\
\leq& \|\boldsymbol{\sigma}_0\|_{L^2(\Omega)}^2 + \left(\|u_0 - \Pi_0 w_{\mathrm{c}}\|_{L^2(\Omega)} + C_F\|\!|w_{\mathrm{c}}|\!\|\right)^2 + (1+C_F^2)\|\!|w_{\mathrm{c}}|\!\|^2 + \|\boldsymbol{q}_{\mathrm{RT}}\|_{H(\mathrm{div},\Omega)}^2 \\
\leq& \|\boldsymbol{\sigma}_0\|_{L^2(\Omega)}^2 + \left(1+C_F^2\right)\|u_0 - \Pi_0 w_{\mathrm{c}}\|_{L^2(\Omega)}^2 + (2+2C_F^2)\|\!|w_{\mathrm{c}}|\!\|^2 + \|\boldsymbol{q}_{\mathrm{RT}}\|_{H(\mathrm{div},\Omega)}^2 .
\end{aligned}$$

Für die Raviart-Thomas Funktion $\boldsymbol{q}_{\mathrm{RT}}$ folgt aus (2.9) und (2.2) die Ungleichung

$$\begin{aligned}
\|\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 =& \|(1-\Pi_0)\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 + \|\Pi_0 \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 \\
\leq& 1/n^2\ \|\bullet - \operatorname{mid}(T)\|_{L^\infty(\Omega)}^2 \|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 + \left(\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}_0\|_{L^2(\Omega)}\right)^2 \\
\leq& h_{\max}^2/(n+1)^2 \|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 + 2\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}^2 + 2\|\boldsymbol{\sigma}_0\|_{L^2(\Omega)}^2 .
\end{aligned}$$

Das Einfügen von $\tilde{\boldsymbol{q}}_{\mathrm{RT}}$ aus (2.20), welches die Anwendung des Tr-Dev-Div Lemmas und der Bemerkung 2.20 erlaubt, sowie die Eigenschaften des Deviators aus Lemma 2.5 ermöglichen $\|\boldsymbol{\sigma}_0\|_{L^2(\Omega)}$ abzuschätzen

$$\begin{aligned}
\|\boldsymbol{\sigma}_0\|_{L^2(\Omega)} \leq& \|\boldsymbol{\sigma}_0 - \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} + \|\tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} \\
\leq& \|\boldsymbol{\sigma}_0 - \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} + \|\operatorname{dev} \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} + \|\operatorname{tr} \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} \\
\leq& \|\boldsymbol{\sigma}_0 - \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} + C_{\mathrm{tdd}}\|\operatorname{div} \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} + (1+C_{\mathrm{tdd}})\|\operatorname{dev} \tilde{\boldsymbol{q}}_{\mathrm{RT}}\|_{L^2(\Omega)} \\
\leq& \|\boldsymbol{\sigma}_0 - \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} + C_{\mathrm{tdd}}\|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} + (1+C_{\mathrm{tdd}})\|\operatorname{dev} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} \\
\leq& \|\boldsymbol{\sigma}_0 - \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} + C_{\mathrm{tdd}}\|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} \\
&+ (1+C_{\mathrm{tdd}})\left(\|\operatorname{dev}(\boldsymbol{\sigma}_0 - \boldsymbol{q}_{\mathrm{RT}})\|_{L^2(\Omega)} + \|\operatorname{dev} \boldsymbol{\sigma}_0\|_{L^2(\Omega)}\right) \\
\leq& (2+C_{\mathrm{tdd}})\|\boldsymbol{\sigma}_0 - \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} + C_{\mathrm{tdd}}\|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} + (1+C_{\mathrm{tdd}})\|\operatorname{dev} \boldsymbol{\sigma}_0\|_{L^2(\Omega)} .
\end{aligned}$$

Weitere Dreiecksungleichungen und eine erneute Anwendung von (2.9) und (2.2) führen zu folgender Darstellung

$$\begin{aligned}
\|\boldsymbol{\sigma}_0\|_{L^2(\Omega)} \leq& (2+C_{\mathrm{tdd}})\left(\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} + \|(1-\Pi_0)\boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)}\right) + C_{\mathrm{tdd}}\|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} \\
&+ (1+C_{\mathrm{tdd}})\left(\|\operatorname{dev} \boldsymbol{\sigma}_0 - \mathrm{D}\, w_{\mathrm{c}}\|_{L^2(\Omega)} + \|\!|w_{\mathrm{c}}|\!\|\right) \\
\leq& (2+C_{\mathrm{tdd}})\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} + (1+C_{\mathrm{tdd}})\left(\|\operatorname{dev} \boldsymbol{\sigma}_0 - \mathrm{D}\, w_{\mathrm{c}}\|_{L^2(\Omega)} + \|\!|w_{\mathrm{c}}|\!\|\right) \\
&+ ((2+C_{\mathrm{tdd}})h_{\max}/(n+1) + C_{\mathrm{tdd}})\|\operatorname{div} \boldsymbol{q}_{\mathrm{RT}}\|_{L^2(\Omega)} .
\end{aligned}$$

Es sind somit alle Terme von $\|\hat{x}_h\|_{\hat{X}}$ durch Terme aus *Schritt 2.* und $\|\!|w_{\mathrm{c}}|\!\|^2$ kontrolliert.

Insgesamt ergibt sich folgende Abschätzung zur Berechnung der Konstanten.

$$\begin{aligned}
\|\hat{x}_h\|_{\hat{X}}^2 \leq& \|\boldsymbol{\sigma}_0\|_{L^2(\Omega)}^2 + \left(1+C_F^2\right)\|u_0 - \Pi_0 w_\mathrm{c}\|_{L^2(\Omega)}^2 + (2+2C_F^2)|\!|\!|w_\mathrm{c}|\!|\!|^2 + \|\boldsymbol{q}_\mathrm{RT}\|_{H(\mathrm{div},\Omega)}^2 \\
\leq& \left(1+C_F^2\right)\|u_0 - \Pi_0 w_\mathrm{c}\|_{L^2(\Omega)}^2 + (2+2C_F^2)|\!|\!|w_\mathrm{c}|\!|\!|^2 + 2\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}^2 \\
&+ (1+h_{\max}^2/(n+1)^2)\|\mathrm{div}\,\boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}^2 + 3\|\boldsymbol{\sigma}_0\|_{L^2(\Omega)}^2 \\
\leq& \left(1+C_F^2\right)\|u_0 - \Pi_0 w_\mathrm{c}\|_{L^2(\Omega)}^2 + (2+2C_F^2)|\!|\!|w_\mathrm{c}|\!|\!|^2 + 2\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}^2 \\
&+ (1+h_{\max}^2/(n+1)^2)\|\mathrm{div}\,\boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}^2 + 3\Big[\,(2+C_\mathrm{tdd})\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)} \\
&+ ((2+C_\mathrm{tdd})h_{\max}/(n+1) + C_\mathrm{tdd})\|\mathrm{div}\,\boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)} \\
&+ (1+C_\mathrm{tdd})\left(\|\mathrm{dev}\,\boldsymbol{\sigma}_0 - \mathrm{D}\,w_\mathrm{c}\|_{L^2(\Omega)} + |\!|\!|w_\mathrm{c}|\!|\!|\right)\Big]^2.
\end{aligned}$$

Der letzte quadrierte Term kann als Skalarprodukt zweier Vektoren im $\mathbb{R}^4$ betrachtet werden. Das Anwenden der Cauchy-Schwarz Ungleichung ergibt dann folgende verkürzte Darstellung

$$\begin{aligned}
3\Big[\,&(2+C_\mathrm{tdd})\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)} + ((2+C_\mathrm{tdd})h_{\max}/(n+1) + C_\mathrm{tdd})\|\mathrm{div}\,\boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)} \\
&\qquad + (1+C_\mathrm{tdd})\left(\|\mathrm{dev}\,\boldsymbol{\sigma}_0 - \mathrm{D}\,w_\mathrm{c}\|_{L^2(\Omega)} + |\!|\!|w_\mathrm{c}|\!|\!|\right)\Big]^2 \\
\leq& 3\left((2+C_\mathrm{tdd})^2 + ((2+C_\mathrm{tdd})h_{\max}/(n+1) + C_\mathrm{tdd})^2 + 2(1+C_\mathrm{tdd})^2\right) \\
&\left(\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}^2 + \|\mathrm{div}\,\boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}^2 + \|\mathrm{dev}\,\boldsymbol{\sigma}_0 - \mathrm{D}\,w_\mathrm{c}\|_{L^2(\Omega)}^2 + |\!|\!|w_\mathrm{c}|\!|\!|^2\right) \\
=& \left(\frac{3\,(C_\mathrm{tdd}+2)^2}{(n+1)^2}h_{\max}^2 + \frac{6\left(C_\mathrm{tdd}^2 + 2C_\mathrm{tdd}\right)}{n+1}h_{\max} + 12\,(C_\mathrm{tdd}+1)^2 + 6\right) \\
&\left(\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}^2 + \|\mathrm{div}\,\boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}^2 + \|\mathrm{dev}\,\boldsymbol{\sigma}_0 - \mathrm{D}\,w_\mathrm{c}\|_{L^2(\Omega)}^2 + |\!|\!|w_\mathrm{c}|\!|\!|^2\right) \\
=:& C_3\left(\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}^2 + \|\mathrm{div}\,\boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}^2 + \|\mathrm{dev}\,\boldsymbol{\sigma}_0 - \mathrm{D}\,w_\mathrm{c}\|_{L^2(\Omega)}^2 + |\!|\!|w_\mathrm{c}|\!|\!|^2\right).
\end{aligned}$$

Eingesetzt in die ursprüngliche Abschätzung ergibt dies

$$\begin{aligned}
\|\hat{x}_h\|_{\hat{X}}^2 \leq& \left(1+C_F^2\right)\|u_0 - \Pi_0 w_\mathrm{c}\|_{L^2(\Omega)}^2 + (2+2C_F^2+C_3)|\!|\!|w_\mathrm{c}|\!|\!|^2 + C_3\|\mathrm{dev}\,\boldsymbol{\sigma}_0 - \mathrm{D}\,w_\mathrm{c}\|_{L^2(\Omega)}^2 \\
&+ (2+C_3)\|\boldsymbol{\sigma}_0 - \Pi_0 \boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}^2 + (1+h_{\max}^2/(n+1)^2 + C_3)\|\mathrm{div}\,\boldsymbol{q}_\mathrm{RT}\|_{L^2(\Omega)}^2 \\
\leq& \max\left\{1+C_F^2,\, C_3 + \max\left\{2,\, 1+h_{\max}^2/(n+1)^2\right\}\right\}\hat{b}(\hat{x}_h, \tilde{y}_h) \\
&+ \left(C_3 + 2C_F^2 + 2\right)|\!|\!|w_\mathrm{c}|\!|\!|^2.
\end{aligned}$$

Mit den Konstanten

$$\begin{aligned}
C_3 :=& h_{\max}^2 3(C_\mathrm{tdd}+2)^2/(n+1)^2 + h_{\max}6(C_\mathrm{tdd}^2 + 2C_\mathrm{tdd})/(n+1) + 12(C_\mathrm{tdd}+1)^2 + 6, \\
C_{x_h} :=& \max\left\{1+C_F^2,\, C_3 + \max\left\{2,\, 1+h_{\max}^2/(n+1)^2\right\}\right\} + \left(C_3 + 2C_F^2 + 2\right)C_{w_\mathrm{c}}
\end{aligned}$$

gilt also

$$\|\hat{x}_h\|_{\hat{X}}^2 \leq C_{x_h} \hat{b}\,(\hat{x}_h, \tilde{y}_h)\,. \qquad \square$$

Schritt 6. Bestimmung der Inf-Sup-Konstante. Für jedes $\hat{x}_h \in \hat{X}_h \setminus \{0\}$ erfüllt das $\tilde{y}_h \in Y_h$ aus *Schritt 1.*

$$0 < \frac{1}{\sqrt{C_{x_h}\,(1 + h_{\max}^2 n^2/(n+1)^2)}} \leq \frac{\hat{b}\,(\hat{x}_h, \tilde{y}_h)}{\|\hat{x}_h\|_{\hat{X}} \|\tilde{y}_h\|_Y} \leq \sup_{y_h \in Y_h} \frac{\hat{b}\,(\hat{x}_h, y_h)}{\|\hat{x}_h\|_{\hat{X}} \|y_h\|_Y},$$

daher gilt $\hat{\beta}_h \geq \left(C_{x_h} + (1 + h_{\max}^2 n^2/(n+1)^2)\right)^{-1/2}$. In (3.15) und (3.17) wird klar, dass das gesuchte $\beta_h \geq \hat{\beta}_h$, denn

$$\begin{aligned} b(x_h, y_h) = \hat{b}\,(\hat{x}_h, y_h) &\geq \hat{\beta}_h \inf_{\hat{x}_h \in \hat{X} \setminus \{0\}} \sup_{y_h \in Y_h} \|\hat{x}_h\|_{\hat{X}_h} \|y_h\|_Y \\ &\geq \beta_h \inf_{x_h \in X_h \setminus \{0\}} \sup_{y_h \in Y_h} \|x_h\|_X \|y_h\|_Y\,. \end{aligned} \qquad \square$$

3.4 Fortin-Interpolator

In [GQ14],[DG11a] und [CDG14] wird der Fortin-Operator explizit konstruiert, der Ansatz stattdessen die diskrete Inf-Sup Bedingung zu beweisen wurde erst in [CGHW14] und [CH15] eingeführt. Die a-posteriori-Abschätzung aus (3.1) für alle $\xi_h \in X_h$

$$\begin{aligned} \beta\|x - \xi_h\|_X \leq \|b\|/\beta_h \, \|F - b\,(\xi_h, \bullet)\|_{Y_h^*} + \|F \circ (1 - \Pi)\|_{Y^*} \\ \leq 2\|b\|^2/\beta_h \|x - \xi_h\|_X\,. \end{aligned}$$

besteht aus einem Datenapproximationsfehler $\|F \circ (1 - \Pi)\|_{Y^*}$ und dem residualen Fehlerschätzer $\|F - b\,(\xi_h, \bullet)\|_{Y_h^*}$. In diesem Abschnitt sollen für den zweidimensionalen, in dieser Arbeit implementierten Fall der Datenapproximationsfehler und seine Konvergenzordnung genauer untersucht werden. Es ist für $n = 2$ möglich auch für die vorgestellte Diskretisierung niedriger Ordnung eine beschränkte, lineare Abbildung $\Pi\colon Y \to Y_h$ mit $b(x_h, (1 - \Pi)y) = 0$ für alle $x_h \in X_h$ und alle $y \in Y$ direkt und einfach anzugeben. Diese Abbildung wird Fortin-Interpolator genannt.

Dafür werden zwei diskrete Helmholtz Zerlegungen verwendet, in denen die folgenden Operatoren und Räume vorkommen. Es werden zunächst der Curl-Operator, der gedrehte Gradienten, und der curl-Operator, dessen Spur, definiert. Für eine Funktion $\beta \in C^1(\Omega; \mathbb{R}^2)$, seien $\operatorname{Curl}\beta := \begin{pmatrix} -\partial\beta_1/\partial x_2 & \partial\beta_1/\partial x_1 \\ -\partial\beta_2/\partial x_2 & \partial\beta_2/\partial x_1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ und $\operatorname{curl}\beta := \operatorname{tr}(\operatorname{Curl}\beta) = \partial\beta_2/\partial x_1 - \partial\beta_1/\partial x_2 \in \mathbb{R}$. Mit $\operatorname{Curl}_{\mathrm{NC}}$ wird die elementweise Anwendung des Curl-Operators bezeichnet.

Damit wird der folgende Teilraum von $S^1(\mathcal{T};\mathbb{R}^2)$ definiert

$$X_{\mathrm{curl}} := \left\{ v_C \in S^1(\mathcal{T};\mathbb{R}^2) : \int_\Omega v_C \,\mathrm{d}x = 0 \text{ und } \int_\Omega \operatorname{curl} v_C \,\mathrm{d}x = 0 \right\}. \tag{3.26}$$

Des Weiteren werden die Crouzeix-Raviart Funktionen mit und ohne homogene Dirichlet Randbedingungen in den Mittelpunkten der Außenkanten

$$\begin{aligned}
\mathrm{CR}^1(\mathcal{T};\mathbb{R}^2) &:= \{v \in P_1(\mathcal{T},\mathbb{R}^2) : v \text{ ist stetig in } \operatorname{mid}(E) \text{ für alle } E \in \mathcal{E}(\Omega)\},\\
\mathrm{CR}^1_0(\mathcal{T};\mathbb{R}^2) &:= \{v \in \mathrm{CR}^1(\mathcal{T};\mathbb{R}^2) : v(\operatorname{mid}(E)) = 0 \text{ für alle } E \in \mathcal{E}(\partial\Omega)\}
\end{aligned}$$

und die diskret divergenzfreien Crouzeix-Raviart Funktionen

$$Z_{\mathrm{CR}} := \{v \in \mathrm{CR}^1_0(\mathcal{T};\mathbb{R}^2) : \operatorname{div}_{\mathrm{NC}} v = 0\} \tag{3.27}$$

benötigt. Die Crouzeix-Raviart Funktionen $w_{\mathrm{CR}} \in \mathrm{CR}^1_0(\mathcal{T};\mathbb{R}^2)$ haben insbesondere die Eigenschaft, dass $⨍_E [w_{\mathrm{CR}}]_E \,\mathrm{d}s = 1/|E| \int_E [w_{\mathrm{CR}}]_E \,\mathrm{d}s = 0$ für alle $E \in \mathcal{E}$ gilt. Um eine ähnliche Eigenschaft zu garantieren wird folgender Interpolant definiert. Für $v \in H^1(\mathcal{T};\mathbb{R}^2)$ sei der lokale nicht konforme Interpolant $I^{\mathrm{pw}}_{\mathrm{NC}} v \in P_1\left(\mathcal{T};\mathbb{R}^2\right)$ für alle $T \in \mathcal{T}$ durch $\left(I^{\mathrm{pw}}_{\mathrm{NC}} v\right)|_T (\operatorname{mid}(E)) = ⨍_E v|_T \,\mathrm{d}x$ definiert. Es ist schnell nachzurechnen, dass für alle $T \in \mathcal{T}$ und alle $E \in \mathcal{E}(T)$ gilt $\int_E I^{\mathrm{pw}}_{\mathrm{NC}} v|_T = \int_E v|_T \,\mathrm{d}s$ und damit auch $\Pi_0 \operatorname{D}_{\mathrm{NC}} v = \operatorname{D}_{\mathrm{NC}} I^{\mathrm{pw}}_{\mathrm{NC}} v$.

Damit können nun die benötigten zweidimensionalen diskreten Helmholtz Zerlegungen zitiert werden. Mit den Räumen (3.26) und (3.27) können stückweise konstante deviatorische Matrizen charakterisiert werden, wie in [CPR13] gezeigt.

Lemma 3.13 (Diskrete Helmholtz Zerlegung für deviatorische Matrizen). *Die Zerlegung*

$$P_0\left(\mathcal{T};\mathbb{R}^{2\times 2}_{\mathrm{dev}}\right) = \operatorname{D}_{NC} Z_{CR} \oplus \operatorname{dev}\operatorname{Curl} X_{\mathrm{curl}}$$

ist orthogonal in $L^2\left(\Omega;\mathbb{R}^{2\times 2}_{\mathrm{dev}}\right)$*, wenn* Ω *einfach zusammenhängend ist.*

Des Weiteren gilt nach [AF89] bzw. [Car09b, Theorem 3.32] die folgende Zerlegung der stückweise konstanten Funktionen.

Lemma 3.14 (Diskrete Helmholtz Zerlegung). *Ist* Ω *einfach zusammenhängend, so ist die Zerlegung*

$$P_0\left(\mathcal{T};\mathbb{R}^{2\times 2}\right) = \operatorname{D} S^1_0(\mathcal{T};\mathbb{R}^2) \oplus \operatorname{Curl}_{NC}\left(\mathrm{CR}^1(\mathcal{T})/\mathbb{R}\right)^2$$

orthogonal in $L^2\left(\Omega;\mathbb{R}^{2\times 2}\right)$.

Dabei bedeutet $\mathrm{CR}^1(\mathcal{T})/\mathbb{R}$ lediglich bis auf Konstante und garantiert die Eindeutigkeit der Zerlegung. Damit diese Helmholtz Zerlegungen verwendet werden können, wird Ω

als einfach zusammenhängend angenommen (im Zweifel sind alle Betrachtungen auf den einzelnen Zusammenhangskomponenten durchzuführen).

Mit dieser Vorarbeit ist es möglich den Fortin-Operator Π und den diskreten Kern N_h anzugeben.

Lemma 3.15. *Zu $y = (\boldsymbol{\tau}, v) \in Y$ erfülle $\alpha_{CR} \in \mathrm{CR}^1_0(\mathcal{T};\mathbb{R}^2)$ für alle $w_{CR} \in \mathrm{CR}^1_0(\mathcal{T};\mathbb{R}^2)$ die folgende Gleichung*

$$\int_\Omega \mathrm{D}_{NC}\,\alpha_{CR} : \mathrm{D}_{NC}\,w_{CR}\,\mathrm{d}x = \int_\Omega w_{CR}\cdot(1-\Pi_0)\,\mathrm{div}_{NC}\,\boldsymbol{\tau}\,\mathrm{d}x + \int_\Omega \Pi_0\boldsymbol{\tau} : \mathrm{D}_{NC}\,w_{CR}\,\mathrm{d}x. \tag{3.28}$$

Außerdem seien $z_{CR} \in Z_{CR}$ und $\beta_c \in X_{\mathrm{curl}}$ die Funktionen aus der nach Lemma 3.13 existierenden Helmholtz Zerlegung

$$\mathrm{dev}\,(\Pi_0\boldsymbol{\tau} - \mathrm{D}_{NC}\,\alpha_{CR}) = \mathrm{D}_{NC}\,z_{CR} + \mathrm{dev}\,\mathrm{Curl}\,\beta_c. \tag{3.29}$$

Dann besitzt die folgende Abbildung

$$\Pi\colon H(\mathrm{div},\mathcal{T};\mathbb{R}^{2\times2})/\mathbb{R}\times H^1(\mathcal{T};\mathbb{R}^2) \to RT_0^{pw}(\mathcal{T})/\mathbb{R}\times P_1(\mathcal{T})$$

mit

$$\begin{aligned}\Pi(\boldsymbol{\tau}) &:= \boldsymbol{\tau}_{RT} := \mathrm{D}_{NC}\,\alpha_{CR} + \mathrm{Curl}\,\beta_c + (\Pi_0\,\mathrm{div}_{NC}\,\boldsymbol{\tau}/2)\otimes(\bullet - \mathrm{mid}(\mathcal{T})),\\ \Pi(v) &:= v_1 := I^{pw}_{NC}v + z_{CR},\end{aligned}$$

die gewünschten Eigenschaften, ist also beschränkt und erfüllt $b(x_h,(1-\Pi)y) = 0$ für alle $x_h \in X_h$ und alle $y \in Y$. Außerdem gilt für den diskreten Kern von b

$$\begin{aligned}N_h :&= \{y_h \in Y_h\colon b(x_h,y_h) = 0 \text{ für alle } x_h \in X_h\}\\ &= \Big\{(\mathrm{Curl}_{NC}\,\beta_{CR}, v_{CR}) \in \mathrm{Curl}_{NC}\left(\mathrm{CR}^1(\mathcal{T})/\mathbb{R}\right)^2 \times Z_{CR}\colon\\ &\qquad\qquad -\mathrm{D}_{NC}\,v_{CR} = \mathrm{dev}\,\mathrm{Curl}_{NC}\,\beta_{CR}\Big\}.\end{aligned}$$

Beweis. Schritt 1. Zunächst wird Π untersucht. Offensichtlich gilt $v_1 \in P_1(\mathcal{T};\mathbb{R}^2)$ und $\boldsymbol{\tau}_{\mathrm{RT}} \in RT_0^{\mathrm{pw}}(\mathcal{T};\mathbb{R}^{2\times2})$. Nach dem Gaußschem Integralsatz, den Eigenschaften der Sprünge von $\mathrm{CR}^1_0(\mathcal{T};\mathbb{R}^2)$ und den Nebenbedingungen in X_{curl} ist

$$\begin{aligned}\int_\Omega \mathrm{tr}\,\boldsymbol{\tau}_{\mathrm{RT}}\,\mathrm{d}x &= \int_\Omega \mathrm{div}_{\mathrm{NC}}\,\alpha_{\mathrm{CR}}\,\mathrm{d}x + \int_\Omega \mathrm{tr}\left((\Pi_0\,\mathrm{div}_{\mathrm{NC}}\,\boldsymbol{\tau}/2)\otimes(\bullet - \mathrm{mid}(\mathcal{T}))\right)\mathrm{d}x\\ &\quad + \int_\Omega \mathrm{curl}\,\beta_c\,\mathrm{d}x\\ &= \sum_{E\in\mathcal{E}}\int_E [\alpha_{\mathrm{CR}}]_E\,\nu_E\,\mathrm{d}s + 0 + 0 = 0\end{aligned}$$

erfüllt. Also ist $\boldsymbol{\tau}_{\mathrm{RT}} \in RT_0^{\mathrm{pw}}\left(\mathcal{T};\mathbb{R}^{2\times2}\right)/\mathbb{R}$.

Schritt 2. Offenbar ist die Forderung $b(x_h, (1-\Pi)y) = 0$ für alle $x_h \in X_h$ und alle $y \in Y$ äquivalent dazu, dass $y = (\boldsymbol{\tau}, v)$ und $\Pi y = (\boldsymbol{\tau}_{\mathrm{RT}}, v_1)$ für alle $x_h = (\boldsymbol{\sigma}_0, u_0, s_1, t_0) \in X_h$ folgendes Gleichungssystem erfüllen

$$0 = b((\boldsymbol{\sigma}_0, 0, 0, 0), y - y_h) = \int_\Omega \boldsymbol{\sigma}_0 : \mathrm{D}_{\mathrm{NC}}(v - v_1)\,\mathrm{d}x + \int_\Omega \operatorname{dev} \boldsymbol{\sigma}_0 : (\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}})\,\mathrm{d}x, \quad (3.30)$$

$$0 = b((0, u_0, 0, 0), y - y_h) = \int_\Omega u_0 \cdot \operatorname{div}_{\mathrm{NC}}(\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}})\,\mathrm{d}x, \quad (3.31)$$

$$0 = b((0, 0, s_1, 0), y - y_h) = -\left\langle \gamma_\nu^{\mathcal{T}}(\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}}), s_1 \right\rangle_{\partial\mathcal{T}}, \quad (3.32)$$

$$0 = b((0, 0, 0, t_0), y - y_h) = -\left\langle t_0, \gamma_0^{\mathcal{T}}(v - v_1) \right\rangle_{\partial\mathcal{T}}. \quad (3.33)$$

Da $\mathrm{D}_{\mathrm{NC}}\,\alpha_{\mathrm{CR}}$, $\operatorname{Curl}\beta_c \in P_0\left(\mathcal{T}; \mathbb{R}^{2\times 2}\right)$ gilt, ist $\operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}_{\mathrm{RT}} = \Pi_0 \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}$. Daraus ergibt sich (3.31), denn nach (2.10) gilt für alle $u_0 \in P_0(\mathcal{T})$

$$\int_\Omega u_0 \cdot \operatorname{div}_{\mathrm{NC}}(\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}})\,\mathrm{d}x = \int_\Omega u_0 \cdot \Pi_0 \operatorname{div}_{\mathrm{NC}}(\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}})\,\mathrm{d}x = 0.$$

Die Gleichung (3.33) lässt sich aus den Eigenschaften von $I_{\mathrm{NC}}^{\mathrm{pw}}$ und den Crouzeix-Raviart Funktionen folgern. Es gilt für alle $t_0 \in P_0(\mathcal{E}; \mathbb{R}^2)$

$$\begin{aligned} -\left\langle t_0, \gamma_0^{\mathcal{T}}(v - v_1) \right\rangle_{\partial\mathcal{T}} &= \sum_{E\in\mathcal{E}} \int_E t_0 \left[v - v_1\right]_E \,\mathrm{d}s \\ &= \sum_{E\in\mathcal{E}} t_0|_E \left(\int_E \left[v - I_{\mathrm{NC}}^{\mathrm{pw}} v\right]_E \,\mathrm{d}s - \int_E \left[z_{\mathrm{CR}}\right]_E \,\mathrm{d}s \right) = 0. \end{aligned}$$

Die Wahl von α_{CR} wie in (3.28) garantiert (3.32). Betrachte zu jedem $s_1 \in S_0^1(\mathcal{E}; \mathbb{R}^2)$ die Fortsetzung $w_{\mathrm{c}} \in S_0^1(\mathcal{T}; \mathbb{R}^2)$ mit $\gamma_0^{\mathcal{T}} w_{\mathrm{c}} = s_1$, dann gilt nach Satz 2.8

$$-\left\langle \gamma_\nu^{\mathcal{T}}(\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}}), s_1 \right\rangle_{\partial\mathcal{T}} = \int_\Omega \mathrm{D}\, w_{\mathrm{c}} : \Pi_0(\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}})\,\mathrm{d}x + \int_\Omega w_{\mathrm{c}} \cdot \operatorname{div}_{\mathrm{NC}}(\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}})\,\mathrm{d}x.$$

Die Orthogonalität in Lemma 3.14, Gleichung (3.28) und die Tatsache $\operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}_{\mathrm{RT}} = \Pi_0 \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}$ zeigen, dass für alle $w_{\mathrm{c}} \in S_0^1(\mathcal{T}; \mathbb{R}^2) \subseteq \mathrm{CR}_0^1(\mathcal{T}; \mathbb{R}^2)$ gilt

$$\begin{aligned} &\int_\Omega \mathrm{D}\, w_{\mathrm{c}} : \Pi_0(\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}})\,\mathrm{d}x + \int_\Omega w_{\mathrm{c}} \cdot \operatorname{div}_{\mathrm{NC}}(\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}})\,\mathrm{d}x \\ &\quad= \int_\Omega \mathrm{D}\, w_{\mathrm{c}} : \Pi_0 \boldsymbol{\tau}\,\mathrm{d}x - \int_\Omega \mathrm{D}\, w_{\mathrm{c}} : \mathrm{D}_{\mathrm{NC}}\,\alpha_{\mathrm{CR}}\,\mathrm{d}x - \int_\Omega \mathrm{D}\, w_{\mathrm{c}} : \operatorname{Curl}\beta_c\,\mathrm{d}x \\ &\qquad+ \int_\Omega w_{\mathrm{c}} \cdot (1 - \Pi_0) \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\,\mathrm{d}x \\ &\quad= \int_\Omega \mathrm{D}\, w_{\mathrm{c}} : \Pi_0 \boldsymbol{\tau}\,\mathrm{d}x - \int_\Omega w_{\mathrm{c}} \cdot (1 - \Pi_0) \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\,\mathrm{d}x - \int_\Omega \Pi_0 \boldsymbol{\tau} : \mathrm{D}\, w_{\mathrm{c}}\,\mathrm{d}x \\ &\qquad+ \int_\Omega w_{\mathrm{c}} \cdot (1 - \Pi_0) \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\,\mathrm{d}x \\ &\quad= 0. \end{aligned}$$

Die linke Seite der Gleichung (3.30) lässt sich mit Lemma 2.5 und (2.10) wie folgt umschreiben

$$\begin{aligned}\int_\Omega \boldsymbol{\sigma}_0 : \mathrm{D_{NC}}(v - v_1)\,\mathrm{d}x + \int_\Omega \operatorname{dev}\boldsymbol{\sigma}_0 : (\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}})\,\mathrm{d}x \\ = \int_\Omega \boldsymbol{\sigma}_0 : \Pi_0\,\mathrm{D_{NC}}(v - v_1)\,\mathrm{d}x + \int_\Omega \boldsymbol{\sigma}_0 : \operatorname{dev}\Pi_0(\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}})\,\mathrm{d}x.\end{aligned}$$

Die Gleichheit $\Pi_0\,\mathrm{D_{NC}}\,v = \mathrm{D_{NC}}\,I_{\mathrm{NC}}^{\mathrm{pw}}v$ ergibt zusammen mit (3.29)

$$\begin{aligned}\int_\Omega \boldsymbol{\sigma}_0 :&\Pi_0\,\mathrm{D_{NC}}(v - v_1)\,\mathrm{d}x + \int_\Omega \boldsymbol{\sigma}_0 : \operatorname{dev}\Pi_0(\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{RT}})\,\mathrm{d}x \\ &= -\int \boldsymbol{\sigma}_0 : \mathrm{D_{NC}}\,z_{\mathrm{CR}}\,\mathrm{d}x + \int_\Omega \boldsymbol{\sigma}_0 : \operatorname{dev}(\Pi_0\boldsymbol{\tau} - \mathrm{D_{NC}}\,\alpha_{\mathrm{CR}} - \operatorname{Curl}\beta_c)\,\mathrm{d}x \\ &= -\int \boldsymbol{\sigma}_0 : \mathrm{D_{NC}}\,z_{\mathrm{CR}}\,\mathrm{d}x + \int_\Omega \boldsymbol{\sigma}_0 : \mathrm{D_{NC}}\,z_{\mathrm{CR}}\,\mathrm{d}x = 0.\end{aligned}$$

Damit gilt $b(x_h, (1-\Pi)y) = 0$ für alle $x_h \in X_h$ und alle $y \in Y$. Es lässt sich direkt verifizieren, dass Π idempotent ist, also $\Pi(\Pi(y)) = \Pi(y)$ gilt.

Schritt 3. Die Beschränktheit der Abbildung lässt sich ebenfalls nachrechnen. Sei $y = (\boldsymbol{\tau}, v) \in Y$ beliebig, dann gilt für $\Pi(y) \in Y_h$ wie in Lemma 3.15

$$\begin{aligned}\|\Pi(y)\|_Y^2 =&\|\Pi(\boldsymbol{\tau})\|_{H(\mathrm{div},\Omega)}^2 + \|\Pi(v)\|_{H^1(\Omega)}^2 \\ =&\|\mathrm{D_{NC}}\,\alpha_{\mathrm{CR}}\|_{L^2(\Omega)}^2 + \|\operatorname{Curl}\beta_c\|_{L^2(\Omega)}^2 + \|\Pi_0\operatorname{div_{NC}}\boldsymbol{\tau}/2\ \otimes(\bullet - \operatorname{mid}(\mathcal{T}))\|_{L^2(\Omega)}^2 \\ &+\|\Pi_0\operatorname{div_{NC}}\boldsymbol{\tau}\|_{L^2(\Omega)} + \left\|I_{\mathrm{NC}}^{\mathrm{pw}}v + z_{\mathrm{CR}}\right\|_{L^2(\Omega)}^2 + \left\|\mathrm{D_{NC}}(I_{\mathrm{NC}}^{\mathrm{pw}}v + z_{\mathrm{CR}})\right\|_{L^2(\Omega)}^2.\end{aligned}$$

Gleichung (3.28), die Cauchy-Schwarz und die diskrete Friedrichs Ungleichung mit Konstante C_{dF} aus [BS08, 10.6.14] ermöglichen die Abschätzung des ersten Terms

$$\begin{aligned}\int_\Omega \mathrm{D_{NC}}\,\alpha_{\mathrm{CR}} : \mathrm{D_{NC}}\,\alpha_{\mathrm{CR}}\,\mathrm{d}x &= \int_\Omega \alpha_{\mathrm{CR}}\cdot(1-\Pi_0)\operatorname{div_{NC}}\boldsymbol{\tau}\,\mathrm{d}x + \int_\Omega \Pi_0\boldsymbol{\tau} : \mathrm{D_{NC}}\,\alpha_{\mathrm{CR}}\,\mathrm{d}x \\ &\le \|\alpha_{\mathrm{CR}}\|_{L^2(\Omega)}\|(1-\Pi_0)\operatorname{div_{NC}}\boldsymbol{\tau}\|_{L^2(\Omega)} + \|\Pi_0\boldsymbol{\tau}\|_{L^2(\Omega)}\,|\!|\!|\alpha_{\mathrm{CR}}|\!|\!|_{\mathrm{NC}} \\ &\le |\!|\!|\alpha_{\mathrm{CR}}|\!|\!|_{\mathrm{NC}}\left(C_{\mathrm{dF}}\|\operatorname{div_{NC}}\boldsymbol{\tau}\|_{L^2(\Omega)} + \|\boldsymbol{\tau}\|_{L^2(\Omega)}\right).\end{aligned}$$

Nach Cauchy-Schwarz und der Normabschätzung in (2.2) gilt

$$\begin{aligned}\|\Pi_0\operatorname{div_{NC}}\boldsymbol{\tau}/2\ \otimes(\bullet - \operatorname{mid}(\mathcal{T}))\|_{L^2(\Omega)} &\le \|\Pi_0\operatorname{div_{NC}}\boldsymbol{\tau}/2\|_{L^2(\Omega)}\|\bullet - \operatorname{mid}(\mathcal{T})\|_{L^\infty(\Omega)} \\ &\le h_{\max}/3\ \|\operatorname{div_{NC}}\boldsymbol{\tau}\|_{L^2(\Omega)}.\end{aligned}$$

Außerdem gilt wegen der bereits erwähnten diskreten Friedrichs Ungleichung und $\Pi_0\,\mathrm{D_{NC}}\,v = \mathrm{D_{NC}}\,I_{\mathrm{NC}}^{\mathrm{pw}}v$, dass

$$\begin{aligned}\left\|I_{\mathrm{NC}}^{\mathrm{pw}}v\right\|_{L^2(\Omega)}^2 + \left\|\mathrm{D_{NC}}\,I_{\mathrm{NC}}^{\mathrm{pw}}v\right\|_{L^2(\Omega)}^2 &\le (1+C_{\mathrm{dF}}^2)\left\|\mathrm{D_{NC}}\,I_{\mathrm{NC}}^{\mathrm{pw}}\right\|_{L^2(\Omega)}^2 = (1+C_{\mathrm{dF}}^2)\|\Pi_0\,\mathrm{D_{NC}}\,v\|_{L^2(\Omega)} \\ &\le (1+C_{\mathrm{dF}}^2)\|\mathrm{D_{NC}}\,v\|_{L^2(\Omega)}.\end{aligned}$$

Auf Grund der diskreten Friedrichs Ungleichung genügt es zuletzt die folgende Summe zu betrachten,

$$\|\mathrm{D}_{\mathrm{NC}}\, z_{\mathrm{CR}}\|^2_{L^2(\Omega)} + \|\mathrm{Curl}\, \beta_c\|^2_{L^2(\Omega)}$$

Für $\beta_c \in X_{\mathrm{curl}}$ gilt $\operatorname{div}\mathrm{Curl}\,\beta_c \equiv 0$ und $\int_\Omega \operatorname{tr}\mathrm{Curl}\,\beta_c \,\mathrm{d}x = 0$, damit liegt $\mathrm{Curl}\,\beta_c \in H(\operatorname{div},\Omega;\mathbb{R}^{2\times 2})/\mathbb{R}$ und das Tr-Div-Dev Lemma, Lemma 2.18, ist anwendbar. Es gilt zusammen mit Lemma 2.5

$$\begin{aligned}
\|\mathrm{Curl}\,\beta_c\|^2_{L^2(\Omega)} &= \|\operatorname{dev}\mathrm{Curl}\,\beta_c\|^2_{L^2(\Omega)} + 1/2\ \|\operatorname{tr}\mathrm{Curl}\,\beta_c\|^2_{L^2(\Omega)} \\
&\leq (1+C^2_{\mathrm{tdd}})\|\operatorname{dev}\mathrm{Curl}\,\beta_c\|^2_{L^2(\Omega)} + C^2_{\mathrm{tdd}}\|\operatorname{div}\mathrm{Curl}\,\beta_c\|^2_{L^2(\Omega)} \\
&= (1+C^2_{\mathrm{tdd}})\|\operatorname{dev}\mathrm{Curl}\,\beta_c\|^2_{L^2(\Omega)}\,.
\end{aligned}$$

Mit der Abkürzung $C := (1+C^2_{\mathrm{tdd}})$, mit (3.29) und der Orthogonalität in eben dieser Helmholtz Zerlegung ergibt folgende Abschätzung

$$\begin{aligned}
\|\mathrm{D}_{\mathrm{NC}}\, z_{\mathrm{CR}}\|^2_{L^2(\Omega)} + \|\mathrm{Curl}\,\beta_c\|^2_{L^2(\Omega)} &\leq \|\mathrm{D}_{\mathrm{NC}}\, z_{\mathrm{CR}}\|^2_{L^2(\Omega)} + C\|\operatorname{dev}\mathrm{Curl}\,\beta_c\|^2_{L^2(\Omega)} \\
&\leq \max\{1,C\}\|\mathrm{D}_{\mathrm{NC}}\, z_{\mathrm{CR}} + \operatorname{dev}\mathrm{Curl}\,\beta_c\|^2_{L^2(\Omega)} \\
&\leq C\|\operatorname{dev}(\Pi_0\boldsymbol{\tau} - \mathrm{D}_{\mathrm{NC}}\,\alpha_{\mathrm{CR}})\|^2_{L^2(\Omega)} \\
&\leq 2C\left(\|\Pi_0\boldsymbol{\tau}\|^2_{L^2(\Omega)} + \|\mathrm{D}_{\mathrm{NC}}\,\alpha_{\mathrm{CR}}\|^2_{L^2(\Omega)}\right) \\
&\leq C\left(6\|\boldsymbol{\tau}\|^2_{L^2(\Omega)} + 4C^2_{\mathrm{dF}}\|\operatorname{div}_{\mathrm{NC}}\boldsymbol{\tau}\|^2_{L^2(\Omega)}\right).
\end{aligned}$$

Damit sind alle Terme durch Teile von $\|y\|_Y$ kontrolliert und die angegebene Abbildung ist beschränkt.

Schritt 4. Es bleibt N_h zu bestimmen. Für alle $\tilde{y}_h = (\tilde{\boldsymbol{\tau}}_{\mathrm{RT}}, \tilde{v}_1) \in N_h$, ergibt sich aus

$$0 = b((0,0,0,t_0),(\tilde{\boldsymbol{\tau}}_{\mathrm{RT}},\tilde{v}_1)) = -\left\langle t_0, \gamma_0^{\mathcal{T}}\tilde{v}_1\right\rangle_{\partial\mathcal{T}} = -\sum_{E\in\mathcal{E}} t_0|_E \int_E [\tilde{v}_1]_E \,\mathrm{d}s$$

für alle $t_0 \in P_0(\mathcal{E};\mathbb{R}^2)$, die Bedingung $v_1 \in \mathrm{CR}^1_0(\mathcal{T};\mathbb{R}^2)$. Aus

$$0 = b((0,u_0,0,0),(\tilde{\boldsymbol{\tau}}_{\mathrm{RT}},\tilde{v}_1)) = \int_\Omega u_0 \cdot \operatorname{div}_{\mathrm{NC}}\tilde{\boldsymbol{\tau}}_{\mathrm{RT}}\,\mathrm{d}x$$

für alle $u_0 \in P_0(\mathcal{T};\mathbb{R}^2)$ folgt $\operatorname{div}_{\mathrm{NC}}\tilde{\boldsymbol{\tau}}_{\mathrm{RT}} = 0$, also $\Pi_0\tilde{\boldsymbol{\tau}}_{\mathrm{RT}} = \tilde{\boldsymbol{\tau}}_{\mathrm{RT}}$. Des Weiteren kann jedes $s_1 \in S^1_0(\mathcal{E};\mathbb{R}^2)$ wieder durch ein $w_{\mathrm{c}} \in S^1_0(\mathcal{T};\mathbb{R}^2)$ mit $\gamma_0^{\mathcal{T}} w_{\mathrm{c}} = s_1$ fortgesetzt werden und es ergibt sich nach (2.8)

$$0 = b((0,0,s_1,0),(\tilde{\boldsymbol{\tau}}_{\mathrm{RT}},\tilde{v}_1) = -\left\langle \gamma_\nu^{\mathcal{T}}(\boldsymbol{\tau}-\boldsymbol{\tau}_{\mathrm{RT}}), s_1\right\rangle_{\partial\mathcal{T}} = \int_\Omega \mathrm{D}\, w_{\mathrm{c}} : \tilde{\boldsymbol{\tau}}_{\mathrm{RT}}\,\mathrm{d}x$$

für alle $w_{\mathrm{c}} \in S^1_0(\mathcal{T};\mathbb{R}^2)$. Nach der Helmholtz Zerlegung Lemma 3.14 gilt $\tilde{\boldsymbol{\tau}}_{\mathrm{RT}} = \mathrm{Curl}_{\mathrm{NC}}\,\beta_{\mathrm{CR}}$

für ein $\beta_{\mathrm{CR}} \in \left(\mathrm{CR}^1(\mathcal{T})/\mathbb{R}\right)^2$. Nun folgt aus

$$0 = b((\boldsymbol{\sigma}_0, 0, 0, 0), (\tilde{\boldsymbol{\tau}}_{\mathrm{RT}}, \tilde{v}_1)) = \int_\Omega \boldsymbol{\sigma}_0 : \mathrm{D}_{\mathrm{NC}}\, \tilde{v}_1 \,\mathrm{d}x + \int_\Omega \operatorname{dev} \boldsymbol{\sigma}_0 : \tilde{\boldsymbol{\tau}}_{\mathrm{RT}} \,\mathrm{d}x$$

für alle $\boldsymbol{\sigma}_0 \in P_0(\mathcal{T}; \mathbb{R}^{2\times 2})/\mathbb{R}$ noch $\operatorname{dev} \tilde{\boldsymbol{\tau}}_{\mathrm{RT}} = -\mathrm{D}_{\mathrm{NC}}\, \tilde{v}_1$, da $\tilde{v}_1 \in \mathrm{CR}^1_0(\mathcal{T}; \mathbb{R}^2)$ divergenzfrei ist. Insgesamt ergibt sich

$$N_h = \Big\{(\operatorname{Curl}_{\mathrm{NC}} \beta_{\mathrm{CR}}, v_{\mathrm{CR}}) \in \operatorname{Curl}_{\mathrm{NC}} \left(\mathrm{CR}^1(\mathcal{T})/\mathbb{R}\right)^2 \times Z_{\mathrm{CR}} : \\ - \mathrm{D}_{\mathrm{NC}}\, v_{\mathrm{CR}} = \operatorname{dev} \operatorname{Curl}_{\mathrm{NC}} \beta_{\mathrm{CR}}\Big\}.$$

Die Dimension dieses Raumes lässt sich bestimmen, in dem eine Zerlegung $Y_h = N_h \oplus M_h$ betrachtet wird. Dabei sei wie in [Wer11, V.3.4]

$$M_h = N_h^\perp = \Big\{m_h \in Y_h :\ \langle m_h, n_h\rangle_Y = 0 \text{ für alle } n_h \in N_h\Big\}$$

als das orthogonale Komplement zu N_h im Hilbertraum Y gewählt. Sei wieder T_h der trial-to-test Operator aus Abschnitt 3.1, dann gilt $T_h(X_h) = M_h$. Denn für alle $x_h \in X_h$ und $n_h \in N_h$ gilt $\langle T_h(x_h), n_h\rangle_Y = b(x_h, n_h) = 0$, also $T_h(X_h) \subseteq M_h$. Umgekehrt definiert $m_h \in M_h$ eine Abbildung $\langle \bullet, m_h\rangle_Y \in M_h^*$ und eine Lösung $x_h \in X_h$ mit $b(x_h, \bullet) = \langle m_h, \bullet\rangle \in M_h^*$. Für alle $\hat{y}_h = \hat{m}_h + \hat{n}_h \in Y_h = N_h \oplus M_h$ gilt $\langle m_h, y_h\rangle_Y = \langle m_h, \hat{m}_h\rangle_Y = b(x_h, \hat{m}_h) = b(x_h, \hat{y}_h)$, also ist $m_h = T_h(x_h)$ und $M_h \subseteq T_h(X_h)$. Nun gilt wegen der eindeutigen Lösbarkeit

$$\begin{aligned}\dim(N_h) &= \dim(Y_h) - \dim(M_h) = \dim(Y_h) - \dim(X_h)\\ &= 12|\mathcal{T}| - 1 - \left(6|\mathcal{T}| - 1 + 2|\mathcal{N}(\Omega)| + 2|\mathcal{E}|\right).\end{aligned}$$

Nach den Euler Formeln [Car09b, Remark 1.21] gilt $2|\mathcal{T}| + 1 = |\mathcal{N}| + |\mathcal{E}(\Omega)|$. Da $|\mathcal{E}(\partial\Omega)| = |\mathcal{N}(\partial\Omega)|$ ergibt sich

$$\begin{aligned}\dim(N_h) &= 6|\mathcal{T}| - 2|\mathcal{N}(\Omega)| - 2|\mathcal{E}| = 6|\mathcal{T}| - 2\left(|\mathcal{N}| - |\mathcal{N}(\partial\Omega)| + |\mathcal{E}|\right)\\ &= 6|\mathcal{T}| - 2\left(|\mathcal{N}| + |\mathcal{E}(\Omega)|\right) = 2(|\mathcal{T}| - 1).\end{aligned}$$ □

Vor der genaueren Untersuchung des Datenapproximationsfehlers, sollen noch einige Überlegungen den Kern N_h betreffend festgehalten werden. Die Untersuchung des Kerns N_h bzw. des Bildes $M_h := T_h(X_h)$ des trial-to-test Operators ist lohnend, da die Kenntnis von M_h erlaubt, die dPG-Methode schlicht als gemischte Methode zu betrachten.

Bemerkung 3.16 Es gilt, wie soeben gezeigt,

$$N_h = \Big\{(\operatorname{Curl}_{\mathrm{NC}} \beta_{\mathrm{CR}}, v_{\mathrm{CR}}) \in \operatorname{Curl}_{\mathrm{NC}} \left(\mathrm{CR}^1(\mathcal{T})/\mathbb{R}\right)^2 \times Z_{\mathrm{CR}} : \\ - \mathrm{D}_{\mathrm{NC}}\, v_{\mathrm{CR}} = \operatorname{dev} \operatorname{Curl}_{\mathrm{NC}} \beta_{\mathrm{CR}}\Big\}.$$

Nach der Helmholtz Zerlegung, Lemma 3.13, ist $\operatorname{dev}\operatorname{Curl}_{\mathrm{NC}}\beta_{\mathrm{CR}} \perp \operatorname{dev}\operatorname{Curl} X_{\mathrm{curl}}$. Sei $b_{\mathrm{CR}} \in (\mathrm{CR}^1(\mathcal{T})/\mathbb{R})^2$ gegeben, so müsste $\operatorname{dev}\operatorname{Curl}_{\mathrm{NC}} b_{\mathrm{CR}}$ auf $\operatorname{dev}\operatorname{Curl} X_{\mathrm{curl}}$ L^2- projiziert werden. Der Fehler dieser Projektion ergibt ein geeignetes β_{CR}. Die genauere Betrachtung von $\beta_{\mathrm{CR}} \in (\mathrm{CR}^1(\mathcal{T})/\mathbb{R})^2$ und $b_c \in X_{\mathrm{curl}}$ verdeutlicht einen Zusammenhang zur linearen Elastizität. Nach Definition des symmetrischen Ableitung $\varepsilon(\bullet) := \operatorname{sym}\mathrm{D}(\bullet) = (\mathrm{D}(\bullet) + \mathrm{D}(\bullet)^\top)/2$ gilt

$$0 = \int_\Omega \operatorname{dev}\operatorname{Curl}_{\mathrm{NC}}\beta_{\mathrm{CR}} : \operatorname{dev}\operatorname{Curl} b_c \,\mathrm{d}x = \int_\Omega \varepsilon_{\mathrm{NC}}(\beta_{\mathrm{CR}}) : \varepsilon(b_c)\,\mathrm{d}x.$$

Gleichungen dieser Art werden in der Elastizität häufiger untersucht.
Nach der anderen Helmholtz Zerlegung, Lemma 3.14, ist $\operatorname{Curl}_{\mathrm{NC}}\beta_{\mathrm{CR}} \in \operatorname{Curl}(\mathrm{CR}^1(\mathcal{T})/\mathbb{R})^2$ bezüglich der L^2-Norm senkrecht zu $\mathrm{D}\,S^1_0(\mathcal{T};\mathbb{R}^2)$. D.h. da $\operatorname{Curl}_{\mathrm{NC}}\beta_{\mathrm{CR}}$ divergenzfrei ist gilt

$$\mathrm{D}\,S^1_0(\mathcal{T};\mathbb{R}^2) \times \{0\} \subset M_h.$$

Im Folgenden soll nun der Datenapproximationsfehler untersucht werden. Handelt es sich um ein Problem mit homogenen Randdaten so lässt sich der Datenapproximationsfehler,

$$\|F \circ (1-\Pi)\|_{Y^*} = \sup_{(v,\boldsymbol{\tau})\in S(Y)} \int_\Omega f \cdot (v - \Pi v)\,\mathrm{d}x,$$

mit Hilfe des folgenden Lemmas, [CG14, Theorem 4], abschätzen.

Lemma 3.17. *Sei $j_{1,1}$ die erste Wurzel der ersten Besselfunktion, dann gilt mit der Konstante $\kappa := \sqrt{1/48 + j_{1,1}^{-2}} = 0.298234942888$ für alle $v \in H^1(\mathcal{T};\mathbb{R}^2)$*

$$\left\|v - I^{pw}_{NC} v\right\|_{L^2(\Omega)} \leq \kappa |\!|\!| h_{\mathcal{T}} \left(v - I^{pw}_{NC} v\right) |\!|\!|_{NC}.$$

Nach der Definition des Fortin-Operators Π in Lemma 3.15, der Cauchy-Schwarz und der diskreten Friedrichs Ungleichung aus [BS08, 10.6.14] mit Konstante C_{dF} gilt für jedes $y = (\boldsymbol{\tau}, v) \in Y$,

$$\begin{aligned} F((1-\Pi)(y)) &= \int_\Omega f \cdot (v - \Pi v)\,\mathrm{d}x \\ &= \int_\Omega f \cdot (v - I^{\mathrm{pw}}_{\mathrm{NC}} v)\,\mathrm{d}x - \int_\Omega f z_{\mathrm{CR}}\,\mathrm{d}x \\ &\leq \|h_{\mathcal{T}} f\|_{L^2(\Omega)} \left\|h_{\mathcal{T}}^{-1}(v - I^{\mathrm{pw}}_{\mathrm{NC}} v)\right\|_{L^2(\Omega)} + C_{\mathrm{dF}} \|f\|_{L^2(\Omega)} |\!|\!| z_{\mathrm{CR}} |\!|\!|_{\mathrm{NC}}. \end{aligned}$$

Der erste Term kann mit Lemma 3.17 abgeschätzt werden

$$\left\|h_{\mathcal{T}}^{-1}(v - I^{\mathrm{pw}}_{\mathrm{NC}} v)\right\|_{L^2(\Omega)} \leq \kappa |\!|\!| v - I^{\mathrm{pw}}_{\mathrm{NC}} v |\!|\!|_{\mathrm{NC}} = \kappa \|(1-\Pi_0)\,\mathrm{D}_{\mathrm{NC}}\, v\|_{L^2(\Omega)} \leq \kappa |\!|\!| v |\!|\!|_{\mathrm{NC}}.$$

Zur Betrachtung des zweiten Summanden sei an die Definition von z_{CR} erinnert. Es gilt nach (3.29), der Orthogonalität in der Helmholtz Zerlegung, Lemma 3.13, und Lem-

ma 2.5

$$\begin{aligned}\|\!|z_{\mathrm{CR}}|\!\|_{\mathrm{NC}}^2 &= \int_\Omega \mathrm{D}_{\mathrm{NC}}\, z_{\mathrm{CR}} : \mathrm{D}_{\mathrm{NC}}\, z_{\mathrm{CR}}\ \mathrm{d}x \\ &= \int_\Omega \mathrm{D}_{\mathrm{NC}}\, z_{\mathrm{CR}} : (\operatorname{dev}(\Pi_0 \boldsymbol{\tau} - \mathrm{D}_{\mathrm{NC}}\, \alpha_{\mathrm{CR}}) + \operatorname{dev}\operatorname{Curl}\beta_c)\ \mathrm{d}x \\ &= \int_\Omega \mathrm{D}_{\mathrm{NC}}\, z_{\mathrm{CR}} : (\Pi_0 \boldsymbol{\tau} - \mathrm{D}_{\mathrm{NC}}\, \alpha_{\mathrm{CR}})\ \mathrm{d}x.\end{aligned}$$

Da insbesondere $z_{\mathrm{CR}} \in \mathrm{CR}_0^1(\mathcal{T};\mathbb{R}^2)$, ergibt die Gleichung (3.28)

$$\begin{aligned}\|\!|z_{\mathrm{CR}}|\!\|_{\mathrm{NC}}^2 &= \int_\Omega \mathrm{D}_{\mathrm{NC}}\, z_{\mathrm{CR}} : \Pi_0 \boldsymbol{\tau}\ \mathrm{d}x - \int_\Omega z_{\mathrm{CR}} \cdot (1-\Pi_0) \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\ \mathrm{d}x - \int_\Omega \Pi_0 \boldsymbol{\tau} : \mathrm{D}_{\mathrm{NC}}\, z_{\mathrm{CR}}\ \mathrm{d}x \\ &= -\int_\Omega z_{\mathrm{CR}} \cdot (1-\Pi_0) \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\ \mathrm{d}x.\end{aligned}$$

Dies kann nun mit der Cauchy-Schwarz-Ungleichung und der Poincaré Ungleichung, Lemma 2.17, abgeschätzt werden

$$\begin{aligned}\|\!|z_{\mathrm{CR}}|\!\|_{\mathrm{NC}}^2 &= -\int_\Omega (1-\Pi_0) z_{\mathrm{CR}} \cdot (1-\Pi_0) \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\ \mathrm{d}x \\ &\leq \|(1-\Pi_0) z_{\mathrm{CR}}\|_{L^2(\Omega)} \|(1-\Pi_0) \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\|_{L^2(\Omega)} \\ &\leq \|(1-\Pi_0) \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\|_{L^2(\Omega)} \sum_{T\in\mathcal{T}} \Big\| z_{\mathrm{CR}} - \fint_T z_{\mathrm{CR}}\ \mathrm{d}x \Big\|_{L^2(T)} \\ &\leq \|(1-\Pi_0) \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\|_{L^2(\Omega)} \sum_{T\in\mathcal{T}} h_T/j_{1,1}\ \|\mathrm{D}\, z_{\mathrm{CR}}\|_{L^2(T)} \\ &\leq h_{\max}/j_{1,1} \|\!|z_{\mathrm{CR}}|\!\|_{\mathrm{NC}} \|\operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\|_{L^2(\Omega)},\end{aligned}$$

wobei $j_{1,1}$ wie in Lemma 3.17 die erste Wurzel der ersten Bessel-Funktion ist. Insgesamt ergibt sich somit

$$\begin{aligned}\|F \circ (1-\Pi)\|_{Y^*} &= \sup_{(v,\boldsymbol{\tau})\in S(Y)} \int_\Omega f \cdot (v - \Pi v)\ \mathrm{d}x \\ &\leq \sup_{(v,\boldsymbol{\tau})\in S(Y)} \Big(h_{\max} C_{\mathrm{dF}}/j_{1,1}\ \|f\|_{L^2(\Omega)} \|\operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}\|_{L^2(\Omega)} \\ &\qquad\qquad + \kappa \|h_{\mathcal{T}} f\|_{L^2(\Omega)}\, \|\!|v|\!\|_{\mathrm{NC}} \Big) \\ &\leq h_{\max} C_{\mathrm{dF}}/j_{1,1}\ \|f\|_{L^2(\Omega)} + \kappa \|h_{\mathcal{T}} f\|_{L^2(\Omega)}.\end{aligned}$$

Dies ist kein Term höherer Ordnung. Für das Poisson-Modell Problem und für die Navier-Lamé Gleichungen in [CH15, 6.2] wurde gezeigt, dass eine leichte Vergrößerung des Testraums, welche keinen Einfluss auf den Beweis der diskreten und kontinuierlichen Inf-Sup Bedingung hat, die höhere Ordnung liefert. Dabei werden beispielsweise bubble-Funktionen bzw. RT_1-Funktionen verwendet. Ein ähnliches Vorgehen würde sich auch für die Stokes Gleichungen anbieten.

Zusammenfassend ist festzustellen, dass der in dieser Arbeit zur Verringerung des Re-

chenaufwands minimal gehaltende Testraum, im Allgemeinen nicht zu einen Datenapproximationsfehler von höherer Ordnung führt. Da die a-posteriori-Abschätzung aus [CDG14] allerdings für jedes Element des Ansatzraumes gilt, ist es möglich die Lösung des Problems mit kleinerem Testraum mit dem Datenfehler einer anderen Methode mit vergrößertem Testraum zu kombinieren. Die diskrete Lösung $x_h \in X_h$ zu Problem 10 kann in (3.1) mit erweitertem Testraum $\hat{Y}_h \supset Y_h$ eingesetzt werden. In diesem Fall wäre als Fehlerschätzer $\|F - b(x_h, \bullet)\|_{\hat{Y}_h^*}$ zu verwenden und der Datenapproximationsfehler $\|F \circ (1 - \hat{\Pi})\|_{Y^*}$ mit dem Fortin-Operator $\hat{\Pi}\colon Y \to \hat{Y}_h$ wäre von höherer Ordnung. Es würde sich also beispielsweise anbieten $\|F - b(x_h, \bullet)\|_{\hat{Y}_h^*}$ bei der adaptiven Verfeinerung zu betrachten.

Diese Fortin-Operatoren und Vergrößerungen des Testraumes sind aktuell Forschungsgegenstand. Es wird untersucht in wie weit die numerischen Lösungen und die Fehlerschätzer davon beeinflusst werden. Interessant wäre in diesem Kontext auch die Betrachtung inhomogener Randdaten, also die Einbeziehung des Terms $\left\langle (1 - \Pi)\gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}, \hat{s} \right\rangle_{\partial\mathcal{T}}$.

4 Implementierung

In diesem Kapitel werden die Rechnungen vorgestellt, die der Implementierung der Methode in Dimension $n = 2$ zu Grunde liegen, und wesentliche Informationen zu den als Erweiterung des AFEM-Software-Paketes [Car09a] implementierten Programmen gegeben. Der entstandene Code befindet sich vollständig auf der dieser Arbeit beiliegenden CD.

Zunächst wird in Abschnitt 4.1 der allgemeine Aufbau des zu lösenden Gleichungssystems und eines impliziten Fehlerschätzers beschrieben. Anschließend werden in Abschnitt 4.2 die verwendeten Basisfunktionen vorgestellt, mit denen in Abschnitt 4.3 die in (4.1) benötigten Matrizen explizit berechnet werden. Diese Rechnungen werden durch die entsprechenden Codezeilen ergänzt. Der Paragraph 4.4 widmet sich dem Problem des exakten Fehlers. Danach beschäftigt sich Abschnitt 4.5 mit der Gebietsskalierung, um verschiedene Gewichtungen der Fehlerterme zu vermeiden. Zuletzt werden in Abschnitt 4.6 die einzelnen implementierten Programme kurz vorgestellt.

4.1 Löser und Fehlerschätzer

Zur Lösung des Problems 10 ist zunächst eine Basis für die endlich dimensionalen Unterräume X_h und Y_h zu wählen. Für die theoretischen Untersuchungen seien $X_h = \operatorname{span}\{\xi_1, \ldots, \xi_N\}$ und $Y_h = \operatorname{span}\{\zeta_1, \ldots, \zeta_M\}$. Bezüglich dieser Basen sei die Steifigkeitsmatrix $\mathbb{B} \in \mathbb{R}^{M\times N}$ definiert durch

$$\mathbb{B}_{m,n} := b(\xi_n, \zeta_m) \quad \text{für } n = 1, \ldots, N \text{ und } m = 1, \ldots, M.$$

Die positiv definite und symmetrische Normmatrix $\mathbb{M} \in \mathbb{R}^{M\times M}$ sei so definiert, dass für jeden Koeffizientenvektor $y \in \mathbb{R}^M$ zu $y_h = \sum_{m=1}^{M} y_m\zeta_m \in Y_h$ gilt

$$\|y_h\|_Y^2 = y \cdot \mathbb{M}y.$$

Da Y_h in der gewählten Diskretisierung ein Hilbertraum ist, ist folgende Definition von $\mathbb{M}$ möglich

$$\mathbb{M}_{m,k} := \langle \zeta_m, \zeta_k \rangle_Y \quad \text{für } m, k = 1, \ldots, M.$$

Außerdem sei der Vektor der rechten Seite $\mathbb{F} \in \mathbb{R}^M$ als

$$\mathbb{F}_m := F(\zeta_m) \quad \text{für } m = 1, \ldots, M$$

festgelegt.

In der in Abschnitt 4.2 gewählten Basis für X_h wird die lineare Nebenbedingung, $\int_\Omega \operatorname{tr} \boldsymbol{\sigma}_0 \, \mathrm{d}x = 0$ für alle $x_h = (\boldsymbol{\sigma}_0, u_0, s_1, t_0) \in X_h$ nicht direkt realisiert. Sei daher

$\Lambda : X_h \to \mathbb{R}$ das lineare Funktional, das diese Nebenbedingung, durch $\Lambda(x_h) = 0$ beschreibt und $\mathbb{L} \in \mathbb{R}^N$ der zugehörige Vektor

$$\mathbb{L}_n := \Lambda(\xi_n) \quad \text{für } n = 1, \ldots, N.$$

Mit Hilfe dieser Definitionen kann das gegebene Minimierungsproblem als Lösung eines linearen Gleichungssystems erkannt werden.

Theorem 4.1. *Es sei $x_h \in X_h$ und $x \in \mathbb{R}^N$ der zugehörige Koeffizientenvektor mit $x_h = \sum_{n=1}^N x_n \xi_n$, dann gilt*

$$x_h \in \operatorname*{argmin}_{\xi_h \in X_h} \|F - b(\xi_h, \bullet)\|_{Y_h^*}$$

genau dann, wenn für ein $z \in \mathbb{R}$ gilt

$$\begin{pmatrix} \mathbb{B}^\top \mathbb{M}^{-1} \mathbb{B} & \mathbb{L} \\ \mathbb{L}^\top & 0 \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} \mathbb{B}^\top \mathbb{M}^{-1} \mathbb{F} \\ 0 \end{pmatrix}. \tag{4.1}$$

Beweis. Dieser Beweis ist ähnlich zu dem in [Hel14, Satz 7.10]. Da die Norm nicht negativ ist, kann statt $x_h \in \operatorname{argmin}_{\xi_h \in X_h} \|F - b(\xi_h, \bullet)\|_{Y_h^*}$ auch

$$x_h \in \operatorname*{argmin}_{\xi_h \in X_h} \|F - b(\xi_h, \bullet)\|_{Y_h^*}^2$$

untersucht werden. Es sei $\xi \in \mathbb{R}^N$ der Koeffizientenvektor zu einem beliebigen $\xi_h \in X_h$ und $y \in \mathbb{R}^M$ der Koeffizientenvektor zu einem beliebigen $y_h \in Y_h$. Dann gilt nach den obigen Definitionen $b(\xi_h, y_h) = y \cdot \mathbb{B}\xi$, $\|y_h\|_Y^2 = y \cdot \mathbb{M} y$ und $F(y_h) = y \cdot \mathbb{F}$. So ergibt sich nach Definition der Dualnorm

$$\|F - b(\xi_h, \bullet)\|_{Y_h^*}^2 = \sup_{\substack{y_h \in Y_h, \\ \|y_h\|_Y = 1}} (F(y_h) - b(\xi_h, y_h))^2 = \sup_{\substack{y \in \mathbb{R}^M, \\ y \cdot \mathbb{M} y = 1}} (y \cdot (\mathbb{F} - \mathbb{M}\xi))^2.$$

Aus Gründen der Übersichtlichkeit sei $t := (\mathbb{F} - \mathbb{M}\xi)$. Es handelt sich also um die Extremalaufgabe $\sup_y (y \cdot t)^2$ zu finden unter der Nebenbedingung $y \in \varphi^{-1}(0)$, wobei $\varphi(y) := y \cdot \mathbb{M} y - 1$. Daher kann die Multiplikatorregel von Lagrange angewendet werden [Kön04, S.124]. Der Gradient der Zielfunktion $f(y) := (y \cdot t)^2$ ist $\nabla f(y) = 2t$ und für die Nebenbedingungsfunktion gilt $\nabla \varphi(y) = 2\mathbb{M} y$. Nach der Multiplikatorregel ist $y_0 \in \mathbb{R}^M$ eine Extremalstelle, wenn $\nabla f(y_0) \in \operatorname{span}\{\nabla \varphi(y_0)\}$, d.h. für ein $\lambda \in \mathbb{R}$ gilt

$$t = \lambda \mathbb{M} y_0. \tag{4.2}$$

Diese Bedingung ist bereits hinreichend für eine Maximalstelle y_0, da die stetige Zielfunktion f ihr Maximum auf der kompakten Menge $\varphi^{-1}(0)$ annimmt.Außerdem erfüllt y_0 die Nebenbedingung, also

$$y_0 \cdot \mathbb{M} y_0 = 1. \tag{4.3}$$

Auf der einen Seite ergibt (4.3) eingesetzt in (4.2) die Gleichung $y_0 \cdot t = \lambda$, d.h.

$$\sup_{\substack{y\in\mathbb{R}^M,\\ y\cdot\mathbb{M}y=1}} f(y) = f(y_0) = (y_0 \cdot t)^2 = \lambda^2. \tag{4.4}$$

Andererseits ist die Normmatrix $\mathbb{M}$ positiv definit und symmetrisch also invertierbar. Daher ergibt sich aus (4.2) $\mathbb{M}^{-1}t = \lambda y_0$ und damit $t \cdot \mathbb{M}^{-1}t = \lambda t \cdot y_0 = \lambda y_0 \cdot t = \lambda^2$ nach Gleichung (4.4)

$$\max_{\substack{y\in\mathbb{R}^M,\\ y\cdot\mathbb{M}y=1}} f(y) = (y_0 \cdot t)^2 = \lambda^2 = t \cdot \mathbb{M}^{-1}t.$$

Die gesuchte Norm kann daher wie folgt geschrieben werden

$$\|F - b(\xi_h, \bullet)\|^2_{Y_h^*} = (\mathbb{F} - \mathbb{B}\xi) \cdot \mathbb{M}^{-1} (\mathbb{F} - \mathbb{B}\xi). \tag{4.5}$$

Durch Ausmultiplizieren ergibt sich nun

$$\|F - b(\xi_h, \bullet)\|^2_{Y_h^*} = \xi \cdot \left(\mathbb{B}^\top\mathbb{M}\mathbb{B}\right)\xi - 2\left(\mathbb{F} \cdot \mathbb{M}^{-1}\mathbb{B}\right)\xi + \mathbb{F} \cdot \mathbb{M}^{-1}\mathbb{F}.$$

Die Definition von $\mathbb{L}$ bedeutet, dass $\mathbb{L} \cdot \xi = \Lambda(\xi)$, daher gilt $\xi_h \in X_h$ sobald für den Koeffizientenvektor $\mathbb{L} \cdot \xi = 0$. Also ergibt sich folgende Äquivalenz

$$\begin{aligned}
&x_h \in \operatorname*{argmin}_{\xi_h\in X_h} \|F - b(\xi_h, \bullet)\|^2_{Y_h^*} \\
&\Leftrightarrow x \in \operatorname*{argmin}_{\substack{\xi\in\mathbb{R}^N,\\ \mathbb{L}\cdot\xi=0}} \xi \cdot \left(\mathbb{B}^\top\mathbb{M}\mathbb{B}\right)\xi - 2\left(\mathbb{F} \cdot \mathbb{M}^{-1}\mathbb{B}\right)\xi + \mathbb{F} \cdot \mathbb{M}^{-1}\mathbb{F} \\
&\Leftrightarrow x \in \operatorname*{argmin}_{\substack{\xi\in\mathbb{R}^N,\\ \mathbb{L}\cdot\xi=0}} 1/2\ \xi \cdot \left(\mathbb{B}^\top\mathbb{M}\mathbb{B}\right)\xi - \left(\mathbb{F} \cdot \mathbb{M}^{-1}\mathbb{B}\right)\xi.
\end{aligned}$$

Eine erneute Anwendung der Multiplikatorregel von Lagrange mit Parameter $z \in \mathbb{R}$ zur Berechnung einer Minimalstelle x unter Nebenbedingung $\mathbb{L} \cdot x = 0$ ergibt

$$\left(\mathbb{B}^\top\mathbb{M}\mathbb{B}\right)x - \mathbb{B}^\top\mathbb{M}^{-1}\mathbb{F} = -z\mathbb{L}.$$

Dies und die Nebenbedingung $\mathbb{L} \cdot x = 0$ ergeben das behauptete Gleichungssystem. □

In Abschnitt 3.1 wurde die folgende Fehlerabschätzung für $\xi_h \in X_h$ und die exakte Lösung $x \in X$ bei einer dPG-Methode vorgestellt

$$\|x - \xi_h\|_X \lesssim \|F - b(\xi_h, \bullet)\|_{Y_h^*} + \|F \circ (1 - \Pi)\|_{Y^*} \lesssim \|x - \xi_h\|_X.$$

Im Allgemeinen ist dabei der Datenapproximationsfehler $\|F \circ (1 - \Pi)\|_{Y^*}$ von höherer Ordnung. In Abschnitt 3.4 konnte dies leider nicht nachgewiesen werden, doch numerische Beispiele zeigen, dass

$$\eta := \|F - b(x_h, \bullet)\|_{Y_h^*}$$

ein angemessener Fehlerschätzer ist, wenn $x_h \in X_h$ die diskrete Lösung bezeichnet.

Bemerkung 4.2 (Globaler und lokaler Fehlerschätzer) Sei $x \in \mathbb{R}^N$ der Koeffizientenvektor der diskreten Lösung $x_h \in X_h$. Aus dem Beweis von Theorem 4.1 ist in Gleichung (4.5) direkt eine Formel zur Berechnung von η abzulesen. Ein geeigneter globaler Fehlerschätzer kann mittels

$$\eta^2 = \|F - b(x_h, \bullet)\|^2_{Y_h^*} = (\mathbb{F} - \mathbb{B}x) \cdot \mathbb{M}^{-1} (\mathbb{F} - \mathbb{B}x)$$

berechnet werden. Bei den Basisfunktionen aus Abschnitt 4.2 existiert für jede der $1 \leq m \leq M$ gebrochenen Testfunktionen genau ein $T \in \mathcal{T}$, so dass $\operatorname{supp} \zeta_m \subseteq T$. Damit gilt für

$$Y_h(T) := \{\zeta_m : \ 1 \leq m \leq M \text{ und } \operatorname{supp} \zeta_m \subseteq T\},$$

dass $Y_h = \bigcup_{T \in \mathcal{T}} Y_h(T)$, und dass die einzelnen $Y_h(T)$ disjunkt sind. Daher zerfällt die Normmatrix $\mathbb{M}$ in lokale Blöcke, wie in Abschnitt 4.3.2 deutlich wird,

$$\mathbb{M}(T) := \begin{bmatrix} \boldsymbol{T}(T) & 0 \\ 0 & \boldsymbol{V}(T) \end{bmatrix} \in \mathbb{R}^{12 \times 12},$$

wobei $\mathbb{M}(T)$, $\boldsymbol{T}(T)$ und $\boldsymbol{V}(T)$ jeweils alle Beiträge von Basisfunktionen aus $Y_h(T)$ enthalten. Damit kann für die gewählte Basis ein lokaler Fehlerschätzer für $T \in \mathcal{T}$ durch

$$\eta^2(T) := \|F - b(x_h, \bullet)\|^2_{Y_h(T)^*} = (\mathbb{F} - \mathbb{B}x) \cdot \mathbb{M}^{-1}(T) (\mathbb{F} - \mathbb{B}x)$$

definiert werden. Aufgrund der Blockstruktur der inversen Normmatrix (siehe Abschnitt 4.3.2) und der erwähnten Eigenschaften von $Y_h(T)$ ergibt sich

$$\eta^2 = \sum_{T \in \mathcal{T}} \eta^2(T).$$

In dem Programm wird dieser lokale Fehlerschätzer zur Adaptiven Verfeinerung verwendet.

Um die Matrizen $\mathbb{B}$, $\mathbb{M}$, $\mathbb{F}$ und $\mathbb{L}$ zur Lösung des linearen Gleichungssystems (4.1) in Abschnitt 4.3 berechnen zu können, wird zunächst im folgenden Paragraphen die gewählte Basis vorgestellt.

4.2 Basiswahl

Bei der Basiswahl für die diskreten Räume X_h und Y_h zu der Triangulierung $\mathcal{T}$ werden die linearen Nebenbedingungen zunächst nicht berücksichtigt. Im Ansatzraum wird die Nebenbedingung, wie in Abschnitt 4.1 erläutert, durch einen Lagrange-Multiplikator mittels des Vektors $\mathbb{L}$ garantiert und im Testraum wird sie laut dem folgenden Lemma 4.3 nicht benötigt.

Lemma 4.3. *Sei* $\hat{Y}_h = RT_0^{pw}\left(\mathcal{T};\mathbb{R}^{2\times 2}\right) \times P_1\left(\mathcal{T};\mathbb{R}^2\right)$ *und* Y_h *wie in (2.12), dann gilt für jedes* $x_h = (\boldsymbol{\sigma}_0, u_0, s_1, t_0) \in X_h$ *wie in (2.11)*

$$\sup_{\hat{y}_h \in \hat{Y}_h \setminus \{0\}} \frac{b(x_h, \hat{y}_h)}{\|\hat{y}_h\|_Y} = \sup_{y_h \in Y_h \setminus \{0\}} \frac{b(x_h, y_h)}{\|y_h\|_Y}.$$

Beweis. Dieser Beweis verläuft analog zu dem von [Hel14, Lemma 5.8]. Zunächst gilt wegen $Y_h \subseteq \hat{Y}_h$, dass

$$\sup_{y_h \in Y_h \setminus \{0\}} \frac{b(x_h, y_h)}{\|y_h\|_Y} \leq \sup_{\hat{y}_h \in \hat{Y}_h \setminus \{0\}} \frac{b(x_h, \hat{y}_h)}{\|\hat{y}_h\|_Y}.$$

Um die andere Abschätzung zu zeigen, sei $\hat{y}_h = (\hat{\boldsymbol{\tau}}_{\mathrm{RT}}, v_1) \in \hat{Y}_h \setminus \{0\}$ beliebig. Wähle $y_h = (\boldsymbol{\tau}_{\mathrm{RT}}, v_1) \in Y_h$ mit $\boldsymbol{\tau}_{\mathrm{RT}} = \hat{\boldsymbol{\tau}}_{\mathrm{RT}} - c\mathrm{I}_{2\times 2}$ wobei $c = \fint_\Omega \operatorname{tr} \hat{\boldsymbol{\tau}}_{\mathrm{RT}} \,\mathrm{d}x/2$. Es gilt

$$\langle y_h, (c\mathrm{I}_{2\times 2}, 0)\rangle_Y = \int_\Omega \boldsymbol{\tau}_{\mathrm{RT}} : c\mathrm{I}_{2\times 2} \,\mathrm{d}x + \int_\Omega \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}_{\mathrm{RT}} \cdot \operatorname{div}_{\mathrm{NC}} \mathrm{I}_{2\times 2} \,\mathrm{d}x = c\int_\Omega \operatorname{tr} \boldsymbol{\tau}_{\mathrm{RT}} \,\mathrm{d}x = 0$$

und damit nach Satz des Pythagoras

$$\|y_h\|_Y^2 \leq \|y_h\|_Y^2 + \|(c\mathrm{I}_{2\times 2}, 0)\|_Y^2 = \|\hat{y}_h\|_Y^2 .$$

Außerdem gilt wegen den Randbedingungen an s_1, der Nebenbedingung an $\boldsymbol{\sigma}_0$ und Lemma 2.5

$$\begin{aligned} b\left(x_h, (c\mathrm{I}_{2\times 2}, 0)\right) &= -\left\langle \gamma_\nu^{\mathcal{T}} c\mathrm{I}_{2\times 2}, s_1\right\rangle_{\partial\mathcal{T}} + \int_\Omega \operatorname{dev} \boldsymbol{\sigma}_0 : c\mathrm{I}_{2\times 2} \,\mathrm{d}x + \int_\Omega u \cdot \operatorname{div}_{\mathrm{NC}} c\mathrm{I}_{2\times 2} \,\mathrm{d}x \\ &= -\int_{\partial\Omega} \gamma_\nu c\mathrm{I}_{2\times 2} \cdot s_1 \,\mathrm{d}s + c\int_\Omega \operatorname{tr} \boldsymbol{\sigma}_0 \,\mathrm{d}x = 0 \end{aligned}$$

und damit $b(x_h, y_h) = b(x_h, \hat{y}_h)$. Da $\hat{y}_h$ beliebig war, ergibt sich

$$\sup_{\hat{y}_h \in \hat{Y}_h \setminus \{0\}} \frac{b(x_h, \hat{y}_h)}{\|\hat{y}_h\|_Y} \leq \sup_{y_h \in Y_h \setminus \{0\}} \frac{b(x_h, y_h)}{\|y_h\|_Y},$$

und somit die Behauptung. □

Im Folgenden werden Basen für die neu definierten Räume

$$\begin{aligned} X_h :=& P_0\left(\mathcal{T};\mathbb{R}^{2\times 2}\right) \times P_0\left(\mathcal{T};\mathbb{R}^2\right) \times S_0^1\left(\mathcal{E};\mathbb{R}^2\right) \times P_0\left(\mathcal{E};\mathbb{R}^2\right), \\ Y_h :=& RT_0^{\mathrm{pw}}\left(\mathcal{T};\mathbb{R}^{2\times 2}\right) \times P_1\left(\mathcal{T};\mathbb{R}^2\right) \end{aligned}$$

vorgestellt.

Dazu werden die Dreiecke $T_1, \ldots, T_J$ mit $J := |\mathcal{T}|$ der regulären Triangulierung $\mathcal{T}$ des polygonal berandeten Lipschitz Gebietes $\Omega \subseteq \mathbb{R}^2$, die Menge der Knoten $\mathcal{N} = \left\{z_1, \ldots, z_{|\mathcal{N}|}\right\}$ und der Kanten $\mathcal{E} = \left\{E_1, \ldots, E_{|\mathcal{E}|}\right\}$ betrachtet. Die lokale Nummerierung

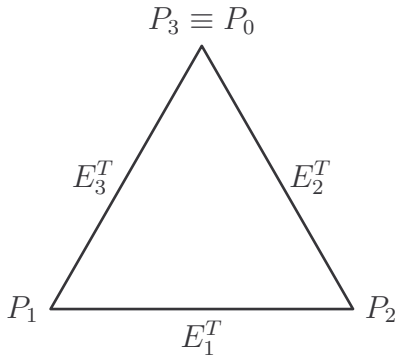


Abbildung 4.1: Lokale Nummerierung der Seiten und Eckpunkte eines Dreiecks $T \in \mathcal{T}$

der Seiten und Eckpunkte in einem Dreieck $T \in \mathcal{T}$ ist dabei stets wie in Abbildung 4.1 dargestellt.

Daneben wird als Anzahl der Freiheitsgrade mit nrDoF$= \dim(X_h)$ die Dimension des diskreten Ansatzraums bezeichnet.
Des Weiteren sei im Folgenden $\chi(T) \in L^2(\Omega)$ die charakteristische Funktion für ein Dreieck $T \in \mathcal{T}$, also gilt $\chi(T) \equiv 1$ auf T und $\chi(T) \equiv 0$ sonst, und $\chi(E) \in L^2(\partial\mathcal{T})$ die charakteristische Funktion für eine Kante $E \in \mathcal{E}$, also gilt $\chi(E) \equiv 1$ auf E und $\chi(E) \equiv 0$ sonst.
Außerdem sei die Standardbasis des $\mathbb{R}^2$ durch $e_1 := (1,0)^\top$ und $e_2 := (0,1)^\top$ gegeben, wohingegen

$$\mathbf{e}_1 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ \mathbf{e}_2 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ \mathbf{e}_3 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \mathbf{e}_4 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

die Basis des $\mathbb{R}^{2\times 2}$ sei.
Ebenfalls benötigt werden φ_k, die konforme nodale Basisfunktion erster Ordnung zum Knoten $z_k \in \mathcal{N}$ für $k = 1, \ldots, |\mathcal{N}|$ respektive der Triangulierung $\mathcal{T}$, d.h. $\varphi_k \in P_1(\mathcal{T})$ mit $\varphi_k(z_j) = \delta_{j,k}$ für $z_j \in \mathcal{N}$. Für jedes $T \in \mathcal{T}$ und $\gamma = 1, 2, 3$ sei die lokale nodale Basisfunktion $\varphi(T,\gamma) := \chi(T)\varphi_k$, wobei γ die lokale Nummer des Knoten z_k in T sei.
Schließlich werden die lokalen kantenbasierten Raviart-Thomas Basisfunktionen für jedes $T \in \mathcal{T}$ und die lokale Kantennummerierung $\alpha = 1, 2, 3$ mit

$$\psi(T,\alpha) := \chi(T)\,(x - P_{\alpha-1})\left|E_\alpha^T\right| / |2T|$$

bezeichnet, wobei nach der lokalen Nummerierung aus Abbildung 4.1 $P_{\alpha-1}$ der Eckpunkt gegenüber der Seite E_α^T in T ist.

Aus Gründen der Übersichtlichkeit werden im Folgenden mit $\boldsymbol{\sigma}^b$ Basisfunktionen zur Lösungskomponente $\boldsymbol{\sigma}_0$ bezeichnet, mit u^b die zu u_0, mit s^b die zu s_1, mit t^b die zu t_0 und genauso mit $\boldsymbol{\tau}^b$ die Basisfunktionen zur Komponente $\boldsymbol{\tau}_{\mathrm{RT}}$ und mit v^b die zu v_1. In der Basis der Produkträume X_h und Y_h werden für jeden Teilraum die Basisfunktionen,

die Träger auf einem Element $T \in \mathcal{T}$ haben, hintereinander sortiert. Dies ermöglicht eine Blockstruktur der Matrix.

Spannungsansatzfunktionen. Für die Elemente $j = 1, \ldots, J$ und die Komponenten $\lambda = 1, \ldots, 4$ sei durch

$$\boldsymbol{\sigma}^b_{4(j-1)+\lambda} := \boldsymbol{\sigma}^B\left(T_j, \lambda\right) := \chi(T_j)\ \mathbf{e}_\lambda$$

die Basis von $P_0\left(\mathcal{T}; \mathbb{R}^{2\times 2}\right) = \operatorname{span}\left\{\boldsymbol{\sigma}^b_1, \ldots, \boldsymbol{\sigma}^b_{4J}\right\}$ definiert.

Verschiebungsansatzfunktionen. Für die Elemente $j = 1, \ldots, J$ und die Komponenten $\kappa = 1, 2$ sei durch

$$u^b_{2(j-1)+\kappa} := u^B\left(T_j, \kappa\right) := \chi(T_j)\ e_\kappa$$

die Basis von $P_0\left(\mathcal{T}; \mathbb{R}^2\right) = \operatorname{span}\left\{u^b_1, \ldots, u^b_{2J}\right\}$ beschrieben.

Oberflächenverschiebungsfunktionen. Für die inneren Knoten $k = 1, \ldots, |\mathcal{N}(\Omega)|$ und die Komponenten $\kappa = 1, 2$ sei mit den nodalen Basisfunktionen durch

$$s^b_{2(k-1)+\kappa} := s^B\left(z_k, \kappa\right) := \gamma_0^{\mathcal{T}}\varphi_k\ \ e_\kappa$$

die Basis von $S^1_0\left(\mathcal{E}; \mathbb{R}^2\right) = \operatorname{span}\left\{s^b_1, \ldots, s^b_{2|\mathcal{N}(\Omega)|}\right\}$ gegeben.

Die Nullrandbedingungen werden im Programm wie üblich realisiert, in dem zwar die Beiträge für alle Knoten $z_k \in \mathcal{N}$ berechnet werden, das entstehende Gleichungssystem (4.1) allerdings nur in den freien Variablen, d.h. nur für die inneren Knoten, gelöst wird.

Traktionansatzfunktionen. Für die Kanten $\ell = 1, \ldots, |\mathcal{E}|$ und die Komponenten $\kappa = 1, 2$ sei durch

$$t^b_{2(\ell-1)+\kappa} := t^B\left(E_\ell, \kappa\right) := \chi(E_\ell)\ e_\kappa$$

die Basis von $P_0\left(\mathcal{E}; \mathbb{R}^2\right) = \operatorname{span}\left\{t^b_1, \ldots, t^b_{2|\mathcal{E}|}\right\}$ definiert.

Spannungstestfunktionen. Für die Elemente $j = 1, \ldots, J$, die lokalen Kanten $\alpha = 1, 2, 3$ und die Komponenten $\kappa = 1, 2$ sei mit den lokalen Raviart-Thomas Basisfunktionen durch

$$\boldsymbol{\tau}^b_{6(j-1)+2(\alpha-1)+\kappa} := \boldsymbol{\tau}^B\left(T_j, \alpha, \kappa\right) := e_\kappa \otimes \psi\left(T_j, \alpha\right)$$

die Basis von $RT^{\mathrm{pw}}_0\left(\mathcal{T}; \mathbb{R}^{2\times 2}\right) = \operatorname{span}\left\{\boldsymbol{\tau}^b_1, \ldots, \boldsymbol{\tau}^b_{6J}\right\}$ gegeben.

Verschiebungstestfunktionen Für die Elemente $j = 1, \ldots, J$, die Knoten $\gamma = 1, 2, 3$ und die Komponenten $\kappa = 1, 2$ definiere durch

$$v^b_{6(j-1)+2(\gamma-1)+\kappa} := v^B\left(T_j, \gamma, \kappa\right) := \varphi\left(T_j, \gamma\right)\ \ e_\kappa$$

die Basis von $P_1\left(\mathcal{T}; \mathbb{R}^2\right) = \operatorname{span}\left\{v^b_1, \ldots, v^b_{6J}\right\}$ mit den lokalen nodalen Basisfunktionen.

Damit sind die Koeffizientenvektoren $x \in \mathbb{R}^N$ mit $N = 6J + 2|\mathcal{N}(\Omega)| + 2|\mathcal{E}|$ und $y \in \mathbb{R}^M$ mit $M = 12J$ so organisiert, dass

$$x_h = \left(\sum_{n=1}^{4J} x_n \boldsymbol{\sigma}_n^b, \sum_{n=1}^{2J} x_{4J+n}\, u_n^b, \sum_{n=1}^{2|\mathcal{N}(\Omega)|} x_{6J+n}\, s_n^b, \sum_{n=1}^{2|\mathcal{E}|} x_{6J+2|\mathcal{N}(\Omega)|+n}\, t_n^b \right),$$
$$y_h = \left(\sum_{m=1}^{6J} x_m \boldsymbol{\tau}_m^b, \sum_{m=1}^{6J} x_{6J+m}\, v_m^b \right).$$

Daraus ergibt sich die in den folgenden Paragraphen erklärte Assemblierung der Koeffizientenmatrix $\mathbb{B}$ sowie der Normmatrix $\mathbb{M}$ und ihrer Inversen $\mathbb{M}^{-1}$ sowie der Vektoren der rechten Seite $\mathbb{F}$ und der linearen Nebenbedingung $\mathbb{L}$.

4.3 Berechnung der Bestandteile des linearen Gleichungssystems

4.3.1 Koeffizientenmatrix

Wie in Abschnitt 4.1 ist die Koeffizientenmatrix $\mathbb{B} \in \mathbb{R}^{M \times N}$ definiert durch

$$\mathbb{B}_{m,n} := b(\xi_n, \zeta_m) \quad \text{für } n = 1, \ldots, N \text{ und } m = 1, \ldots, M.$$

Damit gilt für den Koeffizientenvektor $x \in \mathbb{R}^N$ von $x_h = \sum_{n=1}^N x_n \xi_n \in X_h$ und den $y \in \mathbb{R}^N$ von $y_h = \sum_{m=1}^M y_n \zeta_n \in Y_h$, dass $b(x_h, y_h) = y \cdot \mathbb{B}x$.

Für die Bilinearform aus Gleichung (3.6)

$$\begin{aligned} b(x_h, y_h) := &- \langle t_0, \gamma_0 v_1 \rangle_{\partial\mathcal{T}} - \langle \gamma_\nu \boldsymbol{\tau}_{\mathrm{RT}}, s_1 \rangle_{\partial\mathcal{T}} \\ &+ \int_\Omega \boldsymbol{\sigma}_0 : \mathrm{D}_{\mathrm{NC}}\, v_1 \, \mathrm{d}x + \int_\Omega \operatorname{dev} \boldsymbol{\sigma}_0 : \boldsymbol{\tau}_{\mathrm{RT}} \, \mathrm{d}x + \int_\Omega u_0 \cdot \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}_{\mathrm{RT}} \, \mathrm{d}x. \end{aligned}$$

ergibt sich folgende $2 \times 4-$Blockstruktur

$$\mathbb{B} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} & \boldsymbol{C} & 0 \\ \boldsymbol{D} & 0 & 0 & \boldsymbol{E} \end{bmatrix} \tag{4.6}$$

mit $\boldsymbol{A} \in \mathbb{R}^{6J \times 4J}$, $\boldsymbol{B} \in \mathbb{R}^{6J \times 2J}$, $\boldsymbol{C} \in \mathbb{R}^{6J \times 2|\mathcal{N}(\Omega)|}$, $\boldsymbol{D} \in \mathbb{R}^{6J \times 4J}$ und $\boldsymbol{E} \in \mathbb{R}^{6J \times 2|\mathcal{E}|}$.

Die folgenden Matrizen vereinfachen die Notation dieser Blöcke. Für $T \in \mathcal{T}$ mit den lokalen Seiten E_1^T, E_2^T, E_3^T definiere

$$L(T) := \operatorname{diag}\left(\left|E_1^T\right|, \left|E_1^T\right|, \left|E_2^T\right|, \left|E_2^T\right|, \left|E_3^T\right|, \left|E_3^T\right| \right) \in \mathbb{R}^{6 \times 6} \tag{4.7}$$

und für $\alpha = 1, 2, 3$ sei $\operatorname{sgn}_\alpha^T := \operatorname{sgn}\left(T, E_\alpha^T\right) = \nu_{E_\alpha^T} \cdot \nu_T \in \{\pm 1\}$ wie in (2.1) und

$$S(T) := \operatorname{diag}\left(\operatorname{sgn}_1^T, \operatorname{sgn}_1^T, \operatorname{sgn}_2^T, \operatorname{sgn}_2^T, \operatorname{sgn}_3^T, \operatorname{sgn}_3^T\right) \in \mathbb{R}^{6\times 6}. \tag{4.8}$$

Hilfreich sind auch die folgenden konstanten Matrizen

$$H_1 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad H_2 := \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad H_3 := \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.9}$$

Im Folgenden wird die genaue Berechnung der in (4.6) erwähnten Blöcke erläutert. Die gebrochenen Testfunktionen haben jeweils nur ein Element als Träger, daher müssen die vorkommenden Integrale über das Gebiet Ω bzw. das Skelett $\partial\mathcal{T}$ jeweils nur auf diesem einen Dreieck bzw. seinem Rand ausgewertet werden.

Block A enthält die Anteile der Bilinearform für $\boldsymbol{\sigma}_0 \in P_0\left(\mathcal{T}; \mathbb{R}^{2\times 2}\right)$ und $\boldsymbol{\tau}_{\mathrm{RT}} \in RT_0^{\mathrm{pw}}\left(\mathcal{T}; \mathbb{R}^{2\times 2}\right)$, die Einträge ergeben sich also aus folgendem Integral der Basisfunktionen

$$\boldsymbol{A}_{mn} = \int_\Omega \boldsymbol{\tau}_m^b : \operatorname{dev} \boldsymbol{\sigma}_n^b \, \mathrm{d}x \quad \text{für } m = 1, \ldots, 6J \text{ und } n = 1, \ldots, 4J.$$

Dieses Integral verschwindet, es sei denn $\boldsymbol{\tau}_m^b$ und $\boldsymbol{\sigma}_n^b$ haben als Träger dasselbe Dreieck $T \in \mathcal{T}$, d.h. es gilt $\boldsymbol{\tau}_m^b = \boldsymbol{\tau}^B(T, \alpha, \kappa)$ und $\boldsymbol{\sigma}_n^b = \boldsymbol{\sigma}^B(T, \lambda)$ für ein $T \in \mathcal{T}$, $\alpha = 1, 2, 3$, $\kappa = 1, 2$ und $\lambda = 1, \ldots, 4$.

Daher ist es angebracht die lokalen Blöcke, die Integrale über Basisfunktionen in $Y_h(T)$ bzw. $X_h(T)$ für ein $T \in \mathcal{T}$, zu betrachten.

Lemma 4.4. *Für jedes Dreieck $T \in \mathcal{T}$ sei der lokale Block $\boldsymbol{A}(T)$ definiert via*

$$\boldsymbol{A}\left(T; 2(\alpha-1)+\kappa, \lambda\right) := \int_T \boldsymbol{\tau}^B(T, \alpha, \kappa) : \operatorname{dev} \boldsymbol{\sigma}^B(T, \lambda) \, \mathrm{d}x$$

für $\kappa = 1, 2$, $\alpha = 1, 2, 3$ und $\lambda = 1, 2, 3, 4$. Es gilt

$$\boldsymbol{A}\left(T; 2(\alpha-1)+\kappa, \lambda\right) = \begin{cases} 0 & \text{für } \lambda = 1, \\ \mathbf{e}_\lambda : \left|E_\alpha^T\right| / 2 \left(e_k \otimes \left(\operatorname{mid}(T) - P_{\alpha-1}\right)\right) & \text{für } \lambda = 2, 3, 4, \end{cases}$$

d.h. der Block $\boldsymbol{A}(T)$ hat folgende Form

$$\boldsymbol{A}(T) = \frac{1}{2} \, L(T) \left(H_1 \, \mathrm{K}(T) \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} + H_2 \, \mathrm{K}(T) \begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & -1 & 0 & 0 \end{pmatrix} \right) \in \mathbb{R}^{6\times 4},$$

vermöge der Matrix $L(T)$ aus (4.7) und der konstanten Matrizen aus (4.9) und mit der lokalen Matrix $\mathrm{K}(T) \in \mathbb{R}^{3\times 2}$, die für die lokalen Seiten E_α^T mit $\alpha = 1, 2, 3$ und nach Abbildung 4.1 gegenüberliegenden Eckpunkt $P_{\alpha-1}$ und $\kappa = 1, 2$, die Einträge

$$\mathrm{K}(T;\alpha,\kappa) := (\mathrm{mid}(T) - P_{\alpha-1}) \cdot e_\kappa$$

hat.

Beweis. Auf dem Element $T \in \mathcal{T}$ verschwindet der Deviator von $\boldsymbol{\sigma}^B(T,\lambda)$ für $\lambda = 1$, wohingegen die Funktionen $\boldsymbol{\sigma}^B(T,\lambda)$ für $\lambda = 2, 3, 4$ deviatorisch sind. Außerdem können die stückweise linearen lokalen Raviart-Thomas Funktionen mit der Mittelpunktsformel exakt integriert werden. Also gilt

$$\int_\Omega \psi(T,\alpha) = \int_T (x - P_{\alpha-1})\left|E_\alpha^T\right| / (2|T|)\ \mathrm{d}x = (\mathrm{mid}(T) - P_{\alpha-1})\left|E_\alpha^T\right| /2.$$

Daher ergibt sich für $\lambda = 2, 3, 4$, $\alpha = 1, 2, 3$ und $\kappa = 1, 2$

$$\begin{aligned}
&\int_T \boldsymbol{\tau}(T,\alpha,\kappa) : \mathrm{dev}\,\boldsymbol{\sigma}(T,\lambda)\ \mathrm{d}x \\
&\qquad = \mathbf{e}_\lambda : \left|E_\alpha^T\right| /2 \left(e_\kappa \otimes (\mathrm{mid}(T) - P_{\alpha-1})\right) \\
&\qquad = \begin{cases}
\left|E_\alpha^T\right| /2\,(\mathrm{mid}(T) - P_{\alpha-1}) \cdot e_1 & \text{für } (\lambda = 2, \kappa = 1),\ (\lambda = 3, \kappa = 2), \\
-\left|E_\alpha^T\right| /2\,(\mathrm{mid}(T) - P_{\alpha-1}) \cdot e_1 & \text{für } (\lambda = 4, \kappa = 2), \\
\left|E_\alpha^T\right| /2\,(\mathrm{mid}(T) - P_{\alpha-1}) \cdot e_2 & \text{für } (\lambda = 3, \kappa = 1),\ (\lambda = 4, \kappa = 1), \\
-\left|E_\alpha^T\right| /2\,(\mathrm{mid}(T) - P_{\alpha-1}) \cdot e_2 & \text{für } (\lambda = 2, \kappa = 2).
\end{cases}
\end{aligned}$$

$\boldsymbol{A}(T)$ hat also vermöge der Matrizen $K(T)$ und $L(T)$ die folgende Form

$$\boldsymbol{A}(T) = \frac{1}{2} L(T) \begin{pmatrix}
0 & K(T;1,1) & K(T;1,2) & K(T;1,2) \\
0 & -K(T;1,2) & K(T;1,1) & -K(T;1,1) \\
0 & K(T;2,1) & K(T;2,2) & K(T;2,2) \\
0 & -K(T;2,2) & K(T;2,1) & -K(T;2,1) \\
0 & K(T;3,1) & K(T;3,2) & K(T;3,2) \\
0 & -K(T;3,2) & K(T;3,1) & -K(T;3,1)
\end{pmatrix} \in \mathbb{R}^{6\times 4}.$$

Diese ist identisch zu dem behaupteten Matrixprodukt. □

In dem Code in `computeBlocksStokesDPG` bzw. `computeBlocksStokesDPGFass` bzw. `computeBlocksStokesDPGIntegrate` ist die Berechnung der lokalen Blöcke wie folgt implementiert

```
% initialisation of matrix blocks
A4e = zeros(6,4,nrElems);       % local matrices for block A
% constant auxillary matrices
H1 = sparse([1,3,5],[1,2,3],[1,1,1],6,3);
H2 = sparse([2,4,6],[1,2,3],[1,1,1],6,3);
```

```
%% Parfor-loop to fill local matrices
for elem = 1:nrElems            %elem as current triangle T
    coords  = coord4e(:,:,elem);        % coordinates of three nodes
    mid     = mid4e(elem,:);            % midpoint of T
    slength = sl4e(elem,:)';            % length of sides
    L       = [slength,slength]';  L=diag(L(:));
    K       = [mid;mid;mid]-[coords(3,:);coords(1,:);coords(2,:)];

    % block A int_T(dev(sigma):tau)
    A4e(:,:,elem)=1/2*L*(H1*K*[0,1,0,0;0,0,1,1]+H2*K*[0,0,1,-1;0,-1,0,0]);
end
```

Der Matrixblock $\boldsymbol{A}$ wird aus den lokalen Matrizen für alle $\alpha = 1, 2, 3$, $\kappa = 1, 2$, $\lambda = 1, 2, 3, 4$ und $j = 1, \ldots, J$ wie folgt assembliert

$$\boldsymbol{A}_{6(j-1)+2(\alpha-1)+\kappa,4(j-1)+\lambda} = \boldsymbol{A}\left(T_j; 2(\alpha-1)+\kappa, \lambda\right).$$

In diesem Fall, entsteht also folgende Blockdiagonalform

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}(T_1) & 0 & \ldots & 0 \\ 0 & \boldsymbol{A}(T_2) & \ldots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \ldots & 0 & \boldsymbol{A}(T_J) \end{pmatrix} \in \mathbb{R}^{6J\times 4J}.$$

Die globale Assemblierung geschieht in dem Programm mittels folgender Zeilen

```
%% Constructing global matrix blocks
% with tempI as vector of row indices and tempJ vector of column indices
tempI = permute(reshape(repmat(1:6*nrElems,4,1),4,6,nrElems),[2,1,3]);
tempJ = repmat(1:4*nrElems,6,1);
A=sparse(tempI(:),tempJ(:),A4e(:));     % block A
```

Block B enthält die Beiträge für $u_0 \in P_0\left(\mathcal{T};\mathbb{R}^2\right)$ und $\boldsymbol{\tau}_{\mathrm{RT}} \in RT_0^{\mathrm{pw}}\left(\mathcal{T};\mathbb{R}^{2\times 2}\right)$. Die Einträge ergeben sich also aus folgendem Integral

$$\boldsymbol{B}_{mn} = \int_\Omega \mathrm{div}_{\mathrm{NC}}\left(\boldsymbol{\tau}_m^b\right) \cdot u_n^b \,\mathrm{d}x \quad \text{für } m = 1, \ldots, 6J \text{ und } n = 1, \ldots, 2J.$$

Dieses Integral verschwindet, außer $\boldsymbol{\tau}_m^b$ und u_n^b haben dasselbe Dreieck $T \in \mathcal{T}$ als Träger, d.h. $\boldsymbol{\tau}_m^b = \boldsymbol{\tau}^B(T, \alpha, \kappa)$ und $u_n^b = u^B(T, \varkappa)$ für ein $T \in \mathcal{T}$, $\alpha = 1, 2, 3$ und $\kappa, \varkappa = 1, 2$, und in diesem Fall ist es lediglich ein Integral über dieses Element $T \in \mathcal{T}$

Daher ist es sinnvoll, die im folgenden Lemma berechneten lokalen Blöcke zu definieren.

Lemma 4.5. *Für jedes Dreieck $T \in \mathcal{T}$ sei der lokale Block $\boldsymbol{B}(T)$ definiert via*

$$\boldsymbol{B}(T; 2(\alpha-1)+\kappa, \varkappa) := \int_T \mathrm{div}\,\boldsymbol{\tau}^B(T;\alpha,\kappa) \cdot u^B(T,\varkappa)\,\mathrm{d}x$$

für $\alpha = 1, 2, 3$ und $\kappa, \varkappa = 1, 2$. Es gilt

$$\boldsymbol{B}(T; 2(\alpha-1)+\kappa, \varkappa) = \left|E_\alpha^T\right| \delta_{\kappa,\varkappa},$$

d.h. der Block $\boldsymbol{B}(T)$ lässt sich wie folgt berechnen

$$\boldsymbol{B}(T) = L(T) \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{6\times 2},$$

vermöge der Matrix $L(T)$ aus (4.7).

Beweis. Für die lokalen Raviart-Thomas Basisfunktionen $\psi\,(T,\alpha)$ gilt

$$\int_\Omega \operatorname{div} \psi\,(T,\alpha)\ \mathrm{d}x = \int_T \left|E_\alpha^T\right| / |T|\ \mathrm{d}x = \left|E_\alpha^T\right|.$$

Daher ergibt sich für $\kappa,\ \varkappa = 1,\ 2$ und $\alpha = 1,\ 2,\ 3$

$$\int_T \operatorname{div} \boldsymbol{\tau}^B\,(T,\alpha,\kappa) \cdot u^B\,(T,\varkappa)\ \mathrm{d}x = \left|E_\alpha^T\right| \delta_{\kappa,\varkappa}$$

und $\boldsymbol{B}(T)$ hat die behauptete Form. □

Die entsprechenden Codezeilen für die Berechnung der lokalen Blöcke lauten

```
% initialisation of matrix blocks
B4e = zeros(6,2,nrElems);       % local matrices for block B

%% Parfor-loop to fill local matrices
for elem = 1:nrElems            %elem as current triangle T
    slength = sl4e(elem,:)';              % length of sides
    L       = [slength,slength]';  L=diag(L(:));

    % block B int_T(u.div(tau))
    B4e(:,:,elem)=L*[1,0;0,1;1,0;0,1;1,0;0,1];
end
```

In der Koeffizientenmatrix $\mathbb{B}$ wird der Block $\boldsymbol{B}$ aus den lokalen Matrizen $\boldsymbol{B}(T)$, für alle $j = 1,\ \ldots,\ J$, $\alpha = 1,\ 2,\ 3$ und $\kappa,\ \varkappa = 1,\ 2$ durch

$$\boldsymbol{B}_{6(j-1)+2(\alpha-1)+\kappa,2(j-1)+\varkappa} = \boldsymbol{B}\left(T_j; 2(\alpha-1)+\kappa, \varkappa\right)$$

bestimmt. Auch hier ergibt sich die bekannte Blockdiagonalform

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{B}(T_1) & 0 & \ldots & 0 \\ 0 & \boldsymbol{B}(T_2) & \ldots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \ldots & 0 & \boldsymbol{B}(T_J) \end{pmatrix} \in \mathbb{R}^{6J\times 2J}.$$

In dem Programm geschieht dieses Assemblieren in

```
%% Constructing global matrix blocks
% with tempI as vector of row indices and tempJ vector of column indices
tempI = permute(reshape(repmat(1:6*nrElems,2,1),2,6,nrElems),[2,1,3]);
tempJ = repmat(1:2*nrElems,6,1);
B=sparse(tempI(:),tempJ(:),B4e(:));     % block B
```

Block C enthält die Beiträge für $s_1 \in S_0^2\left(\partial\mathcal{T};\mathbb{R}^2\right)$ und $\boldsymbol{\tau}_{\mathrm{RT}} \in RT_0^{\mathrm{pw}}\left(\mathcal{T};\mathbb{R}^{2\times 2}\right)$. Die Einträge sind also durch folgendes Integral bestimmt

$$\boldsymbol{C}_{mn} = -\left\langle \gamma_\nu \boldsymbol{\tau}_m^b, s_n^b \right\rangle_{\partial\mathcal{T}} \quad \text{für } m = 1, \ldots, 6J \text{ und } n = 1, \ldots, 2|\mathcal{N}(\Omega)|\,.$$

Auch dieses verschwindet, es sei denn die Träger von $\boldsymbol{\tau}_m^b$ und s_n^b sind Teil desselben Dreiecks $T \in \mathcal{T}$, d.h. $\boldsymbol{\tau}_m^b = \boldsymbol{\tau}^B(T,\alpha,\kappa)$ und $s_n^b = s^B(z,\varkappa)$ mit $z \in \mathcal{N}(T)$ für ein $T \in \mathcal{T}$, $\alpha = 1, 2, 3$ und $\kappa, \varkappa = 1, 2$ und in diesem Fall wird lediglich über den Rand dieses $T \in \mathcal{T}$ integriert.

Die im folgenden Lemma bestimmten lokalen Blöcke sind daher nützlich zur Berechnung von $\boldsymbol{C}$.

Lemma 4.6. *Für jedes Element $T \in \mathcal{T}$, wird der lokale Block $\boldsymbol{C}(T)$ definiert durch*

$$\boldsymbol{C}(T; 2(\alpha-1)+\kappa,\, 2(\gamma-1)+\varkappa) := -\int_{\partial T} \gamma_\nu \left(\boldsymbol{\tau}^B(T,\alpha,\kappa)\right) \cdot s^B\left(P_\gamma,\varkappa\right) \mathrm{d}s,$$

für $\kappa, \varkappa = 1, 2$, die lokale Nummerierung der Kanten durch $\alpha = 1, 2, 3$ und die lokale Nummerierung der Knoten durch $\gamma = 1, 2, 3$. Es ergeben sich folgende Einträge

$$\boldsymbol{C}(T; 2(\alpha-1)+\kappa,\, 2(\gamma-1)+\varkappa) = \begin{cases} -\delta_{\kappa,\varkappa}\left|E_\alpha^T\right|/2 & \text{für } P_\gamma \in \mathcal{N}\left(E_\alpha^T\right), \\ 0 & \text{sonst.} \end{cases}$$

und damit folgende Form des lokalen Blocks

$$\boldsymbol{C}(T) = -\frac{1}{2} L(T)\, H_3 \in \mathbb{R}^{6\times 6},$$

wobei $L(T)$ wie in (4.7) und H_3 wie in in (4.9) seien.

Beweis. Die lokalen Raviart-Thomas Funktionen $\psi(T,\alpha)$ verschwinden entlang von Kanten $F \notin \mathcal{E}(T)$. Außerdem ist $(x - P_{\alpha-1})$ entlang der anliegenden Kanten, also für $x \in F \in \mathcal{E}(T) \setminus \left\{E_\alpha^T\right\}$, tangential zu ∂T und entlang der gegenüberliegenden Kante, also für $x \in E_\alpha^T$, ist $(x - P_{\alpha-1}) \cdot \nu_T$ genau die Höhe des Dreiecks, d.h. die Konstante $2|T|/|E_\alpha^T|$. Daher gilt

$$\left(\psi(T,\alpha)\cdot\nu_T\right)|_F = \left|E_\alpha^T\right| / \left(2|T|\right)\left((x-P_{\alpha-1})\cdot\nu_T|_F\right) = \begin{cases} 1 & \text{für } F = E_\alpha^T, \\ 0 & \text{für } F \neq E_\alpha^T. \end{cases}$$

Dies ist auch in [BC05, Lemma 4.1] bewiesen. Zusätzlich gilt für die Nodalen Basisfunktion auf dem Rand $\gamma_0^{\mathcal{T}}\varphi_k$, dass

$$\int_E \gamma_0\varphi_k \,\mathrm{d}s = \begin{cases} |E|/2 & \text{für } z_k \in \mathcal{N}(E), \\ 0 & \text{sonst.} \end{cases}$$

Damit ergibt sich für die lokale Nummerierung der Seiten mit $\alpha = 1, 2, 3$, $\kappa, \varkappa = 1, 2$, und die lokale Knotenummerierung $\gamma = 1, 2, 3$,

$$\int_{\partial T} \left(\boldsymbol{\tau}^B(T,\alpha,\kappa)\,\nu_T\right) \cdot s^B\left(P_\gamma, \varkappa\right) \mathrm{d}x = \begin{cases} \delta_{\kappa,\varkappa}\left|E_\alpha^T\right|/2 & \text{für } P_\gamma \in \mathcal{N}\left(E_\alpha^T\right), \\ 0 & \text{sonst.} \end{cases}$$

In Abbildung 4.1 wird deutlich, dass sich daraus die behauptete Form ergibt

$$\boldsymbol{C}(T) = -\frac{1}{2}L(T)\begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} = -\frac{1}{2}L(T)\,H_3 \in \mathbb{R}^{6\times 6}.$$

□

In dem Programmcode werden die folgenden lokalen Rechnungen durchgeführt

```
% initialisation of matrix blocks
C4e = zeros(6,6,nrElems);       % local matrices for block C
IndJC4e=zeros(6,6,nrElems);     % columns for block C
% constant auxillary matrices
H3 = sparse([1:6,1:6],[1:6,3:6,1:2],ones(12,1),6,6);

%% Parfor-loop to fill local matrices
for elem = 1:nrElems            %elem as current triangle T
    nodes   = n4e(elem,:);              % nodes for element
    slength = sl4e(elem,:)';            % length of sides
    L       = [slength,slength]';  L=diag(L(:));

    % block C int_dT(gamma_nu(tau).s)
    C4e(:,:,elem)=-1/2*L*H3;
        % index transformation (local to global)
        temp = [2*(nodes-1)+1;2*(nodes-1)+2];
        IndJC4e(:,:,elem)=repmat(temp(:)',6,1);
end
```

Aus diesen lokalen Matrizen wird der Matrixblock $\boldsymbol{C}$, für alle $j = 1, \ldots, J$, $\alpha = 1, 2, 3$, $\kappa, \varkappa = 1, 2$, und die globale Nummerierung $k \in \left\{1, \ldots, |\mathcal{N}(\Omega)|\right\}$ der Knoten P_γ für $\gamma = 1, 2, 3$ wie folgt assembliert

$$\boldsymbol{C}_{6(j-1)+2(\alpha-1)+\kappa,2(k-1)+\varkappa} = \boldsymbol{C}\left(T_j; 2(\alpha-1)+\kappa,\, 2(\gamma-1)+\varkappa\right).$$

In diesem Fall, werden die lokalen Blöcke also aufgeteilt, indem die einzelnen Spalten der jeweiligen globalen Knotennummer zugeordnet und so umsortiert werden. Dies geschieht in folgenden Zeilen

```
%% Constructing global matrix blocks
% with tempI as vector of row indices and tempJ vector of column indices
tempI = permute(reshape(repmat(1:6*nrElems,6,1),6,6,nrElems),[2,1,3]);
C=sparse(tempI(:),IndJC4e(:),C4e(:));   % block C
```

Block D enthält die Anteile der Bilinearform für $\boldsymbol{\sigma}_0 \in P_0\left(\mathcal{T}; \mathbb{R}^{2\times 2}\right)$ und $v_1 \in P_1\left(\mathcal{T}; \mathbb{R}^2\right)$, die Einträge ergeben sich also durch

$$\boldsymbol{D}_{mn} = \int_\Omega \mathrm{D}_{\mathrm{NC}}\, v_m^b : \boldsymbol{\sigma}_n^b \,\mathrm{d}x \quad \text{für } m = 1, \ldots, 6J \text{ und } n = 1, \ldots, 4J.$$

Sie verschwinden, es sei denn v_m^b und $\boldsymbol{\sigma}_n^b$ haben dasselbe $T \in \mathcal{T}$ als Träger, d.h. $v_m^b = v^B(T, \gamma, \kappa)$ und $\boldsymbol{\sigma}_n^b = \boldsymbol{\sigma}^B(T, \lambda)$ für ein $T \in \mathcal{T}$, $\gamma = 1, 2, 3$, $\kappa = 1, 2$ und $\lambda = 1, \ldots, 4$. In diesem Fall ist lediglich über dieses Dreieck $T \in \mathcal{T}$ zu integrieren.

Es erweist sich daher als hilfreich, die lokalen Blöcke zu untersuchen.

Lemma 4.7. *Für jedes Dreieck $T \in \mathcal{T}$ wird der lokale Block $\boldsymbol{D}(T)$ definiert als*

$$\boldsymbol{D}(T; 2(\gamma-1)+\kappa, \lambda) := \int_T \mathrm{D}\, v^B(T, \gamma, \kappa) : \boldsymbol{\sigma}^B(T, \lambda)\,\mathrm{d}x$$

für $\kappa = 1, 2$, $\gamma = 1, 2, 3$, und $\lambda = 1, 2, 3, 4$. Es ergibt sich

$$\boldsymbol{D}(T; 2(\gamma-1)+\kappa, \lambda) = |T| \left(\left(e_\kappa \otimes \nabla \varphi_{P_\gamma} \right) : \mathbf{e}_\lambda \right)$$

und damit folgende Form des Blockes

$$\boldsymbol{D}(T) = |T| \left(H_1\, \mathrm{G}(T) \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} + H_2\, \mathrm{G}(T) \begin{pmatrix} 0 & 0 & 1 & -1 \\ 1 & -1 & 0 & 0 \end{pmatrix} \right) \in \mathbb{R}^{6\times 4},$$

dabei sei die Matrix $\mathrm{G}(T)$ für die lokale Nummerierung P_1, P_2, P_3 der Eckpunkte von T

$$\mathrm{G}(T) := \begin{pmatrix} 1 & 1 & 1 \\ P_1 & P_2 & P_3 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 2}$$

und H_1, H_2 seien die konstanten Matrizen aus (4.9).

Beweis. Wie in [ACF99, S.122] sind die Ableitungen der Nodalen Basisfunktionen zu P_1, P_2, P_3 auf T durch die folgenden Konstanten beschrieben

$$\begin{pmatrix} \nabla\varphi_{P_1} \\ \nabla\varphi_{P_2} \\ \nabla\varphi_{P_3} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ P_1 & P_2 & P_3 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Daher ergibt sich für $\kappa = 1, 2$, $\lambda = 1, \ldots, 4$, und $\gamma = 1, 2, 3$

$$\begin{aligned}\int_T \mathrm{D}\, v^B(T,\gamma,\kappa) : \boldsymbol{\sigma}^B(T,\lambda)\,\mathrm{d}x \\ &= |T| \left(\left(e_\kappa \otimes \nabla \varphi_{P_\gamma} \right) : \mathbf{e}_\lambda \right) \\ &= \begin{cases} |T|\,\nabla\varphi_{P_\gamma}\cdot e_1 & \text{für } (\lambda=1,\kappa=1),\ (\lambda=2,\kappa=1),\ (\lambda=3,\kappa=2), \\ -|T|\,\nabla\varphi_{P_\gamma}\cdot e_1 & \text{für } (\lambda=4,\kappa=2), \\ |T|\,\nabla\varphi_{P_\gamma}\cdot e_2 & \text{für } (\lambda=1,\kappa=2),\ (\lambda=3,\kappa=1),\ (\lambda=4,\kappa=1), \\ -|T|\,\nabla\varphi_{P_\gamma}\cdot e_2 & \text{für } (\lambda=2,\kappa=2). \end{cases}\end{aligned}$$

Vermöge der Matrix $G(T)$ ergibt sich

$$\boldsymbol{D}(T) = |T| \begin{pmatrix} G(T;1,1) & G(T;1,1) & G(T;1,2) & G(T;1,2) \\ G(T;1,2) & -G(T;1,2) & G(T;1,1) & -G(T;1,1) \\ G(T;2,1) & G(T;2,1) & G(T;2,2) & G(T;2,2) \\ G(T;2,2) & -G(T;2,2) & G(T;2,1) & -G(T;2,1) \\ G(T;3,1) & G(T;3,1) & G(T;3,2) & G(T;3,2) \\ G(T;3,2) & -G(T;3,2) & G(T;3,1) & -G(T;3,1) \end{pmatrix} \in \mathbb{R}^{6\times 4},$$

somit hat $\boldsymbol{D}(T)$ die behauptete Form. □

In dem Programm werden die lokalen Blöcke wie folgt bestimmt

```
% initialisation of matrix blocks
D4e = zeros(6,4,nrElems);       % local matrices for block D
% constant auxillary matrices
H1 = sparse([1,3,5],[1,2,3],[1,1,1],6,3);
H2 = sparse([2,4,6],[1,2,3],[1,1,1],6,3);

%% Parfor-loop to fill local matrices
for elem = 1:nrElems            %elem as current triangle T
    coords  = coord4e(:,:,elem);         % coordinates of three nodes
    area    = area4e(elem);              % area of T
    G       = [1, 1, 1; coords']\[0, 0; 1, 0; 0, 1];

    % block D int_T(sigma:Dv)
    D4e(:,:,elem)=area*(H1*G*[1,1,0,0;0,0,1,1]+H2*G*[0,0,1,-1;1,-1,0,0]);
end
```

Aus diesen lokalen Matrixblöcken wird der Block $\boldsymbol{D}$ für alle $j = 1, \ldots, J$, $\gamma = 1, 2, 3$, $\kappa = 1, 2$ und $\lambda = 1, 2, 3, 4$ wie folgt assembliert

$$\boldsymbol{D}_{6(j-1)+2(\gamma-1)+\kappa,4(j-1)+\lambda} = \boldsymbol{D}\left(T_j; 2(\gamma-1)+\kappa,\ \lambda\right).$$

Es ergibt sich also wieder eine Diagonalstruktur

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{D}(T_1) & 0 & \dots & 0 \\ 0 & \boldsymbol{D}(T_2) & \dots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \dots & 0 & \boldsymbol{D}(T_J) \end{pmatrix} \in \mathbb{R}^{6J \times 4J}.$$

Dies wird realisiert via

```
%% Constructing global matrix blocks
% with tempI as vector of row indices and tempJ vector of column indices
tempI = permute(reshape(repmat(1:6*nrElems,4,1),4,6,nrElems),[2,1,3]);
tempJ = repmat(1:4*nrElems,6,1);
D=sparse(tempI(:),tempJ(:),D4e(:));     % block D
```

Block E enthält die Anteile der Bilinearform für $t_0 \in P_0\left(\mathcal{T}; \mathbb{R}^2\right)$ und $v_1 \in P_1\left(\mathcal{T}; \mathbb{R}^2\right)$ und daher sind die Einträge via

$$\boldsymbol{E}_{mn} = -\left\langle t_n^b, \gamma_0 v_m^b \right\rangle_{\partial\mathcal{T}} \quad \text{für } m = 1, \dots, 6J \text{ und } n = 1, \dots, 2|\mathcal{E}|$$

definiert. Sie verschwinden, es sei denn die Träger der beiden Funktionen v_m^b und t_n^b liegen im selben Dreieck $T \in \mathcal{T}$, d.h. $v_m^b = v^B(T, \gamma, \kappa)$ und $t_n^b = t^B(E_\ell, \varkappa)$ mit $E_\ell \in \mathcal{E}(T)$ für ein $T \in \mathcal{T}$, $\gamma = 1, 2, 3$, $\kappa, \varkappa = 1, 2$ und $\ell = 1, \dots, |\mathcal{E}|$ und in diesem Fall ist jeweils nur über den Rand dieses $T \in \mathcal{T}$ zu integrieren.

Im Folgenden wird die Berechnung der zu betrachtenden lokalen Blöcke genauer analysiert.

Lemma 4.8. *Für jedes Element $T \in \mathcal{T}$ wird der lokale Block $\boldsymbol{E}(T)$ definiert via*

$$\boldsymbol{E}(T; 2(\gamma-1)+\kappa, 2(\alpha-1)+\varkappa) := -\int_{\partial T} \gamma_0 \left(v^B(T, \gamma, \kappa)\right) \cdot t^B\left(E_\alpha^T, \varkappa\right) \mathrm{d}s$$

für $\kappa, \varkappa = 1, 2$, $\gamma = 1, 2, 3$ und die lokale Nummerierung $\alpha = 1, 2, 3$ der Kanten von T. Damit ergibt sich

$$\boldsymbol{E}(T; 2(\gamma-1)+\kappa, 2(\alpha-1)+\varkappa) = \begin{cases} -\delta_{\kappa,\varkappa} \operatorname{sgn}_\alpha^T \left|E_\alpha^T\right| / 2 & \text{für } P_\gamma \in \mathcal{N}\left(E_\alpha^T\right), \\ 0 & \text{sonst.} \end{cases}$$

und die Blockform

$$\boldsymbol{E}(T) = -\frac{1}{2} H_3^\top L(T)\, S(T) \in \mathbb{R}^{6\times 6},$$

vermöge $S(T)$, $L(T)$ und H_3 wie in (4.7), (4.8) und (4.9), respektive.

Beweis. Es sei erinnert, dass das Integral der Nodalen Basisfunktion φ_k entlang einer Kante den Wert $\int_E \gamma_0^{\mathcal{T}} \varphi_k \, \mathrm{d}s = |E|/2$ hat, wenn $z_k \in \mathcal{N}(E)$, und sonst verschwindet. Damit ergibt sich für die lokale Nummerierung $\gamma = 1, 2, 3$ der Eckpunkte, $\alpha = 1, 2, 3$

der Seiten und κ, $\varkappa = 1, 2$ vermöge der Einbettung $P_0\left(\mathcal{E};\mathbb{R}^2\right) \hookrightarrow H^{-1/2}\left(\partial\mathcal{T};\mathbb{R}^2\right)$ aus Paragraph 2.4 folgende Form

$$\int_{\partial T} v^B\left(T,\gamma,\kappa\right) t^B\left(E_\alpha^T,\varkappa\right)\,\mathrm{d}x = \begin{cases} \delta_{\kappa,\varkappa}\,\mathrm{sgn}_\alpha^T\left|E_\alpha^T\right|/2 & \text{für } P_\gamma \in \mathcal{N}\left(E_\alpha^T\right), \\ 0 & \text{sonst.} \end{cases}$$

Die Nummerierungskonvention aus Abbildung 4.1 zeigt, dass

$$\boldsymbol{E}(T) = -\frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix} L(T)\,S(T) = -\frac{1}{2}H_3^\top\,L(T)\,S(T) \in \mathbb{R}^{6\times 6},$$

also die Behauptung. □

In dem Programm werden folgende lokale Rechnungen durchgeführt

```
% initialisation of matrix blocks
E4e = zeros(6,6,nrElems);       % local matrices for block E
IndJE4e=zeros(6,6,nrElems);     % columns for block E
% constant auxillary matrices
H3 = sparse([1:6,1:6],[1:6,3:6,1:2],ones(12,1),6,6);

%% Parfor-loop to fill local matrices
for elem = 1:nrElems            %elem as current triangle T
    sides   = s4e(elem,:);              % sides of T
    slength = sl4e(elem,:)';            % length of sides
    sign    = diag(ns4e(:,:,elem)*normal4e(:,:,elem)');
                                        % vector with sgn(sides,elem)
    L       = [slength,slength]';  L=diag(L(:));
    S       = [sign,sign]';  S=diag(S(:));

    % block E int_dT(t.gamma_0(v))
    E4e(:,:,elem)=-1/2*H3'*S*L;
        % index transformation (local to global)
        temp = [2*(sides-1)+1;2*(sides-1)+2];
        IndJE4e(:,:,elem)=repmat(temp(:)',6,1);
end
```

Die lokalen Blöcke werden in $\boldsymbol{E}$ für alle $j = 1, \ldots, J$, $\gamma = 1, 2, 3$, κ, $\varkappa = 1, 2$, und den globalen Nummern $\ell \in \{1, \ldots, |\mathcal{E}|\}$ der lokalen Seiten E_α^T für $\alpha = 1, 2, 3$ assembliert via

$$\boldsymbol{E}_{6(j-1)+2(\gamma-1)+\kappa,2(\ell-1)+\varkappa} = \boldsymbol{E}\left(T_j; 2(\gamma-1)+\kappa,\, 2(\alpha-1)+\varkappa\right).$$

Auch hier werden die lokalen Blöcke also aufgespalten und die Spalten den jeweiligen globalen Kantennummern zugeordnet. Dies geschieht durch

```
%% Constructing global matrix blocks
% with tempI as vector of row indices and tempJ vector of column indices
tempI = permute(reshape(repmat(1:6*nrElems,6,1),6,6,nrElems),[2,1,3]);
E=sparse(tempI(:),IndJE4e(:),E4e(:));   % block E
```

Damit kann die Koeffizientenmatrix $\mathbb{B}$ nun vollständig assembliert werden. Die lokalen Träger der Testfunktionen garantieren dabei, dass keine Einträge aufsummiert werden, die lokalen Beiträge müssen nur wie beschrieben geeignet in die globalen Matrixblöcke einsortiert werden. Es ergibt sich, dann

```
% coefficient matrix BB=sparse(dimY,dimX)
BB = [A,B,C,sparse(6*nrElems,2*nrSides); ...
           D,sparse(6*nrElems,2*nrElems+2*nrNodes),E];
```

Im nächsten Paragraphen wird nun der Aufbau der Normmatrix für den Testraum Y_h beschrieben.

4.3.2 Normmatrix

Im Hilbertraum $Y_h = \operatorname{span}\{\zeta_1, \ldots, \zeta_M\}$ kann die Normmatrix $\mathbb{M} \in \mathbb{R}^{M\times M}$, die für den Koeffizientenvektor $y \in \mathbb{R}^M$ von $y_h = \sum_{m=1}^M y_m\zeta_m \in Y_h$ erfüllt, dass $\|y_h\|_Y^2 = y \cdot \mathbb{M}y$, insbesondere wie folgt definiert werden

$$\mathbb{M}_{m,k} := \langle \zeta_m, \zeta_k \rangle_Y \quad \text{für } m,\, k = 1, \ldots, M.$$

Für das Skalarprodukt aus Bemerkung 2.14 bzw. die Norm von $y_h = (\boldsymbol{\tau}_{\mathrm{RT}}, v_1) \in Y_h$

$$\|y_h\|_Y^2 = \|\boldsymbol{\tau}_{\mathrm{RT}}\|_{H(\mathrm{div},\mathcal{T})}^2 + \|v_1\|_{H_0^1(\mathcal{T})}^2$$

ergibt sich in der gegebenen Basis folgende Blockstruktur

$$\mathbb{M} = \begin{bmatrix} \boldsymbol{T} & 0 \\ 0 & \boldsymbol{V} \end{bmatrix}$$

mit $\boldsymbol{T}$, $\boldsymbol{V} \in \mathbb{R}^{6J\times 6J}$.

Die in (4.7), (4.8) und (4.9) definierten Matrizen sind erneut nützlich, um die Blöcke $\boldsymbol{V}$ und $\boldsymbol{T}$ genauer zu beschreiben. In (4.1) wird lediglich die Inverse der Normmatrix verwendet, daher sind die entsprechenden Codezeilen aus `computeBlocksStokesDPG` bzw.`computeBlocksStokesDPGFass` bzw.`computeBlocksStokesDPGIntegrate` erst in diesem Zusammenhang zu finden.

Block T beschreibt den Norm Anteil von $\boldsymbol{\tau}_{\mathrm{RT}} \in RT_0^{\mathrm{pw}}\left(\mathcal{T}; \mathbb{R}^{2\times 2}\right)$. Seine Einträge ergeben sich als

$$\boldsymbol{T}_{mk} = \int_\Omega \boldsymbol{\tau}_m^b : \boldsymbol{\tau}_k^b \,\mathrm{d}x + \int_\Omega \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}_m^b \cdot \operatorname{div}_{\mathrm{NC}} \boldsymbol{\tau}_k^b \,\mathrm{d}x \quad \text{für } m,\, k = 1, \ldots, 6J.$$

Diese Integrale verschwinden außer $\boldsymbol{\tau}_m^b$ und $\boldsymbol{\tau}_k^b$ haben dasselbe Dreieck $T \in \mathcal{T}$ als Träger, d.h. $\boldsymbol{\tau}_m^b = \boldsymbol{\tau}^B(T, \alpha, \kappa)$ und $\boldsymbol{\tau}_k^b = \boldsymbol{\tau}^B(T, \varkappa, \gamma)$, für ein $T \in \mathcal{T}$, $\alpha, \gamma = 1, 2, 3$ und $\kappa, \varkappa = 1, 2$ und in diesem Fall muss nur über dieses $T \in \mathcal{T}$ integriert werden.

Das folgende Lemma beschreibt daher die Berechnung der lokalen Blöcke.

Lemma 4.9. *Für jedes Element $T \in \mathcal{T}$ sei der lokale Beitrag $\boldsymbol{T}(T)$ wie folgt definiert*

$$\begin{aligned}\boldsymbol{T}\left(T; 2(\alpha-1)+\kappa,\, 2(\gamma-1)+\varkappa\right) := &\int_T \boldsymbol{\tau}^B\left(T,\alpha,\kappa\right) : \boldsymbol{\tau}^B\left(T,\gamma,\varkappa\right)\,\mathrm{d}x \\ &+ \int_T \operatorname{div}\boldsymbol{\tau}^B\left(T,\alpha,\kappa\right)\cdot\operatorname{div}\boldsymbol{\tau}^B\left(T,\gamma,\varkappa\right)\,\mathrm{d}x\end{aligned}$$

für $\alpha,\gamma = 1,\,2,\,3$ und $\kappa,\,\varkappa = 1,\,2$. Dann gilt

$$\begin{aligned}\boldsymbol{T}\left(T; 2(\alpha-1)+\kappa,\, 2(\gamma-1)+\varkappa\right) = \delta_{\kappa,\varkappa}\frac{\left|E_\alpha^T\right|\left|E_\gamma^T\right|}{4|T|}\Big(&\left(\operatorname{mid}(T)-P_{\alpha-1}\right)\cdot\Big(\operatorname{mid}(T)-P_{\gamma-1}\Big) \\ &+ s(T)^2/36+4\Big)\end{aligned}$$

mit

$$s(T)^2 := \left|E_1^T\right|^2 + \left|E_2^T\right|^2 + \left|E_3^T\right|^2 \in \mathbb{R},$$

wobei die lokalen Seiten E_1^T, E_2^T, E_3^T und die lokalen Eckpunkte P_1, P_2, $P_3 \equiv P_0$ wie in Abbildung 4.1 angeordnet sind, d.h. so dass $P_{\alpha-1}$ gegenüber von E_α^T liegt. Damit ergibt sich folgende Blockform

$$\boldsymbol{T}(T) = H_1\,\tilde{\boldsymbol{T}}(T)\,H_1^\top + H_2\,\tilde{\boldsymbol{T}}(T)\,H_2^\top \in \mathbb{R}^{6\times 6},$$

mit

$$\tilde{\boldsymbol{T}}(T) := \frac{1}{|T|}\left(1+\frac{s(T)^2}{144}\right)\tilde{\mathrm{L}}(T)\otimes\tilde{\mathrm{L}}(T) + \frac{1}{4|T|}\operatorname{diag}(\tilde{\mathrm{L}}(T))\,\mathrm{K}(T)^\top\,\mathrm{K}(T)\,\operatorname{diag}(\tilde{\mathrm{L}}(T)) \in \mathbb{R}^{3\times 3}$$

sowie

$$\tilde{\mathrm{L}}(T) := \left(\left|E_1^T\right|,\,\left|E_2^T\right|,\,\left|E_3^T\right|\right)^\top \in \mathbb{R}^3,$$

und $\mathrm{K}(T) \in \mathbb{R}^{3\times 2}$ wie in Lemma 4.4

$$\mathrm{K}(T;\alpha,\kappa) = \left(\operatorname{mid}(T)-P_{\alpha-1}\right)\cdot e_\kappa \qquad \text{für } \alpha = 1,\,2,\,3 \text{ und } \kappa = 1,\,2.$$

Beweis. Für die lokalen Raviart-Thomas Basisfunktionen $\psi\left(T,\alpha\right)$ and $\psi\left(T,\gamma\right)$ auf einem Element $T \in \mathcal{T}$ gilt, nach Gleichung (2.2)

$$\begin{aligned}&\int_T \psi\left(T,\alpha\right)\cdot\psi\left(T,\gamma\right) \\ &\quad= \frac{\left|E_\alpha^T\right|\left|E_\gamma^T\right|}{4|T|^2}\int_T\Big(\left(x-\operatorname{mid}(T)+\operatorname{mid}(T)-P_{\alpha-1}\right)\cdot\Big(x-\operatorname{mid}(T)+\operatorname{mid}(T)-P_{\gamma-1}\Big)\Big)\,\mathrm{d}x \\ &\quad= \frac{\left|E_\alpha^T\right|\left|E_\gamma^T\right|}{4|T|^2}\Big(\|\bullet-\operatorname{mid}(T)\|_{L^2(T)}^2 + |T|\left(\operatorname{mid}(T)-P_{\alpha-1}\right)\cdot\Big(\operatorname{mid}(T)-P_{\gamma-1}\Big)\Big).\end{aligned}$$

Da $\operatorname{div}\psi(T,\alpha) = \left|E_\alpha^T\right|/|T|$, ergibt sich

$$\int_T \operatorname{div}\psi(T,\alpha)\operatorname{div}\psi(T,\gamma) = \frac{\left|E_\alpha^T\right|\left|E_\gamma^T\right|}{|T|}.$$

Mit $s(T)^2 := \left|E_1^T\right|^2 + \left|E_2^T\right|^2 + \left|E_3^T\right|^2$ gilt nach [Car09b, Ex. 3.13] $\|\bullet - \operatorname{mid}(T)\|_{L^2(T)}^2 = s(T)^2|T|/36$. Also ergibt sich, für α, $\gamma = 1, 2, 3$ und κ, $\varkappa = 1, 2$

$$\begin{aligned}&\int_T \left(\boldsymbol{\tau}^B(T,\alpha,\kappa) : \boldsymbol{\tau}^B(T,\gamma,\varkappa) + \operatorname{div}\boldsymbol{\tau}^B(T,\alpha,\kappa)\cdot\operatorname{div}\boldsymbol{\tau}^B(T,\gamma,\varkappa)\right)\mathrm{d}x\\ &= \delta_{\kappa,\varkappa}\frac{\left|E_\alpha^T\right|\left|E_\gamma^T\right|}{4|T|}\left(s(T)^2/36 + (\operatorname{mid}(T) - P_{\alpha-1})\cdot\left(\operatorname{mid}(T) - P_{\gamma-1}\right) + 4\right).\end{aligned}$$

Daher führt

$$\begin{aligned}\tilde{\boldsymbol{T}}(T) =& \frac{1}{4|T|}\left(4 + \frac{s(T)^2}{36}\right)\begin{pmatrix}|E_\alpha^T|\\ |E_\beta^T|\\ |E_\gamma^T|\end{pmatrix}\left(\left|E_\alpha^T\right| \;\; \left|E_\beta^T\right| \;\; \left|E_\gamma^T\right|\right) + \frac{1}{4|T|}\begin{pmatrix}\left|E_\alpha^T\right|(\operatorname{mid}(T) - P_{\alpha-1})^\top\\ \left|E_\beta^T\right|\left(\operatorname{mid}(T) - P_{\beta-1}\right)^\top\\ \left|E_\gamma^T\right|\left(\operatorname{mid}(T) - P_{\gamma-1}\right)^\top\end{pmatrix}\\ &\times\left(\left|E_\alpha^T\right|(\operatorname{mid}(T) - P_{\alpha-1}) \;\; \left|E_\beta^T\right|\left(\operatorname{mid}(T) - P_{\beta-1}\right) \;\; \left|E_\gamma^T\right|\left(\operatorname{mid}(T) - P_{\gamma-1}\right)\right),\end{aligned}$$

zu folgender Blockform von $\boldsymbol{T}(T)$

$$\boldsymbol{T}(T) = \begin{pmatrix}\tilde{\boldsymbol{T}}(T;1,1) & 0 & \tilde{\boldsymbol{T}}(T;1,2) & 0 & \tilde{\boldsymbol{T}}(T;1,3) & 0\\ 0 & \tilde{\boldsymbol{T}}(T;1,1) & 0 & \tilde{\boldsymbol{T}}(T;1,2) & 0 & \tilde{\boldsymbol{T}}(T;1,3)\\ \tilde{\boldsymbol{T}}(T;2,1) & 0 & \tilde{\boldsymbol{T}}(T;2,2) & 0 & \tilde{\boldsymbol{T}}(T;2,3) & 0\\ 0 & \tilde{\boldsymbol{T}}(T;2,1) & 0 & \tilde{\boldsymbol{T}}(T;2,2) & 0 & \tilde{\boldsymbol{T}}(T;2,3)\\ \tilde{\boldsymbol{T}}(T;3,1) & 0 & \tilde{\boldsymbol{T}}(T;3,2) & 0 & \tilde{\boldsymbol{T}}(T;3,3) & 0\\ 0 & \tilde{\boldsymbol{T}}(T;3,1) & 0 & \tilde{\boldsymbol{T}}(T;3,2) & 0 & \tilde{\boldsymbol{T}}(T;3,3)\end{pmatrix} \in \mathbb{R}^{6\times 6}.$$

Dies ist das behauptete Matrixprodukt. □

Aus diesen lokalen Blöcken wird Block $\boldsymbol{T}$ für alle $j = 1, \ldots, J$, α, $\gamma = 1, 2, 3$, und κ, $\varkappa = 1, 2$ via

$$\boldsymbol{T}_{6(j-1)+2(\alpha-1)+\kappa,6(j-1)+2(\gamma-1)+\varkappa} = \boldsymbol{T}\left(T_j; 2(\alpha-1)+\kappa,\, 2(\gamma-1)+\varkappa\right).$$

assembliert. Damit ergibt sich auch hier eine Blockdiagonalstruktur

$$\boldsymbol{T} = \begin{pmatrix}\boldsymbol{T}(T_1) & 0 & \ldots & 0\\ 0 & \boldsymbol{T}(T_2) & \ldots & 0\\ \vdots & 0 & \ddots & \vdots\\ 0 & \ldots & 0 & \boldsymbol{T}(T_J)\end{pmatrix} \in \mathbb{R}^{6J\times 6J}.$$

Block V beschreibt den Norm Anteil von $v_1 \in P_1\left(\mathcal{T};\mathbb{R}^2\right)$, damit ergeben sich die Einträge durch

$$\boldsymbol{V}_{mk} = \int_\Omega v_m^b \cdot v_k^b \,\mathrm{d}x + \int_\Omega \mathrm{D_{NC}}\, v_m^b : \mathrm{D_{NC}}\, v_k^b \,\mathrm{d}x \quad \text{für } m,\, k = 1,\, \ldots,\, 6J.$$

Diese Integrale verschwinden es sei denn v_m^b und v_k^b haben dasselbe Dreieck $T \in \mathcal{T}$ als Träger, d.h. $v_m^b = v^B\left(T,\gamma,\kappa\right)$ und $v_k^b = v^B\left(T,\varkappa,\alpha\right)$, für ein $T \in \mathcal{T}$, $\gamma,\, \alpha = 1,\, 2,\, 3$ und $\kappa,\, \varkappa = 1,\, 2$ und in diesem Fall wird lediglich über dieses $T \in \mathcal{T}$ integriert.

Lemma 4.10. *Für jedes Dreieck $T \in \mathcal{T}$ sei der lokale Block $\boldsymbol{V}(T)$ definiert durch*

$$\begin{aligned}\boldsymbol{V}\left(T; 2(\gamma-1)+\kappa,\, 2(\alpha-1)+\varkappa\right) :=& \int_T v^B\left(T,\gamma,\kappa\right)\cdot v^B\left(T,\alpha,\varkappa\right)\,\mathrm{d}x \\ &+ \int_T \mathrm{D}\, v^B\left(T,\gamma,\kappa\right) : \mathrm{D}\, v^B\left(T,\alpha,\varkappa\right)\,\mathrm{d}x\end{aligned}$$

für $\gamma,\, \alpha = 1,\, 2,\, 3$ und $\kappa,\, \varkappa = 1,\, 2$. Es gilt

$$\boldsymbol{V}\left(T; 2(\gamma-1)+\kappa,\, 2(\alpha-1)+\varkappa\right) = |T|\,\delta_{\kappa,\varkappa}\left(\nabla\varphi_\alpha\cdot\nabla\varphi_\gamma + 1/12\left(1+\delta_{\gamma,\alpha}\right)\right)$$

und es ergibt sich damit folgende Struktur

$$\boldsymbol{V}(T) = \left(H_1\,\tilde{\boldsymbol{V}}(T)\,H_1^\top + H_2\,\tilde{\boldsymbol{V}}(T)\,H_2^\top\right) \in \mathbb{R}^{6\times 6},$$

wobei

$$\tilde{\boldsymbol{V}}(T) := |T|\,\mathrm{G}(T)\,\mathrm{G}(T)^\top + \frac{|T|}{12}\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix} \in \mathbb{R}^{3\times 3}$$

und $\mathrm{G}(T)$ sei wie in Lemma 4.7 definiert durch

$$\mathrm{G}(T) := \begin{pmatrix}1&1&1\\P_1&P_2&P_3\end{pmatrix}^{-1}\begin{pmatrix}0&0\\1&0\\0&1\end{pmatrix} \in \mathbb{R}^{3\times 2},$$

für die lokale Nummerierung P_1, P_2, P_3 der Eckpunkte von T .

Beweis. Es gilt

$$\begin{aligned}\int_T \left(\varphi\left(T,\gamma\right)\cdot\varphi\left(T,\alpha\right) + \nabla\varphi\left(T,\gamma\right) : \nabla\varphi\left(T,\alpha\right)\right)\,\mathrm{d}x =& |T|\,/12\left(1+\delta_{\gamma,\alpha}\right) + |T|\,\nabla\varphi_\alpha\cdot\nabla\varphi_\gamma \\ =& \tilde{B}(T;\gamma,\alpha) + \tilde{M}(T;\gamma,\alpha),\end{aligned}$$

wobei $\tilde{B}(T) := |T|\,\mathrm{G}(T)\otimes\mathrm{G}(T)$ die lokale Steifigkeitsmatrix und $\tilde{M}(T) := \frac{|T|}{12}\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}$ die lokale Massematrix für die nodalen Basisfunktionen im Poisson Modell Problem sind.

Dies wird in [ACF99, S.122/127] hergeleitet. Da $\tilde{V}(T) = \tilde{B}(T;\gamma,\alpha) + \tilde{M}(T;\gamma,\alpha)$, hat $\boldsymbol{V}(T)$ auf Grund der Basisanordnung die folgende Form

$$\boldsymbol{V}(T) = \begin{pmatrix} \tilde{\boldsymbol{V}}(T;1,1) & 0 & \tilde{\boldsymbol{V}}(T;1,2) & 0 & \tilde{\boldsymbol{V}}(T;1,3) & 0 \\ 0 & \tilde{\boldsymbol{V}}(T;1,1) & 0 & \tilde{\boldsymbol{V}}(T;1,2) & 0 & \tilde{\boldsymbol{V}}(T;1,3) \\ \tilde{\boldsymbol{V}}(T;2,1) & 0 & \tilde{\boldsymbol{V}}(T;2,2) & 0 & \tilde{\boldsymbol{V}}(T;2,3) & 0 \\ 0 & \tilde{\boldsymbol{V}}(T;2,1) & 0 & \tilde{\boldsymbol{V}}(T;2,2) & 0 & \tilde{\boldsymbol{V}}(T;2,3) \\ \tilde{\boldsymbol{V}}(T;3,1) & 0 & \tilde{\boldsymbol{V}}(T;3,2) & 0 & \tilde{\boldsymbol{V}}(T;3,3) & 0 \\ 0 & \tilde{\boldsymbol{V}}(T;3,1) & 0 & \tilde{\boldsymbol{V}}(T;3,2) & 0 & \tilde{\boldsymbol{V}}(T;3,3) \end{pmatrix} \in \mathbb{R}^{6\times 6},$$

wird also durch das behauptete Matrixprodukt beschrieben. □

Aus diesen lokalen Blöcken wird der Matrixblock $\boldsymbol{V}$ für alle $j = 1, \ldots, J$, γ, $\alpha = 1, 2, 3$, und κ, $\varkappa = 1, 2$ via

$$\boldsymbol{V}_{6(j-1)+2(\gamma-1)+\kappa,6(j-1)+2(\alpha-1)+\varkappa} = \boldsymbol{V}\left(T_j; 2(\gamma-1)+\kappa,\, 2(\alpha-1)+\varkappa\right)$$

assembliert. Erneut ergibt sich also eine Blockdiagonalform

$$\boldsymbol{V} = \begin{pmatrix} \boldsymbol{V}(T_1) & 0 & \ldots & 0 \\ 0 & \boldsymbol{V}(T_2) & \ldots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \ldots & 0 & \boldsymbol{V}(T_J) \end{pmatrix} \in \mathbb{R}^{6J\times 6J}.$$

Um Blockdiagonalmatrizen zu invertieren müssen nur die einzelnen Blöcke invertiert werden. D.h. die Blockdiagonalform von $\boldsymbol{T}$ und $\boldsymbol{V}$ erleichtert die Berechnung von $\mathbb{M}^{-1}$, es gilt nämlich

$$\begin{aligned} &\mathbb{M}^{-1} = \begin{pmatrix} \boldsymbol{T}^{-1} & 0 \\ 0 & \boldsymbol{V}^{-1} \end{pmatrix} \text{ mit} \\ &\boldsymbol{T}^{-1} = \operatorname{diag}\left(\boldsymbol{T}(T_1)^{-1}, \boldsymbol{T}(T_2)^{-1}, \ldots, \boldsymbol{T}(T_J)^{-1}\right) \in \mathbb{R}^{6J} \\ &\text{und } \boldsymbol{V}^{-1} = \operatorname{diag}\left(\boldsymbol{V}(T_1)^{-1}, \boldsymbol{V}(T_2)^{-1}, \ldots, \boldsymbol{V}(T_J)^{-1}\right) \in \mathbb{R}^{6J}. \end{aligned}$$

Da für A, B, $C \in \mathbb{R}^{3\times 3}$ und die Hilfsmatrizen aus (4.9) gilt, wenn $A\,B = C$, dann

$$\left(H_1\,A\,H_1^\top + H_2\,A\,H_2^\top\right)\,\left(H_1\,B\,H_1^\top + H_2\,B\,H_2^\top\right) = \left(H_1\,C\,H_1^\top + H_2\,C\,H_2^\top\right),$$

genügt es statt der $12J\times 12J$-Matrix $\mathbb{M}$ für alle $j = 1, \ldots, J$ die 3×3-Blöcke $\tilde{\boldsymbol{T}}(T_j)$ und $\tilde{\boldsymbol{V}}(T_j)$ numerisch zu invertieren. Die Inverse Matrix wird dann mit Hilfe von $\boldsymbol{T}(T_j)^{-1} = H_1\,\tilde{\boldsymbol{T}}(T_j)^{-1}\,H_1^\top + H_2\,\tilde{\boldsymbol{T}}(T_j)^{-1}\,H_2^\top$ und $\boldsymbol{V}(T_j)^{-1} = H_1\,\tilde{\boldsymbol{V}}(T_j)^{-1}\,H_1^\top + H_2\,\tilde{\boldsymbol{V}}(T_j)^{-1}\,H_2^\top$ bestimmt.

Die lokalen Blöcke werden in dem Programm wie folgt assembliert

```
% initialisation of matrix blocks
Tinv4e = zeros(6,6,nrElems);    % local matrices for block T^(-1)
Vinv4e = zeros(6,6,nrElems);    % local matrices for block V^(-1)
% constant auxillary matrices
H1 = sparse([1,3,5],[1,2,3],[1,1,1],6,3);
H2 = sparse([2,4,6],[1,2,3],[1,1,1],6,3);

%% Parfor-loop to fill local matrices
for elem = 1:nrElems            %elem as current triangle T
    coords  = coord4e(:,:,elem);        % coordinates of three nodes
    mid     = mid4e(elem,:);            % midpoint of T
    area    = area4e(elem);             % area of T
    slength = sl4e(elem,:)';            % length of sides
    G       = [1, 1, 1; coords']\[0, 0; 1, 0; 0, 1];
    K       = [mid;mid;mid]-[coords(3,:);coords(1,:);coords(2,:)];
    ssq     = sum(slength.^2);          % s^2

    % block T ||tau||_(H(div))
    temp = 1/(4*area)*(diag(slength)*K*K'*diag(slength))...
                    +1/area*(1+ssq/144)*(slength*slength');
    % inverse of T
    temp = temp\eye(3);
    Tinv4e(:,:,elem) = H1*temp*H1'+H2*temp*H2';

    % block V ||v||_(H^1)
    temp = area*G*G'+area/12*(ones(3,3)+eye(3));
    % inverse of V
    temp = temp\eye(3);
    Vinv4e(:,:,elem) = H1*temp*H1'+H2*temp*H2';
end
```

Daraus ergeben sich die globalen Blöcke via

```
%% Constructing global matrix blocks
% with tempI as vector of row indices and tempJ vector of column indices
tempI = permute(reshape(repmat(1:6*nrElems,6,1),6,6,nrElems),[2,1,3]);
tempJ = repmat(1:6*nrElems,6,1);
Tinv=sparse(tempI(:),tempJ(:),Tinv4e(:));   % block T^{-1}
Vinv=sparse(tempI(:),tempJ(:),Vinv4e(:));   % block V^{-1}
```

Mit diesen Bestandteilen wird in dem Programm $\mathbb{M}^{-1}$ wie folgt assembliert

```
% norm matrix MMinv=sparse(dimY,dimY)
MMinv = [Tinv,sparse(6*nrElems,6*nrElems);...
            sparse(6*nrElems,6*nrElems),Vinv];
```

Im Folgenden werden nun noch die Vektoren $\mathbb{F}$ der rechten Seite und $\mathbb{L}$ der linearen Nebenbedingung bestimmt.

4.3.3 Vektor zum Funktional F und zur linearen Nebenbedingung Λ

Zunächst wird das Funktional F aus Definition 3.2 betrachtet. Zu gegebenen Daten $f \in L^2(\Omega;\mathbb{R}^n)$ und $\hat{s} \in H^{1/2}(\partial\mathcal{T};\mathbb{R}^n)$ mit $\hat{s} = g$ auf Γ ist für alle $y = (\boldsymbol{\tau}, v) \in Y$

$$F(y) = \int_\Omega f \cdot v \, \mathrm{d}x + \left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}, \hat{s} \right\rangle_{\partial\mathcal{T}}.$$

Für $Y_h = \operatorname{span}\{\zeta_1, \ldots, \zeta_M\}$ ist

$$\mathbb{F} := (F(\zeta_m))_{1\le m\le M} \in \mathbb{R}^M.$$

Die Basisanordnung führt zu folgender Blockstruktur

$$\mathbb{F} = \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{G} \end{pmatrix} \in \mathbb{R}^{12J}$$

mit den Blöcken $\boldsymbol{F}$, $\boldsymbol{G} \in \mathbb{R}^{6J}$.

Block F enthält die Beiträge, die sich durch die Testfunktionen $\boldsymbol{\tau}_{\mathrm{RT}} \in RT_0^{\mathrm{pw}}\left(\mathcal{T}; \mathbb{R}^{2\times 2}\right)$ ergeben. Daher sei

$$\boldsymbol{F}_m = \left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}_m^b, \hat{s} \right\rangle_{\partial\mathcal{T}} \quad \text{für } m = 1, \ldots, 6J.$$

Da die Testfunktionen als Träger immer nur ein Element $T \in \mathcal{T}$ haben, ist dieses Integral stets nur auf den Kanten eines Dreiecks auszuwerten. Vor der Berechnung der sich ergebenden lokalen Blöcke, gilt es zunächst eine Funktion $\hat{s}$ zu wählen, die den Bedingungen aus Definition 3.2 genügt. Für die reguläre Triangulierung $\mathcal{T}$ von Ω definiere $\hat{s}$ so, dass $\hat{s}|_E = g|_E$ für alle Randkanten $E \in \mathcal{E}(\partial\Omega)$, also $\hat{s}(z) = g(z)$ für alle Randknoten $z \in \mathcal{N}(\partial\Omega)$, des Weiteren sei $\hat{s}(z) = 0$ für alle Innenknoten $z \in \mathcal{N}(\Omega)$ und $\hat{s}|_E \in P_1(E)$ für alle inneren Kanten $E \in \mathcal{E}(\Omega)$. In diesem Setting berechnet sich Block **F** wie folgt.

Lemma 4.11. *Für jedes Dreieck $T \in \mathcal{T}$ sei der lokale Block $\boldsymbol{F}(T)$ definiert als*

$$\boldsymbol{F}(T; 2(\alpha-1)+\kappa) := \int_{\partial T} \hat{s} \cdot \boldsymbol{\tau}^B(T, \alpha, \kappa)\, \nu_T \, \mathrm{d}s$$

Die Einträge genügen folgender Fallunterscheidung

$$\boldsymbol{F}(T; 2(\alpha-1)+\kappa) = \begin{cases} \int_{E_\alpha^T} g \cdot e_\kappa \, \mathrm{d}s & \text{für } E_\alpha^T \in \mathcal{E}(\partial\Omega) \\ \left(\hat{s}(P_\alpha) + \hat{s}(P_{\alpha+1})\right) \cdot e_\kappa \, |E_\alpha^T|/2 & \text{für } E_\alpha^T \in \mathcal{E}(\Omega). \end{cases}$$

Damit ist das Aussehen dieses lokalen Blocks abhängig von der Art der Kanten.

Beweis. Für die lokalen Raviart-Thomas Funktionen gilt wie bereits in Lemma 4.6 erklärt,

$$(\psi(T,\alpha) \cdot \nu_T)\,|_F = \left|E_\alpha^T\right| / (2|T|)\, ((x - P_{\alpha-1}) \cdot \nu_T|_F) = \begin{cases} 1 & \text{für } F = E_\alpha^T, \\ 0 & \text{für } F \neq E_\alpha^T, \end{cases},$$

d.h. es gilt für $\alpha = 1, 2, 3$ und $\kappa = 1, 2$

$$\boldsymbol{F}(T; 2(\alpha-1)+\kappa) = \int_{E_\alpha^T} \hat{s} \cdot e_\kappa \, \mathrm{d}s.$$

Es ist also lediglich $\hat{s}$ zu integrieren. Für die Randkanten, $E_\alpha^T \in \mathcal{E}(\partial\Omega)$, gilt nach Definition von $\hat{s}$

$$\boldsymbol{F}(T; 2(\alpha-1)+\kappa) = \int_{E_\alpha^T} g \cdot e_\kappa \, \mathrm{d}s.$$

Für die inneren Kanten, $E_\alpha^T \in \mathcal{E}(\Omega)$ ist $\hat{s}|_{E_\alpha^T} \in P_1\left(E_\alpha^T\right)$. Da nach der Nummerierung aus Abbildung 4.1 $E_\alpha^T = \operatorname{conv}\{P_\alpha, P_{\alpha+1}\}$ gilt, folgt also

$$\boldsymbol{F}(T; 2(\alpha-1)+\kappa) = \left(\hat{s}(P_\alpha) + \hat{s}(P_{\alpha+1})\right) \cdot e_\kappa \, |E_\alpha^T|/2.$$

Für die Auswertung entlang innerer Kanten ist eine Unterscheidung nach Eckpunkten zu treffen

$$\boldsymbol{F}(T; 2(\alpha-1)+\kappa) = \begin{cases} 0 & \text{für } P_\alpha,\, P_{\alpha+1} \in \mathcal{N}(\Omega), \\ g(P_{\alpha+1}) \cdot e_\kappa \, |E_\alpha^T|/2 & \text{für } P_\alpha \in \mathcal{N}(\Omega),\, P_{\alpha+1} \in \mathcal{N}(\partial\Omega), \\ g(P_\alpha) \cdot e_\kappa \, |E_\alpha^T|/2 & \text{für } P_\alpha \in \mathcal{N}(\partial\Omega),\, P_{\alpha+1} \in \mathcal{N}(\Omega), \\ \left(g(P_\alpha) + g(P_{\alpha+1})\right) \cdot e_\kappa \, |E_\alpha^T|/2 & \text{für } P_\alpha,\, P_{\alpha+1} \in \mathcal{N}(\partial\Omega). \end{cases}$$

□

Aus den lokalen Vektoren wird der Matrixblock $\boldsymbol{F}$, für alle $j = 1, \ldots, J$, $\alpha = 1, 2, 3$ und $\kappa = 1, 2$ wie folgt assembliert

$$\boldsymbol{F}_{6(j-1)+2(\alpha-1)+\kappa} = \boldsymbol{F}\left(T_j; 2(\alpha-1)+\kappa\right).$$

Also werden die einzelnen Blöcke einfach aneinander gehängt

$$\boldsymbol{F} = \begin{pmatrix} \boldsymbol{F}(T_1) \\ \boldsymbol{F}(T_2) \\ \vdots \\ \boldsymbol{F}(T_J) \end{pmatrix} \in \mathbb{R}^{6J}.$$

In dem Code liest sich das wie folgt

```
%% Constructing global matrix blocks
F=sparse(1:6*nrElems,ones(6*nrElems,1),F4e(:),6*nrElems,1);  % block F
```

Um das numerische Integrieren über die Randkanten zu vermeiden, kann g auf allen Kanten linear approximiert werden, d.h. $\int_E g \cdot e_\kappa \approx (g(z_k) + g(z_\ell)) \cdot e_\kappa |E|/2$ für $E = \operatorname{conv}\{z_k, z_\ell\} \in \mathcal{E}$. In diesem Fall ist $\mathbf{F}$ leicht zu bestimmen. Dazu definiere den Vektor $sh \in \mathbb{R}^{|\mathcal{N}|\times 2}$ durch

$$sh(k) = \begin{cases} g(z_k)^\top & \text{für } z_k \in \mathcal{N}(\partial\Omega), \\ 0 & \text{für } z_k \in \mathcal{N}(\Omega). \end{cases}$$

Für ein Dreieck $T \in \mathcal{T}$, $\alpha = 1, 2, 3$ die lokalen Knotennummern und $\kappa = 1, 2$ seien die lokalen Werte von $\hat{s}$ in den Knoten durch

$$sh(T; 2(\alpha-1)+\kappa) := sh(k) \cdot e_\kappa,$$

wobei k die globale Nummer des Knoten P_α sei. Des Weiteren sei gegeben

$$Sh(T) := sh(T) + \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} sh(T) \in \mathbb{R}^6.$$

Dann ergibt sich vermöge L(T) aus (4.7) folgende approximierte Version von Block **F**

$$\boldsymbol{F}(T) = 1/2\, L(T)\, Sh(T).$$

Diese lokalen Blöcke werden in `computeBlocksStokesDPG` wie folgt assembliert

```
% initialisation of matrix blocks
F4e = zeros(6,nrElems);          % local matrices for block F

%% Parfor-loop to fill local matrices
for elem = 1:nrElems             %elem as current triangle T
    slength = sl4e(elem,:)';              % length of sides
    sh = hats4e(:,:,elem)'; sh = sh(:);  % values of hat(s) in nodes
    L       = [slength,slength]';  L=diag(L(:));
    Sh      = sh+ (diag(ones(4,1),2)+diag(ones(2,1),-4))*sh;

    %% Blocks RHS (right hand side)
    % block F int_dT(gamma_nu tau.hat s)
        F4e(:,elem)= 0.5*L*Sh;
end
```

Diese lineare Approximation erweist sich jedoch nicht immer als praktikabel. Bessere Ergebnisse sind zu erzielen, wenn die Integrale über die Randkanten beispielsweise mit der Fassregel, $\int_E g\, \mathrm{d}s \approx |E|/6\ \ (g(z_k) + g(z_\ell) + 4g((z_k + z_\ell)/2))$ für $E = \mathrm{conv}\,\{z_k, z_\ell\} \in \mathcal{E}\,(\partial\Omega)$, approximiert werden. In `computeBlocksStokesDPGFass` werden daher zunächst die Nummern der Seiten auf dem Dirichlet Rand bestimmt und anschließend die lokalen Blöcke wie folgt assembliert

```
%sides on Dirichlet boundary
DbSides=zeros(size(n4sDb,1),1);
for i=1:size(n4sDb,1)
   DbSides(i)=s4n(n4sDb(i,1), n4sDb(i,2));
end
% initialisation of matrix blocks
F4e = zeros(6,nrElems);          % local matrices for block F

%% Parfor-loop to fill local matrices
for elem = 1:nrElems             %elem as current triangle T
    slength = sl4e(elem,:)';          % length of sides
    sh = hats4e(:,:,elem)'; sh = sh(:);  % values of hat(s) in nodes
    L  = [slength,slength]';  L=diag(L(:));
    Sh = sh+ (diag(ones(4,1),2)+diag(ones(2,1),-4))*sh;

    % block F int_dT(gamma_nu tau.hat s)
        F4e(:,elem)= 0.5*L*Sh;

        %treatment of boundary sides
        Db=ismember(sides,DbSides); % boolean, if side is on boundary
        if max(Db)~=0
```

```
            Db=repmat(Db,2,1); Db=Db(:);
            %integration of along Dirichlet boundary sides applying Fass
            %rule
            smid = mids4e(:,:,elem); % midpoints of sides
            temp = hats4e(:,:,elem)+4.*u4Db(c.*smid)+hats4e([2,3,1],:,elem);
            temp = temp'; temp = temp(:);

            F4e(:,elem)=(ones(6,1)-Db).*F4e(:,elem)+Db.*diag(L)./6.*temp;
        end
end
```

Dabei geht der Skalierungsparameter `c` ein wie in Abschnitt 4.5 erläutert.
In `computeBlocksStokesDPGIntegrate` wird lediglich vor der Schleife mit Hilfe der AFEM-Funktion `integrate` ein Vektor mit den Integralen von g entlang der Randseiten erzeugt und die `if`-Anweisung ein wenig verändert

```
%sides on Dirichlet boundary
DbSides=zeros(size(n4sDb,1),1);
for i=1:size(n4sDb,1)
   DbSides(i)=s4n(n4sDb(i,1), n4sDb(i,2));
end

%numerical integration of g along Dirichlet boundary
int4DbSides = integrate(c4n,n4sDb,@(n4p,Gpts4p,Gpts4ref)(u4Db(c.*Gpts4p)),12);
int4sides   = zeros(nrSides,2);
int4sides(DbSides,:)=int4DbSides;
% slicing for parfor loop
ints4e   = permute(reshape(int4sides(s4e',:)',2,3,nrElems),[2,1,3]);
                                % integral of g along boundary sides
% initialisation of matrix blocks
F4e = zeros(6,nrElems)          % local matrices for block F

%% Parfor-loop to fill local matrices
for elem = 1:nrElems            %elem as current triangle T
    slength = sl4e(elem,:)';        % length of sides
    sh = hats4e(:,:,elem)'; sh = sh(:);  % values of hat(s) in nodes
    L  = [slength,slength]';  L=diag(L(:));
    Sh = sh+ (diag(ones(4,1),2)+diag(ones(2,1),-4))*sh;

    % block F int_dT(gamma_nu tau.hat s)
        F4e(:,elem)= 0.5*L*Sh;

        %treatment of boundary sides
        Db=ismember(sides,DbSides); % boolean, if side is on boundary
        if max(Db)~=0
            Db = repmat(Db,2,1); Db=Db(:);
            %integration of along Dirichlet boundary sides insert
            %numerical integrals of g
            sint = ints4e(:,:,elem);     % integral along sides
            sint = sint';    sint = sint(:);
            F4e(:,elem) = (ones(6,1)-Db).*F4e(:,elem)+Db.*sint;
        end
end
```

Block G enthält die Beiträge, die sich durch die Testfunktionen $v_1 \in P_1^{\mathrm{pw}}\left(\mathcal{T};\mathbb{R}^2\right)$ ergeben. Seine Einträge lauten

$$\boldsymbol{G}_m = \int_{\mathcal{T}} f \cdot v_m^b \,\mathrm{d}x \text{ für } m = 1, \ldots, 6J.$$

Auch hier wird jeweils nur über das zu v_m^b assoziierte $T \in \mathcal{T}$ integriert.

Lemma 4.12. *Der lokale Block $\boldsymbol{G}(T)$ sei definiert als*

$$\boldsymbol{G}(T, 2(\gamma-1)+\kappa) := \int_T f \cdot v^B(T,\gamma,\kappa)\,\mathrm{d}x$$

für $\gamma = 1, 2, 3$ und $\kappa = 1, 2$, ergibt sich

$$\boldsymbol{G}(T, 2(\gamma-1)+\kappa) = \int_T f \cdot \varphi(T;\gamma)\, e_\kappa\,\mathrm{d}x.$$

Beweis. Dies ist lediglich die Definition von $v^B(T,\gamma,\kappa)$. □

Für alle $j = 1, \ldots, J$, $\gamma = 1, 2, 3$, und $\kappa = 1, 2$ wird der Block $\boldsymbol{G}$ aus diesen lokalen Blöcken via

$$\boldsymbol{G}_{6(j-1)+2(\gamma-1)+\kappa} = \boldsymbol{G}\left(T_j; 2(\gamma-1)+\kappa\right).$$

assembliert. Also werden die einzelnen Blöcke einfach aneinander gehängt

$$\boldsymbol{G} = \begin{pmatrix} \boldsymbol{G}(T_1) \\ \boldsymbol{G}(T_2) \\ \vdots \\ \boldsymbol{G}(T_J) \end{pmatrix} \in \mathbb{R}^{6J}.$$

Diese Integrale werden wie üblich nicht durch numerische Integration sondern mit der Mittelpunkts Regel, $\int_T g\,\mathrm{d}x \approx |T| g(\mathrm{mid}(T))$ für $g \in L^2(\Omega)$, approximiert. Es gilt $\varphi_k(\mathrm{mid}(T)) = 1/3$, falls $z_k \in \mathcal{N}(T)$, und Null sonst. Daraus ergibt sich die folgende Form des approximierten Blocks

$$\boldsymbol{G}(T) = \frac{|T|}{3}\begin{pmatrix} f(\mathrm{mid}(T)) \\ f(\mathrm{mid}(T)) \\ f(\mathrm{mid}(T)) \end{pmatrix} \in \mathbb{R}^{6\times 1}.$$

Dieser Block kann global in folgenden Zeilen bestimmt werden

```
nrElems = size(n4e,1);                  % number of elements
mid4e   = computeMid4e(c4n,n4e);        % midpoints of elements
area4e  = computeArea4e(c4n,n4e);       % area of elements

%% Constructing global matrix blocks
temp = repmat(c^2.*[area4e,area4e].*f(c*mid4e)/3,1,3)';
G=sparse(1:6*nrElems,ones(6*nrElems,1),temp(:),6*nrElems,1); % block G
```

Dabei geht der Skalierungsparameter `c` ein, der in Abschnitt 4.5 genauer erläutert wird. Der gesamte Vektor $\mathbb{F}$ ergibt sich in allen drei Fällen als

```
% vector of right hand side FF=sparse(dimY,1)
FF  = [F;G];
```

Es bleibt lediglich die lineare Nebenbedingung zu untersuchen. Das lineare Funktional $\Lambda : X_h \to \mathbb{R}$, das diese Nebenbedingung $\int_\Omega \mathrm{tr}\,\boldsymbol{\sigma}_0\,\mathrm{d}x = 0$ für alle $x_h = (\boldsymbol{\sigma}_0, u_0, s_1, t_0) \in X_h$

durch $\Lambda(x_h) = 0$ beschreibt, lautet

$$\Lambda(x_h) = \int_\Omega \operatorname{tr} \boldsymbol{\sigma}_0 \, \mathrm{d}x \quad \text{für alle } x_h = (\boldsymbol{\sigma}_0, u_0, s_1, t_0) \in X_h.$$

Damit kann der zugehörige Vektor $\mathbb{L} \in \mathbb{R}^N$ mit

$$\mathbb{L}_n := \Lambda(\xi_n) \quad \text{für } n = 1, \ldots, N,$$

wie im folgenden Lemma berechnet werden.

Lemma 4.13. *Der Vektor $\mathbb{L}$ hat folgende Einträge*

$$\mathbb{L}_n = \begin{cases} \int_{T_j} \operatorname{tr}\left(\boldsymbol{\sigma}^b_{4(j-1)+\lambda}\right) \mathrm{d}x & \text{für } 1 \le n \le 4J,\ n = 4(j-1)+\lambda \text{ mit } \lambda = 1, \ldots, 4, \\ 0 & \text{sonst.} \end{cases}$$

Außerdem gilt für alle $T \in \mathcal{T}$ und $\lambda = 1, \ldots, 4$

$$\int_T \operatorname{tr}\left(\boldsymbol{\sigma}^B(T, \lambda)\right) = \begin{cases} 2|T| & \text{für } \lambda = 1, \\ 0 & \text{sonst.} \end{cases}$$

Also

$$\mathbb{L} = \Big(2|T_1| \;\; 0 \;\; 0 \;\; 0 \;\; 2|T_2| \;\; 0 \;\; 0 \;\; 0 \;\; 2|T_3| \;\; \ldots \;\; 2|T_J| \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; \ldots \;\; 0\Big).$$

Beweis. Diese Darstellung geht direkt aus der Definition hervor, da $\operatorname{tr}(\mathbf{e}_1) = 2$ und $\operatorname{tr}(\mathbf{e}_\lambda) = 0$ für $\lambda = 2, 3, 4$ gilt. □

Die entsprechenden Codezeilen zur Assemblierung lauten in `computeBlocksStokesDPG` bzw. `computeBlocksStokesDPGFass` bzw `.computeBlocksStokesDPGIntegrate`

```
nrNodes = size(c4n,1);                  % number of nodes
nrElems = size(n4e,1);                  % number of elements
nrSides = size(n4s,1);                  % number of sides
area4e  = computeArea4e(c4n,n4e);       % area of elements
dimX     = 6*nrElems+2*nrNodes+2*nrSides;   % N, dimension of X_h

% vector for linear side condition LL=sparse(dimX,1)
LL   = sparse(1:4:4*nrElems, ones(nrElems,1), 2*area4e,dimX,1);
```

4.4 Exakter Fehler

Für den Fall, dass eine exakte Lösung $\boldsymbol{\sigma} \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{2\times 2}\right)/\mathbb{R}$, $u \in H^1\left(\Omega; \mathbb{R}^2\right)$ zu Problem 8 bekannt ist, soll der exakte Fehler die Abweichung von $\left(\boldsymbol{\sigma}, u, \gamma_0^{\mathcal{T}} u, \gamma_\nu^{\mathcal{T}} \boldsymbol{\sigma}\right)$ von der assoziierten numerischen Lösung (vgl. Satz 3.3) in der Norm von X angeben.

Hier ist zu beachten, dass die wie in Abschnitt 4.1 bestimmte Lösung $(\boldsymbol{\sigma}_0, u_0, s_1, t_0) \in X_h$, eine Lösung zu Problem 10 ist. Die Lösungskomponente s_1 besitzt also Nullrandbedingungen. In Satz 3.3 ist für den kontinuierlichen Fall zu erkennen, dass $\gamma_0^{\mathcal{T}} u = s + \hat{s}$ gilt, wobei $\hat{s} \in H^{1/2}\left(\partial\mathcal{T}; \mathbb{R}^2\right)$ die in der Definition der rechten Seite verwendete Funktion mit $\hat{s} = g$ auf Γ ist und $s \in H_0^{1/2}\left(\partial\mathcal{T}; \mathbb{R}^2\right)$ die Lösungskomponente. Um wie gewünscht die nodale P_1-Approximation $\tilde{s}_1$ von $\gamma_0^{\mathcal{T}} u$ zu erhalten, werden lediglich in den Randknoten die Werte von g statt den Nullwerten von s_1 gesetzt, d.h.

$$\tilde{s}_1(z_k) = \begin{cases} s_1(z_k) & \text{für } z_k \in \mathcal{N}(\Omega), \\ g(z_k) & \text{für } z_k \in \mathcal{N}(\partial\Omega). \end{cases}$$

Also gilt, wenn $x \in \mathbb{R}^N$ der berechnete Koeffizientenvektor ist,

$$\tilde{s}_1 = \sum_{n=1}^{2\left|\mathcal{N}(\Omega)\right|} x_{6J+n} s_n^b + \sum_{\substack{k \in \mathcal{N}(\partial\Omega), \\ \kappa = 1,\, 2}} \left(g(z_k) \cdot e_\kappa\right) \gamma_0^{\mathcal{T}} \varphi_k \, e_\kappa.$$

Des Weiteren werden zur Berechnung des exakten Fehlers die Fortsetzungen von $\tilde{s}_1$, also $w_{\mathrm{c}} \in S_1\left(\Omega; \mathbb{R}^2\right)$ mit $\gamma_0^{\mathcal{T}} w_{\mathrm{c}} = \tilde{s}_1$, und von t_0, also $\boldsymbol{q}_{\mathrm{RT}} \in RT_0\left(\Omega; \mathbb{R}^{2\times 2}\right)$ mit $\gamma_\nu^{\mathcal{T}} \boldsymbol{q}_{\mathrm{RT}} = t_0$, betrachtet, schließlich sind die Normen der Spurräume über solche Fortsetzungen definiert und in Lemma 3.9 bzw. Lemma 3.10 wird die jeweilige Normäquivalenz bewiesen. Dazu sei für $k = 1, \ldots, |\mathcal{N}(\Omega)|$ und $\kappa = 1, 2$

$$w^b_{2(k-1)+\kappa} := \varphi_k \, e_\kappa$$

die Fortsetzung von $s^b_{2(k-1)+\kappa}$, wobei φ_k die nodale Basisfunktion zum Knoten k sei. Die Basis für $\boldsymbol{q}_{\mathrm{RT}}$ wird mit Hilfe der globalen kantenbasierten Raviart-Thomas Funktionen bestimmt, diese sind für alle $E_\ell \in \mathcal{E}$ wie folgt definiert

$$\psi_\ell = \begin{cases} \operatorname{sgn}(T, \alpha)\psi(T, \alpha) & \text{falls } E_\ell \in \mathcal{E}(T) \text{ und } E_\alpha^T = E_\ell, \\ 0 & \text{sonst,} \end{cases}$$

wobei wie in Abschnitt 4.2 $\psi(T, \alpha)$ die lokale Raviart-Thomas Funktion und $\operatorname{sgn}(T, \alpha) = \nu_{E_\alpha^T} \cdot \nu_T \in \{\pm 1\}$. Es sei für $\ell = 1, \ldots, |\mathcal{E}|$ und $\kappa = 1, 2$

$$q^b_{2(\ell-1)+\kappa} := \psi_\ell \otimes e_\kappa$$

die Fortsetzung für $t^b_{2(\ell-1)+\kappa}$.

Sind also die exakte Spannung $\boldsymbol{\sigma} \in H\left(\operatorname{div}, \Omega; \mathbb{R}^{2\times 2}\right)/\mathbb{R}$ und die exakte Verschiebung $u \in H^1\left(\Omega; \mathbb{R}^2\right)$ bekannt, wird der exakte Fehler wie folgt bestimmt

$$E^2 := \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_0\|^2_{L^2(\Omega)} + \|u - u_0\|^2_{L^2(\Omega)} + \|u - w_{\mathrm{c}}\|^2_{H^1(\Omega)} + \|\boldsymbol{\sigma} - \boldsymbol{q}_{\mathrm{RT}}\|^2_{H(\operatorname{div},\Omega)},$$

wobei aus dem Koeffizientenvektor $x \in \mathbb{R}^N$, der Lösung zu (4.1), die Komponenten wie folgt berechnet werden

$$\boldsymbol{\sigma}_0 := \sum_{n=1}^{4J} x_n \boldsymbol{\sigma}_n^b, \; u_0 := \sum_{n=1}^{2J} x_{4J+n}\, u_n^b, \; \boldsymbol{q}_{\mathrm{RT}} := \sum_{n=1}^{2|\mathcal{E}|} x_{6J+2|\mathcal{N}(\Omega)|+n}\, q_n^b,$$
$$w_{\mathrm{c}} := \sum_{n=1}^{2|\mathcal{N}(\Omega)|} x_{6J+n}\, w_n^b + \sum_{\substack{k \in \mathcal{N}(\partial\Omega), \\ \kappa = 1,2}} (g(z_k) \cdot e_\kappa)\, \varphi_k\, e_\kappa.$$

In der `matlab`-Methode `SolveStokesDPG` werden der Koeffizientenvektor `x` der Lösung zu (4.1) und der Koeffizientenvektor `xh` zu der Approximation von $(\boldsymbol{\sigma}, u, \gamma_0^{\mathcal{T}} u, \gamma_\nu^{\mathcal{T}} \boldsymbol{\sigma})$ bestimmt

```
nrNodes   = size(c4n,1);              % number of nodes
nrElems   = size(n4e,1);              % number of elements
nrSides   = size(e4s,1);              % number of sides
dimX = 6*nrElems+2*nrNodes+2*nrSides; % dimension of X_h
                                      % (surpressing boundary conditions)
% computation of coefficient matrix, norm matrix, vektor for F and linear
% side condition
if Int==0
  [B,Minv, F, D] = computeBlocksStokesDPG(f,u4Db,c4n,n4e,n4sDb,c);
elseif Int ==1
  [B,Minv, F, D] = computeBlocksStokesDPGFass(f,u4Db,c4n,n4e,n4sDb,c);
else
  [B,Minv, F, D] = computeBlocksStokesDPGIntegrate(f,u4Db,c4n,n4e,n4sDb,c);
end
%% Linear system Ax=b corresponding to discrete problem
x = zeros(dimX+1,1);

% compute free components and components fixed by boundary condition for s
DbNodes = unique(n4sDb);                      % Dirichlet boundary nodes
fixed   = [2*(DbNodes'-1)+1;2*(DbNodes'-1)+2];
fixed   = 6*nrElems+fixed(:);
free    = setdiff(1:dimX,fixed);
% define matrix A
A=[B(:,free)'*Minv*B(:,free),D(free);D(free)',0];
% define vector b
b=[B(:,free)'*Minv*F;0];
% compute coefficient vector in free components
x([free,dimX+1]) = A\b;
% remove component from Lagrange multiplicator for linear side condition
x=x(1:dimX);

%% Solution corresponding to original problem
% computing boundary data
DbCoords        = c4n(DbNodes,:);       % coordinates for DbNodes
values4DbNodes  = u4Db(c.*DbCoords)';   % values in DbNodes

% computing coefficient vector corresponding to original problem
xh        = x;
%boundary values
xh(fixed) = values4DbNodes(:);
```

Der Methode `computeErrorStokesDPG` zur Fehlerberechnung wird nur der Koeffizientenvektor `xh` übergeben.

4.5 Skalierung

In dem `matlab`-Code ist eine Skalierung des Gebietes Ω mit dem Faktor $c := \max_{x\in\Omega}|x|$ möglich, damit wird garantiert, dass $\hat{\Omega} := 1/c\,\Omega \subseteq [-1,1]^2$. Denn in die Norm des Testraums geht die Gebietsgröße in beiden Komponenten mit unterschiedlicher Gewichtung ein, dies kann dazu führen, dass bei der Minimierung eine der Komponenten bevorzugt wird.

Diese Skalierung wird vermöge der Transformation

$$\Phi = c\,\mathrm{id} : \hat{\Omega} \to \Omega$$

vorgenommen. Dazu wird mit $\hat{\Gamma} := \Phi^{-1}(\Gamma)$ der Rand von $\hat{\Omega}$ bezeichnet und es werden die Variablen

$$\hat{u} := u \circ \Phi : \hat{\Omega} \to \mathbb{R}^2 \quad \text{und} \quad \hat{\boldsymbol{\sigma}} := \boldsymbol{\sigma} \circ \Phi : \hat{\Omega} \to \mathbb{R}^{2\times 2}$$

sowie die Daten

$$\hat{f} := f \circ \Phi : \hat{\Omega} \to \mathbb{R}^2 \quad \text{und} \quad \hat{g} := g \circ \Phi : \hat{\Gamma} \to \mathbb{R}^2$$

betrachtet. Mit Hilfe der Kettenregel können die mit $\mathrm{D}_{\hat{x}}$ und $\mathrm{div}_{\hat{x}}$ bezeichneten Ableitungen in den neuen Koordinaten $\hat{x} := \Phi^{-1}(x)$ berechnet werden

$$\begin{aligned} \mathrm{D}_{\hat{x}}\,\hat{u} &= \mathrm{D}_{\hat{x}}\,(u \circ \Phi) = \mathrm{D}\,u \circ \mathrm{D}_{\hat{x}}\,\Phi = \mathrm{D}\,u \circ (c\,\mathrm{id}) = c\,\mathrm{D}\,u, \\ \mathrm{div}_{\hat{x}}\,\hat{\boldsymbol{\sigma}} &= c\,\mathrm{div}\,\sigma. \end{aligned}$$

Hier bezeichnen D und div die Differentialoperatoren bezüglich der ursprünglichen Koordinaten x. Betrachte nun das ursprüngliche System aus Problem 8, zu gegebenem $f \in L^2\left(\Omega;\mathbb{R}^2\right)$ und $g \in H^{-1/2}\left(\partial\Omega;\mathbb{R}^2\right)$ finde $u \in H^1\left(\Omega;\mathbb{R}^2\right)$ und $\boldsymbol{\sigma} \in H\left(\mathrm{div},\Omega;\mathbb{R}^{2\times 2}\right)/\mathbb{R}$ mit

$$\begin{aligned} \mathrm{div}\,\boldsymbol{\sigma} &= -f && \text{in } \Omega, \\ \mathrm{dev}\,\boldsymbol{\sigma} &= \mathrm{D}\,u && \text{in } \Omega, \\ u &= g && \text{entlang } \Gamma. \end{aligned}$$

Falls $\boldsymbol{\sigma}$ und u also Problem 8 lösen, so gilt für $\hat{\boldsymbol{\sigma}}$ und $\hat{u}$ für jedes $\hat{x} \in \hat{\Omega}$

$$\begin{aligned} \mathrm{div}_{\hat{x}}\,\hat{\boldsymbol{\sigma}}\,(\hat{x}) &= c\,\mathrm{div}\,\boldsymbol{\sigma}\,(x) = -cf(x) = -c\hat{f}\,(\hat{x})\,, \\ \mathrm{dev}\,\hat{\boldsymbol{\sigma}}(\hat{x}) &= \mathrm{dev}\,\boldsymbol{\sigma}\,(x) = \mathrm{D}\,(u)\,(x) = 1/c\,\mathrm{D}\,(\hat{u})\,(\hat{x})\,. \end{aligned}$$

Also lösen $\hat{\boldsymbol{\sigma}}$ und $\hat{u}$ das folgende System auf $\hat{\Omega}$. Finde zu $\hat{f} \in L^2\left(\hat{\Omega};\mathbb{R}^2\right)$ und $\hat{g} \in$

$H^{-1/2}\left(\partial\hat{\Omega};\mathbb{R}^2\right)$ Funktionen $\hat{u} \in H^1\left(\hat{\Omega};\mathbb{R}^2\right)$ und $\hat{\boldsymbol{\sigma}} \in H\left(\mathrm{div},\hat{\Omega};\mathbb{R}^{2\times 2}\right)/\mathbb{R}$ mit

$$\begin{aligned} \mathrm{div}_{\hat{x}}\,\hat{\boldsymbol{\sigma}} &= -c\hat{f} && \text{in } \hat{\Omega},\\ \mathrm{dev}\,\hat{\boldsymbol{\sigma}} &= 1/c\,\mathrm{D}_{\hat{x}}\,\hat{u} && \text{in } \hat{\Omega},\\ \hat{u} &= \hat{g} && \text{entlang } \hat{\Gamma}. \end{aligned}$$

Werden die ersten beiden Gleichungen mit c multipliziert und die folgenden Substitutionen vorgenommen $\tilde{f} = c^2 f$ und $\tilde{\boldsymbol{\sigma}} = c\hat{\boldsymbol{\sigma}}$ so entsteht eine schönere Darstellung des Problems. Zu $\tilde{f} \in L^2\left(\hat{\Omega};\mathbb{R}^2\right)$ und $\hat{g} \in H^{-1/2}\left(\partial\hat{\Omega};\mathbb{R}^2\right)$ finde $\tilde{\boldsymbol{\sigma}} \in H\left(\mathrm{div},\hat{\Omega};\mathbb{R}^{2\times 2}\right)/\mathbb{R}$ und $\hat{u} \in H^1\left(\hat{\Omega};\mathbb{R}^2\right)$, so dass

$$\begin{aligned} \mathrm{div}_{\hat{x}}\,\tilde{\boldsymbol{\sigma}} &= -\tilde{f} && \text{in } \hat{\Omega},\\ \mathrm{dev}\,\tilde{\boldsymbol{\sigma}} &= \mathrm{D}_{\hat{x}}\,\hat{u} && \text{in } \hat{\Omega},\\ \hat{u} &= \hat{g} && \text{entlang } \hat{\Gamma}. \end{aligned}$$

Die Resubstitution $\tilde{\boldsymbol{\sigma}} = 1/c\,\hat{\boldsymbol{\sigma}}$ ergibt die Lösung auf $\hat{\Omega}$ und $\boldsymbol{\sigma} = \hat{\boldsymbol{\sigma}}\,(1/c\,\bullet)$ bzw. $u = \hat{u}\,(1/c\,\bullet)$ die auf Ω. Da die Spannung $\boldsymbol{\sigma}$ in dem Lösungstupel $(\boldsymbol{\sigma}_0, u_0, s_1, t_0) \in X_h$ durch die Größen $\boldsymbol{\sigma}_0$ und t_0 beschrieben wird, müssen diese beiden Größen nach dem Lösen auf $\hat{\Omega}$ mit $1/c$ auf Ω zurückskaliert werden.

In der Methode `solveStokesDPG` findet die Skalierung statt. Vor der Berechnung der in (4.1) benötigten Matrizen werden die Koordinaten der Knoten skaliert

```
c4n = 1/c.*c4n;  % scaling
```

Während der Berechnung in `computeBlocksStokesDPG`, `computeBlocksStokesDPGFass` bzw. `computeBlocksStokesDPGInetgrate` ist lediglich zu beachten, dass die Auswertungen der Funktionen g und f in den nicht skalierten Punkten erfolgt und die Auswertung von f wie beschrieben zu behandeln ist

```
DbNodes         = unique(n4sDb);        % Dirichlet boundary nodes
DbCoords        = c4n(DbNodes,:);       % Coordinates for DbNodes
values4DbNodes  = u4Db(c.*DbCoords);    % Values in DbNodes

temp = repmat(c^2.*[area4e,area4e].*f(c*mid4e)/3,1,3)';
G=sparse(1:6*nrElems,ones(6*nrElems,1),temp(:),6*nrElems,1); % block G
```

Anschließend wird die zum Originalproblem korrespondierende Lösungkomponente $\boldsymbol{\sigma}$ in `solveStokesDPG` wie beschrieben resubstituiert

```
% rescale
xh(1:6*nrElems,6*nrElems+2*nrNodes:end)= ...
                    1/c.* xh(1:6*nrElems,6*nrElems+2*nrNodes:end);
```

4.6 Realisierung

Als Erweiterung des AFEM-Software-Paketes [Car09a] wurden die im Folgenden erläuterten `matlab`-Methoden für die Dimension $n = 2$ implementiert. In dem Programm

`StokesDPG` läuft im Wesentlichen der folgende Standard AFEM-Algorithmus ab.

Algorithm 1: AFEM

input : Reguläre Anfangstriangulierung $\mathcal{T}_0$ und Bulk-Parameter $0 < \theta \leq 1$

for $\ell = 0, 1, 2, \ldots$ **do**

- **solve** *Berechnung der Lösung x_ℓ zu dem diskreten Problem auf $\mathcal{T}_\ell$*
- **estimate** *Berechnung der lokalen Beiträge $\eta_\ell^2(T)$ für alle $T \in \mathcal{T}_\ell$ und des globalen Fehlerschätzers $\eta_\ell^2 = \sum_{T\in\mathcal{T}} \eta_\ell^2(T)$*
- **mark** *Auswahl einer Menge $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ von minimaler Kardinalität $|\mathcal{M}_\ell|$ mit $\theta \sum_{T\in\mathcal{T}_\ell} \eta_\ell^2(T) \leq \sum_{T\in\mathcal{M}_\ell} \eta_\ell^2(T)$*
- **refine** *Erzeugen einer regulären Triangulierung $\mathcal{T}_{\ell+1}$ minimaler Kardinalität, die eine Verfeinerung von $\mathcal{T}_\ell$ ist mit $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$*

output : Folge von Triangulierungen $\mathcal{T}_\ell$, diskreten Lösungen x_ℓ und Fehlerschätzern η_ℓ

Für den Schritt SOLVE wurde die Methode `solveDPGStokes` zum Aufstellen und Lösen des linearen Gleichungssystems (4.1) implementiert. Die einzelnen Matrixblöcke werden dabei, wie in Abschnitt 4.3 erläutert, in dem Unterprogramm `computeBlocksStokesDPG` berechnet. Der Schritt ESTIMATE erfolgt in `estimateDPGStokes`. Dabei wird der Fehlerschätzer aus Bemerkung 4.2 verwendet. Die Schritte MARK und REFINE werden mit Hilfe der bereits existierenden AFEM-Methoden `markBulk` zum Dörfler Markieren und `refineRGB` zur Rot-Grün-Blau-Verfeinerung bzw. im uniformen Fall, $\theta = 1$, mit `refineUniformRed` realisiert.

Ist zu einem Problem die exakte Lösung bekannt, so steht zur Berechnung der Abweichung dieser Lösung von der numerischen die Methode `computeErrorStokesDPG` zur Verfügung, wie in Abschnitt 4.4 erläutert. Außerdem sind die Programme `visualisation` und `visualisationExSol` zur Ausgabe von Plots der einzelnen Variablen bzw. der exakten Lösungskomponenten, sofern bekannt, implementiert.

Die verwendeten Datenstrukturen entsprechen denen im AFEM-Software-Paket. Die Triangulierung wird daher mit folgenden Matrizen beschrieben

$\texttt{n4e} \in \mathbb{R}^{\lvert\mathcal{T}\rvert\times 3}$	Knoten der Elemente,
$\texttt{c4n} \in \mathbb{R}^{\lvert\mathcal{N}\rvert\times 2}$	Koordinaten der Knoten,
$\texttt{n4sDb} \in \mathbb{R}^{\lvert\mathcal{E}(\Gamma)\rvert\times 2}$	Knoten der Seiten, die zum Dirichlet Rand gehören,
$\texttt{n4sNb} = \emptyset$	Knoten der Seiten, die zum Neumann Rand gehören (die letzte Matrix wird als leer angenommen, da Probleme mit reinem Dirichlet Rand betrachtet werden).

Im Folgenden werden die einzelnen Programme mit Eingabeparametern und Ausgaben kurz erläutert.

`StokesDPG(problem,minNrDoF,theta,scale,int,plotdata,savedata,foldername)`

Über `problem` wird im Hauptprogramm der Datensatz zu einem Standard Stokes Problem ausgewählt. Zur Verfügung stehen `'bf'` ("backward facing step" Beispiel), `'cavity'` ("lid driven cavity" Beispiel), `'colliding'` ("colliding flow" Beispiel), `'Lshape'` (Standardbeispiel im L-Gebiet), `'LshapeHighOsc'` (Beispiel im L-Gebiet mit stark oszillierender rechter Seite), `'slit'` (Standardbeispiel im Schlitzgebiet) und `'noPress'` (ein Poisson-Problem Beispiel). Einige dieser Probleme werden detaillierter in Kapitel 5 vorgestellt. Die zugehörigen Datenfiles wurden aus [Bri12] bzw. [Bri14] entnommen und angepasst. Die minimale Zahl der Freiheitsgrade, `minNrDoF`, legt fest, bis zu welchem Level ℓ verfeinert wird. Der Bulkparameter `theta` wird im Dörfler Marking verwendet. Wird der Parameter `scale=1` gesetzt, wird eine problemgerechte Skalierung, wie in Abschnitt 4.5 erläutert, durchgeführt.
Die Eingabe `int` bestimmt mit welcher Genauigkeit die Funktion $\hat{s} \in H^{1/2}(\partial\mathcal{T};\mathbb{R}^2)$ aus Definition 3.2 über den Rand von Ω integriert wird. Bei `int=0` wird mit linearen Funktionen approximiert, bei `int=1` wird mit der Fassregel integriert und bei `int=2` wird die AFEM-Funktion `integrate` mit `deg=12` verwendet.
Der Parameter `plotdata=1` führt zur Ausgabe von Plots der im letzten Schritt berechneten Lösung und sofern bekannt der exakten Lösungskomponenten. Bei der Belegung `plotdata=2` werden die entsprechenden Plots in jedem Verfeinerungsschritt erzeugt. Außerdem wird in diesen beiden Fällen ein Konvergenzplot des Fehlerschätzers und soweit bekannt des exakten Fehlers und seiner Komponenten erzeugt.
Mit `savedata=1` werden die Fehlerdaten (Fehlerschätzer und falls bekannt exakter Fehler und seine Komponenten) in `Error4lvl.dat` nach jedem Verfeinerungsschritt gespeichert. Bei `savedata=2` werden zusätzlich in jedem Schritt die Koeffizientenvektoren der Lösung (`x.dat` und `xh.dat`) sowie die Triangulierungsdaten (`c4n.dat` und `n4e.dat`) gespeichert. Durch `foldername` wird der Speicherort spezialisiert. Die Daten werden in `tmp/StokesDPG/problem-foldername` oder falls `foldername` leer ist in `tmp/StokesDPG/problem-date` gespeichert.

```
[x,xh,nrDoF,B,Minv,F]=solveStokesDPG(f,u4Db,c4n,n4e,n4sDb,c,int)
```

In diesem Fall werden in `f` das Funktionshändel zu $f \in L^2(\Omega;\mathbb{R}^2)$ und in `u4Db` das zu $g \in L^2(\Gamma;\mathbb{R}^2)$, den Randbedingungen für $u|_\Gamma$ aus Problem 8, übergeben. Wie bereits erläutert wird mit `c4n`, `n4e` und `n4sDb` die Triangulierung beschrieben. Der Parameter `c` ist, soweit eine Skalierung stattfindet, der Skalierungsparameter. Mit diesen Daten werden abhängig vom Parameter `int`, also dem gewünschten Genauigkeitsgrad der Randintegrale, in `computeBlocksStokesDPG`, `computeBlocksStokesDPGFass` bzw. `computeBlocksStokesDPGIntegrate` die Bestandteile des linearen Gleichungssystems (4.1) bestimmt. Dieses wird dann in den freien Komponenten (unter der Beachtung der Nullrandbedingungen für die Komponente s) im unter Umständen skalierten Gebiet gelöst. Die Koeffizientenmatrix `B`, die inverse Normmatrix `Minv` und `F`, der Vektor zum Funktional F, werden zurückgegeben. Außerdem wird die Lösung in Form des Koeffizientenvektors `x` zurückgegeben. Anschließend wird die Lösung des zu Problem 10 korrespondierenden Orginalproblems 8, wie in Satz 3.3 erläutert, bestimmt. Der entsprechende Koeffizientenvektor wird in `xh` übergeben. Die Anzahl der Freiheitsgrade wird als `nrDoF` zurückgegeben.

`[B,Minv,F,L]=computeBlocksStokesDPG(f,u4Db,c4n,n4e,n4sDb,c)`

Mit denselben Eingabedaten wie in `solvesStokesDPG` werden zunächst u.a. mit Hilfe bereits existierender AFEM-Programme gewisse Triangulierungs- und Problemdaten berechnet und die benötigten lokalen Matrizen initialisiert. Anschließend werden in einer `parfor`-Schleife über die $j = 1, \ldots, J$ Elemente T_j, die lokalen Beiträge zu den in Abschnitt 4.3 beschriebenen Matrixblöcken berechnet. Dabei werden die Randintegrale durch lineare Approximation von `u4Db` bestimmt. Die globalen Blöcke werden dann mit Hilfe des `sparse`-Befehls assembliert. Anschließend werden die Koeffizientenmatrix `B`, die inverse Normmatrix `Minv`, `F`, der Vektor zum Funktional F, und der Vektor der linearen Nebenbedingung `L` aus diesen globalen Blöcken zusammengefügt und zurückgegeben.

`[B,Minv,F,L]=computeBlocksStokesDPGFass(f,u4Db,c4n,n4e, n4sDb,c)`

Der Unterschied zu `computeBlocksStokesDPG` besteht lediglich darin, dass die Seitennummern der Dirichlet Seiten bestimmt werden und die Funktion `u4Db` entlang dieser mit Hilfe der Fassregel integriert wird. Dadurch kann der Vektor zum Funktional F besser approximiert werden.

`[B,Minv,F,L]=computeBlocksStokesDPGIntegrate(f,u4Db,c4n,n4e,n4sDb,c)`

Im Unterschied zu `computeBlocksStokesDPG` werden hier die Seitennummern der Dirichlet Seiten und mit Hilfe der AFEM-Funktion `integrate` die Integrale von `u4Db` (exakt bis Polynomgrad 12) entlang dieser Seiten bestimmt. Diese Ergebnisse fließen in die Berechnung des Vektors zum Funktional F ein.

`[eta,eta4e]=estimateStokesDPG(n4e,B,Minv,F,x)`

Mit Hilfe von `n4e`, Knoten der Elemente, `B`, der Koeffizientenmatrix, `Minv`, der inversen Normmatrix, und `F`, des Vektors zum Funktional F, wird der in Bemerkung 4.2 beschriebene residuale Fehlerschätzer für den Koeffizientenvektor `x` zu der Lösung von Problem 10 berechnet. Nachdem die inverse Normmatrix und der residual Vektor entsprechend vorbereitet wurden, wird in einer `parfor`-Schleife der lokale Fehlerschätzer bestimmt. In `eta` wird der globale und in `eta4e` wird der Vektor aller lokalen Fehlerschätzer zurückgeben.

```
[Eu,Esigma,Es,Et,E]=computeErrorStokesDPG(c4n,n4e,xh,f,uExact1,uExact2,...
                                  sigmaExact1,sigmaExact2,gradExact1,gradExact2)
```

Hier müssen die Triangulierungsdaten `c4n` und `n4e`, der Koeffizientenvektor zum Originalproblem `xh`, das Funktionshändel `f` zur rechten Seite $f \in L^2(\Omega;\mathbb{R}^2)$ und die Funktionshändel zu den bekannten Komponenten der Lösung übergeben werden. Dabei seien `uExact1` die erste Komponente der Verschiebungsfunktion $u \in H^1(\Omega;\mathbb{R}^2)$, und `uExact2` die zweite. Mit `sigmaExact1` bzw. `sigmaExact2` werden die ersten beiden bzw. letzten

beiden Komponenten (zeilenweise) des Pseudospannungstensors $\sigma \in H(\mathrm{div}, \Omega; \mathbb{R}^{2\times 2})$ bezeichnet. Ebenso zeilenweise werden die Komponenten des Gradienten des Geschwindigkeitsfeldes $\mathrm{D}\, u \in L^2(\Omega; \mathbb{R}^{2\times 2})$ mit `gradExact1` bzw. `gradExact2` benannt. Nachdem die Koeffizienten der einzelnen Lösungskomponenten aus `xh` ausgelesen wurden, werden die Komponenten des Fehlers mit Hilfe der `integrate`-Funktion und bereits existierender Fehler-Funktionen aus dem AFEM-Software-Paket bestimmt. Die P_0-Fehler $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_0\|^2_{L^2(\Omega)}$, `Esigma`, und $\|u - u_0\|^2_{L^2(\Omega)}$, `Eu`, werden in `error4eP0L2` mit `integrate` bestimmt. Der P_1-Fehler $\|u - w_\mathrm{c}\|^2_{H^1(\Omega)} = \|u - w_\mathrm{c}\|^2_{L^2(\Omega)} + |\!|\!| u - w_\mathrm{c} |\!|\!|^2$,`Es`, wird mit Hilfe der AFEM-Funktionen `error4eP1L2` und `error4eP1Energy` berechnet. Zuletzt wird der RT-Fehler $\|\boldsymbol{\sigma} - \boldsymbol{q}_\mathrm{RT}\|^2_{H(\mathrm{div},\Omega)} = \|\boldsymbol{\sigma} - \boldsymbol{q}_\mathrm{RT}\|^2_{L^2(\Omega)} + \|\mathrm{div}\, \boldsymbol{\sigma} - \mathrm{div}\, \boldsymbol{q}_\mathrm{RT}\|^2_{L^2(\Omega)}$, `Et`, mit Hilfe einer modifizierten Methode aus [Bri12] bestimmt. Diese Fehlerkomponenten und der gesamte Fehler, `E`, werden zurückgeben.

`visualisation(c4n,n4e,xh,eta4e,nrDoF)`

Mit Hilfe der Triangulierungsdaten, `c4n` und `n4e`, des Koeffizientenvektors zum Originalproblem `xh`, des Fehlerschätzers pro Element `eta4n` und der Zahl der Freiheitsgrade `nrDoF` werden in mehreren Figuren Plots erzeugt. Zunächst werden die Koeffizienten der einzelnen Lösungskomponenten aus `xh` ausgelesen. Dann wird in einer Figur ein `quiver`-plot der Komponente u_0 erzeugt, in einer anderen werden beide Komponenten von u_0 einzeln als `P04e`-Plot dargestellt. Auch die vier Komponenten von σ_0 werden in je einem `P04e`-Plot dargestellt. Ein weiterer `P04e`-Plot enthält den Fehlerschätzer pro Element `eta4n`. Außerdem gibt es noch einen `P1`-Plot der Fortsetzungen der beiden Komponenten von s_1. Eine weitere Figur enthält einen Plot der Triangulierung.

`visualisationExSol(c4n,n4e,xh,eta4e,nrDoF,uExact1,uExact2,sigmaExact1,...`
`sigmaExact2)`

Zusätzlich zu den Eingabedaten aus `visualisation` müssen hier noch die Funktionshändel zur exakten Lösung, `uExact1` bzw. `uExact2` für die Komponenten der Verschiebungsfunktion $u \in H^1(\Omega; \mathbb{R}^2)$ sowie `sigmaExact1` bzw. `sigmaExact2` für die Zeilen des Pseudospannungstensors $\sigma \in H(\mathrm{div}, \Omega; \mathbb{R}^{2\times 2})$, übergeben werden. Dann werden dieselben Figuren wie in `visualisation` erzeugt, nur das diese neben den Plots der berechneten Lösung auch die der entsprechenden Approximation der exakten Lösung zum Vergleich enthalten.

`convergenceExSol(nrDoF4lvl,eta4lvl,E4lvl,Eu,Esigma,Es,Et)`
Diese Funktion erzeugt einen Konvergenzplot des Fehlerschätzers, `eta4lvl`, des exakten Fehlers, `E4lvl` sowie der Fehleranteile `Eu`, `Esigma`, `Es`, `Et`. Wie üblich werden diese Größen gegen die Anzahl der Freiheitsgrade, `nrDoF4lvl`, geplottet.

5 Numerische Experimente

In diesem Kapitel werden einige numerische Experimente vorgestellt, die mit dem Programm aus Abschnitt 4 durchgeführt wurden. Es handelt sich um Benchmark Probleme, deren exakte Lösung z.T. bekannt ist und zur Validierung der Methode herangezogen werden kann. Zunächst werden zwei Beispiele auf dem Einheitsquadrat $(-1,1)^2$ vorgestellt, danach werden Beispiele auf nicht konvexen Gebieten untersucht, dabei kann zu dem letzten keine analytische Lösung bestimmt werden.

Die Experimente, bei denen eine exakte Lösung bekannt ist, zeigen, dass der verwendete Fehlerschätzer den Fehler zwar um einen gewissen Faktor unterschätzt, aber die richtige Konvergenzrate aufweist. Damit wird die Annahme unterstützt, dass der Term $\|F - b(x_h, \bullet)\|_{Y_h^*}$ ein angemessener Schätzer ist. Denn im Allgemeinen verschwindet der Datenapproximationsfehler, $\|F \circ (1-\Pi)\|_{Y^*}$, nicht, obwohl z.T. konstante rechte Seite f angesetzt werden. In das Funktional F aus Definition 3.2 fließen schließlich auch die nicht konstanten Dirichlet Randbedingungen g durch den Beitrag $\left\langle \gamma_\nu^{\mathcal{T}} \boldsymbol{\tau}, \hat{s} \right\rangle_{\partial\mathcal{T}}$ ein.

5.1 Das "colliding flow" Beispiel

Das erste Benchmark Problem verwendet $f \equiv 0$ als rechte Seite auf dem Einheitsquadrat $\Omega = (-1,1)^2$. Durch die exakte Lösung für das Geschwindigkeitsfeld

$$u\left((x_1,x_2)^\top\right) = \left(20x_1x_2^4 - 4x_1^5,\, 20x_1^4x_2 - 4x_2^5\right)^\top \quad \text{für alle } (x_1,x_2)^\top \in \Omega,$$

sind die Randdaten $g(x) = u(x)$ für alle $x \in \Gamma$ vorgegeben. Die Druckverteilung lautet

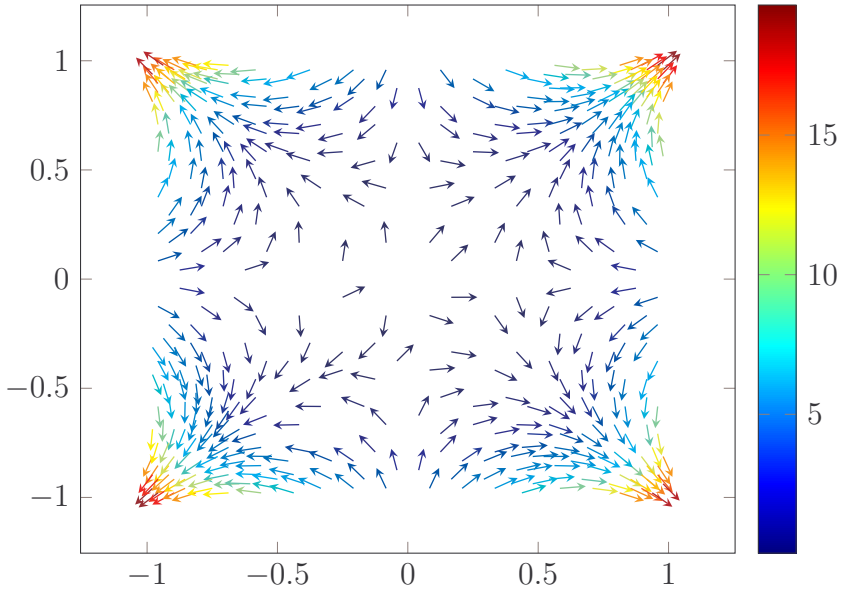
$$p\left((x_1,x_2)^\top\right) = 120x_1^2x_2^2 - 20x_1^4 - 20x_2^4 - 16/3 \quad \text{für alle } (x_1,x_2)^\top \in \Omega.$$

Wie in Abschnitt 2.6 erläutert, kann aus der diskreten Lösung $x_h = (\boldsymbol{\sigma}_0, u_0, s_1, t_0) \in X_h$ eine P_0-Projektion der diskreten Druckverteilung mittels

$$p_h = -1/2 \ \operatorname{tr}(\boldsymbol{\sigma}_0) \in P_0(\mathcal{T})$$

bestimmt werden. In Abbildung 5.1 sind diese Druckverteilung und das Geschwindigkeitsfeld, u_0, auf einer adaptiv mit Bulkparameter $\theta = 0.3$ bestimmten Triangulierung $\mathcal{T}_\ell$ mit 436 Dreiecken dargestellt. Die stärkste Verfeinerung ist in den Ecken des Gebietes zu erkennen, da dort die Funktion g die größte Steigung hat. In Abbildung 5.2 sind der Fehlerschätzer und die einzelnen Fehleranteile bei uniformer und adaptiver Verfeinerung mit $\theta = 0.3$ dargestellt. Wie auf einem konvexen Gebiet zu erwarten, hat die Verfeinerungsstrategie keinen Einfluss auf die Konvergenzrate. In beiden Fällen konvergiert der Fehlerschätzer von Beginn an mit der optimalen Rate von 0.5. Eine Betrachtung der einzelnen Fehleranteile zeigt, dass die Approximation der Pseudospannung zu Beginn etwas schlechter ist. Dies führt bei der uniformen Methode zu einem leichten vorasymptotischen

(a) Geschwindigkeitsfeld u_ℓ



(b) Druckverteilung p_ℓ

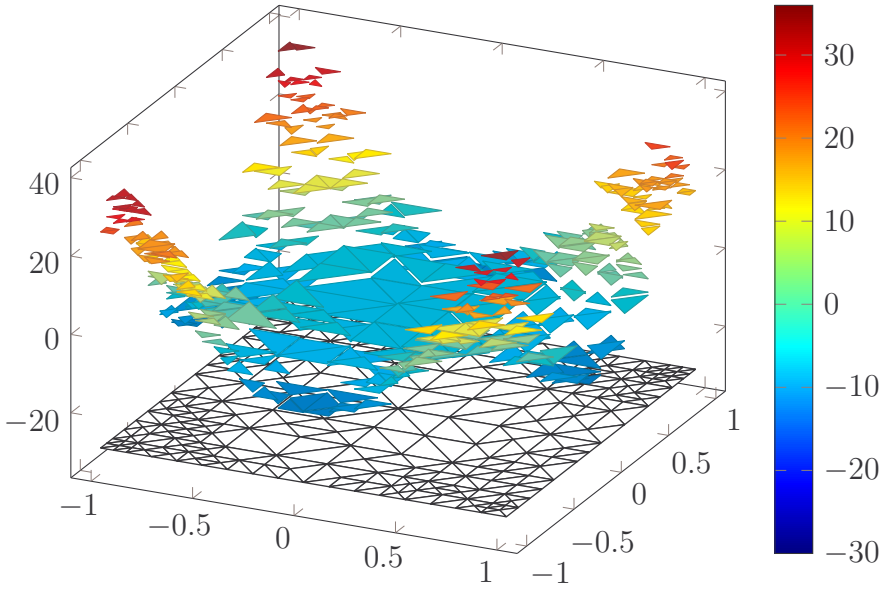
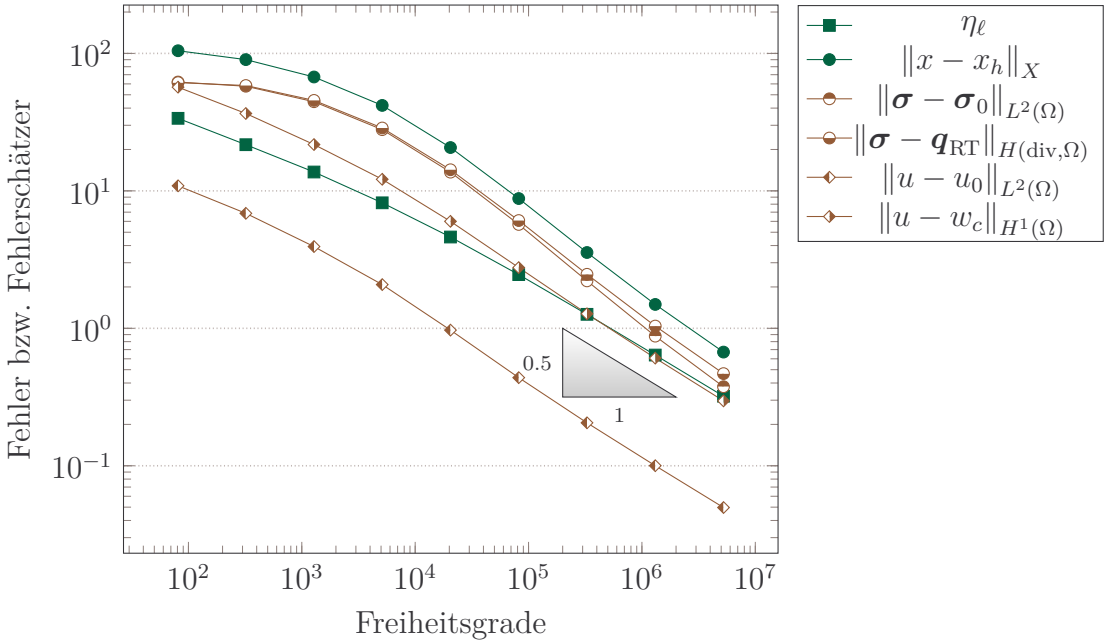


Abbildung 5.1: Plot des Geschwindigkeitsfeldes und der Druckverteilung für das "colliding flow" Beispiel auf einer Triangulierung mit 436 Elementen (4361 Freiheitsgrade), die bei adaptiver Verfeinerung mit $\theta = 0.3$ entstanden ist

(a) Konvergenzplots bei uniformer Verfeinerung



(b) Konvergenzplots bei adaptiver Verfeinerung mit $\theta = 0.3$

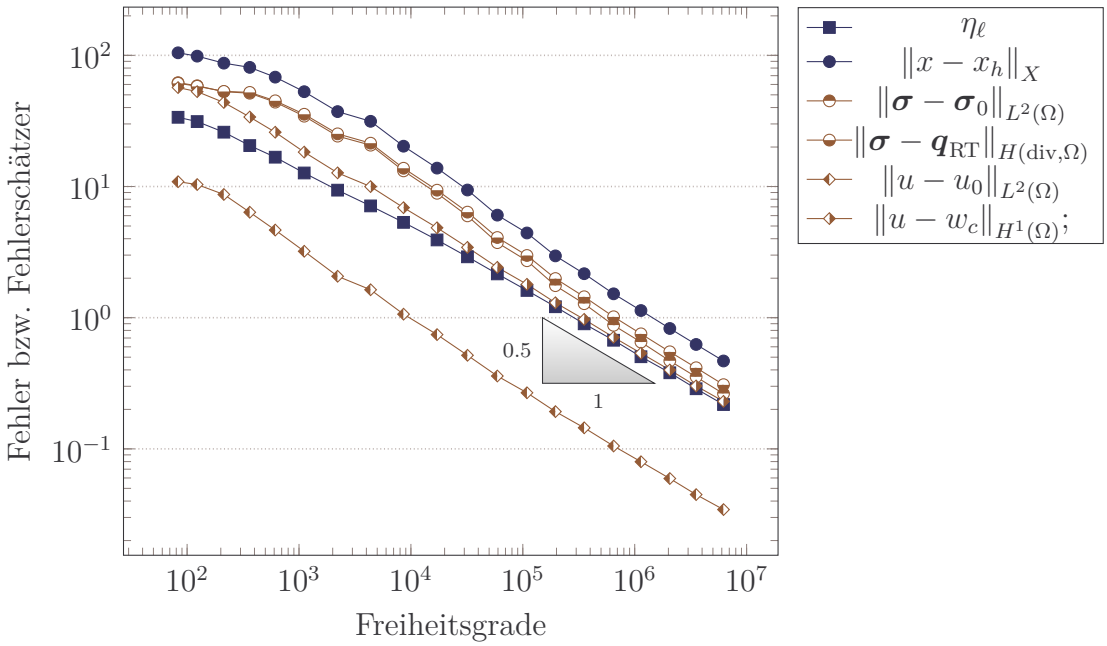


Abbildung 5.2: Konvergenzplots des Fehlerschätzers und der einzelnen Fehleranteile für das "colliding flow" Beispiel bei uniformer und adaptiver Verfeinerung

Effekt beim exakten Fehler. Eine Betrachtung der Komponenten von $\boldsymbol{\sigma}_0$ im Vergleich zur exakten Lösung veranschaulicht, dass zu Beginn die Extremalstellen in den Ecken des Gebietes nur schlecht approximiert werden können. Durch die adaptive Verfeinerung wird der Effekt daher deutlich schwächer. Eine genauere Approximation der Randdaten g der Geschwindigkeit beispielsweise durch Verwendung der Fassregel beeinflusst dieses Verhalten nicht. Die Fehler in den Lösungskomponenten, die das Geschwindigkeitsfeld approximieren, konvergieren auch von Beginn an wie erwartet.

5.2 Ein Poisson-Problem Beispiel

Auch hier wird als Gebiet das Einheitsquadrat $\Omega = (-1, 1)^2$ betrachtet. Die rechte Seite ist definiert als

$$f\left((x_1, x_2)^\top\right) = (2x_2, -2x_1) \quad \text{für alle } (x_1, x_2)^\top \in \Omega.$$

Damit ergibt sich als exakte Lösung zu den implizierten Randdaten für das Geschwindigkeitsfeld

$$u\left((x_1, x_2)^\top\right) = \left(-x_1^2 x_2, x_1 x_2^2\right)^\top \quad \text{für alle } (x_1, x_2)^\top \in \Omega.$$

Da hier der Druck $p \equiv 0$ ist, kann dieses Problem als Poisson Problem in zwei Komponenten betrachtet werden. Abbildung 5.3 zeigt, dass sowohl im uniformen als auch im adaptiven Fall die Konvergenzrate des Fehlers und des Fehlerschätzers mit 0.5 optimal sind. Es sind allerdings bei der adpativen Verfeinerung leichte Abweichungen von der optimalen Kurve zu erkennen. Diese können durch eine exaktere Integration der Randdaten verringert, aber nicht komplett aufgehoben werden. Eine genauere Betrachtung des lokalen Fehlers zeigt, dass dieser im ersten Level bei sechs der acht Elemente beinah identisch ist, diese Dreiecke aber nicht alle gleichermaßen verfeinert werden. Eine uniforme Verfeinerung vor Beginn des adaptiven Verfeinerns genügt, diesen Effekt zu verhindern.

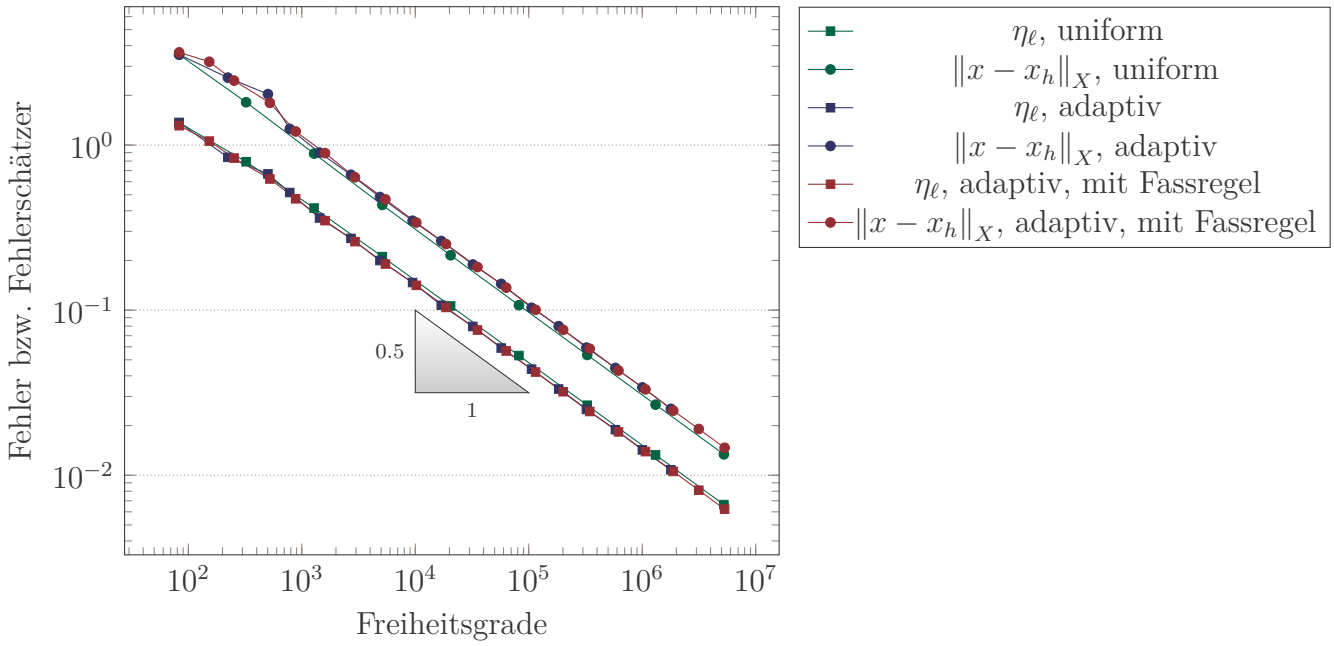


Abbildung 5.3: Konvergenzplot des Fehlerschätzers und exakten Fehlers in dem Poisson-Problem Beispiel bei uniformer Verfeinerung sowie adaptiver Verfeinerung mit $\theta = 0.3$ und verschiedener Behandlung der Randdaten

5.3 Ein Standardbeispiel im L-Gebiet

Dieses Benchmark Problem mit konstanter rechter Seite $f \equiv 0$ ist auf dem L-Gebiet, $\Omega = (-1,1)^2 \setminus ([0,1] \times [-1,0])$, zu lösen. Für die passenden Randdaten $g(x) = u(x)$ für alle $x \in \Gamma$ ist in [Ver89, S.324] eine exakte Lösung in Polarkoordinaten angegeben. Für alle $(r,\varphi) \in [0,\infty) \times [0, 3\pi/2]$ ist mit den Konstanten $\omega = 3\pi/2$ und $\alpha = 856399/1572864$ und der Hilfsfunktion

$$w(\varphi) = \frac{\sin((1+\alpha)\varphi)\cos(\alpha\omega)}{1+\alpha} - \cos((1+\alpha)\varphi) + \frac{\sin((\alpha-1)\varphi)\cos(\alpha\omega)}{1-\alpha} + \cos((\alpha-1)\varphi)$$

das exakte Geschwindigkeitsfeld gegeben durch

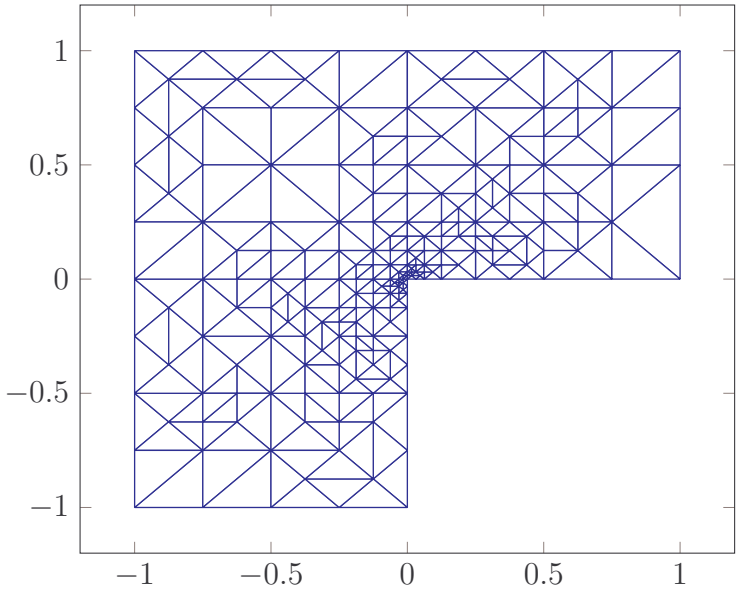
$$u(r,\varphi) = r^\alpha \Big((1+\alpha)\sin(\varphi)w(\varphi) + \cos(\varphi)w'(\varphi),\ -(1+\alpha)\cos(\varphi)w(\varphi) + \sin(\varphi)w'(\varphi)\Big)^\top.$$

Als Druckverteilung ergibt sich

$$p = -r^{\alpha-1}\Big((1+\alpha)^2 w'(\varphi) + w'''(\varphi)\Big)/(1-\alpha).$$

In Abbildung 5.4 ist ein mit adaptiver Verfeinerung mit $\theta = 0.3$ erzeugtes Gitter mit 371 Dreiecken zu sehen. Wie zu erwarten, tritt die stärkste Verfeinerung an der einspringenden Ecke auf. Außerdem ist eine Verfeinerung der Ränder zu erkennen, für die

(a) Triangulierung $\mathcal{T}_\ell$



(b) Geschwindigkeitsfeld u_ℓ

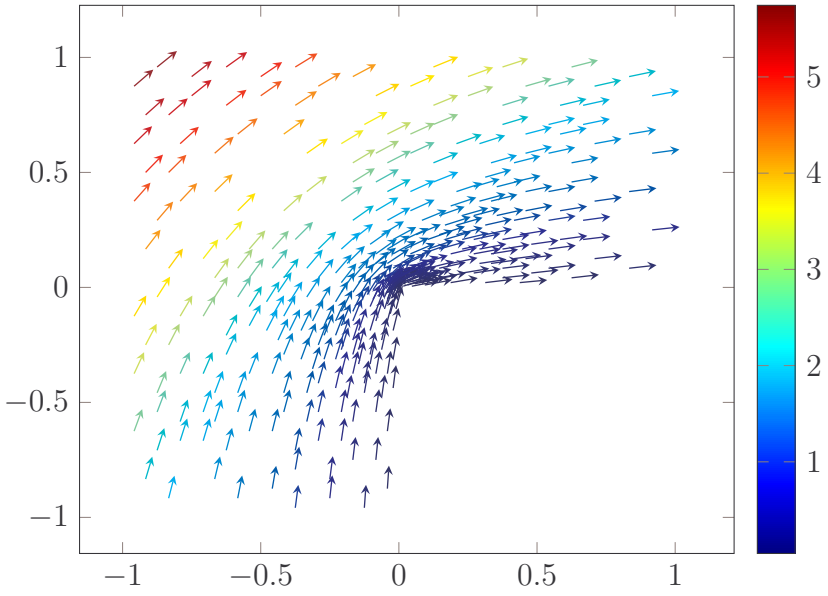
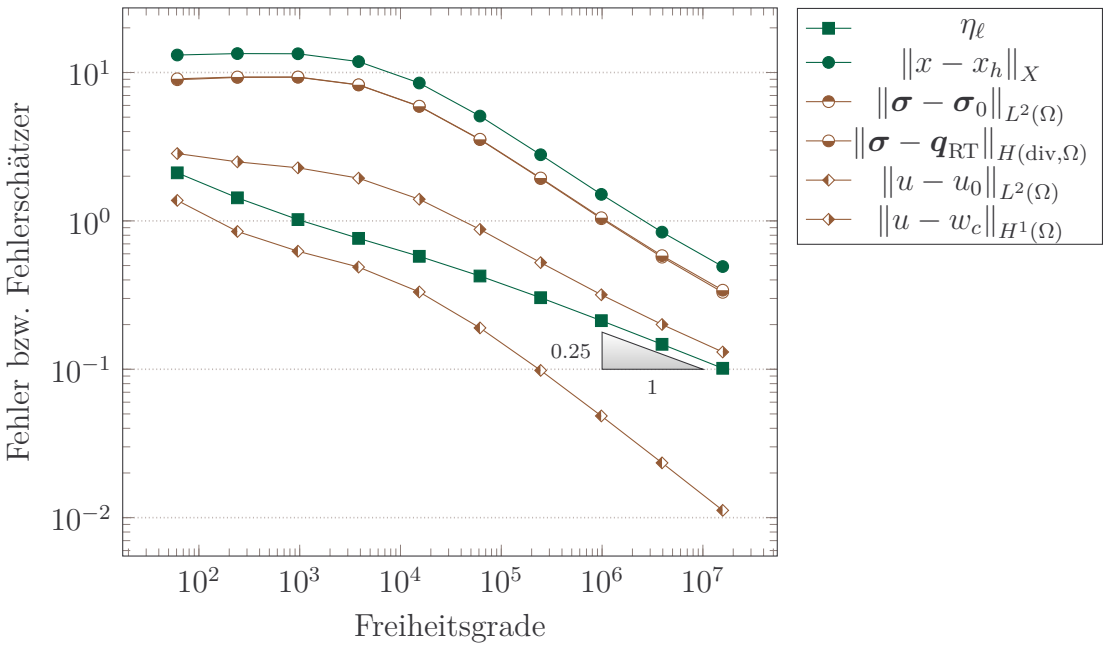


Abbildung 5.4: Triangulierung mit 371 Elementen (3711 Freiheitsgraden) und zugehöriges Geschwindigkeitsfeld in dem Standardbeispiel im L-Gebiet bei adaptiver Verfeinerung mit $\theta = 0.3$

inhomogene Dirichlet Daten gegeben sind. Die Konvergenzplots in Abbildung 5.5 bestätigen, dass die adaptive Verfeinerungsstrategie der uniformen überlegen ist. Im adaptiven Fall wird die optimale Konvergenzrate von 0.5 erreicht. Auf dem nicht konvexen L-Gebiet mit Singularität in der einspringenden Ecke liegt die zu erwartende Konvergenzrate für uniforme Verfeinerung bei 0.25. Diesen Wert erreicht die Methode. Beim exakten Fehler ist ähnlich wie im "colliding flow" Beispiel aus Abschnitt 5.1 ein vorasymptotischer Bereich zu erkennen, der bei der uniformen Methode stärker ausgeprägt ist. Ursache ist wieder die offenbar schlechte Approximation der Pseudospannung. Diese Größe mit Singularität in der einspringenden Ecke ist auf groben Gittern natürlich nur schlecht mit stückweise konstanten bzw. linearen Polynomen zu approximieren. Eine Verbesserung der Randdatenapproximation für die Geschwindigkeit beeinflusst dieses Ergebnis daher nicht. Nach dem vorasymptotischen Bereich scheinen die exakten Fehler zu gut zu konvergieren. Allerdings werden diese durch numerische Integration bestimmt und daher möglicherweise nahe der Singularität unterschätzt. Bei der Betrachtung des gesamten Graphs stimmt auch die Konvergenzrate des Fehlers mit den Erwartungen überein. Insgesamt werden ähnliche Ergebnisse erzielt wie in [Bri12] mit der konformen Least-Squares Methode.

(a) Konvergenzplots bei uniformer Verfeinerung



(b) Konvergenzplots bei adaptiver Verfeinerung mit $\theta = 0.3$

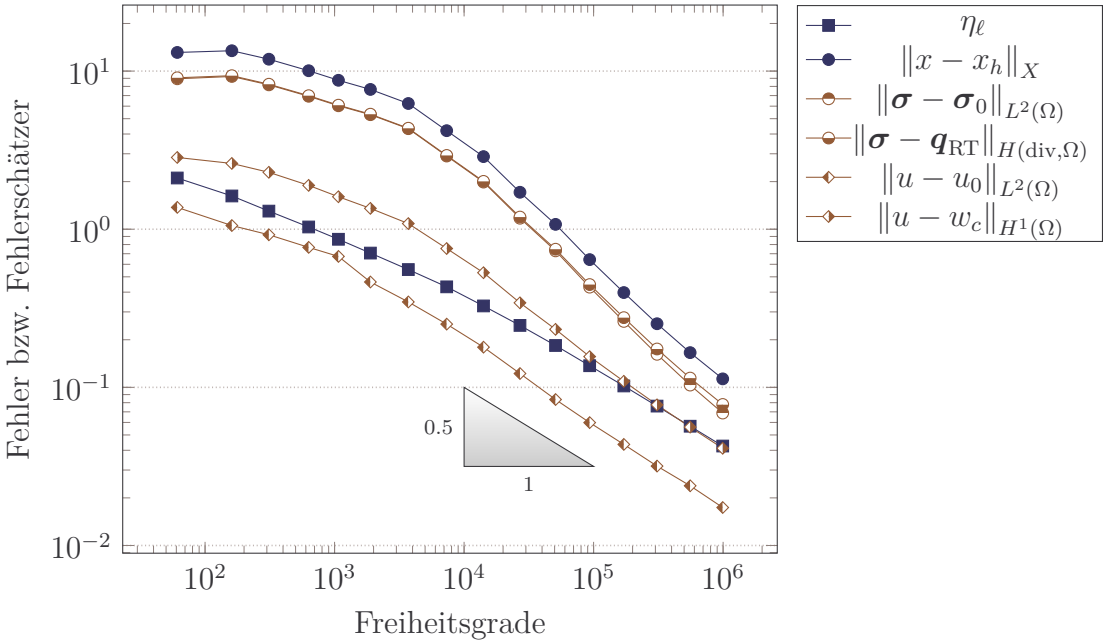


Abbildung 5.5: Konvergenzplot des Fehlerschätzers und der einzelnen Fehleranteile in dem Standardbeispiel im L-Gebiet bei uniformer und adaptiver Verfeinerung

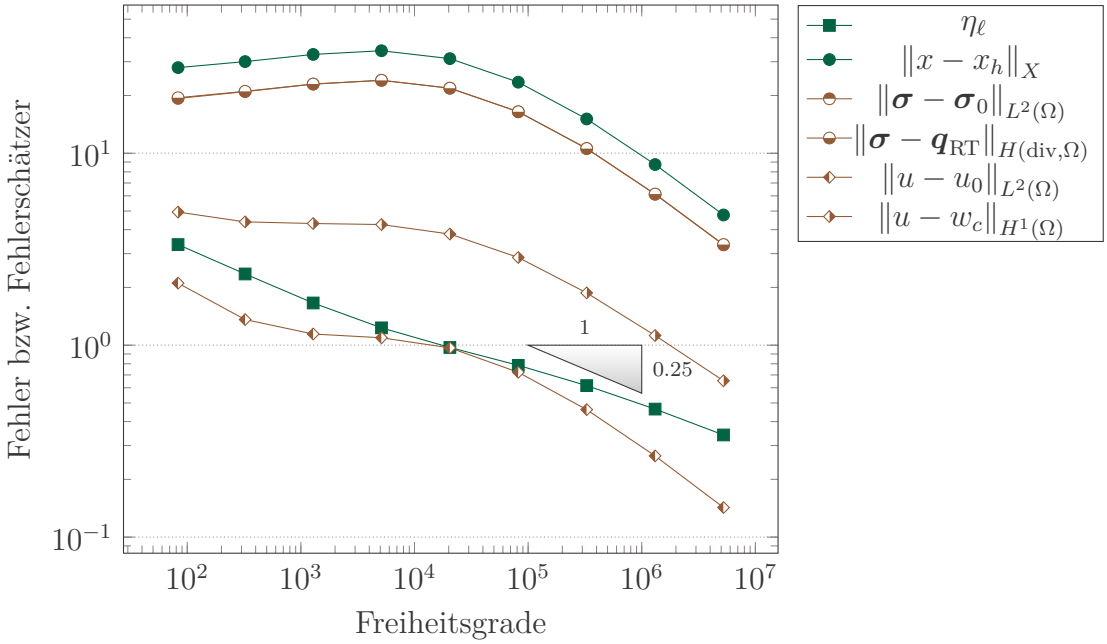
5.4 Ein Beispiel im Schlitzgebiet

Sei das Schlitzgebiet $\Omega = (-1,1)^2 \setminus ([0,1] \times 0)$ vorgeben. Zu konstanter rechter Seite $f \equiv 0$ und entsprechenden Randdaten kann die exakte Lösung in Polarkoordinaten angegeben werden. Für alle $(r,\varphi) \in [0,\infty) \times [0,2\pi)$ gilt

$$u(r,\varphi) = \frac{3\sqrt{r}}{2}\left(\cos(\varphi/2) - \cos(3\varphi/2),\ 3\sin(\varphi/2) - \sin(3\varphi/2)\right),$$
$$p(r,\varphi) = \frac{6}{\sqrt{r}}\cos(\varphi/2).$$

In den Konvergenzplots Abbildung 5.6 ist zu erkennen, dass der Fehlerschätzer wieder mit den erwarteten Raten von 0.25 bzw. 0.5 konvergiert. Im Unterschied dazu weist der exakte Fehler einen stark ausgeprägten vorasymptotischen Bereich auf, der auch durch adaptives Verfeinern nicht völlig verschwindet. Danach scheint der exakte Fehler wieder beinah zu gut zu konvergieren. Am Ursprung liegt hier wieder ein Singularität vor, die durch Polynome niedrigen Grades nur schlecht zu approximieren ist. Vergleichbare Ergebnisse werden auch mit der uniformen Least-Squares Methode (siehe [Bri12]) erzielt.

(a) Konvergenzplots bei uniformer Verfeinerung



(b) Konvergenzplots bei adaptiver Verfeinerung mit $\theta = 0.3$

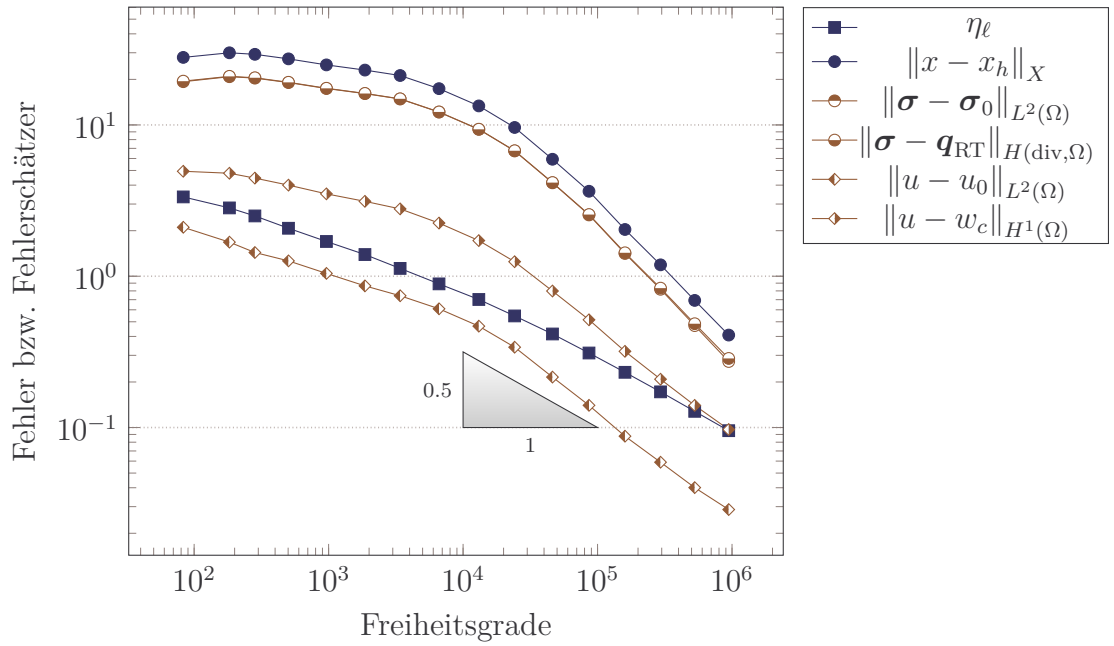


Abbildung 5.6: Konvergenzplot des Fehlerschätzers und der einzelnen Fehleranteile für das Beispiel im Schlitzgebiet bei uniformer und adaptiver Verfeinerung

5.5 Das "backward facing step" Beispiel

Dieses Benchmarkproblem ist anwendungsorientierter als die vorangegangenen Beispiele. Es beschreibt den Fluss einer Newtonschen Flüssigkeit in einer Röhre mit Verengung. Bei diesem Standardbeispiel ist die exakte Lösung nicht bekannt. Im Modell wird die Röhre durch ein verzerrtes L-Gebiet, $\Omega = ((-2,8)\times(-1,1)) \setminus ((-2,0)\times(-1,0))$ beschrieben, Ein- bzw. Ausfluss werden durch folgende Dirichlet Randdaten für $(x_1,x_2)^\top \in \Gamma$ gegeben

$$g\left((x_1,x_2)^\top\right) = \begin{cases} 1/10\,(-x_2(x_2-1),0)^\top & \text{für } x_1 = -2, \\ 1/80\,(-(x_2-1)(x_2+1),0)^\top & \text{für } x_1 = 8, \\ (0,0)^\top & \text{sonst,} \end{cases}$$

und die rechte Seite ist konstant $f \equiv 0$. Bei der adaptiven Verfeinerung mit $\theta = 0.3$ entsteht das in Abbildung 5.7 abgebildete Gitter mit 1551 Elementen. Wie zu erwarten, wird an Ein- bzw. Ausfluss besonders stark verfeinert, stärker noch an der kürzeren Seite, wo die Funktion g stärker ansteigt bzw. abfällt. Die einspringende Ecke wird erst relativ spät verfeinert. In Abbildung 5.8 sind die Fehlerschätzer bei adaptiver Verfeinerung mit variierendem Bulkparameter θ dargestellt. Es wird deutlich, dass gilt, je kleiner der Bulkparameter desto besser wird die Konvergenzrate. Auf der anderen Seite wird auch die Anzahl der Rechenschritte größer, die benötigt wird, um eine bestimmte Anzahl von Freiheitsgraden zu erreichen. Die optimalen Konvergenzrate von 0.5 ist auf jeden Fall bereits bei $\theta = 0.3$ gegeben. Der in den anderen Beispielen verwendete Parameter zur adaptiven Verfeinerung erscheint also geeignet. Dieses Gebiet ist leicht verzerrt, eine Skalierung wie in Abschnitt 4.5 ist somit möglich. Der Skalierungsfaktor $c = 10$ ist allerdings so klein, dass er keine sichtbaren Auswirkungen auf die Ergebnisse hat. Ebenso verhält es sich, wenn die Randintegrale mit Hilfe der Fassregel bzw. mit der Methode `integrate` exakter berechnet werden.

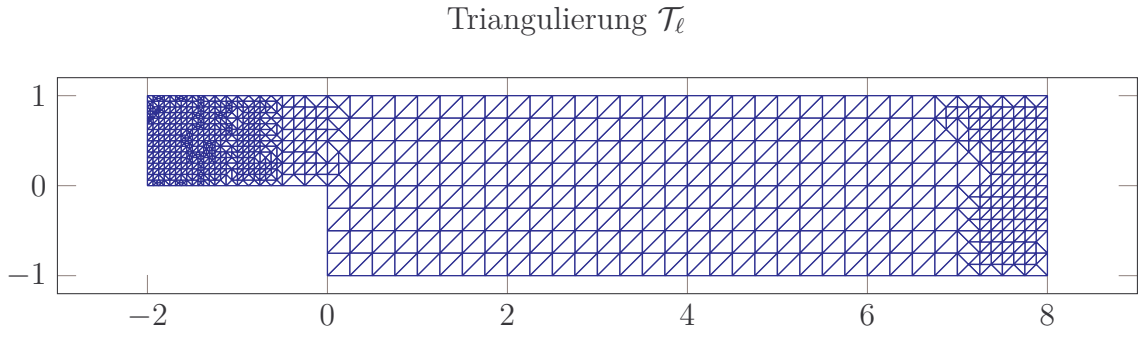


Abbildung 5.7: Triangulierung mit 1551 Elementen (15511 Freiheitsgraden) im "backward facing step" Beispiel bei adaptiver Verfeinerung mit $\theta = 0.3$

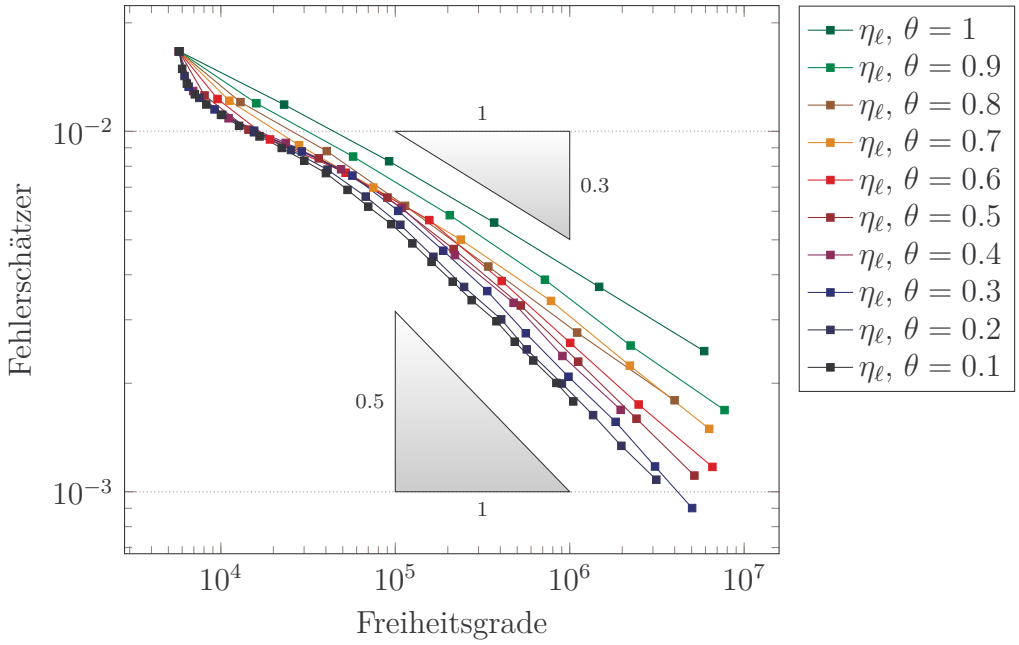


Abbildung 5.8: Konvergenzplot des Fehlerschätzers für das "backward facing step" Beispiel bei verschiedenen Bulkparametern θ

6 Zusammenfassung

In dieser Arbeit wurde eine neue dPG-Methode zum Lösen der Stokes Gleichungen untersucht, die zur Modellierung von Problemen in der Strömungsmechanik verwendet werden. Ausgangspunkt der Untersuchung sind bekannte Ergebnisse zur a-priori- und a-posteriori-Stabilität dieser Methoden, die in Abschnitt 3.1 komprimiert dargestellt werden. Sind Test- und Ansatzraum Hilberträume und ist die zugehörige Bilinearform beschränkt mit trivialen Kern, folgt diese Stabilität bereits aus der kontinuierlichen und diskreten Inf-Sup-Bedingung. Dieser Ansatz ermöglicht die Wahl deutlich kleinerer Räume als bisher in dPG-Methoden für die Stokes Gleichungen vorgeschlagen. Die alternative Herangehensweise basiert auf der Bestimmung eines Fortin-Interpolators. Da dieser Operator in der a-posteriori-Abschätzung im Datenapproximationsfehler auftritt, wird er für den implementierten zweidimensionalen Fall bestimmt und untersucht. Es ist leider nicht möglich zu beweisen, dass dieser Term von höherer Ordnung ist. Die numerischen Experimente bestätigen allerdings, dass der residuale Fehlerschätzer zur adaptiven Verfeinerung genügt.

Die verwendete Bilinearform ergibt sich aus der ultra-schwachen Formulierung der Stokes Gleichungen. In dieser werden die durch die gebrochenen Testfunktionen entstehenden Spurterme auf dem Skelett durch neue Variablen ersetzt. Die Behandlung beliebiger Dirichlet Randdaten wird durch die Einführung eines Spurterms im Funktional der rechten Seite realisiert. Zum Beweis der kontinuierlichen Inf-Sup-Bedingung werden das Splitting Lemma, zwei Dualitätslemmata und die eindeutige Lösbarkeit der gemischten Formulierung verwendet. Die diskrete Inf-Sup-Bedingung wird gezeigt, in dem zu gegebenen Ansatzfunktionen konkrete Testfunktionen gewählt werden. Dabei kommt eine Zerlegung stückweise konstanter deviatorischer Funktionen zum Einsatz, die auf der Lösbarkeit des Stokes Problems beruht. Außerdem erweist sich das Tr-Div-Dev Lemma als äußerst nützlich. Die Beweise werden für beliebige Raumdimensionen durchgeführt und verfolgen die Konstanten explizit.

Das Verfahren wurde für den zweidimensionalen Fall implementiert. Eine ausführliche Dokumentation ist in Kapitel 4 zu finden. Die numerische Experimente zeigen die Stabilität der Methode und optimale Konvergenzraten des verwendeten Fehlerschätzers. In den akademischen Beispielen mit bekannter exakter Lösung wird deutlich, dass der Fehlerschätzer asymptotisch die richtige Konvergenzrate aufweist. Daneben wurden Anwendungsbeispiele untersucht, die analytisch nicht lösbar sind. In allen Versuchen auf nicht konvexen Gebieten ist die adaptive Verfeinerungsstrategie der uniformen deutlich überlegen, wie die Konvergenzgraphen des Schätzers und des tatsächlichen Fehlers zeigen.

Im direkten Anschluss an diese Arbeit wären die Implementierung für höhere Raumdimensionen und Betrachtung höhere Polynomgrade lohnend. Die aktuelle Forschung beschäftigt sich mit dem Nachweis der optimalen Konvergenz für die adaptiven dPG-Methoden. Auf diesem Gebiet liegen noch keine analytischen Ergebnisse vor, daher ist dies äußerst interessant. In diesem Zusammenhang ist sicher auch eine genauere Untersuchung des Einflusses des Datenapproximationsfehlers aufschlussreich.

Literatur

[ACF99] J. Alberty, C. Carstensen, and S.A. Funken. Remarks around 50 lines of Matlab: short finite element implementation. *Numer. Algorithms*, 20(2–3):117–137, 1999.

[AF89] D. N. Arnold and R. S. Falk. A uniformly accurate finite element method for the Reissner-Mindlin plate. *SIAM Journal on Numerical Analysis*, 26(6):1276–1290, 1989.

[Alt06] H. W. Alt. Lineare Funktionalanalysis: Eine anwendungsorientierte Einführung, 5. Auflg, 2006.

[BC05] C. Bahriawati and C. Carstensen. Three Matlab implementations of the lowest-order Raviart-Thomas MFEM with a posteriori error control. *CMAM*, 5(4):333–361, 2005.

[BMS02] C L Bottasso, S Micheletti, and R Sacco. The discontinuous Petrov–Galerkin method for elliptic problems. *Computer Methods in Applied Mechanics and Engineering*, 191(31):3391–3409, 2002.

[Bra13] D. Braess. *Finite Elemente: Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*. Springer-Verlag, 2013.

[Bri12] P. Bringmann. *Least-Squares Finite-Elemente-Methoden für die Stokes Gleichungen*. Bachelorarbeit, Humboldt-Universität zu Berlin, 2012.

[Bri14] P. Bringmann. *Optimal Mesh-Refinement for Incompressible Fluid Dynamics*. Masterarbeit, Humboldt-Universität zu Berlin, 2014.

[BS08] S.C. Brenner and R. Scott. *The Mathematical Theory of Finite Element Methods*. Texts in Applied Mathematics. Springer, 2008.

[Car09a] C. Carstensen. AFEM Software-Paket. Humboldt Universität zu Berlin, 2009.

[Car09b] C. Carstensen. Yonsei Lectures at the WCU Department Computional Science and Engineering on Finite Element Method. Lokal verfügbar, 2009.

[CDG14] C. Carstensen, L. Demkowicz, and J. Gopalakrishnan. A posteriori error control for DPG methods. *SIAM Journal on Numerical Analysis*, 52(3):1335–1353, 2014.

[CDG15a] C Carstensen, L Demkowicz, and J Gopalakrishnan. Breaking spaces and forms for the DPG method and applications including Maxwell equations. *arXiv preprint arXiv:1507.05428*, 2015.

[CDG15b] C. Carstensen, L. Demkowicz, and J. Gopalakrishnan. The paradigmen of broken test-functions in DPG discretisations of elliptic second-order PDEs. in Vorbereitung, 2015.

[CG14] C. Carstensen and D. Gallistl. Guaranteed lower eigenvalue bounds for the biharmonic equation. *Numer. Math.*, 126(1):33–51, 2014.

[CGHW14] C. Carstensen, D. Gallistl, F. Hellwig, and L. Weggler. Low-order dPG-FEM for an elliptic PDE. *Comput. Math. Appl.*, 68(11):1503–1512, 2014.

[CGS14] C. Carstensen, D. Gallistl, and M. Schedensack. L2 Best-Approximation of the Elastic Stress in the Arnold-Winther FEM. *Preprint 2014-15, Humboldt-Universität zu Berlin, Institut für Mathematik*, 2014.

[CH15] C. Carstensen and F. Hellwig. Low-order DPG-FEMs for linear elasticity. *SIAM Journal on Numerical Analysis*, eingereicht, 2015.

[CPR13] C. Carstensen, D. Peterseim, and H. Rabus. Optimal adaptive nonconforming FEM for the Stokes problem. *Numer. Math.*, 123(2):291–308, 2013.

[CTVW10] Z. Cai, C. Tong, P. S Vassilevski, and C. Wang. Mixed finite element methods for incompressible flow: Stationary Stokes equations. *Numerical Methods for Partial Differential Equations*, 26(4):957–978, 2010.

[CW07] Z. Cai and Y. Wang. A multigrid method for the pseudostress formulation of Stokes problems. *SIAM Journal on Scientific Computing*, 29(5):2078–2095, 2007.

[DG10] L. Demkowicz and J. Gopalakrishnan. A class of discontinuous Petrov–Galerkin methods. Part I: The transport equation. *Computer Methods in Applied Mechanics and Engineering*, 199(23):1558–1572, 2010.

[DG11a] L. Demkowicz and J. Gopalakrishnan. Analysis of the DPG method for the Poisson equation. *SIAM Journal on Numerical Analysis*, 49(5):1788–1809, 2011.

[DG11b] L. Demkowicz and J. Gopalakrishnan. A class of discontinuous Petrov–Galerkin methods. Part II. Optimal test functions. *Numerical Methods for Partial Differential Equations*, 27(1):70–105, 2011.

[DGN12] L. Demkowicz, J. Gopalakrishnan, and A. H. Niemi. A class of discontinuous Petrov–Galerkin methods. Part III: adaptivity. *Applied numerical mathematics*, 62(4):396–427, 2012.

[GQ14] J. Gopalakrishnan and W. Qiu. An analysis of the practical DPG method. *Mathematics of Computation*, 83(286):537–552, 2014.

[GR86] Vivette Girault and Pierre-Arnaud Raviart. *Finite element methods for Navier-Stokes equations: theory and algorithms*. Springer, 1986.

[Hel14] F. Hellwig. *Drei dPG-Methoden niedriger Ordnung für Lineare Elastizität*. Masterarbeit, Humboldt-Universität zu Berlin, 2014.

[Kön04] K. Königsberger. *Analysis 2*. Springer-Verlag Berlin Heidelberg New York, 4. auflage edition, 2004.

[LS10] R. S. Laugesen and B. A. Siudeja. Minimizing Neumann fundamental tones of triangles: an optimal Poincaré inequality. *Journal of Differential Equations*, 249(1):118–135, 2010.

[PW60] L. E. Payne and H. F. Weinberger. An optimal Poincaré inequality for convex domains. *Archive for Rational Mechanics and Analysis*, 5(1):286–292, 1960.

[RBTD14] N. V. Roberts, T. Bui-Thanh, and L. Demkowicz. The DPG method for the Stokes problem. *Computers & Mathematics with Applications*, 67(4):966 – 995, 2014. High-order Finite Element Approximation for Partial Differential Equations.

[Ver89] R Verfürth. A posteriori error estimators for the Stokes equations. *Numerische Mathematik*, 55(3):309–325, 1989.

[Wer11] D. Werner. *Funktionalanalysis*. Springer Berlin Heidelberg, 2011.

Danksagung

Es gibt ohne Frage viele Menschen, die mir bei meinem Studium und dieser Arbeit geholfen haben.

Besonders möchte ich mich an dieser Stelle bei Prof. Carstensen für die umfangreiche Betreuung bedanken. Die vielen Gespräche und Diskussionen haben mir sehr geholfen. Außerden möchte ich den Mitgliedern der Arbeitsgruppe vom Prof. Carstensen danken. Sie alle hatten immer ein offenes Ohr für Fragen und haben Probleme mit mir diskutiert. Friedrike Hellwig konnte mir bei vielen Problemen mit den dPG-Methoden helfen, Philipp Bringmann hat sich meinen `matlab`-Fragen angenommen und Karoline Köhler hat mir zugehört, wann immer ich nicht weiter kam.

Meiner Familie kann ich gar nicht genug danken. Es ist schön zu wissen, dass ich immer kommen kann und ihr versuchen werdet mir zu helfen. Papa, ich freue mich, wenn du auch zukünftig meine Arbeiten liest und mir sagst an welcher Stelle nur noch Eingeweihte verstehen, was ich meine. Ein großes Danke geht an Giordana Tornow, Linda Nguyen und Maren Strobl. Mit euch hat sogar das Studieren Spaß gemacht. Lucian, bei dir möchte ich mich auch bedanken, nicht nur für das Finden von gefühlt tausenden Tippfehlern, aber das weißt du.

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 Masterarbeit in diesem Studiengang einreiche.

Berlin, den 29. September 2015 Unterschrift