

Eingereicht von
Gobrial Mario

Angefertigt am
**Institut für Numerische
Mathematik**

Betreuer
Dr. Martin Neumüller

September 2018

Ein äquilibrierter Fehlerschätzer für die Poisson-Gleichung



Bachelorarbeit
zur Erlangung des akademischen Grades
Bachelor of Science
im Bachelorstudium
Technische Mathematik

Zusammenfassung

In den letzten Jahren gewannen a-posteriori Fehlerschätzer immer mehr an Popularität und es wurden eine Vielzahl von Verfahren vorgestellt, um den Approximationsfehler zu kalkulieren. Äquilibrierte Fehlerschätzer sind eine Art von a-posteriori Fehlerschätzer, die den Fehler basierend auf dem Residuum berechnen. Die Resultate von äquilibrierten Techniken sind konstant-freie Fehlerabschätzungen bezüglich der Energienorm, die in Folge dessen zu exakte Fehlerberechnungen führen. In dieser Arbeit wird der äquilibrierte Fehlerschätzer bezüglich der Poisson-Gleichung behandelt. Zunächst werden Grundlagen eingeführt, die in weiterer Folge als Basis für die Theorie der äquilibrierten Fehlerschätzer dienen. Im zweiten Abschnitt der Arbeit wird der residuale Fehlerschätzer betrachtet, der sich als geeignetes Hilfsmittel für die Berechnung des äquilibrierten Fehlerschätzers herausstellt. Unter anderem werden zwei wichtige Eigenschaften für den residualen Fehlerschätzer erwähnt. Zum einen die Effizienz und zum anderen die Zuverlässigkeit eines Schätzers. In weiterer Folge werden notwendige Markierungs- und Verfeinerungsstrategien vorgestellt, die dazu beitragen, dass die Näherungslösung nur an Stellen mit großem Fehlerbeitrag verbessert wird. Schließlich wird der Satz von Prager-Syngé eingeführt, der den äquilibrierten Fehlerschätzer über das duale Variationsproblem berechnet, und dadurch sich ein Verfahren zur Konstruktion von äquilibrierten Flüssen beschreiben lässt. Mit einer Implementierung im eindimensionalen Fall werden praktische Fälle ausgetestet.

Danksagung

Ich möchte mich im Rahmen dieser Arbeit bei meinem Betreuer, Herrn Martin Neumüller, für die Möglichkeit des Verfassens dieser Bachelorarbeit und der Bereitstellung jeglicher Nachschlagewerke, sowie für die Unterstützung und die präzise Beantwortung all meiner Fragen herzlichst bedanken.

Ebenso möchte ich meiner Familie einen Dank aussprechen, die mich während meinem Studium begleitet, tatkräftig unterstützt und stets motiviert hat.

Ein großes Dankeschön gilt insbesondere meiner Schwester Marina Gobrial, die sich für das Korrekturlesen dieser Arbeit bereitgestellt hat.

Mario Gobrial
Linz, September 2018

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen	3
2.1	Die Poisson-Gleichung	3
2.2	Klassische Formulierung	3
2.3	Variationsformulierung	4
2.4	Schwache Ableitung	5
2.5	Sobolev-Räume	5
2.6	Der Spursatz	6
2.7	Das endgültige Variationsproblem	6
2.8	Der Satz von Lax-Milgram	7
2.8.1	Abstraktes Variationsproblem	8
2.8.2	Die Friedrichs-Ungleichung	9
2.8.3	Der Riesz'sche Darstellungssatz	10
2.9	Das symmetrische Variationsproblem	11
2.10	Die Galerkin-Prinzip Approximation	12
2.11	Finite-Elemente-Räume	13
2.12	Das lineare Gleichungssystem	14
3	A-posteriori Fehlerschätzer und adaptive Verfahren	16
3.1	Residuale Fehlerschätzer	17
3.1.1	Eine obere Schranke für $\ u - u_N\ _a$	19
3.2	Effizienz des residualen Fehlerschätzers η_R	22
3.3	Adaptive Gittersteuerung	23
3.3.1	Markierungsstrategien	23
3.3.2	Lokale Verfeinerung und Gitterabschluss	24
4	Ein äquilibrierter Fehlerschätzer	27
4.1	Das gemischte Variationsproblem	28
4.2	Gemischte Finite-Elemente-Methoden	30
4.3	Gemischte Formulierung des Poisson-Problems	31
4.3.1	Satz von Prager-Synge und Anwendung	35
5	Numerisches Experiment	42
6	Konklusion	46
	Literaturverzeichnis	47

Kapitel 1

Einleitung

Das Ziel dieser Bakkalaureatsarbeit ist es, einen äquilibrierten Fehlerschätzer für die zwei- bzw. eindimensionale Poisson-Gleichung zu konstruieren und im Zuge dessen ein Programm zu entwickeln, das diesen Fehlerschätzer im eindimensionalen Fall berechnen kann. Grundsätzlich wird zwischen zwei Arten von Fehlerschätzern unterschieden. Zum einen der a-priori Fehlerschätzer und zum anderen der a-posteriori Fehlerschätzer, wobei der äquilibrierte Fehlerschätzer der zweit genannten Art zugeordnet wird.

Im Laufe der Jahre gewannen a-posteriori Fehlerschätzer immer mehr an Popularität und spielen nun eine wesentliche Rolle in der Mathematik sowie in der Industrie. Zum Beispiel werden physikalische Phänomene im Engineering Bereich, wie in etwa das Wärmeleitproblem, als Modelle aufgefasst, die durch partielle Differentialgleichungen beschrieben werden. Im Allgemeinen sind diese Gleichungen jedoch nicht analytisch lösbar, weshalb numerische Methoden entwickelt wurden, um diese Gleichungen annähernd lösen zu können.

Eine der am häufigst angewendeten Methoden ist die Finite-Elemente-Methode, die eine Vielfalt an Anwendungsmöglichkeiten anbietet. Das Ziel ist nun, eine numerische Näherungslösung zu finden, die die exakte Lösung der partiellen Differentialgleichung möglichst genau approximiert. Der äquilibrierte Fehlerschätzer soll das Verfahren nun so beeinflussen, dass die Näherungslösung die analytische Lösung möglichst gut approximiert und zugleich das numerische Verfahren steuert, sodass möglichst wenige Operationen benötigt werden.

In Kapitel 2 werden zunächst die Grundlagen der Poisson-Gleichung und die Herleitung der Finite-Elemente-Methode formuliert. Basierend auf den Gaußschen Integralsatz, wird die Variationsformulierung eingeführt. Ein weiterer wichtiger Begriff stellt die schwache Ableitung dar. Der Sobolev-Raum erster Ordnung H^1 dient als geeigneter Arbeitsraum, da die Existenz der schwachen Ableitung jeder H^1 -Funktion sichergestellt ist. Der zentrale Satz der Variationsformulierung ist der Satz von Lax-Milgram, der die Existenz und Eindeutigkeit einer Lösung garantiert. Zuletzt wird die Finite-Elemente-Methode eingeführt, mit der die numerische Berechnung einer Lösung des zum Variationsproblem äquivalenten diskreten Problems möglich ist.

Der Begriff a-posteriori Fehlerschätzer wird in Kapitel 3 eingeführt. Außerdem wird in diesem Kapitel der residuale Fehlerschätzer genauer betrachtet. Dieser beruht auf die Berechnung des Fehlers anhand des Residuums. Wichtige Eigenschaften die ein Fehlerschätzer besitzen kann, sind die sogenannte Effizienz und Zuverlässigkeit. Es wird sich herausstellen, dass die Effizienz und die Zuverlässigkeit für den residualen Fehlerschätzer gelten werden. Ein weiterer Aspekt, welcher in Kapitel 3 behandelt wird, ist die adaptive Gittersteuerung. Der Zweck dieser Gittersteuerung ist das Gitter des numerischen Verfahrens an jenen Stellen zu verfeinern, in dem der Diskretisierungsfehler groß ist. Verschiedene und geeignete Markierungs- und Verfeinerungsstrategien werden an dieser Stelle explizit vorgestellt.

In Kapitel 4 betrachtet man schließlich den a-posteriori Fehlerschätzer über duale Variationsprobleme hinweg. Die Intention dieses Kapitels ist die völlige Zuverlässigkeit des Fehlerschätzers zu zeigen. Völlig zuverlässige Fehlerschätzer lassen sich als zuverlässige Fehlerschätzer, dessen Abschätzungen von keiner Konstante, sprich Konstante ist gleich 1, abhängen, definieren. Dies wird erreicht, indem das gemischte duale Problem des Poisson-Problems aufgestellt wird. Der zentrale Satz, der schließlich das gewünschte Ziel liefern wird, ist der Satz von Prager-Synge. Durch die Einführung von Flusskorrekturen kann der äquilibrierte Fehlerschätzer letztendlich berechnet werden.

Im letzten Kapitel, wird ein Beispiel vorgestellt, das den äquilibrierten Fehlerschätzer, anhand von den in den vorigen Kapiteln beschriebenen Methoden umsetzt. Wie am Anfang bereits erwähnt, wird nur das eindimensionale Poisson-Problem betrachtet. Das Programm berechnet im Intervall $(0, 1)$ ein eindimensionales Poisson-Problem mit homogenen Dirichlet-Randbedingungen. Das äquivalente lineare Gleichungssystem wird über die Finite-Elemente-Methode aufgestellt und anschließend gelöst. Das Gleichungssystem hängt von der Diskretisierung des Intervalls ab, weshalb der äquilibrierte Fehlerschätzer elementweise berechnet wird. Die in Kapitel 3 beschriebenen Markierungs- und Verfeinerungsstrategien bestimmen gewisse Gitterelemente, in der weitere Verfeinerungen durchgeführt werden müssen, bis schließlich der Fehler klein genug und dadurch die Näherungslösung möglichst genau approximiert worden ist.

Kapitel 2

Grundlagen

2.1 Die Poisson-Gleichung

In diesem Kapitel wird die Poisson-Gleichung, die Variationsformulierung, sowie die Finite-Elemente-Methode vorgestellt. Speziell erwähnt wird die Poisson-Gleichung im zweidimensionalen bzw. im eindimensionalen Fall, da diese die Theorie der äquilibrierten Fehlerschätzer bzw. die Implementierung eines eindimensionalen Beispiels vereinfacht. Dieses Kapitel dient hauptsächlich als Vorbereitung und Einführung, um die Theorie der äquilibrierten Fehlerschätzer verstehen zu können.

Definition 2.1.1 Sei $U \subset \mathbb{R}^n$ eine offene Menge und $f : U \rightarrow \mathbb{R}$ eine vorgegebene Funktion, dann bezeichnet man die partielle Differentialgleichung

$$-\Delta u(x) = f(x) \quad \forall x \in U \quad (2.1)$$

als Poisson-Gleichung bzw. die inhomogene Laplace-Gleichung. Der Operator Δ wird als Laplace Operator bezeichnet und ist definiert als

$$\begin{aligned} -\Delta u &= \operatorname{div}(\nabla u), \\ -\Delta u &= \nabla \cdot \nabla u. \end{aligned}$$

Speziell lautet diese für die zweidimensionale Poisson-Gleichung ($n = 2$), $\Omega \subset \mathbb{R}^2$ offen mit Rand $\Gamma := \partial\Omega$ und $f : \Omega \rightarrow \mathbb{R}$,

$$-\Delta u(x) = f(x) \quad \forall x \in \Omega. \quad (2.2)$$

Die eindimensionale Poisson-Gleichung auf dem Intervall $(0, 1)$ ist gegeben durch

$$-u''(x) = f(x) \quad \forall x \in (0, 1). \quad (2.3)$$

Die Poisson-Gleichung kann schließlich als Randwertproblem klassisch formuliert werden.

2.2 Klassische Formulierung

Gesucht ist eine Funktion $u : \Omega \rightarrow \mathbb{R}$, sodass

$$-\Delta u(x) = f(x) \quad \forall x \in \Omega, \quad (2.4)$$

$$u(x) = 0 \quad \forall x \in \Gamma, \quad (2.5)$$

oder im eindimensionalen Fall

$$-u''(x) = f(x) \quad \forall x \in (0, 1) , \quad (2.6)$$

$$u(0) = u(1) = 0 , \quad (2.7)$$

erfüllt ist.

Die Randbedingungen (2.5) und (2.7) werden homogene Dirichlet-Randbedingungen genannt. Im Allgemeinen müssen die Randbedingungen nicht 0 sein, sondern können den Wert einer vorgegebenen Funktion annehmen.

Für numerische Verfahren ist die klassische Formulierung nicht genügend, deswegen wird ein schwacher Lösungsbegriff eingeführt. Dies erfolgt über die sogenannte Variationsformulierung.

2.3 Variationsformulierung

Für die Herleitung der Variationsformulierung wird die klassische Formulierung benötigt und es werden bestimmte Integrabilitätsbedingungen angenommen. Die Herleitung erfolgt in mehreren Schritten.

- **Schritt 1:** Multiplikation der Differentialgleichung (2.4) mit einer Testfunktion $v : \Omega \rightarrow \mathbb{R}$ und Integration über das Rechengebiet Ω ,

$$\int_{\Omega} -\Delta u(x)v(x) \, dx = \int_{\Omega} f(x)v(x) \, dx .$$

- **Schritt 2:** Partielle Integration des Hauptteils (dies erfolgt über den Gauß'schen Integralsatz)

$$\int_{\Omega} -\Delta u(x)v(x) \, dx = \int_{\Omega} \nabla u(x) \cdot \nabla v(x) \, dx - \int_{\Gamma} \nabla u(x) \cdot n(x)v(x) \, ds_x ,$$

wobei $n(x)$ der Normalvektor auf dem Gebiet Γ ist.

- **Schritt 3:** Einarbeitung der Randbedingungen.

Da die Ableitung von u am Dirichlet-Rand nicht bekannt ist und u am Dirichlet-Rand nach Voraussetzung 0 sein soll, wählt man eine Testfunktion v , die am Dirichlet-Rand ebenso 0 ist. Man erhält schließlich

$$\int_{\Omega} -\Delta u(x)v(x) \, dx = \int_{\Omega} \nabla u(x) \cdot \nabla v(x) \, dx .$$

- **Zusammenfassung:** Nun lässt sich das Variationsproblem wie folgt aufschreiben: Gesucht ist $u : \Omega \rightarrow \mathbb{R}$, sodass

$$\int_{\Omega} \nabla u(x) \cdot \nabla v(x) \, dx = \int_{\Omega} f(x)v(x) \, dx \quad (2.8)$$

für alle Testfunktionen $v : \Omega \rightarrow \mathbb{R}$ mit $v|_{\Gamma} = 0$ erfüllt ist.

Da die Ableitungen nur unter einem Integral auftauchen, führt man einen neuen Ableitungsbegriff ein, die sogenannte **schwache Ableitung**.

2.4 Schwache Ableitung

Definition 2.4.1 (Schwacher Gradient) Sei $\Omega \subset \mathbb{R}^2$ und $v \in L^2(\Omega)$. Eine Funktion $w \in [L^2(\Omega)]^2$ heißt schwacher Gradient von v , falls für alle Testfunktionen mit kompaktem Träger $\varphi \in C_0^\infty(\Omega)$ gilt

$$\int_{\Omega} w(x)\varphi(x) \, dx = - \int_{\Omega} v(x)\nabla\varphi \, dx .$$

Falls $v \in C^1(\bar{\Omega})$ ist, so stimmt die klassische Ableitung mit der schwachen Ableitung überein.

Lemma 2.4.2 Eine stückweise differenzierbare Funktion hat genau dann eine schwache Ableitung, wenn diese stetig ist.

Beweis: Vgl. [7] Lemma 1.9 .

□

2.5 Sobolev-Räume

Für die vorkommenden Funktionen des Variationsproblems (2.8) möchte man die Existenz der schwachen Ableitung sicherstellen. Der dazu geeignete Raum ist der sogenannte Sobolev-Raum.

Definition 2.5.1 (Sobolev-Raum) Der Sobolev-Raum der Ordnung 1 ist gegeben durch

$$H^1(\Omega) := \{v \in L^2(\Omega) : \nabla v \in [L^2(\Omega)]^2\} ,$$

mit dem inneren Produkt

$$(u, v)_{H^1(\Omega)} := (u, v)_{L^2(\Omega)} + (\nabla u, \nabla v)_{L^2(\Omega)} = \int_{\Omega} u(x)v(x) \, dx + \int_{\Omega} \nabla u(x) \cdot \nabla v(x) \, dx ,$$

der Norm

$$\|u\|_{H^1(\Omega)} := [(u, u)_{H^1(\Omega)}]^{1/2} = \left[\int_{\Omega} |u(x)|^2 \, dx + \int_{\Omega} |\nabla u(x)|^2 \, dx \right]^{1/2}$$

und der Semi-Norm

$$|u|_{H^1(\Omega)} := \|\nabla u\|_{L^2(\Omega)} = \left[\int_{\Omega} |\nabla u(x)|^2 \, dx \right]^{1/2}$$

für alle $u, v \in H^1(\Omega)$.

Bemerkung Der Raum $(H^1(\Omega), (\cdot, \cdot)_{H^1(\Omega)})$ ist ein Hilbertraum. Die Funktionenräume $C^1(\bar{\Omega})$ und $C^\infty(\bar{\Omega})$ liegen dicht in $H^1(\Omega)$.

2.6 Der Spursatz

Für allgemeine $L^2(\Omega)$ -Funktionen sind Auswertungen am Rand Γ , das heißt $u|_{\Gamma}$, schwer zu definieren. Der Grund dafür ist, dass der Dirichlet-Rand bezüglich Ω nur eine Nullmenge und die Funktion u für Nullmengen nicht definiert ist. Der Spursatz liefert einen geeigneten Ausweg um wohldefinierte Randwerte ausdrücken zu können.

Theorem 2.6.1 (*Spursatz*) Sei Ω ein beschränktes Gebiet mit Lipschitzstetigem Rand. Dann existiert ein linearer stetiger Operator, auch genannt Spuroperator,

$$T_0 : H^1(\Omega) \rightarrow H^{\frac{1}{2}}(\Gamma)$$

mit der Eigenschaft, falls $u \in C(\bar{\Omega}) \cap H^1(\Omega)$, dann gilt

$$T_0(u) = u|_{\Gamma} .$$

Beweis: Vgl. [2] Satz 3.1 .

□

2.7 Das endgültige Variationsproblem

Aufbauend auf die bereits in den vorherigen Kapiteln bearbeitete Theoriegrundlage, ist es nun möglich das endgültige Variationsproblem zu betrachten. Es werden dafür neue Bezeichnungen für den Sobolev-Raum und den Raum der Testfunktionen eingeführt. Dabei sei

$$V := H^1(\Omega)$$

der allgemeine Arbeitsraum und

$$V_0 := \{v \in V : v = 0 \text{ auf } \Gamma\}$$

ein linearer Unterraum. Das Variationsproblem (2.8) lässt sich schließlich schreiben als: Gesucht ist $u \in V_0$, sodass

$$a(u, v) = \langle F, v \rangle \quad \forall v \in V_0 \tag{2.9}$$

erfüllt ist. Dabei sind

$$a(u, v) = \int_{\Omega} \nabla u(x) \cdot \nabla v(x) \, dx ,$$

$$\langle F, v \rangle = \int_{\Omega} f(x)v(x) \, dx ,$$

für alle $u, v \in V$.

Bemerkung Die Abbildung $a(u, v)$ wird auch als Bilinearform bezeichnet und das Funktional $\langle F, v \rangle = F(v)$ ist linear.

Definition 2.7.1 (Lineares Funktional auf \mathbb{R}) Sei V ein \mathbb{R} -Vektorraum. Ein Funktional F ist eine Abbildung $F : V \rightarrow \mathbb{R}$. Das Funktional ist linear, wenn die Addition und die Multiplikation mit einem Skalar punktweise definiert ist. Das heißt

$$\begin{aligned} (F + G)(v) &:= F(v) + G(v) \quad \forall v \in V, \\ (\lambda F)(v) &:= \lambda F(v) \quad \forall v \in V. \end{aligned}$$

Definition 2.7.2 (Bilinearform) Eine Abbildung $a : V \times V \rightarrow \mathbb{R}$ heißt Bilinearform, wenn diese linear in beiden Argumenten ist, das heißt

$$\begin{aligned} a(\lambda u + \mu v, w) &= \lambda a(u, w) + \mu a(v, w), \\ a(u, \lambda v + \mu w) &= \lambda a(u, v) + \mu a(u, w), \end{aligned}$$

für alle $u, v \in V$ und für alle $\lambda, \mu \in \mathbb{R}$.

Bemerkung Die Lösungen des Variationsproblems (2.9) werden als schwache Lösungen bezeichnet. Im Allgemeinen bedingt eine schwache Lösung nicht immer einer klassischen Lösung. Durch bestimmte Glattheitsbedingungen werden schwache Lösungen auch zu klassische Lösungen. Sei daher $u \in V_0$ eine schwache Lösung mit den zusätzlichen Voraussetzungen

$$u \in C^2(\Omega) \cap C^1(\bar{\Omega}) \cap C(\bar{\Omega})$$

und

$$f \in C(\Omega).$$

Dann gilt für den Hauptteil für $v \in C_0^\infty(\Omega)$

$$\int_{\Omega} \nabla u(x) \cdot \nabla v(x) \, dx = - \int_{\Omega} \underbrace{\nabla \cdot (\nabla u(x))}_{=H_u(x)} v(x) \, dx + \int_{\Gamma} \nabla u \cdot n \underbrace{v}_{=0, \text{ auf } \Gamma} \, ds_x.$$

Setzt man diesen Hauptteil nun in das ursprüngliche Variationsproblem (2.9) ergibt sich

$$\int_{\Omega} -H_u(x)v(x) \, dx = \int_{\Omega} f(x)v(x) \, dx \quad \forall v \in C_0^\infty(\Omega). \quad (2.10)$$

Somit ergibt sich für die Variationsgleichung (2.10)

$$\int_{\Omega} \underbrace{[-H_u(x) - f(x)]}_{\substack{\in C(\Omega), \text{ da jeder} \\ \text{einzelne Term stetig ist}}} v(x) \, dx = 0 \quad \forall v \in C_0^\infty(\Omega).$$

Daher gilt

$$\Rightarrow -H_u(x) - f(x) = 0 \quad \forall x \in \Omega.$$

Somit erfüllt diese Variationsgleichung die Differentialgleichung punktweise und unter den obigen Voraussetzungen ist die schwache Lösung auch eine klassische Lösung.

2.8 Der Satz von Lax-Milgram

Im folgenden Abschnitt wird das Variationsproblem (2.9) auf Lösbarkeit untersucht. Hierbei werden folgende Punkte genauer betrachtet:

- Existenz einer Lösung.
- Eindeutigkeit einer Lösung.
- Stetige Abhängigkeit der Lösung an den Daten.

Um das Variationsproblem auf Lösbarkeit zu untersuchen, werden bestimmte Annahmen getroffen und ein abstraktes Variationsproblem aufgestellt.

2.8.1 Abstraktes Variationsproblem

Für das abstrakte Variationsproblem sei V ein Hilbertraum mit dem inneren Produkt $(\cdot, \cdot)_V$ und der induzierten Norm $\|\cdot\|_V$, $a : V \times V \rightarrow \mathbb{R}$ eine Bilinearform und $F : V \rightarrow \mathbb{R}$ ein lineares Funktional. Weiteres sei die Bilinearform $a : V \times V \rightarrow \mathbb{R}$ beschränkt auf V und V -elliptisch und das lineare Funktional $F : V \rightarrow \mathbb{R}$ sei ebenfalls beschränkt auf V .

Definition 2.8.1 (*Beschränktheit und Elliptizität*) Sei $a : V \times V \rightarrow \mathbb{R}$ eine Bilinearform und $F : V \rightarrow \mathbb{R}$ ein lineares Funktional. Die Bilinearform $a(\cdot, \cdot)$ heißt beschränkt auf V , falls gilt

$$\exists \alpha > 0 : |a(u, v)| \leq \alpha \|u\|_V \|v\|_V \quad \forall u, v \in V .$$

Weiter heißt $a(\cdot, \cdot)$ V -elliptisch, falls gilt

$$\exists \beta > 0 : |a(v, v)| \geq \beta \|v\|_V^2 \quad \forall v \in V .$$

Das Funktional F heißt beschränkt, falls gilt

$$\exists c_F > 0 : |F(v)| \leq c_F \|v\|_V \quad \forall v \in V .$$

Theorem 2.8.2 Sei F ein lineares Funktional auf einem normierten Raum V . F ist genau dann stetig auf V , wenn F auch beschränkt ist.

Beweis: Sei F ein lineares Funktional.

“ \Leftarrow “ Sei F beschränkt. Dann gibt es ein $\alpha > 0$, sodass $|F(v)| \leq \alpha \|v\|$ für alle $v \in V$.

Sei außerdem v_n eine Nullfolge. Dann gilt

$$|F(v_n)| \leq \alpha \|v_n\| \rightarrow 0 .$$

Somit ist f stetig in Null und aufgrund der Linearität auch in jedem anderen Punkt stetig.

“ \Rightarrow “ Sei nun f stetig, somit ist f auch stetig in 0. Dann existiert ein $\delta > 0$, sodass $|F(v)| < \epsilon = 1$ für $\|v\| \leq \delta$. Für $v \neq 0$ gilt schließlich

$$|F(v)| = \left| F \left(\frac{\delta v}{\|v\|} \frac{\|v\|}{\delta} \right) \right| = \underbrace{\left| F \left(\frac{\delta v}{\|v\|} \right) \right|}_{< 1} \frac{\|v\|}{\delta} < \frac{1}{\delta} \|v\| .$$

Somit ist F beschränkt mit $\alpha = \frac{1}{\delta}$.

□

Definition 2.8.3 (*Dualraum*) Sei V ein Hilbertraum. Dann ist der Dualraum V^* von V gegeben durch

$$V^* := \{F : V \rightarrow \mathbb{R} : F \text{ ist linear und beschränkt}\}$$

mit der dualen Norm

$$\|F\|_{V^*} := \sup_{0 \neq v \in V} \frac{\langle F, v \rangle}{\|v\|_V} .$$

Das abstrakte Variationsproblem lässt sich schließlich schreiben als: Gesucht ist $u \in V$, sodass

$$a(u, v) = \langle F, v \rangle \quad \forall v \in V \tag{2.11}$$

erfüllt ist, wobei $a(\cdot, \cdot)$ eine V -elliptische und beschränkte Bilinearform auf V darstellt und F ein lineares und beschränktes Funktional ist.

Theorem 2.8.4 (Satz von Lax-Milgram) Sei V ein Hilbertraum und $F \in V^*$ ein lineares und beschränktes Funktional. Außerdem sei $a : V \times V \rightarrow \mathbb{R}$ eine V -elliptische und auf V beschränkte Bilinearform. Dann ist das Variationsproblem (2.11) eindeutig lösbar. Außerdem hängt die Lösung $u \in V$ stetig von den Daten ab und es gilt

$$\frac{1}{\alpha} \|F\|_{V^*} \leq \|u\|_V \leq \frac{1}{\beta} \|F\|_{V^*} .$$

Beweis: Vgl. [7] Kapitel 1.2.8 .

□

Das Ziel ist nun den Satz von Lax-Milgram auf das Variationsproblem (2.11) anzuwenden, um die Eindeutigkeit der Lösung sicherstellen zu können. Dazu müssen folgenden Aussagen überprüft werden:

- Der Raum V ist ein Hilbertraum.
- Das Funktional F ist linear und beschränkt.
- Die Bilinearform $a(\cdot, \cdot)$ ist V -elliptisch und beschränkt.

Der Spuroperator ist beschränkt und linear, deswegen ist V_0 ein abgeschlossener Unterraum von $H^1(\Omega)$. Somit ist V_0 mit dem inneren Produkt $(u, v)_V = (u, v)_{L^2(\Omega)} + (\nabla u, \nabla v)_{L^2(\Omega)}$ ein Hilbertraum.

Um die Beschränktheit von F zu zeigen, wird folgendes betrachtet

$$\langle F, v \rangle = \int_{\Omega} f(x)v(x) \, dx \leq \int_{\Omega} |f(x)v(x)| \, dx \leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \leq \underbrace{\|f\|_{L^2(\Omega)}}_{=: \gamma > 0} \|v\|_{H^1(\Omega)} = \gamma \|v\|_{H^1(\Omega)} .$$

Dies gilt für alle $v \in H^1(\Omega)$. Daraus folgt, dass das lineare Funktional F auf V_0 beschränkt ist.

Die Beschränktheit der Bilinearform $a(\cdot, \cdot)$ erfolgt auf ähnlicher Weise und es gilt

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx \leq \|\nabla u\| \|\nabla v\| \leq \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} = 1 \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} .$$

Dies gilt für alle $u, v \in H^1(\Omega)$. Daraus folgt die Beschränktheit der Bilinearform $a(\cdot, \cdot)$ auf V_0 .

Für die Elliptizität der Bilinearform $a(\cdot, \cdot)$ wird der folgende Satz benötigt.

2.8.2 Die Friedrichs-Ungleichung

Theorem 2.8.5 (Friedrichs-Ungleichung) Es existiert eine Konstante $c_F > 0$, sodass gilt

$$\|v\|_{L^2(\Omega)} \leq c_F \|v\|_{H^1(\Omega)} \quad \forall v \in V_0 .$$

Beweis: Vgl. [2] Satz 1.5 .

□

Mit diesem Satz ist es nun möglich die V_0 -Elliptizität des Variationsproblems (2.11) zu zeigen.

$$a(u, u) = \int_{\Omega} \nabla u \cdot \nabla u \, dx = \int_{\Omega} |\nabla u|^2 \, dx = |u|_{H^1(\Omega)}^2 \quad \forall u \in H^1(\Omega) .$$

Mithilfe der Friedrichs-Ungleichung gilt für $u \in V_0$

$$\|u\|_{H^1(\Omega)}^2 = \underbrace{\|u\|_{L^2(\Omega)}^2}_{\leq c_F^2 |u|_{H^1(\Omega)}^2} + |u|_{H^1(\Omega)}^2 \leq (c_F^2 + 1) |u|_{H^1(\Omega)}^2 .$$

Somit folgt die V_0 -Elliptizität mit

$$a(u, u) = |u|_{H^1(\Omega)}^2 \geq \frac{1}{1 + c_F^2} \|u\|_{H^1(\Omega)}^2 \quad \forall u \in V_0 .$$

Es sind nun für das Variationsproblem (2.11) alle Voraussetzungen vom Satz von Lax-Milgram erfüllt. Dadurch ist die Existenz einer eindeutigen Lösung sichergestellt, welche in stetiger Abhängigkeit zu den Daten steht.

2.8.3 Der Riesz'sche Darstellungssatz

Für ein lineares Funktional $F \in V^*$ in einem Hilbertraum V mit dem inneren Produkt $(\cdot, \cdot)_V$ kann folgendes Variationsproblem betrachtet werden. Gesucht ist $u \in V$, sodass

$$(u, v)_V = \langle F, v \rangle \quad \forall v \in V \tag{2.12}$$

erfüllt ist.

Lemma 2.8.6 *Sei V ein Hilbertraum und $F \in V^*$. Dann ist das Variationsproblem (2.12) äquivalent zum Minimierungsproblem: Gesucht ist $u \in V$, sodass*

$$J(u) = \min_{v \in V} J(v) \tag{2.13}$$

mit dem quadratischen Funktional

$$J(v) := \frac{1}{2}(v, v)_V - \langle F, v \rangle \quad \forall v \in V .$$

Beweis: Vgl. [2] Satz 2.5 bzw. [7] Lemma 1.35 .

□

Theorem 2.8.7 (Riesz) *Sei V ein Hilbertraum und $F \in V^*$. Dann gibt es genau eine Lösung des Minimierungsproblems (2.13) bzw. (2.12) und für diese Lösung $u \in V$ gilt*

$$\|u\|_V = \|F\|_{V^*} .$$

Beweis: Vgl. [2] Satz 2.5 .

□

Definition 2.8.8 (Riesz-Abbildung) *Sei V ein Hilbertraum. Die Abbildung $R: V^* \rightarrow V$, $F \mapsto u$ heißt Riesz-Abbildung, wobei u die Lösung von (2.12) bzw. (2.13) mit rechter Seite $F \in V^*$ ist.*

Bemerkung Die Riesz-Abbildung ist linear und isometrisch, da folgendes gilt

$$\|RF\|_V \stackrel{\text{Def}}{=} \|u\|_V \stackrel{\text{Theorem}}{=} \|F\|_{V^*} \quad \forall F \in V^* .$$

Für die Umkehrabbildung gilt

$$R^{-1} : V \rightarrow V^*, \quad u \mapsto F_u$$

mit

$$\langle F_u, v \rangle := (u, v)_V \quad \forall v \in V .$$

Somit ist die Riesz-Abbildung ein isometrischer Isomorphismus.

Man betrachtet nun für eine beschränkte Bilinearform $a : V \times V \rightarrow \mathbb{R}$ den Operator $A : V \rightarrow V^*$ mit $u \mapsto Au$, sodass gilt

$$\langle Au, v \rangle := a(u, v) \quad \forall v \in V .$$

Da die Bilinearform $a(\cdot, \cdot)$ beschränkt ist, gilt auch $Au \in V^*$. Außerdem ist der Operator linear und beschränkt, in diesem Fall gilt

$$\|Au\|_{V^*} = \sup_{0 \neq v \in V} \frac{\langle Au, v \rangle}{\|v\|_V} = \sup_{0 \neq v \in V} \frac{a(u, v)}{\|v\|_V} \leq \alpha \sup_{0 \neq v \in V} \frac{\|u\|_V \|v\|_V}{\|v\|_V} = \alpha \|u\|_V \quad \forall u \in V .$$

Somit ist das Variationsproblem (2.11) äquivalent zur Operatorgleichung

$$Au = F \quad \text{in } V^* .$$

2.9 Das symmetrische Variationsproblem

In manchen Fällen lässt sich das Variationsproblem vereinfachen, indem man die Eigenschaft der Symmetrie ausnutzt.

Lemma 2.9.1 Sei V ein Hilbertraum, $F \in V^*$ und $a : V \times V \rightarrow \mathbb{R}$ eine beschränkte, symmetrische und nicht-negative Bilinearform, also

$$\begin{aligned} a(u, v) &= a(v, u) & \forall u, v \in V & & (\text{Symmetrie}), \\ a(u, u) &\geq 0 & \forall u \in V & & (\text{Nicht-Negativität}). \end{aligned}$$

Dann ist das Variationsproblem: Gesucht $u \in V$, sodass

$$a(u, v) = \langle F, v \rangle \quad \forall v \in V$$

erfüllt ist, äquivalent zum Minimierungsproblem: Gesucht ist $u \in V$, sodass

$$J_a(u) = \min_{v \in V} J_a(v)$$

erfüllt ist, mit dem Ritz-Funktional

$$J_a(v) := \frac{1}{2} a(v, v) - \langle F, v \rangle \quad \forall v \in V .$$

Beweis: Analog zum Lemma 2.8.7 .

□

2.10 Die Galerkin-Prinzip Approximation

Das große Ziel dieses Kapitels ist es eine gute Näherung der Lösung u des Variationsproblems (2.11) zu finden. Die Idee der Galerkin-Methode basiert auf der Aufspaltung des Raumes V in mehreren endlich dimensionalen und abgeschlossenen Teilräumen $V_N \subset V$. In diesen endlich dimensionalen Teilräumen wird die Lösung bestmöglich approximiert.

Zusammengefasst betrachtet man also den endlich dimensionalen Teilraum

$$V_N \subset V$$

und das abstrakte Variationsproblem: Gesucht ist $u \in V$, sodass

$$a(u, v) = \langle F, v \rangle \quad \forall v \in V$$

erfüllt ist. Ein diskretes Problem wird daher von der Galerkin-Approximation abgeleitet, welches für den Raum $V_N \subset V$ mit $\dim V_N = N$ lautet: Gesucht ist $u_N \in V_N$, sodass

$$a(u_N, v_N) = \langle F, v_N \rangle \quad \forall v_N \in V_N \tag{2.14}$$

erfüllt ist.

Mögliche Lösungen des diskreten Problems (2.14) nennt man Näherungslösungen. Im Allgemeinen erwartet man für $\dim(V_N) \rightarrow \infty$, dass die Näherungslösung u_N gegen die exakte Lösung u konvergiert. Wie auch für das abstrakte Variationsproblem, möchte man die Existenz und Eindeutigkeit einer Lösung für das diskrete Problem gewährleisten. Dazu wird wieder das Lemma von Lax-Milgram verwendet, jedoch werden die Voraussetzungen explizit für das diskrete Variationsproblem gefordert. Dies führt nun zum folgenden Lemma.

Lemma 2.10.1 *Sei $V_N \subset V$ ein abgeschlossener Teilraum, $a : V_N \times V_N \rightarrow \mathbb{R}$ eine beschränkte und V_N -elliptische Bilinearform und $F : V_N \rightarrow \mathbb{R}$ ein beschränktes lineares Funktional. Dann existiert eine eindeutige Lösung $u_N \in V_N$ des Variationsproblems (2.14).*

Beweis: Analog zum Beweis des Satzes von Lax-Milgram. □

Eine bestmögliche Näherungslösung kann durch die Verwendung der Orthogonalität erreicht werden. Der Fehler $u - u_N$ liegt bezüglich der Bilinearform orthogonal auf allen Testfunktionen $v_N \in V_N$. Diese Eigenschaft wird Galerkin-Orthogonalität genannt und lässt sich schreiben als

$$a(u - u_N, v_N) = 0 \quad \forall v_N \in V_N . \tag{2.15}$$

Ein sehr populärer Satz ist das Lemma von Ceá, welcher eine Aussage über einen a-priori Fehlerschätzer trifft. In dieser Arbeit wird diese Fehlerschätzung kurz vorgestellt, aber nicht genau behandelt, da der Fokus auf einen a-posteriori Fehlerschätzer liegt.

Lemma 2.10.2 *(Lemma von Ceá) Sei V ein Hilbertraum, V_N ein abgeschlossener Teilraum von V und $F \in V^*$. Weiteres sei $a : V \times V \rightarrow \mathbb{R}$ eine Bilinearform, die die Voraussetzungen des Satzes von Lax-Milgram erfüllt. Außerdem sei*

$$\begin{aligned} u \in V : \quad a(u, v) &= \langle F, v \rangle & \forall v \in V, \\ u_N \in V_N : \quad a(u_N, v_N) &= \langle F, v_N \rangle & \forall v_N \in V_N . \end{aligned}$$

Dann gilt die Fehlerschätzung

$$\|u - u_N\|_V \leq \frac{\alpha}{\beta} \inf_{v_N \in V_N} \|u - v_N\|_V .$$

Beweis: Vgl. [7] Lemma 1.49 .

□

Bemerkung Im vorherigen Satz wird die Bestapproximation

$$w_N \in V_N : \|u - w_N\|_V = \inf_{v_N \in V_N} \|u - v_N\|_V$$

als geeignete Fehlerschranke verwendet. Für die Berechnung der Bestapproximation ist jedoch das Kennen der Lösung $u \in V$ eine nötige Voraussetzung.

Das Lemma von Ceá besagt also, dass der Fehler der Näherungslösung $u_N \in V_N$ mit dem Fehler der Bestapproximation bis auf eine Konstante abgeschätzt werden kann. Man spricht deshalb von einer *quasi-optimalen* Fehlerschätzung.

Bemerkung Für symmetrische Probleme ist die obere Schranke des Fehlers der Näherungslösung $u_N \in V_N$ der Galerkin-Approximation gleich der Bestapproximation in der Energienorm, das heißt

$$\|u - u_N\|_a \leq \inf_{v_N \in V_N} \|u - v_N\|_a .$$

2.11 Finite-Elemente-Räume

Definition 2.11.1 (*Gitter, Knoten*) Ein Gitter oder eine Zerlegung \mathcal{T}_h eines Gebiets Ω ist eine Approximation eines gegebenen Intervalles, einer gegebenen Oberfläche oder eines gegebenen Raumvolumens durch meist sehr einfache Elemente. Für das Gebiet Ω seien Knoten oder Gitterpunkte $\{V_j\}_{j=0}^N$ vorgegeben, welche die Elemente $\{T_j\}_{j=1}^N$ der Zerlegung definieren. Die Zerlegung lässt sich schließlich schreiben als

$$\mathcal{T}_h = \{T_1, T_2, \dots, T_N\} .$$

Eine Kante E bezeichnet die Strecke zwischen zwei Knoten V_1 und V_2 . Des Weiteren wird der Durchmesser eines Elements T mit h_T und die Länge einer Kante E mit h_E bezeichnet.

Bemerkung Eine Triangulierung ist ein Gitter bestehend aus Dreieckselementen, wie in Abbildung 2.1 zu sehen ist.

Definition 2.11.2 (*Zulässigkeit*) Eine Zerlegung \mathcal{T}_h heißt **zulässig**, wenn zwei verschiedene Elemente aus \mathcal{T}_h entweder disjunkt oder eine gemeinsame Kante oder einen gemeinsamen Elementeckpunkt vorweisen.

Definition 2.11.3 (*Form-regulär*) Eine Familie von Zerlegungen \mathcal{T}_h heißt **form-regulär**, wenn es eine Konstante $c > 0$ gibt, sodass jedes $T \in \mathcal{T}_h$ eine Kugel mit Radius $\rho_T \geq c \dim(T)$ enthält.

Definition 2.11.4 Sei \mathcal{T}_h eine reguläre Zerlegung von Ω . Dann bezeichnet man ν_h als die Menge aller Knoten in \mathcal{T}_h ,
 \mathcal{E}_h als die Menge aller Kanten in \mathcal{T}_h ,
 und F_h als die Menge aller Flächen in \mathcal{T}_h .

Die Mengen \mathcal{E}_h und F_h sind exklusiv der Kanten bzw. Flächen auf dem Rand Γ .

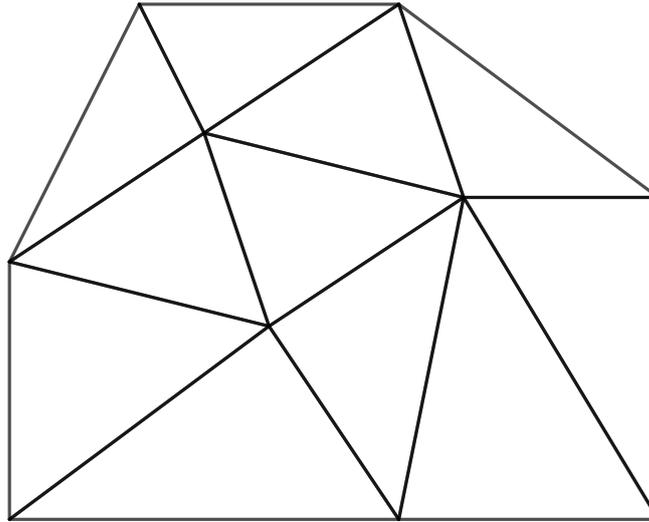


Abbildung 2.1: Triangulierung

Nun sei \mathcal{T}_h eine affine, form-reguläre und zulässige Triangulierung von Ω . Es lässt sich ein konformer, stückweise polynomialer Approximationsraum $V_N \subset V$ der Ordnung $k \geq 1$ bilden als

$$V_N = S_h^{1,k} := \{v \in H^1(\Omega) : v|_T \in P^k(T) \quad \forall T \in \mathcal{T}_h\} .$$

Im weiteren Verlauf der Arbeit genügt der Approximationsraum $S_h^{1,1}$ als diskreter Arbeitsraum, das heißt $k = 1$. Um die Bedingung $v \in H^1(\Omega)$ zu garantieren, benötigt man eine bestimmte Glattheit über die Elementgrenzen hinweg. Dazu werden die sogenannten „Konformitätsbedingungen“ bestimmt:

Lemma 2.11.5 *Sei $\Omega \subset \mathbb{R}^2$ beschränkt. Eine stückweise definierte $H^1(\Omega)$ -Funktion $v : \bar{\Omega} \rightarrow \mathbb{R}$ liegt genau dann in $H^1(\Omega)$, wenn $v \in C^0(\Omega)$.*

Beweis: Vgl. [2] Kapitel 2.

□

2.12 Das lineare Gleichungssystem

Sei $\{\phi_i\}_{i=1}^N$ die nodale Basis vom Raum V_N mit

$$\phi_i(V_j) := \begin{cases} 1, & i = j \\ 0, & \text{sonst} \end{cases} \quad \forall i, j = 1, \dots, N$$

wobei V_j , für $j = 1, \dots, N$, die Gitterpunkte eines Gitters sind. Die Näherungslösung $u_N \in V_N$ lässt sich schreiben als

$$u_N = \sum_{j=1}^N u_j \phi_j ,$$

wobei N für die Anzahl der Knoten steht. Außerdem gilt $u_j = u_N(V_j)$ für alle $j = 1, \dots, N$.

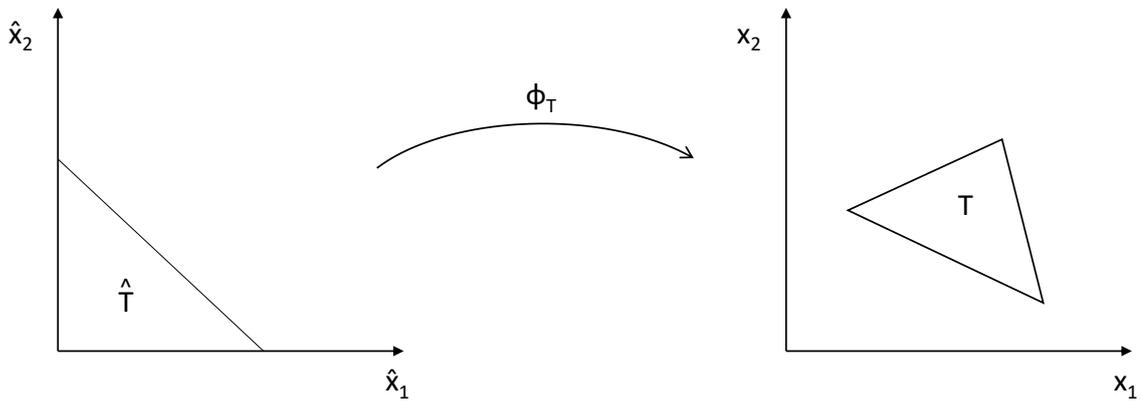


Abbildung 2.2: Referenzgebiet

Definiert man nun

$$\begin{aligned} \underline{u}_N &:= [u_j]_{j=1}^N, \\ \underline{f}_N &:= [\langle F, \phi_i \rangle]_{i=1}^N, \\ K_N &:= [a(\phi_j, \phi_i)]_{j,i=1}^N, \end{aligned}$$

dann ist das lineare Gleichungssystem

$$K_N \underline{u}_N = \underline{f}_N$$

äquivalent zum diskreten Variationsproblem (2.11). Dabei kann das lineare Gleichungssystem effizient über lokale Elementmatrizen und Elementvektoren assembliert werden. Die Berechnung der lokalen Einträge erfolgt über ein Referenzelement, siehe Abbildung 2.2.

Jedes Element $T \in \mathcal{T}_h$ lässt sich als eine affine lineare Abbildung eines Referenzelements \hat{T} beschreiben

$$\phi_T : \hat{T} \rightarrow T \quad \text{mit} \quad \hat{V} \mapsto B_T \hat{V} + b_T$$

für $B_T \in \mathbb{R}^{2 \times 2}$ als reguläre Matrix und $b_T \in \mathbb{R}^2$.

Für form-reguläre Netze gilt

$$\begin{aligned} \|B_T\| &= \sqrt{\lambda_{\max}(B^T B)} \leq c_1 h_T, \\ \|B_T^{-1}\| &= \sqrt{\lambda_{\min}(B^T B)} \leq c_2 h_T^{-1}, \\ |\det B_T| &\approx h_T^2, \end{aligned}$$

wobei h_T der Elementdurchmesser des Elements T ist.

Aufgrund der Elliptizität und Symmetrie der Bilinearform, ist die Matrix K_N positiv definit und ebenfalls symmetrisch.

Kapitel 3

A-posteriori Fehlerschätzer und adaptive Verfahren

Das Ziel dieses Kapitels ist es einen sogenannten a-posteriori Fehlerschätzer zu finden, der die Größe des tatsächlichen Fehlers, nach Berechnung der numerisch approximierten Lösung geeignet schätzt. Folgende Aspekte werden dabei genauer betrachtet.

- Wie groß ist der tatsächliche Fehler eines bestimmten Szenarios (d.h. bei einer konkret gewählten Triangulierung, konkreter Polynomgrad, gegebene Daten wie f)?
- Kann man die Diskretisierung so steuern, um mit möglichst wenig Aufwand eine bestimmte Genauigkeit für

$$\|u - u_N\| \leq TOL$$

zu erreichen?

Im Allgemeinen ist die exakte Lösung $u \in H^1(\Omega)$ nicht bekannt, deswegen erwartet man für den Fehler $u - u_N \in H^1(\Omega)$ ebenfalls kein Ergebnis. Die Aufgabe eines Fehlerschätzers ist es daher den Fehler $\|u - u_N\|$ so zu bestimmen, dass dieser möglichst klein bleibt. Damit kann in weiterer Folge ein geeignetes Abbruchkriterium für adaptive Verfahren festgelegt werden.

Die Kernidee eines adaptiven Verfahrens beruht auf die Anpassung der Input-Daten, sodass bestimmte Ziele als Output erzwungen werden. Die Genauigkeit einer Näherungslösung u_N wird durch die Größe des Fehlers $\|u - u_N\|$ gemessen. Ist dieser größer als eine bestimmte Toleranzgrenze TOL , so wird das Verfahren erneut angewendet, jedoch mit verfeinerten Elementen. Das Ziel eines adaptiven Verfahrens liegt darin, die Elemente so oft zu verfeinern, bis der Fehler möglichst klein wird. Die Abbildung 3.1 zeigt den Verlauf eines adaptiven Verfahrens.

Definition 3.0.1 *Unter einem a-posteriori Fehlerschätzer versteht man eine berechenbare Größe η , die sich durch vorhandene Informationen wie f und der Näherungslösung u_N bestimmen lässt. Um zu verdeutlichen, dass dieser Fehlerschätzer von bestimmten Daten abhängt, wird dieser auch geschrieben als*

$$\eta(f, u_N) .$$

Die Qualität des Fehlerschätzers lässt sich durch folgende zwei Eigenschaften bestimmen.

Definition 3.0.2 *(Zuverlässigkeit und Effizienz) Sei $\|\cdot\|$ eine bestimmte Norm in einem Raum. Ein Fehlerschätzer η heißt **zuverlässig (Englisch: reliable)**, falls eine Konstante $c_r > 0$ existiert, sodass*

$$\|u - u_N\| \leq c_r \eta(u_N, f) .$$

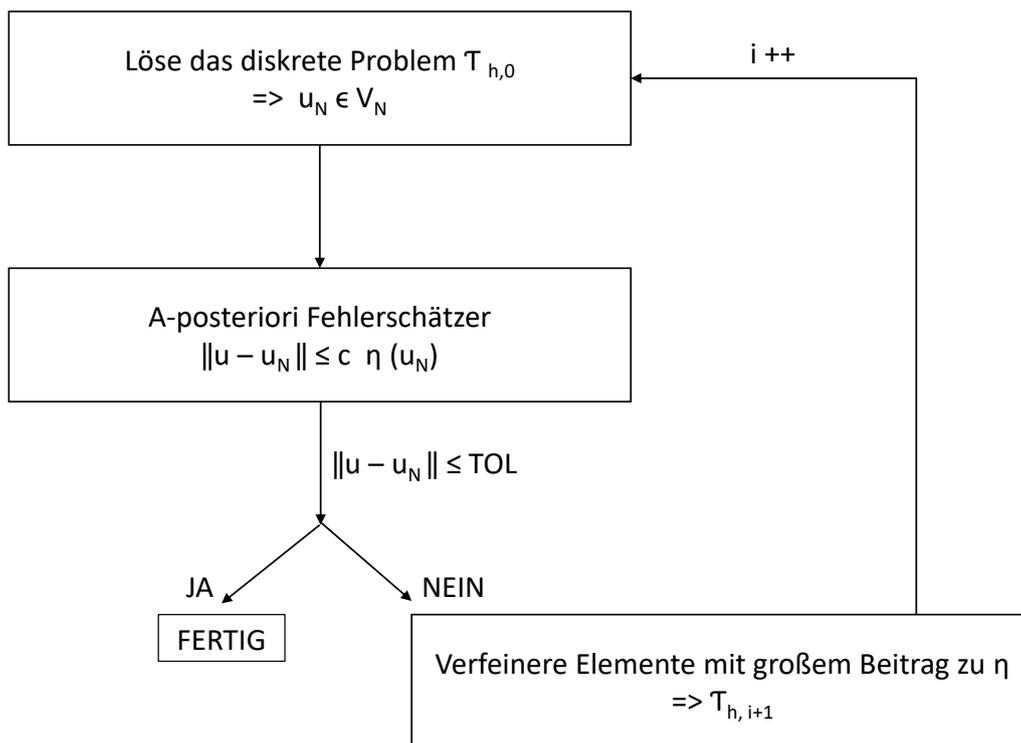


Abbildung 3.1: Adaptive Verfahren

Ein Fehlerschätzer η heißt **effizient (Englisch: efficient)**, falls eine Konstante $c_e > 0$ existiert, sodass

$$\|u - u_N\| \geq c_e \eta(u_N, f) .$$

Das Ziel ist es nun einen Fehlerschätzer zu finden, der sowohl zuverlässig als auch effizient ist. Diese Eigenschaften des Fehlerschätzers dienen als Hilfestellung, um das Abbruchkriterium von adaptiven Verfahren zu bestimmen. Beispielsweise ist der zuverlässige Fehlerschätzer eine geeignete obere Schranke für das Abbruchkriterium.

Bemerkung Die Norm in dem der Fehler gemessen wird, spielt eine wesentliche Rolle. Unterschiedliche Normen führen zu unterschiedlichen Messungen von Fehlern. Die Wahl der Norm hängt hierbei von der konkreten praktischen Problemstellung ab.

3.1 Residuale Fehlerschätzer

Residuale Fehlerschätzer sind klassische Fehlerschätzer und beruhen auf der Schätzung des Fehlers anhand des **Residuums**.

Man bezeichnet den residualen Fehler auch als $R(u_N, f)$ und es gilt

$$\begin{aligned} \langle R(u_N, f), v \rangle &:= \langle F, v \rangle - a(u_N, v) \quad \forall v \in V , \\ &= (f, v)_V - (\nabla u, \nabla v)_V . \end{aligned} \tag{3.1}$$

Durch elementweise partielle Integration erhält man eine berechenbare Darstellung des Residu-

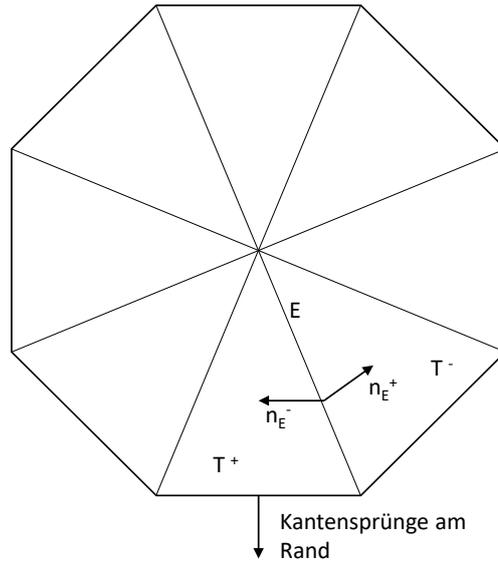


Abbildung 3.2: Kantensprünge

ums aus (3.1), sodass

$$\begin{aligned}
 \langle R(u_N, f), v \rangle &= \sum_{T \in \mathcal{T}_h} \underbrace{(f, v)_T}_{\int_T f v \, dx} - (\nabla u_N, \nabla v)_T \\
 &\stackrel{\text{part. Integration}}{=} \sum_{T \in \mathcal{T}_h} (f, v)_T + (\Delta u_N, v)_T - \underbrace{\langle \nabla u_N \cdot \underline{n}, v \rangle_{\partial T}}_{= \frac{\partial u_N}{\partial n}} \\
 &= \sum_{T \in \mathcal{T}_h} (f + \Delta u_N, v)_T + \sum_{E \in \mathcal{E}_h} \langle -\nabla u_N^+ \cdot \underline{n}_E^+ - \nabla u_N^- \cdot \underline{n}_E^-, v \rangle_E .
 \end{aligned}$$

Der hintere Ausdruck $\langle -\nabla u_N^+ \cdot \underline{n}_E^+ - \nabla u_N^- \cdot \underline{n}_E^-, v \rangle_E$ bezeichnet den Sprung der Normalableitung $n \cdot \nabla u_N$ über die Kante E , siehe Abbildung 3.2.

Schließlich lässt sich der residuale Fehler schreiben als

$$\langle R(u_N, f), v \rangle = \sum_{T \in \mathcal{T}_h} \underbrace{(f + \Delta u_N, v)_T}_{=: \langle R_T(u_N, f), v \rangle} + \sum_{E \in \mathcal{E}_h} \underbrace{\langle \llbracket \nabla u_N \rrbracket_E, v \rangle_E}_{=: \langle R_E(u_N), v \rangle} ,$$

mit den Kantensprüngen

$$\llbracket v \rrbracket_E = \lim_{t \rightarrow 0} v(x + t n_E) - v(x - t n_E)$$

bzw.

$$\llbracket \nabla u_N \rrbracket_E = (\nabla u_h|_{T^+} \cdot n^+ + \nabla u_h|_{T^-} \cdot n^-)|_E .$$

Für die äußere Kante gilt demnach

$$\llbracket \nabla u_N \rrbracket_E = ((\nabla u_h)_T \cdot n)|_E .$$

Die zwei verschiedenen Residuen $R_T(u_N, f)$ und $R_E(u_N)$ stehen für die **elementweisen** bzw. die **kantenbezogenen** Residuen. Zusammenfassen ergibt sich folgendes Lemma.

Lemma 3.1.1 (*Eigenschaften des Residuums*)

1. $R(u_N, f) \in V^*$ besitzt eine berechenbare Darstellung, welche sich aus der Summe der Elementen- und Kantenbeiträge zusammensetzen lässt

$$\begin{aligned} \langle R(u_N, f), v \rangle &= \sum_{T \in \mathcal{T}_h} \langle R_T(u_N, f), v \rangle + \sum_{E \in \mathcal{E}_h} \langle R_E(u_N), v \rangle \\ &= \sum_{T \in \mathcal{T}_h} (f + \Delta u_N, v)_T + \sum_{E \in \mathcal{E}_h} \langle [[\nabla u_N]]_E, v \rangle_E . \end{aligned}$$

2. Das Residuum verschwindet aufgrund der Galerkin-Orthogonalität auf V_N , das heißt

$$\langle R(u_N, f), v_N \rangle = 0 \quad \forall v_N \in V_N .$$

Beweis: Punkt 1 folgt aus der vorherigen Herleitung. Punkt 2 folgt aus

$$\begin{aligned} \langle R(u_N, f), v_N \rangle &\stackrel{(3.1)}{=} F(v_N) - a(u_N, v_N) \\ &= a(u, v_N) - a(u_N, v_N) \\ &= a(u - u_N, v_N) = 0 , \end{aligned}$$

wegen der Galerkin-Orthogonalität. □

Das Residuum lässt sich nun schreiben als

$$a(u - u_N, v) = \langle R(u_N, f), v \rangle .$$

Anschließend wird nun ein Fehlerschätzer bezüglich der Energienorm gesucht, sodass gilt

$$\|u - u_N\|_a = \sup_{0 \neq v \in V} \frac{a(u - u_N, v)}{\|v\|_a} = \sup_{0 \neq v \in V} \frac{\langle R(u_N, f), v \rangle}{\|v\|_a} \leq c \eta ,$$

wobei η den gesuchten Fehlerschätzer darstellt.

3.1.1 Eine obere Schranke für $\|u - u_N\|_a$

Wie schon vorher erwähnt, ist der zuverlässige Fehlerschätzer eine mögliche Wahl für das Aufstellen einer oberen Schranke. Wird zudem die Eigenschaft der Galerkin-Orthogonalität genutzt, so erhält man eine geeignete obere Schranke des residualen Fehlerschätzers. Es gilt also

$$\langle R(u_N, f), v \rangle = \langle R(u_N, f), v - v_N \rangle .$$

Im Bereich der Elemente und an den Kanten wird nun eine Abschätzung für $\|v - v_N\|$ benötigt. Die Verwendung des Standards FE-Interpolationsoperators $I_{\mathcal{T}}$ ist in diesem Fall nicht möglich, da dieser der Regularität $v \in H^2(\Omega)$ bedingt. Um dieses Problem zu umgehen, werden sogenannte **Quasi-Interpolations-Operatoren** verwendet. Diese haben ähnliche Eigenschaften wie $I_{\mathcal{T}}$, besitzen eine „fast“ lokale Konstruktion und sind für Funktionen $v \in H^1(\Omega)$ definiert.

Theorem 3.1.2 (*Quasi-Interpolation und Fehlerabschätzung*) Sei \mathcal{T}_h eine zulässige und form-reguläre Zerlegung von $\Omega \subset \mathbb{R}^2$. Dann existiert ein Quasi-Interpolationsoperator $\Pi_N : H^1(\Omega) \rightarrow S_h^{1,1}$, sodass folgende lokale Abschätzungen gelten

$$\begin{aligned} \|v - \Pi_N v\|_{0,T} &\leq c h_T |v|_{1, \tilde{w}_T} \quad \forall v \in H^1(\Omega) , \\ \|v - \Pi_N v\|_{0,E} &\leq c h_E^{\frac{1}{2}} |v|_{1, \tilde{w}_E} \quad \forall v \in H^1(\Omega) , \end{aligned}$$

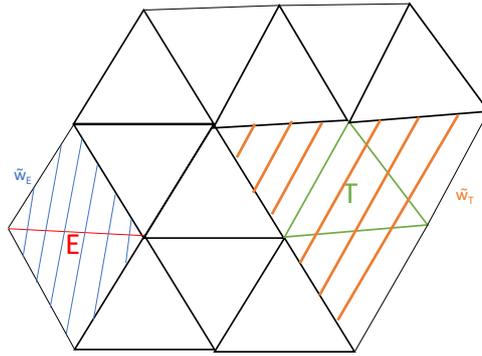


Abbildung 3.3: Element- und Kantenpatch

wobei

$$\tilde{w}_T = \{T' : T \cap T' \neq \emptyset, T' \in \mathcal{T}_h\}$$

der in Abbildung 3.3 abgebildete Elementpatch und

$$\tilde{w}_E = \{T' : E \cap T' \neq \emptyset, T' \in \mathcal{T}_h\}$$

der Kantenpatch ist.

Beweis: Vgl. [8] Theorem 2.1 .

□

Definition 3.1.3 (Clement-Quasi-Interpolationsoperator) Der Clement-Quasi- Interpolationsoperator ist ein beschränkter Operator und ist definiert als

$$\Pi_N : H_0^1(\Omega) \rightarrow S_{h,0}^{1,1} \quad \text{mit} \quad v \mapsto \Pi_N(v) = \sum_{i=1}^N \pi_i(v) \varphi_i ,$$

wobei $\{\varphi_i\}$ die Knotenbasisfunktionen und π_i die Mittelungsoperatoren (L^2 -orthogonale Projektionen) über dem Kantenpatch w_i sind. Die Mittelungsoperatoren sind definiert als

$$\pi_i : H^1(w_i) \rightarrow \mathbb{R} \quad \text{mit} \quad \pi_i(v) = \frac{1}{|w_i|} \int_{w_i} v \, dx .$$

Der Vertexpatch w_i um den Knoten $v_i \in V_h$ ist definiert als

$$w_i = \{T : v_i \in T, T \in \mathcal{T}_h\} .$$

Mithilfe des vorherigen Theorems ist es schließlich möglich, den residualen Fehlerschätzer zu

bestimmen. Es gilt mit der Cauchy-Schwarz (CS) Ungleichung

$$\begin{aligned}
 |a(u - u_N, v)| &= |(R(u_N, f), v)| = |(R(u_N, f), v - \Pi_N v)| \\
 &= \left| \sum_{T \in \mathcal{T}_h} (R_T(u_N, f), v - \Pi_N v)_T + \sum_{E \in \mathcal{E}_h} (R_E(u_N, f), v - \Pi_N v)_E \right| \\
 &\stackrel{CS}{\leq} \sum_{T \in \mathcal{T}_h} \|R_T(u_N, f)\|_{0,T} \|v - \Pi_N v\|_{0,T} + \sum_{E \in \mathcal{E}_h} \|R_E(u_N, f)\|_{0,E} \|v - \Pi_N v\|_{0,E} \\
 &\leq \sum_{T \in \mathcal{T}_h} \|R_T(u_N, f)\|_{0,T} c_1 h_T |v|_{1, \tilde{w}_T} + \sum_{E \in \mathcal{E}_h} \|R_E(u_N, f)\|_{0,E} c_2 h_E^{\frac{1}{2}} |v|_{1, \tilde{w}_E} \\
 &\stackrel{CS}{\leq} \left(\sum_{T \in \mathcal{T}_h} \|R_T(u_N, f)\|_{0,T}^2 h_T^2 \right)^{\frac{1}{2}} c_1 \left(\sum_{T \in \mathcal{T}_h} |v|_{1, \tilde{w}_T}^2 \right)^{\frac{1}{2}} \\
 &\quad + \left(\sum_{E \in \mathcal{E}_h} \|R_E(u_N, f)\|_{0,E}^2 h_E \right)^{\frac{1}{2}} c_2 \left(\sum_{E \in \mathcal{E}_h} |v|_{1, \tilde{w}_E}^2 \right)^{\frac{1}{2}} \\
 &\leq \left(\sum_{T \in \mathcal{T}_h} \|R_T(u_N, f)\|_{0,T}^2 h_T^2 \right)^{\frac{1}{2}} 4 c_1 |v|_{1, \Omega} + \left(\sum_{E \in \mathcal{E}_h} \|R_E(u_N, f)\|_{0,E}^2 h_E \right)^{\frac{1}{2}} 2 c_2 |v|_{1, \Omega}.
 \end{aligned}$$

Patches überlappen sich nur endlich oft mit Konstanten $4 c_1$ bzw. $2 c_2$.

$$\begin{aligned}
 &= 2 (2 c_1 + c_2) |v|_{1, \Omega} \left(\sum_{T \in \mathcal{T}_h} \|R_T(u_N, f)\|_{0,T}^2 h_T^2 + \sum_{E \in \mathcal{E}_h} \|R_E(u_N, f)\|_{0,E}^2 h_E \right)^{\frac{1}{2}} \\
 &= \tilde{c} |v|_{1, \Omega} \left(\sum_{T \in \mathcal{T}_h} \left[\|R_T(u_N, f)\|_{0,T}^2 h_T^2 + \frac{1}{2} \sum_{E \in \mathcal{E}(T)} \|R_E(u_N, f)\|_{0,E}^2 h_E \right] \right)^{\frac{1}{2}}.
 \end{aligned}$$

Insgesamt erhält man das Ergebnis

$$\begin{aligned}
 \|\nabla u - \nabla u_N\|_0 &= \sup_{\substack{v \in H_0^1(\Omega) \\ |v|_{1, \Omega} \neq 0}} \frac{(\nabla u - \nabla u_N, \nabla v)}{|v|_{1, \Omega}} = \sup_{\substack{v \in H_0^1(\Omega) \\ |v|_{1, \Omega} \neq 0}} \frac{a(u - u_N, v)}{|v|_{1, \Omega}} \\
 &\leq \tilde{c} \left(\sum_{T \in \mathcal{T}_h} \left[\|R_T(u_N, f)\|_{0,T}^2 h_T^2 + \frac{1}{2} \sum_{E \in \mathcal{E}(T)} \|R_E(u_N, f)\|_{0,E}^2 h_E \right] \right)^{\frac{1}{2}}.
 \end{aligned}$$

Der **elementweise Fehlerindikator** η_T wird demnach definiert als

$$\eta_T^2 := \left[\|R_T(u_N, f)\|_{0,T}^2 h_T^2 + \frac{1}{2} \sum_{E \in \mathcal{E}(T)} \|R_E(u_N, f)\|_{0,E}^2 h_E \right].$$

Das Ziel des äquilibrierten Fehlerschätzers ist es, eine berechenbare Darstellung zu ermitteln, die dem zuletzt eingeführten Fehlerindikator möglichst nahe kommt. Die Fehlerabschätzung soll dabei konstant-frei sein, sprich $\tilde{c} = 1$. Im Allgemeinen sind konstant-freie Abschätzungen nicht einfach zu ermitteln, weshalb Umwege über gemischte Variationsprobleme gemacht werden müssen. Vorerst kann die Zuverlässigkeit des residualen Fehlerschätzers als Theorem notiert werden.

Theorem 3.1.4 (Zuverlässigkeit des residualen Fehlerschätzers) Sei \mathcal{T}_h eine zulässige und form-reguläre Zerlegung. Dann ist der residuale Fehlerschätzer η_R definiert als

$$\eta_R(u_N, f) := \left(\sum_{T \in \mathcal{T}_h} \eta_T^2(u_N, f) \right)^{\frac{1}{2}}$$

mit den elementweisen Fehlerindikatoren

$$\eta_T^2(u_N, f) := \left(\|R_T(u_N, f)\|_{0,T}^2 h_T^2 + \frac{1}{2} \sum_{E \in \mathcal{E}(T)} \|R_E(u_N, f)\|_{0,E}^2 h_E \right)$$

zuverlässig, das heißt

$$\exists c_r > 0 : \|u - u_N\|_a = \|\nabla u - \nabla u_N\|_{0,\Omega} \leq c_r \eta_R(u_N, f) .$$

Beweis: Siehe vorherige Herleitung. □

3.2 Effizienz des residualen Fehlerschätzers η_R

Theorem 3.2.1 Es gelten die selben Voraussetzungen wie in Theorem (3.1.4). Dann erfüllt der residuale Fehlerschätzer η_R die Abschätzung

$$\exists c > 0 : \eta_R \leq c \left(\|\nabla u - \nabla u_N\|_0^2 + \sum_{T \in \mathcal{T}_h} h_T^2 \|f - f_N\|_{0,T}^2 \right)^{\frac{1}{2}} ,$$

wobei f_N eine stückweise polynomiale Approximation von f darstellt, mit $f_N|_T \in P^{p-1}(T)$.

Es gilt sogar die lokale Fehlerabschätzung

$$\eta_T \leq c \left[\|\nabla u - \nabla u_N\|_{L^2(w_T)}^2 + h_T^2 \|f - f_N\|_{0,w_T}^2 \right]^{\frac{1}{2}} .$$

Beweis: Vgl. [2] Satz 8.23 . □

Bemerkung Der Datenanteil $\sum_{T \in \mathcal{T}_h} h_T^2 \|f - f_N\|_{0,w_T}^2$ ist ein Term höherer Ordnung bezüglich h_T und somit klein im Vergleich zu $\|\nabla u - \nabla u_N\|_0^2$. Außerdem gilt für $p=1$, $f_N|_T = \frac{1}{|T|} \int_T f \, dx$.

Die Abschätzung von Theorem 3.2.1 kann lokalisiert werden, indem man jedes Element explizit betrachtet. Dadurch ergibt sich aus

$$a(u - u_N, v) = \langle R(u_N, f), v \rangle \quad \forall v \in H_0^1(\Omega)$$

die lokale Eigenschaft

$$\begin{aligned} a(u - u_N, v_T) &= \langle R(u_N, f), v_T \rangle \quad \forall v_T \in H_0^1(T) \\ &= \langle R_T(u_N, f), v_T \rangle \end{aligned}$$

und

$$\begin{aligned} a(u - u_N, v_E) &= \langle R(u_N, f), v_E \rangle \quad \forall v_E \in H_0^1(w_E) \\ &= \sum_{T \in w_E} \langle R_T(u_N, f), v_E \rangle + \langle R_E(u_N, f), v_E \rangle . \end{aligned}$$

Die untere Abschätzung des Theorems 3.2.1 liefert also Informationen über lokale Eigenschaften der Diskretisierung. Man erhält sie über Testfunktionen mit lokalem Träger. Wesentliche Hilfsmittel bilden die sogenannten Abschneidefunktionen ψ_T und ψ_E . Die Funktionen ψ_T werden auch „Blasenfunktionen“ bzw. „Bubble-Funktionen“ genannt.

3.3 Adaptive Gittersteuerung

Das Gebiet auf welchem die Lösung der Finite-Elementen-Methode berechnet wird, wird durch eine Triangulierung diskretisiert. Je nach Feinheit dieser Diskretisierung, verändert sich der Diskretisierungsfehler. Um diesen Fehler möglichst gering zu halten, wird das Gitter verfeinert, wobei der Rechenaufwand mit der Verfeinerung des Gebietes zunimmt. Das Ziel ist die Verfeinerung des Gitters zu steuern und an jenen Stellen zu verfeinern, an denen der Diskretisierungsfehler groß ist. Dabei soll der Rechenaufwand möglichst klein bleiben. Im folgenden wird ein Algorithmus zur adaptiven Gittersteuerung vorgestellt.

Algorithmus: Adaptive Gittersteuerung

Start mit globaler Triangulierung \mathcal{T}_0 von Ω .

$l = 0$

DO

(1) Lösung des diskretisierten Problems \mathcal{T}_l .

(2) Berechnung der Fehlerindikatoren η_T .

(3) IF $\left(\left(\sum_{T \in \mathcal{T}_l} \eta_T^2 \right)^{\frac{1}{2}} \leq \text{TOL} \right)$ THEN break();

(4) Markiere Element mit großem Fehlerbeitrag.

(5) Verfeinere markierte Elemente und zulässigen Gitterabschluss
(Eliminierung hängender Knoten).
 $\Rightarrow \mathcal{T}_l$ ist zulässig

(6) $l++$

END

3.3.1 Markierungsstrategien

Markierungsstrategien werden in adaptiven Verfahren angewendet, um jene Elemente, deren Fehlerrate sehr hoch ist, zu markieren. Im folgenden werden zwei verschiedenen Markierungsstrategien vorgestellt. Zum einen die Maximumsstrategie und zum anderen die Äquilibrierungsstrategie. Beide Strategien können im vorherig vorgestellten Algorithmus der adaptiven Gittersteuerung eingesetzt werden.

Maximumsstrategie

Diese Strategie wählt den größten Fehlerindikator. Es werden jene Elemente markiert, dessen Fehlerindikatoren größer als der skalierte maximale Fehlerindikator sind. Das heißt,

$$\bar{\eta} := \max_{T \in \mathcal{T}_l} \eta_T .$$

Es werden jene Elemente $T \in \mathcal{T}_l$ markiert, für die gilt

$$\eta_T \geq \theta \bar{\eta} \quad \text{für } 0 \leq \theta \leq 1 .$$

Äquilibrierungsstrategie

Die Äquilibrierungsstrategie ist auch als Dörfler-Marking bekannt. Gegeben ist hier $0 \leq \theta \leq 1$ und eine Menge M von Elementen mit $M = \emptyset$. Der Markierungsalgorithmus lautet wie folgt

DO

$$\text{IF} \left(\sum_{T \in M} \eta_T^2 \right)^{\frac{1}{2}} \geq (1 - \theta) \left(\sum_{T \in \mathcal{T}_l} \eta_T^2 \right)^{\frac{1}{2}} \quad \text{THEN break}()$$

$$\mu := \max_{T \in \mathcal{T}_l \setminus M} \eta_T$$

erweitere M um alle $T \in \mathcal{T}_l \setminus M$ mit $\eta_T = \mu$

END

Eine praktische Wahl für θ wäre

- (i) $\theta \in [0.5, 0.75]$ oder
- (ii) $\theta \approx 0.75$.

3.3.2 Lokale Verfeinerung und Gitterabschluss

Dieser Abschnitt bezieht sich auf den fünften Schritt des Algorithmus, also die Verfeinerung der markierten Elemente. Das Ziel ist es, die markierten Elemente zu verfeinern, sodass die verfeinerte Zerlegung \mathcal{T}_{l+1} zulässig und form-regulär bleibt. Folgende Verfeinerungen sind anwendbar.

Schwerpunkt einfügen

Diese Verfeinerung beruht auf die Bildung des Schwerpunktes. Aus einem groben Element entstehen drei verfeinerte Elemente, wie in Abbildung 3.4 zu sehen ist.

ROT-GRÜN-Verfeinerung

Die ROT-GRÜN-Verfeinerung ist eine weitere Gitterverfeinerungsmethode. Diese beruht auf folgenden Schritten.

(a) ROT-Verfeinerung

Man fügt weitere Knoten auf den Kantenmittelpunkten eines Elements ein und verbindet diese. Dadurch erhält man 4 kongruente Dreiecke. Die ROT-Verfeinerung ist eine sogenannte reguläre Verfeinerung. Die Verbindungstrecken werden dabei rot markiert. Abbildung 3.5 zeigt eine ROT-Verfeinerung.

(b) GRÜN-Verfeinerung

Im Gegenzug zur ROT-Verfeinerung spricht man bei der GRÜN-Verfeinerung von einer irregulären Verfeinerung. Hier wird der Mittelpunkt einer Kante mit dem gegenüberliegenden Punkt verbunden. Die GRÜN-Verfeinerung wird verwendet, um Dreiecke mit einem hängenden Knoten aufzuheben. Hängende Knoten sind Mittelpunkte, die Eckpunkte eines

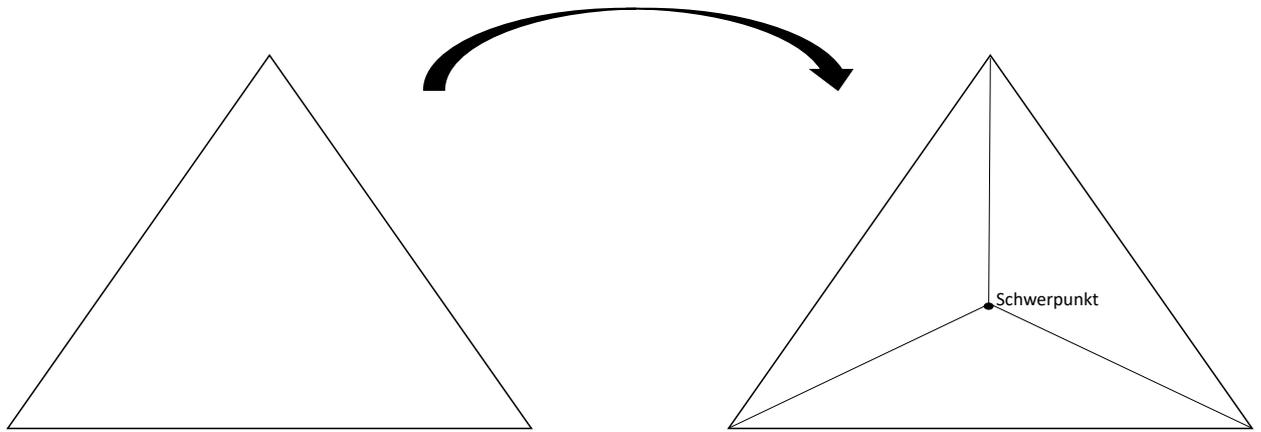


Abbildung 3.4: Verfeinerung durch den Schwerpunkt

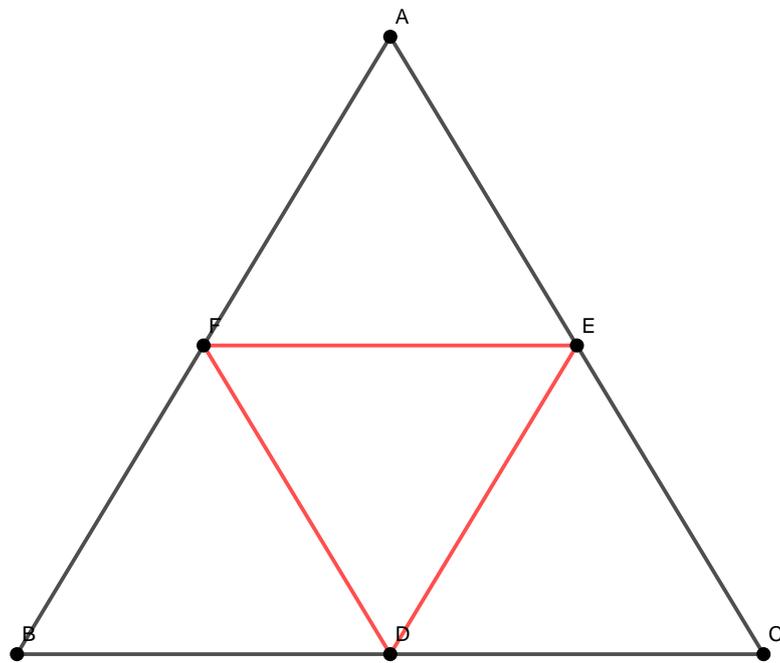


Abbildung 3.5: ROT-Verfeinerung

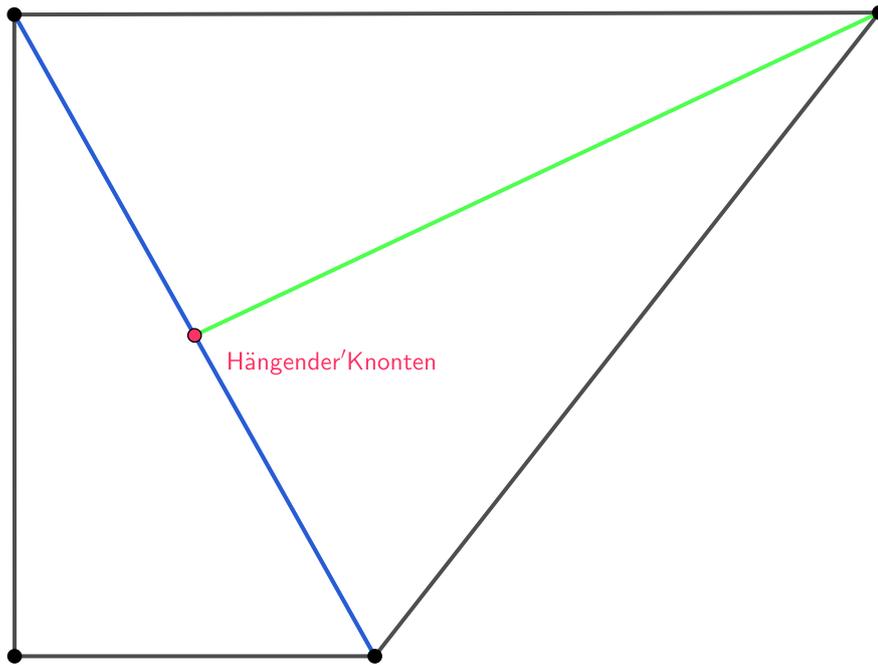


Abbildung 3.6: GRÜN-Verfeinerung

Sohnes T_1 und T'_1 , welche aber nicht Eckpunkte eines Sohnes eines Nachbardreiecks sind. Der Knoten wird von dem Nachbardreieck nicht „gestützt“, sie „hängen“ gewissermaßen an der Kante. Das heißt, die GRÜN-Verfeinerung verbindet hängende Knoten mit dem gegenüberliegenden Eckpunkt. Dabei werden die Kanten der verfeinerten Elemente grün markiert. Abbildung 3.6 zeigt eine GRÜN-Verfeinerung.

- (c) Bei Dreiecken mit mehr als einem hängenden Knoten wird die ROT-Verfeinerung angewendet.

BISEKTION VON MARKIERTEN KANTEN

- Markiere in \mathcal{T}_l für jedes Dreieck T die längste Kante.
- Verfeinerung eines markierten Dreiecks:
Bisektion durch Verbindung des Mittelpunktes der markierten Kanten mit dem gegenüberliegenden Eckpunkt.
- Markiere in den neuen Dreiecken, die alten gleichbleibenden Kanten.
- Elemente mit hängendem Knoten werden wieder durch Bisektion der markierten Kante unterteilt.

Bei beiden Strategien, sowohl bei der Verfeinerung über dem Schwerpunkt, als auch bei der ROT-GRÜN-Verfeinerung, wird der kleinste Winkel höchstens halbiert, deswegen sind die Zerlegungen $\{\mathcal{T}_l\}$ form-regulär. Ebenfalls kann gezeigt werden, dass der Algorithmus terminiert, das heißt, alle hängenden Knoten werden eliminiert und die nächste Zerlegung \mathcal{T}_{l+1} ist zulässig.

Kapitel 4

Ein äquilibrierter Fehlerschätzer

Im vorherigen Kapitel wurde die Effizienz und Zuverlässigkeit für den residualen Fehlerschätzer gezeigt. Außerdem ist man zum Entschluss gekommen, dass die Zuverlässigkeit eines Fehlerschätzers als obere Schranke der Fehlerabschätzung dient. Demnach kann man den Fehler in folgender Form anschreiben

$$\|u - u_N\|_a \leq c \eta .$$

Der residuale Fehlerschätzer ist dabei definiert als

$$\eta_R(u_N, f) := \left(\sum_{T \in \mathcal{T}_h} \eta_T^2(u_N, f) \right)^{\frac{1}{2}} ,$$

mit den elementweisen Fehlerindikatoren

$$\eta_T^2(u_N, f) := \|R_T(u_N, f)\|_{0,T}^2 h_T^2 + \frac{1}{2} \sum_{E \in \mathcal{E}(T)} \|R_E(u_N, f)\|_{0,E}^2 h_E .$$

Die Zuverlässigkeit des residualen Fehlerschätzers wird für das betrachtete Poisson-Problem folgendermaßen angeschrieben,

$$\exists c_r > 0 : \|u - u_N\|_a = \|\nabla u - \nabla u_N\|_{0,\Omega} \leq c_r \eta_R(u_N, f) .$$

Das zentrale Ziel dieses Kapitels ist es nun, einen **völlig zuverlässigen** Fehlerschätzer zu finden, das heißt

$$\|u - u_N\|_a \leq 1 \eta \leq \text{TOL} .$$

Um diese konstant-freie Fehlerabschätzung konstruieren zu können, betrachtet man das sogenannte duale gemischte Variationsproblem. Der Satz von Prager-Synge liefert schließlich das gewünschte Ergebnis. Abschließend wird mithilfe der Flusskorrekturen ein Weg gefunden, um die Fehlerabschätzung mit wenig Aufwand berechnen zu können.

4.1 Das gemischte Variationsproblem

In dieser Arbeit wurde bisher jenes Variationsproblem betrachtet, welches durch den Hilbertraum V definiert ist. Nun ist es möglich, sich dem sogenannten gemischten Variationsproblem zu widmen. Dafür sind zwei reelle Hilberträume $(X, (\cdot, \cdot)_X)$ und $(M, (\cdot, \cdot)_M)$ nötig. Bei einem gemischten Variationsproblem ist die Bilinearform $a(\cdot, \cdot)$ über dem Produktraum $X \times M$ definiert. Die Variationsformulierung lautet hierzu: Gesucht ist $\delta \in X$, sodass

$$a(\delta, \lambda) = f(\lambda)$$

für alle $\lambda \in M$ erfüllt ist. Diese Art von gemischten Variationsproblemen wird jedoch in dieser Arbeit nicht direkt benötigt.

Es gibt andere Formen von gemischten Variationsproblemen, die in weiterer Folge für diese Arbeit eine wesentliche Rolle spielen. Dazu seien X und M wieder zwei reelle Hilberträume, $a : X \times X \rightarrow \mathbb{R}$ und $b : X \times M \rightarrow \mathbb{R}$ zwei stetige (beschränkte) Bilinearformen, das heißt es existieren $\alpha, \beta > 0$, sodass

$$\begin{aligned} |a(\delta, \tau)| &\leq \alpha \|\delta\|_X \|\tau\|_X & \forall \delta, \tau \in X, \\ |b(\delta, \lambda)| &\leq \beta \|\delta\|_X \|\lambda\|_M & \forall \delta \in X, \forall \lambda \in M. \end{aligned}$$

Für die Formulierung des abstrakten gemischten Variationsproblems, werden noch zusätzliche Voraussetzungen an die Bilinearformen gestellt. Man fordert die sogenannten LBB-Bedingungen (Ladyzhenskaja-Babuška-Brezzi), oder auch inf – sup – Bedingungen von $b(\cdot, \cdot)$ und die ker B –Elliptizität von $a(\cdot, \cdot)$.

- LBB-Bedingungen (oder inf – sup – Bedingungen) von $b(\cdot, \cdot)$

$$\exists \beta_b > 0 : \inf_{\mu \in M} \sup_{0 \neq \delta \in X} \frac{b(\delta, \mu)}{\|\delta\|_X \|\mu\|_M} \geq \beta_b \iff \exists \beta_b > 0 : \sup_{0 \neq \delta \in X} \frac{b(\delta, \mu)}{\|\delta\|_X} \geq \beta_b \|\mu\|_M \quad \forall \mu \in M.$$

- ker B –Elliptizität von $a(\cdot, \cdot)$

$$\exists \beta_a > 0 : a(\delta, \delta) \geq \beta_a \|\delta\|_X^2 \quad \forall \delta \in \ker B,$$

wobei

$$\ker B := \{\tau \in X : b(\tau, \mu) = 0 \quad \forall \mu \in M\}.$$

Das abstrakte gemischte Variationsproblem lässt sich nun schreiben als: Gesucht ist $(\delta, \lambda) \in X \times M$, sodass

$$\begin{aligned} a(\delta, \tau) + b(\tau, \lambda) &= f(\tau) & \forall \tau \in X, \\ b(\delta, \mu) &= g(\mu) & \forall \mu \in M, \end{aligned} \tag{4.1}$$

wobei $f \in X^*$, $g \in M^*$, die Bilinearform $a(\cdot, \cdot) : X \times X \rightarrow \mathbb{R}$ symmetrisch, beschränkt auf $X \times X$ und ker B –elliptisch ist, die Bilinearform $b(\cdot, \cdot) : X \times M \rightarrow \mathbb{R}$ beschränkt auf $X \times M$ und die LBB-Bedingung für $b(\cdot, \cdot)$ erfüllt ist.

Bemerkung Das abstrakte gemischte Variationsproblem lässt sich als ein Sattelpunktproblem auffassen. Allgemein ist ein Sattelpunktproblem eine lineares Gleichungssystem in Blockgestalt der Form

$$M = \begin{pmatrix} A & B^* \\ B & 0 \end{pmatrix},$$

wobei im diskreten Fall A eine $n \times n$ -Matrix und B eine $n \times m$ -Matrix ist. Der 0-Block ist von der Größe $m \times m$ und somit ist M eine $(n+m) \times (n+m)$ -Matrix. Das Sattelpunktproblem (4.1) wird oft übersichtlicher, wenn es in die äquivalente Operatorgleichung umgeformt wird. Dazu wird der Bilinearform $a(\cdot, \cdot)$ die Abbildung

$$\begin{aligned} A : X &\rightarrow X^* , \\ \langle A\delta, \tau \rangle &= a(\delta, \tau) \quad \text{für } \tau \in X \end{aligned}$$

zugeordnet. Die Abbildung $A\delta \in X^*$ ist also ein Funktional, das Elemente $\tau \in X$ nach \mathbb{R} abbildet. Ebenso wird der Bilinearform $b(\cdot, \cdot)$ der Operator B zugeordnet und der adjungierte Operator B^* hat folgende Form.

$$\begin{aligned} B : X &\rightarrow M^* , & B^* : M &\rightarrow X^* , \\ \langle B\delta, \mu \rangle &= b(\delta, \mu) \quad \text{für } \mu \in M . & \langle B^*\lambda, \tau \rangle &= b(\tau, \lambda) \quad \text{für } \tau \in X . \end{aligned}$$

Das Problem (4.1) ist demnach äquivalent zu

$$\begin{aligned} A\delta + B^*\lambda &= f , \\ B\delta &= g . \end{aligned}$$

Nun ist es möglich die Existenz und Eindeutigkeit einer Lösung des gemischtes Problems zu zeigen. Dazu wird jedoch noch folgender Hilfssatz benötigt und es wird der Raum V eingeführt als

$$V := \{ \delta \in X : b(\delta, \mu) = 0 \text{ für } \mu \in M \} .$$

Theorem 4.1.1 *Die folgenden Aussagen sind äquivalent.*

(1) (LBB-Bedingung) *Es existiert eine Zahl $\beta > 0$ mit*

$$\inf_{\mu \in M} \sup_{0 \neq \tau \in X} \frac{b(\tau, \mu)}{\|\tau\|_X \|\mu\|_M} \geq \beta .$$

(2) *Der Operator $B : V^\perp \rightarrow M^*$ ist ein Isomorphismus und es gilt*

$$\|B\tau\|_{M^*} \geq \beta \|\tau\|_X \quad \text{für } \tau \in V^\perp ,$$

wobei

$$V^\perp := \{ x \in X : (x, v)_V = 0 \quad \forall v \in V \}$$

das orthogonale Komplement von V ist.

(3) *Der Operator $B^* : M \rightarrow V^\circ \subset X^*$ ist ein Isomorphismus und es ist*

$$\|B^*\mu\|_{V^\circ} \geq \beta \|\mu\|_M \quad \text{für } \mu \in M ,$$

wobei

$$V^\circ := \{ l \in X^* : \langle l, v \rangle = 0 \quad \forall v \in V \}$$

die Polare des abgeschlossenen Unterraumes V von X ist.

Beweis: Vgl. [2] Hilfssatz 4.2 .

□

Mit diesen Hilfssatz ist es möglich den Hauptsatz der Sattelpunktprobleme nach Brezzi aufzustellen. Dazu sei V der $\ker B$ von X ein abgeschlossener Unterraum.

Theorem 4.1.2 (*Brezzis Splitting Theorem*) Die durch das Sattelpunktproblem (4.1) definierte lineare Abbildung

$$L : X \times M \longrightarrow X^* \times M^* \quad \text{mit} \quad (\delta, \lambda) \longmapsto (f, g)$$

ist genau dann ein Isomorphismus, wenn die beiden folgenden Bedingungen,

(i) Die Bilinearform $a(\cdot, \cdot)$ ist $\ker B$ -elliptisch, das heißt, es gibt ein $\alpha > 0$, sodass

$$a(\tau, \tau) \geq \alpha \|\tau\|_X^2 \quad \forall \tau \in V = \ker B .$$

(ii) Die Bilinearform $b(\cdot, \cdot)$ erfüllt die LBB-Bedingung des vorherigen Hilfssatzes.

erfüllt sind.

Beweis: Vgl. [2] Satz 4.3 .

□

Bemerkung Dieses Theorem garantiert die Existenz und Eindeutigkeit einer Lösung $(u, \lambda) \in X \times M$, die stabil von den Daten (f, g) abhängt.

Bemerkung Lokal berechnete Näherungslösungen von Sattelpunktproblemen liefern untere Abschätzungen bezüglich der Energienorm und damit a-posteriori Schätzer.

Das nächste Ziel ist nun, das gemischte Variationsproblem für die Poisson-Gleichung aufzustellen und diese genauer zu analysieren.

4.2 Gemischte Finite-Elemente-Methoden

Die Intention dieses Abschnittes ist es eine numerische Lösung des Sattelpunktproblems (4.1) zu berechnen. Dafür bietet sich wiederum eine Diskretisierung der Funktionenräume an. Dazu wählt man endlich dimensionale Teilräume $X_N \subset X$ und $M_N \subset M$ und stellt das folgende Modellproblem auf: Gesucht ist $(u_N, \lambda_N) \in X_N \times M_N$, sodass

$$\begin{aligned} a(u_N, v_N) + b(v_N, \lambda_N) &= \langle f, v_N \rangle & \text{für } v_N \in X_N , \\ b(u_N, \mu_N) &= \langle g, \mu_N \rangle & \text{für } \mu_N \in M_N \end{aligned} \quad (4.2)$$

erfüllt ist.

Um nun die Existenz und Eindeutigkeit einer Lösung des diskreten Problems sicherstellen zu können, werden ähnliche Voraussetzungen wie beim abstrakten Problem gefordert. In der Regel ist dies nicht einfach zu erreichen, deswegen werden die diskreten Finite-Elemente-Räume $X_N \subset X$ und $M_N \subset M$ „ausbalanciert“, um die Stabilität der Finite-Elemente Rechnungen sicherstellen zu können.

Analog zu Abschnitt 4.1, steht V_N wieder für den Kern von B , das heißt

$$V_N := \{v_N \in X_N : b(v_N, \mu_N) = 0 \quad \text{für alle } \mu_N \in M_N\} .$$

Schließlich fordert man erneut die V_N -Elliptizität von $a(\cdot, \cdot)$ und die inf-sup-Bedingungen für $b(\cdot, \cdot)$, das heißt es gilt

(i) V_N -Elliptizität von $a(\cdot, \cdot)$

$$\exists \alpha > 0 : a(\delta_N, \delta_N) \geq \alpha \|\delta_N\|_{X_N}^2 \quad \forall \delta \in V_N .$$

(ii) Inf-sup-Bedingungen für $b(\cdot, \cdot)$

$$\exists \beta > 0 : \sup_{0 \neq v_N \in X_N} \frac{b(v_N, \lambda_N)}{\|v_N\|_{X_N}} \geq \beta \|\lambda_N\|_{M_N} \quad \forall \lambda_N \in M_N .$$

Bemerkung Im Allgemeinen ist $V_N \not\subset V$, da die Bedingungen im Raum V_N abgeschwächt werden und somit mehr $v_N \in V_N$ zugelassen sind als $v \in V$. Jedoch ist $V_N \subset V$ ein Sonderfall der konformen Approximation. Gilt nun $V_N \subset V$, dann ist laut Definition von V , jedes $v_N \in X_N : b(v_N, \mu_N) = 0$ für alle $\mu_N \in M_N$. Trotzdem werden die LBB-Bedingungen und die ker B -Bedingung explizit für die diskreten Räume gefordert, um die Existenz und Eindeutigkeit des diskreten Problems sicherstellen zu können.

Theorem 4.2.1 *Seien die Voraussetzungen des Brezzis Splitting Theorems erfüllt und seien die oben erwähnten Bedingungen an die Bilinearformen $a(\cdot, \cdot)$ und $b(\cdot, \cdot)$ in den diskreten Räumen X_N und M_N erfüllt. Außerdem gelte $V_N \subset V$, dann gilt für die Lösung des diskreten Sattelpunktproblems (4.2)*

$$\|u - u_N\| \leq c \inf_{v_N \in X_N} \|u - v_N\| .$$

Beweis: Sei $v_N \in X_N$, sodass $b(v_N, \mu_N) = \langle g, \mu_N \rangle$ für $\mu_N \in M_N$ gilt. Für $v_N \in V_N$ erhält man

$$\begin{aligned} a(u_N - v_N, v_N) &= a(u_N, v_N) - a(u, v_N) + a(u - v_N, v_N) \\ &= b(v_N, \lambda - \lambda_N) + a(u - v_N, v_N) \\ &\leq C \|u - v_N\| \|v_N\| , \end{aligned}$$

weil $b(v_N, \lambda - \lambda_N)$ wegen der Bedingung $V_N \subset V$ verschwindet. Setzt man nun $v_N := u_N - v_N$, so folgt $\|u_N - v_N\|^2 \leq \alpha^{-1} C \|u_N - v_N\| \|u - v_N\|$ und nach dem Kürzen und der Dreiecksungleichung die Behauptung.

□

4.3 Gemischte Formulierung des Poisson-Problems

In diesem Abschnitt wird die gemischte Formulierung des Poisson-Problems aufgestellt. Es gibt zwei verschiedene stabile Paarungen von Räumen. Man unterscheidet die sogenannten *dualen gemischten Methoden* und die *primalen gemischten Methoden*. Interessant ist jedoch, dass sich diese zwei Methoden durch ihre unterschiedlichen natürlichen Randbedingungen unterscheiden.

Man betrachte dazu das folgende Poisson-Problem,

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega , \\ \frac{\partial u}{\partial n} &= 0 && \text{auf } \Gamma_N , \\ u &= 0 && \text{auf } \Gamma_D = \partial\Omega \setminus \Gamma_N . \end{aligned} \tag{4.3}$$

Durch die Einführung der Hilfsfunktion $\underline{\delta} = \nabla u$ kann man das Problem auf ein System 1. Ordnung umformen, also

$$\begin{aligned}
 & \underline{\delta} = \nabla u && \text{in } \Omega , \\
 -\Delta u = -\operatorname{div}(\nabla u) = -\operatorname{div} \underline{\delta} = f & \iff \operatorname{div} \underline{\delta} = -f && \text{in } \Omega , \\
 & \underline{\delta} \cdot \underline{n} = 0 && \text{auf } \Gamma_N , \\
 & u = 0 && \text{auf } \Gamma_D .
 \end{aligned} \tag{4.4}$$

Aus diesem System erhält man direkt das folgende Sattelpunktproblem: Gesucht ist $(\underline{\delta}, u) \in [L_2(\Omega)]^d \times H_0^1(\Omega)$, sodass

$$\begin{aligned}
 (\underline{\delta}, \tau)_{0,\Omega} - (\nabla u, \tau)_{0,\Omega} &= 0 && \forall \tau \in [L_2(\Omega)]^d , \\
 -(\underline{\delta}, \nabla v)_{0,\Omega} &= -(f, v)_{0,\Omega} && \forall v \in H_0^1(\Omega) .
 \end{aligned}$$

In diesem Fall ist der Raum $X = [L_2(\Omega)]^d$ und $M = H_0^1(\Omega)$. Außerdem sind die Bilinearformen definiert als $a(\underline{\delta}, \tau) = (\underline{\delta}, \tau)_{0,\Omega}$ und $b(\tau, v) = -(\tau, \nabla v)_{0,\Omega}$.

Die primale gemischte Formulierung

Die oben erwähnte Sattelpunktgleichung ist die primale Formulierung der Poisson-Gleichung. Nun möchte man die Herleitung dieses Sattelpunktproblems genauer betrachten. Die gemischte Formulierung der Poisson-Gleichung (4.4) kann wiederum über partielle Integration auf eine Variationsformulierung gebracht werden. Die schwache Ableitung von $\operatorname{div} \underline{\delta}$ lautet

$$-\int_{\Omega} \operatorname{div} \underline{\delta} \psi \, dx = \int_{\Omega} \underline{\delta} \cdot \nabla \psi \, dx - \underbrace{\int_{\partial\Omega} \underbrace{(\underline{\delta} \cdot \underline{n})}_{= 0 \text{ auf } \Gamma_N} \psi \, ds_x}_{= 0} \quad \text{mit } \psi \in C_0^\infty(\Omega) : \psi = 0 \text{ auf } \Gamma_D .$$

Durch partielle Integration folgt daher

$$(\operatorname{div} \underline{\delta}, v) = -(\underline{\delta}, \nabla v) + \underbrace{\int_{\partial\Omega} (\underline{\delta} \cdot \underline{n}) v \, ds_x}_{= 0} ,$$

wobei der letztere Ausdruck gleich Null ist, da $(\underline{\delta} \cdot \underline{n}) = 0$ auf Γ_N ist und $v \in M$ so wählt, dass $v = 0$ auf Γ_D gilt.

Somit erhält man, wie schon oben genannt, die variationelle Formulierung: Gesucht ist $(\underline{\delta}, u) \in X \times M = [L_2(\Omega)]^d \times H_{0,D}^1(\Omega)$, sodass

$$\begin{aligned}
 (\underline{\delta}, \underline{\tau})_{0,\Omega} - (\nabla u, \underline{\tau})_{0,\Omega} &= 0 && \forall \underline{\tau} \in X , \\
 -(\underline{\delta}, \nabla v)_{0,\Omega} &= -(f, v)_{0,\Omega} && \forall v \in M
 \end{aligned} \tag{4.5}$$

erfüllt ist.

Definition 4.3.1 ($H_{0,D}^1(\Omega)$ -Raum und $H_{0,N}^1(\Omega)$ -Raum) Die Räume $H_{0,D}^1(\Omega)$ und $H_{0,N}^1(\Omega)$ sind definiert als

$$\begin{aligned}
 H_{0,D}^1(\Omega) &:= \{v \in H^1(\Omega) : v = 0 \text{ auf } \Gamma_D\} , \\
 H_{0,N}^1(\Omega) &:= \{v \in H^1(\Omega) : v = 0 \text{ auf } \Gamma_N\} .
 \end{aligned}$$

Theorem 4.3.2 *Das primale gemischte Variationsproblem (4.5) für $X = [L_2(\Omega)]^d$ und $M = H_{0,D}^1(\Omega)$ ist stabil lösbar.*

Beweis: Die Bilinearformen sind stetig und offensichtlich erfüllt $a(\cdot, \cdot)$ die L_2 -Elliptizität. Für die inf-sup-Bedingungen ist die Friedrichs-Ungleichung notwendig. Zu einem gegebenen $v \in H_{0,D}^1(\Omega)$ wertet man den Quotienten auf Basis der Bedingung $\underline{\tau} := -\nabla v \in [L_2(\Omega)]^d$ aus, das heißt

$$\frac{b(\underline{\tau}, v)}{\|\underline{\tau}\|_0} = \frac{-(\underline{\tau}, \nabla v)_{0,\Omega}}{\|\underline{\tau}\|_0} = \frac{-(-\nabla v, \nabla v)_{0,\Omega}}{\|\nabla v\|_0} = |v|_1 \geq \frac{1}{c} \|v\|_1 .$$

Da die Konstante der Friedrichs-Ungleichung nur von Ω abhängt, ist das Sattelpunktproblem bzw. die primale Formulierung (4.5) stabil. □

Die duale gemischte Formulierung

Für die Aufstellung der dualen gemischten Formulierung wird der vektorwertige Funktionenraum $H(\operatorname{div}, \Omega)$ oder kurz $H(\operatorname{div})$ eingeführt.

Definition 4.3.3 *Der Raum $H(\operatorname{div}, \Omega)$ ist definiert als*

$$H(\operatorname{div}, \Omega) := \{\underline{\tau} \in [L_2(\Omega)]^d : \operatorname{div} \underline{\tau} \in L_2(\Omega)\}$$

mit der Norm

$$\|\underline{\tau}\|_{H(\operatorname{div})} := \sqrt{\|\underline{\tau}\|_{0,\Omega}^2 + \|\operatorname{div} \underline{\tau}\|_{0,\Omega}^2} .$$

Außerdem bildet der Raum $(H(\operatorname{div}, \Omega), \|\cdot\|_{H(\operatorname{div})})$ einen Hilbertraum.

Dieser Raum weist gewisse Eigenschaften auf.

- $(H^1(\Omega))^d \subset H(\operatorname{div})$.
- Der Unterraum $H_{0,N}(\operatorname{div})$ ist definiert als

$$H_{0,N}(\operatorname{div}) := \{\underline{\tau} \in H(\operatorname{div}) : \underline{\tau} \cdot \underline{n} = 0 \text{ auf } \Gamma_N\}$$

und $\operatorname{meas}(\Gamma_N)_{d-1} > 0$.

Für $v \in H_0^1(\Omega)$ und $\underline{\delta} \in H(\operatorname{div}, \Omega)$ gelten aufgrund partieller Integration und dem Wegfallen der Randterme,

$$(\operatorname{div} \underline{\delta}, v)_{0,\Omega} = \int_{\Omega} \operatorname{div} \underline{\delta} v \, dx = - \int_{\Omega} \underline{\delta} \cdot \nabla v \, dx + \underbrace{\int_{\partial\Omega} (\underline{\delta} \cdot \underline{n}) v \, ds_x}_{= 0} = (\underline{\delta}, \nabla v)_{0,\Omega} .$$

Die duale gemischte Formulierung der Poisson-Gleichung lautet schließlich: Gesucht ist $(\underline{\delta}, u) \in X \times M$, sodass

$$\begin{aligned} (\underline{\delta}, \underline{\tau})_{0,\Omega} + (\operatorname{div} \underline{\tau}, u)_{0,\Omega} &= 0 & \forall \underline{\tau} \in X , \\ (\operatorname{div} \underline{\delta}, v)_{0,\Omega} &= -(f, v)_{0,\Omega} & \forall v \in M , \end{aligned}$$

mit $X = H_{0,N}(\operatorname{div})$ und $M = L_2(\Omega)$ erfüllt ist, wobei die Bilinearformen definiert sind als $a(\underline{\delta}, \underline{\tau}) = (\underline{\delta}, \underline{\tau})_{0,\Omega}$ und $b(\underline{\tau}, v) = (\operatorname{div} \underline{\tau}, v)_{0,\Omega}$.

Lemma 4.3.4 *Das duale gemischte Problem der Poisson-Gleichung ist ein stabiles Sattelpunktproblem und es existiert eine eindeutige Lösung $(\underline{\delta}, u) \in H_{0,N}(\text{div}) \times L_2(\Omega)$.*

Beweis: Vgl. [7] Lemma 2.17 .

□

Wie bereits erwähnt, sind sich die primale und die duale gemischte Formulierung des Poisson-Problems ähnlich. Das folgende Lemma zeigt, dass die Lösung des dualen gemischten Poisson-Problems auch immer eine Lösung des primalen Poisson-Problems ist. Es sei darauf hingewiesen, dass die Umkehrung dieser Aussage im Allgemeinen jedoch nicht gilt.

Lemma 4.3.5 *Sei $(\underline{\delta}, w) \in H_{0,N}(\text{div}) \times L_2(\Omega)$ die Lösung des dualen gemischten Poisson-Problems. Dann ist $w \in H_{0,D}^1(\Omega)$ und stimmt mit der Lösung $u \in H_{0,D}^1(\Omega)$ des primalen Poisson-Problems überein.*

Beweis: Vgl. [7] Lemma 2.18 .

□

In Kapitel 2 wurde bereits gezeigt, dass das abstrakte Variationsproblem (2.12) äquivalent zum Minimierungsproblem (2.13) ist. Ebenfalls ist es möglich, das primale und das duale gemischte Problem auf andere äquivalente Energieproblemstellungen umzuformen.

Das primale gemischte Problem ist äquivalent zur Minimierung des Ritz' schen Funktionals

$$J(u) = \frac{1}{2}(\nabla u, \nabla u)_0 - f(u)$$

$$J(u) \xrightarrow[u \in H_{0,D}^1(\Omega)]{} \min !$$

Das folgende Lemma gibt darüber Auskunft, wie man das duale gemischte Problem äquivalent als eine Maximierungsaufgabe umschreiben kann.

Lemma 4.3.6 *(Komplementäre Energie)*

(a) *Das duale gemischte Problem ist äquivalent zur Maximierung der komplementären Energie*

$$J^*(\underline{\tau}) = - \frac{1}{2}(\underline{\tau}, \underline{\tau})_0 ,$$

unter der Nebenbedingung

$$\underline{\tau} \in H_{0,N}(\text{div}) : \text{div } \underline{\tau} + f = 0 .$$

Das heißt, die Lösung $(\underline{\delta}, u) \in H_{0,N}(\text{div}) \times L_2(\Omega)$ des dualen gemischten Problems ist zugleich die Lösung des Problems

$$J^*(\underline{\delta}) \longrightarrow \max !$$

$$\text{div } \underline{\delta} + f = 0 \quad \underline{\delta} \in H_{0,N}(\text{div}) .$$

(b) *In den jeweiligen Extrempunkten stimmen die Energiefunktionale überein, so wie in Abbildung 4.1 zu sehen ist. Das heißt*

$$J(u) = J^*(\underline{\delta}) .$$

Beweis: Vgl. [7] Lemma 2.19 .

□

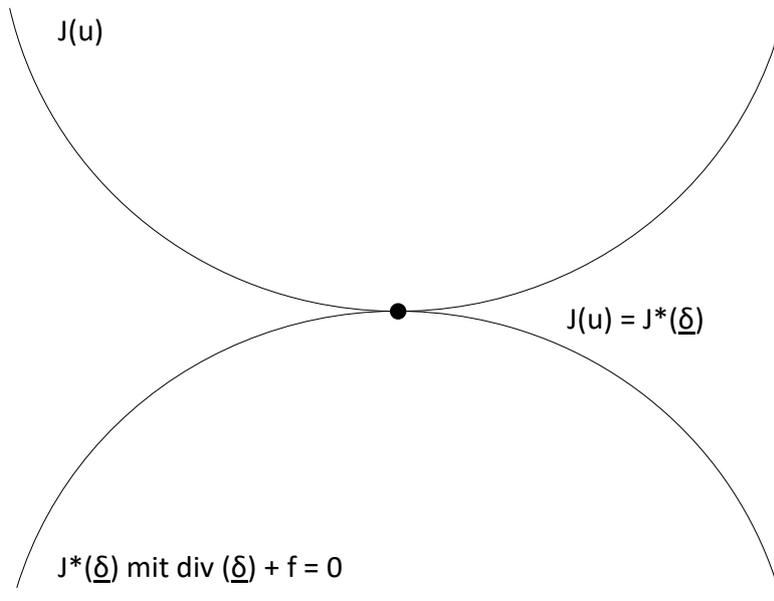


Abbildung 4.1: Übereinstimmung der Extremalpunkte

4.3.1 Satz von Prager-Synge und Anwendung

Der Satz von Prager-Synge liefert schlussendlich eine konstant-freie Fehlerabschätzung des a-posteriori Fehlerschätzers. Die Vorarbeit des gesamten vierten Kapitels diente dazu, um diesen bedeutenden Satz auf das Poisson-Problem anwenden zu können.

Theorem 4.3.7 (Satz von Prager-Synge) Sei $u \in H_{0,D}^1(\Omega)$ eine Lösung des primalen Problems der Poisson-Gleichung (4.3). Sei $\underline{\delta}$ eine vektorwertige Funktion aus $H(\operatorname{div})$ mit $\underline{\delta} \cdot \underline{n} = 0$ auf Γ_N , welche die Gleichung $\operatorname{div} \underline{\delta} + f = 0$ erfüllt. Außerdem sei $v \in H_{0,D}^1(\Omega)$, dann gilt

$$\|\underline{\delta} - \nabla u\|_0^2 + \|\nabla u - \nabla v\|_0^2 = \|\underline{\delta} - \nabla v\|_0^2 .$$

Beweis: Es gilt,

$$\begin{aligned} \|\underline{\delta} - \nabla v\|_0^2 &= \|\underline{\delta} - \nabla u + \nabla u - \nabla v\|_0^2 \\ &= \|\underline{\delta} - \nabla u\|_0^2 + \|\nabla u - \nabla v\|_0^2 + 2(\underline{\delta} - \nabla u, \nabla u - \nabla v) . \end{aligned}$$

Zu zeigen ist nun $(\underline{\delta} - \nabla u, \nabla u - \nabla v) = 0$. Dies liefert die Green'sche Formel.

$$\begin{aligned} \int_{\Omega} (\underline{\delta} - \nabla u) \cdot (\nabla u - \nabla v) \, dx &= \int_{\Omega} \underline{\delta} \cdot (\nabla u - \nabla v) \, dx - \int_{\Omega} \nabla u \cdot (\nabla u - \nabla v) \, dx \\ &= - \int_{\Omega} \operatorname{div} \underline{\delta} \cdot (u - v) + \int_{\partial\Omega} \underbrace{(\underline{\delta} \cdot \underline{n})}_{=0 \text{ auf } \Gamma_N} \underbrace{(u - v)}_{=0 \text{ auf } \Gamma_D} \, ds_x - \underbrace{\int_{\Omega} \nabla u \cdot (\nabla u - \nabla v) \, dx}_{(f, u-v)} \\ &= (f, u - v) - (f, u - v) = 0 . \end{aligned}$$

□

Dieser Satz wird schließlich angewendet, um äquilibrierte Fehlerschätzer mit der Zuverlässigkeitskonstante 1 zu bestimmen. Dazu sei nun $\underline{\tau} \in H_{0,N}^1(\operatorname{div})$, sodass $\operatorname{div} \underline{\tau} + f = 0$ erfüllt ist. Dann gilt

(i)

$$\begin{aligned}
 J(u_N) - J^*(\underline{\tau}) &= \frac{1}{2}(\nabla u_N, \nabla u_N) - (f, u_N) + \frac{1}{2}(\underline{\tau}, \underline{\tau}) \\
 &= \frac{1}{2}(\nabla u_N, \nabla u_N) - (f, u_N) + \frac{1}{2}(\underline{\tau}, \underline{\tau}) - (\nabla u_N, \underline{\tau}) + (\nabla u_N, \underline{\tau}) \\
 &= \frac{1}{2}(\nabla u_N, \nabla u_N) - (\nabla u_N, \underline{\tau}) + \frac{1}{2}(\underline{\tau}, \underline{\tau}) - (f, u_N) + (\nabla u_N, \underline{\tau}) \\
 &\quad \underbrace{\hspace{10em}}_{=\frac{1}{2}(\nabla u_N - \underline{\tau}, \nabla u_N - \underline{\tau})} \\
 &= \frac{1}{2}\|\nabla u_N - \underline{\tau}\|_0^2 - (f, u_N) + \underbrace{(\nabla u_N, \underline{\tau})}_{\text{part. Integration}} \\
 &= \frac{1}{2}\|\nabla u_N - \underline{\tau}\|_0^2 - \underbrace{(f, u_N) - (\operatorname{div} \underline{\tau}, u_N)}_{-(f + \operatorname{div} \underline{\tau}, u_N) = 0} + \underbrace{(u_N, \underline{\tau} \cdot \underline{n})_{\partial\Omega}}_{= 0} \\
 &= \frac{1}{2}\|\nabla u_N - \underline{\tau}\|_0^2 \\
 &= \frac{1}{2}\|\underline{\tau} - \nabla u\|_0^2 + \frac{1}{2}\|\nabla u - \nabla u_N\|_0^2 \quad \text{mit Satz von Prager-Synge.}
 \end{aligned}$$

Daraus folgt

$$2 (J(u_N) - J^*(\underline{\tau})) \geq \|\nabla u - \nabla u_N\|_0^2 .$$

(ii) Durch das Einsetzen von $v = u_N \in H_{0,D}^1(\Omega)$ der diskreten Lösung des Poisson-Problems erhält man für beliebiges $\underline{\delta} \in H_{0,N}^1(\operatorname{div})$ mit $\operatorname{div} \underline{\delta} + f = 0$ somit

$$\|\nabla u - \nabla u_N\|_0 \leq 1 \|\nabla u_N - \underline{\delta}\|_0 .$$

Das heißt, es existiert eine völlig zuverlässige a-posteriori Fehlerschranke.

$\underline{\delta}$ wird auch als äquilibrierter Fluss bezeichnet.

Es stellt sich nun die Frage, wie der äquilibrierte Fluss $\underline{\delta}$ gewählt und konstruiert werden kann. Im Allgemeinen ist der äquilibrierte Fluss nicht eindeutig, da

$$\begin{aligned}
 \operatorname{div} \underline{\delta} + f &= 0 & \text{für } \tilde{\underline{\delta}} = \underline{\delta} + \nabla \times \varphi \quad \text{mit } \varphi \in [H_{0,N}^1(\Omega)]^d \text{ gilt:} \\
 \operatorname{div} \tilde{\underline{\delta}} + f &= 0 \iff \operatorname{div} \underline{\delta} + \underbrace{\operatorname{div}(\nabla \times \varphi)}_{= 0} + f = 0 \\
 \underline{\delta} \cdot \underline{n} &= 0 \quad \text{auf } \Gamma_N & \implies \operatorname{div} \tilde{\underline{\delta}} + f = \operatorname{div} \underline{\delta} + f = 0 \implies \operatorname{div} \tilde{\underline{\delta}} = \operatorname{div} \underline{\delta} .
 \end{aligned}$$

Zur Bestimmung des äquilibrierten Flusses betrachtet man schließlich die Diskretisierung des Raumes $H(\operatorname{div})$. Hier wird im Speziellen die Diskretisierung im eindimensionalen bzw. im zweidimensionalen Raum erwähnt. Sei nun also $\Omega \subset \mathbb{R}^d$. Dann gilt für

- $d = 1$: Im Eindimensionalen ist der Raum $H(\operatorname{div}) = H^1(\Omega)$.

Die Räume X_N und M_N sind:

$$X_N = S_h^{1,1}, \quad M_N = \{v \in L_2(\Omega) : v|_T \in P^0(T)\} .$$

Die Diskretisierung des dualen gemischten Problems auf $X_N \times M_N$ ist stabil.

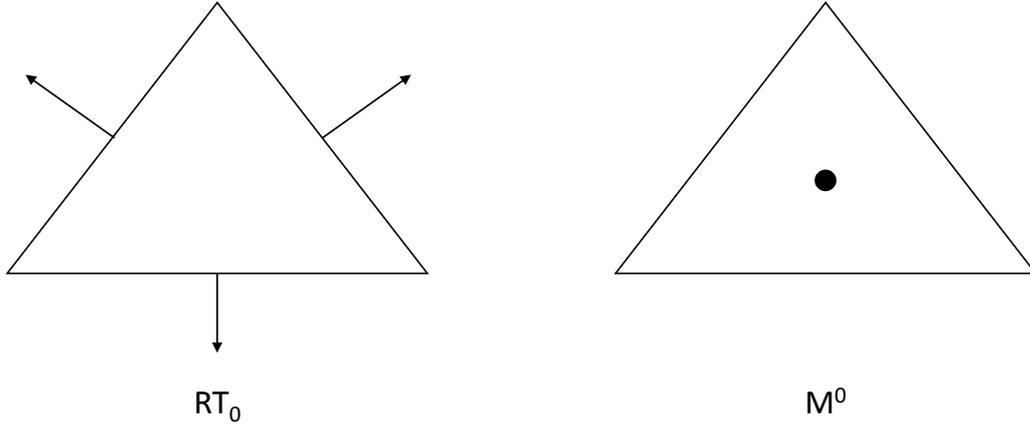


Abbildung 4.2: Raviart-Thomas-Element

- $d = 2$: Im Zweidimensionalen ist der Raum $H(\text{div})$ bereits in Definition 4.3.3 eingeführt worden:

$$H(\text{div}) = \{ \tau \in [L_2(\Omega)]^2 : \text{div } \tau \in L_2(\Omega) \} .$$

Für den diskreten Raum X_N eignen sich die sogenannten Raviart-Thomas (RT) Elemente. Diese sind definiert durch

$$\begin{aligned} X_N = RT_0(\mathcal{T}_h) &:= \{ \underline{\tau} \in H(\text{div}) : \underline{\tau}|_T = \begin{pmatrix} a_T \\ b_T \end{pmatrix} + c_T \begin{pmatrix} x \\ y \end{pmatrix} \quad \forall T \in \mathcal{T}_h \text{ und } a_T, b_T, c_T \in \mathbb{R} \} \\ &= \{ \underline{\tau} \in [L_2(\Omega)]^2 : \underline{\tau}|_T = \begin{pmatrix} a_T \\ b_T \end{pmatrix} + c_T \begin{pmatrix} x \\ y \end{pmatrix} \quad \forall T \in \mathcal{T}_h; \\ &\quad \underline{\tau} \cdot \underline{n} \text{ ist stetig an den Elementgrenzen} \} . \end{aligned}$$

Der Raum $RT_0(\mathcal{T}_h)$ wird auch Raviart-Thomas Raum der nullten Ordnung genannt.

Der Raum $M_N := \mathcal{M}^0(\mathcal{T}_h)$ ist definiert als

$$\mathcal{M}^0(\mathcal{T}_h) = \{ v \in L_2(\Omega) : v|_T = a_T \quad \text{mit } a_T \in \mathbb{R} \} .$$

Die Stetigkeit der Normalkomponenten an den Elementgrenzen sichert die Konformität von $X_N \subset H(\text{div})$. Die Raviart-Thomas Elemente sind dadurch charakterisiert, dass auf jeder Gerade $\alpha x + \beta y = \text{const}$ die Normalableitung konstant ist, siehe dazu Abbildung 4.2. Das heißt es gilt unter anderem, $\underline{\tau} \cdot \underline{n}|_E = \text{const}$ für alle $E \in \mathcal{E}(T)$.

Theorem 4.3.8 Sei $u_N \in \mathcal{M}_0^1(\mathcal{T}_h)$, wobei $\mathcal{M}_0^1(\mathcal{T}_h)$ der linear polynomiale Finite-Element-Raum in $H^1(\Omega)$ ist, und sei $f_N \in \mathcal{M}^0(\mathcal{T}_h)$. Außerdem sei $(\underline{\delta}_N, w_N) \in RT_0(\mathcal{T}_h) \times \mathcal{M}^0(\mathcal{T}_h)$ die Lösung der gemischten Methode mit Raviart-Thomas-Elementen niedrigster Ordnung. Dann liefert der Satz von Prager-Synge

$$\| \nabla u - \nabla u_N \|_0 \leq 1 \quad \| \underline{\delta}_N - \nabla u_N \|_0 .$$

Weiteres gilt

$$\| \nabla u_N - \underline{\delta}_N \|_0 = \min \left\{ \| \nabla u_N - \underline{\tau}_N \|_0 : \underline{\tau}_N \in RT_0(\mathcal{T}_h), \text{div } \underline{\tau}_N + f_N = 0 \right\} .$$

Beweis: Folgt direkt aus dem Satz von Prager-Synge. □

Dieser Satz benötigt jedoch die globale Lösung des dualen gemischten Problems. Die Berechnung der globalen Lösung ist allerdings viel zu teuer und mit hohem Rechenaufwand verbunden. Ein Hilfsmittel, welches die Berechnung der globalen Lösung des dualen Problems umgeht, ist die sogenannte Flusskorrektur.

Äquilibrierter Fluss durch direkte Konstruktion einer Flusskorrektur $\underline{\delta}^\Delta := \nabla u_N - \underline{\delta}$

Theorem 4.3.9 Sei $R(u_N)$ das Residuum der diskreten Lösung $u_N \in S_h^{1,1}$ des primalen Poisson-Problems und $f \in L_2(\Omega)$ stückweise konstant. Außerdem sei $\underline{\delta} = \nabla u_N - \underline{\delta}^\Delta$ ein äquilibrierter Fluss und es gelte für die Korrektur $\underline{\delta}^\Delta \in [L_2(\Omega)]^d$ mit $\underline{\delta}^\Delta|_T \in RT_0(\mathcal{T}_h)$

$$\begin{aligned} \operatorname{div} \underline{\delta}^\Delta|_T &= R_T(u_N) & \forall T \in \mathcal{T}_h, \\ \llbracket \underline{\delta}^\Delta \cdot \underline{n} \rrbracket|_E &= R_E(u_N) & \forall E \in \mathcal{E}_h \setminus \Gamma_h, \\ \underline{\delta}^\Delta \cdot \underline{n} &= 0 & \text{auf } \Gamma_h. \end{aligned}$$

Dann gilt die Abschätzung

$$\|\nabla u - \nabla u_N\|_0 \leq \|\underline{\delta}^\Delta\|_0.$$

Beweis: Sei $\underline{\delta} = \nabla u_N - \underline{\delta}^\Delta \in [L_2(\Omega)]^d$. Die schwache Divergenz von $\underline{\delta}$ lautet für alle $v \in C_0^\infty(\Omega)$,

$$\begin{aligned} -(\underline{\delta}, \nabla v) &= -(\nabla u_N, \nabla v) + (\underline{\delta}^\Delta, \nabla v) \\ &= -(\nabla u_N, \nabla v) - \sum_{T \in \mathcal{T}_h} \left[(\operatorname{div} \underline{\delta}^\Delta|_T, v) - \sum_{\partial T} \langle \underline{\delta}^\Delta \cdot \underline{n}, v \rangle_{\partial T} \right] \\ &= -(\nabla u_N, \nabla v) - \sum_T (R_T(u_N), v) - \sum_E \langle R_E(u_N), v \rangle \\ &\quad \underbrace{=}_{(R(u_N), v) = (f, v) - (\nabla u_N, \nabla v)} -(f, v) \end{aligned}$$

für $f \in L_2(\Omega)$. Somit gilt insgesamt

$$\begin{aligned} -(\underline{\delta}, \nabla v) &= (\operatorname{div} \underline{\delta}, v) = -(f, v) \\ \implies \operatorname{div} \underline{\delta} &= -f \quad \in L_2(\Omega). \end{aligned}$$

□

Mit diesem Satz ist es nun möglich den äquilibrierten Fluss $\underline{\delta}$ aus der Differenz $\underline{\delta}^\Delta := \nabla u_N - \underline{\delta}$ zu berechnen, anstatt ihn aus der Eigenschaft $\operatorname{div} \underline{\delta} = -f$ direkt ermitteln zu müssen. Die Konstruktion der Korrektur $\underline{\delta}^\Delta$ erfolgt über das sogenannte lokale Vertex-Patch Problem.

Berechnung der Korrektur Korrektur in 1D

Man betrachte dazu das eindimensionale Gebiet $\Omega = (a, b)$ und die Näherungslösung $u_N \in S_h^{1,1}$. Das Residuum ist definiert als

$$\begin{aligned} \langle R(u_N), v_N \rangle &= \sum_{T \in \mathcal{T}_h} (R_T(u_N), v_N) + \sum_{V \in \mathcal{V}_h} \langle R_V(u_N), v_N \rangle \\ &= \sum_{T \in \mathcal{T}_h} (f_T, v_N) + \sum_{V \in \mathcal{V}_h} \langle \llbracket u'_N \rrbracket_V, v_N \rangle, \end{aligned}$$

da gilt

$$R_T(u_N) = |f_T + \underbrace{\Delta u_N}_0 \text{ in 1D}|_T \quad \text{und} \quad R_E(u_N) = R_V(u_N) = \llbracket \nabla u_N \cdot \underline{n}_E \rrbracket \stackrel{\text{in 1D}}{=} \llbracket u'_N \rrbracket.$$

Durch die Anwendung des vorherigen Theorems lautet das lokale Vertex-Patch Problem schließlich. Für jedes $V \in \mathcal{V}_h$ (die Menge aller Knoten) sei $\delta_V \in L_2(w_V)$, sodass gilt

$$\begin{aligned} \operatorname{div} \delta_V|_T &= \frac{1}{2} f_T && \text{auf } T \in w_V, \\ \llbracket \delta_V \rrbracket_V &= \llbracket u'_N \rrbracket_V, \\ \delta_V &= 0 && \text{auf } \partial w_V. \end{aligned}$$

Definiert man δ^Δ als

$$\delta^\Delta := \sum_{V \in \mathcal{V}_h} \delta_V,$$

so folgt

$$\begin{aligned} \operatorname{div} \delta^\Delta &= \operatorname{div} \sum_{V \in \mathcal{V}_h} \delta_V|_T = \frac{1}{2} f_T + \frac{1}{2} f_T = f_T, \\ \llbracket \sum_{V \in \mathcal{V}_h} \delta_V \rrbracket_{V'} &= \llbracket \delta_{V'} \rrbracket_{V'} = \llbracket u'_N \rrbracket_{V'}, \\ \sum_{V \in \mathcal{V}_h} \delta_V &= 0 && \text{auf } \partial w_V. \end{aligned}$$

Die Flusskorrektur kann daher über das lokale Vertex-Patch Problem einfach und mit wenig Aufwand konstruiert werden. Die zweite Gleichung des lokalen Vertex-Patch Problems lautet

$$\llbracket \delta_V \rrbracket_V = \llbracket u'_N \rrbracket_V.$$

Es stellt sich nun die Frage, wie sich diese Gleichung lösen lässt.

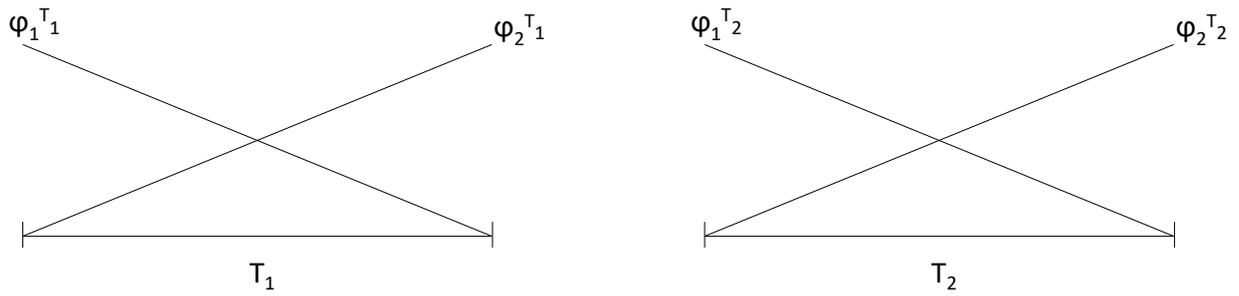


Abbildung 4.3: Elementfunktionen φ_i

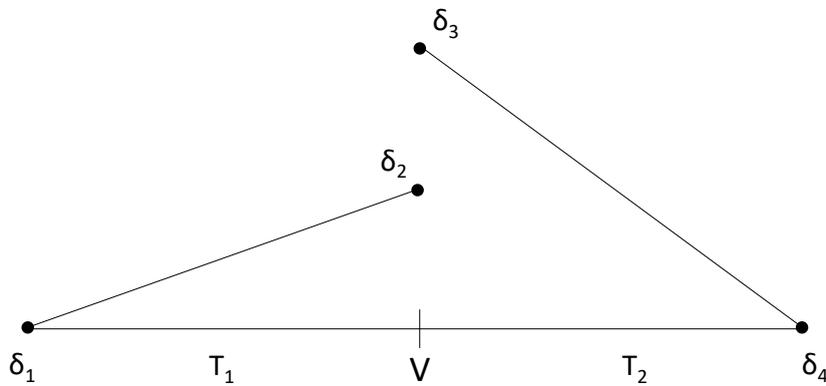


Abbildung 4.4: Sprung der Elementfunktionen von 2 verschiedenen Elementen T_1 und T_2

Dazu wählt man $\delta_V \in \{\tau : \tau|_T \in P^1(T), \quad \forall T \in w_V\}$. Dadurch ergibt sich, wie in Abbildung 4.3 und 4.4 zu sehen ist,

$$\begin{aligned}\delta_V|_{T_1} &= \delta_1 \varphi_1^{T_1} + \delta_2 \varphi_2^{T_1} , \\ \delta_V|_{T_2} &= \delta_3 \varphi_1^{T_2} + \delta_4 \varphi_2^{T_2} , \\ \operatorname{div} \delta_V|_{T_1} &= \frac{1}{2} f_T ,\end{aligned}$$

wobei $\varphi_1^{T_1}, \varphi_2^{T_1}, \varphi_1^{T_2}, \varphi_2^{T_2} \in S_h^{1,1}$.

Das Ziel ist es, die Koeffizienten $\delta_1, \delta_2, \delta_3, \delta_4$ anhand der vorhandenen Informationen des lokalen Vertex-Patch Problems zu bestimmen. Offensichtlich sind δ_1 und δ_4 gleich Null, da diese am Rand der Elemente vorgegeben sind. Zur Bestimmung von δ_2 und δ_3 werden die Informationen des lokalen Vertex-Patch Problems benötigt.

Es gelten folgende Eigenschaften.

(i)

$$\operatorname{div} \delta_V|_{T_1} \stackrel{\underbrace{\quad}_{\text{Ableitung einer linearen Funktion}}}{=} \frac{\delta_2 - \delta_1}{|T_1|} = \frac{1}{2} f_{T_1} \implies \delta_2 = \frac{1}{2} |T_1| f_{T_1} + \underbrace{\delta_1}_{=0} .$$

(ii)

$$\underbrace{[[\delta_V]]_V}_{\text{Sprung im Knoten}} = [[u'_N]]_V ,$$

$$\delta_3 - \delta_2 = [[u'_N]]_V \iff \delta_3 = [[u'_N]]_V + \delta_2 \implies \delta_3 = [[u'_N]]_V + \frac{1}{2} |T_1| f_{T_1} .$$

(iii)

$$\operatorname{div} \delta_V|_{T_2} \stackrel{\underbrace{\quad}_{\text{Ableitung einer linearen Funktion}}}{=} \frac{\delta_4 - \delta_3}{|T_2|} = \frac{1}{2} f_{T_2} \implies \delta_4 = \frac{1}{2} |T_2| f_{T_2} + [[u'_N]]_V + \frac{1}{2} |T_1| f_{T_1} .$$

Wie vorher erwähnt, soll der vierte Koeffizient δ_4 ebenfalls gleich Null sein. Deswegen sind gewisse Kompatibilitätsbedingungen nötig. Es wird also überprüft, ob $\delta_4 = \frac{1}{2} |T_2| f_{T_2} + [[u'_N]]_V + \frac{1}{2} |T_1| f_{T_1} = 0$ ist.

Da die Galerkin-Orthogonalität für den residualen Fehlerschätzer $\langle R(u_N), v_N \rangle = 0$ für alle $v_N \in S_h^{1,1}$ gilt, so gilt auch für die Hutfunktion ϕ_V mit $\phi_V(V') = \delta_{V,V'}$ bezüglich dem Knoten V

$$\begin{aligned} 0 &= \langle R(u_N), \phi_V \rangle = \sum_{T \in \mathcal{T}_h} (R_T(u_N), \phi_V)_{0,T} + \sum_{V' \in \mathcal{V}_h} R_V(u_N) \phi_V(V') \\ &= \int_{T_1} f_{T_1} \phi_V \, dx + \int_{T_2} f_{T_2} \phi_V \, dx + [[u'_N]]_V \cdot 1 \\ &= \frac{1}{2} f_{T_1} |T_1| + \frac{1}{2} f_{T_2} |T_2| + [[u'_N]]_V = \delta_4 . \end{aligned}$$

Somit ist die Lösbarkeit von $[[\delta_V]]_V = [[u'_N]]_V$ in $\{\tau : \tau|_V \in P^1(T), \forall T \in \mathcal{W}_V\}$ garantiert und es folgt

$$\|\nabla u - \nabla u_N\| \leq 1 \left\| \sum_{V \in \mathcal{V}_h} \delta_V \right\|_0 = \|\delta^\Delta\| .$$

Das Ziel dieses Kapitels ist somit erreicht und das Ergebnis liefert einen konstant-freien äquili-brierten a-posteriori Fehlerschätzer, der über die Flusskorrektur mit geringem Rechenaufwand berechenbar ist.

Kapitel 5

Numerisches Experiment

Für die Realisierung des C++ Programmes wurde die folgende eindimensionale Poisson-Gleichung gewählt. Gesucht ist $u : [0, 1] \rightarrow \mathbb{R}$, sodass

$$\begin{aligned} u''(x) &= \begin{cases} 1, & x \in (0, \frac{1}{2}) \\ -2, & x \in [\frac{1}{2}, 1) \end{cases}, \\ u(0) &= 0, \\ u(1) &= 0, \end{aligned} \quad (5.1)$$

erfüllt ist. Somit ist die exakte Lösung gegeben durch

$$u(x) = \begin{cases} \frac{1}{8} (x - 4x^2), & x \in (0, \frac{1}{2}) \\ \frac{3}{8} - \frac{11}{8}x + x^2, & x \in [\frac{1}{2}, 1) \end{cases}.$$

Abbildung 5.1 stellt die Lösung dieser Poisson-Gleichung als Graphik dar. Demnach gilt für die Ableitung,

$$u'(x) = \begin{cases} \frac{1}{8} (1 - 8x), & x \in (0, \frac{1}{2}) \\ -\frac{11}{8} + 2x, & x \in [\frac{1}{2}, 1) \end{cases}.$$

Am Anfang besteht das Intervall $(0, 1)$ aus einer sehr groben Zerlegung, beispielsweise eine Aufteilung in vier gleichgroßen Elementen. Das Programm berechnet für jedes Element den lokalen Fehler des äquilibrierten Flusses. Ist der globale Fehler größer als eine fix gewählte Toleranzgrenze, so werden durch die Maximum-Markierungsstrategie jene Elemente markiert, deren lokaler Fehler größer als der skalierte größte Fehler ist. Die markierten Elemente werden verfeinert. Die Verfeinerung beruht auf eine Halbierung des Elements. Das Programm berechnet nun wieder die lokalen Fehler und wiederholt sich, bis schließlich der globale Fehler kleiner als die Toleranzgrenze ist. In Abbildung 5.2 ist die Verfeinerung des Intervalls $(0, 1)$ nach gewissen Iteration abgebildet. Die Grafiken 5.3 bis 5.7 stellen die Näherungslösung bei bestimmter Zerlegung dar. Es ist gut zu erkennen, dass die Näherungslösung bei einer groben Zerlegung die exakte Lösung sehr schlecht approximiert. Je feiner die Zerlegung, desto besser stimmt die Näherungslösung mit der exakten Lösung überein. Schließlich wird der Fehler des äquilibrierten Flusses mit dem Fehler bestehend aus der Differenz der exakten Lösung und Näherungslösung verglichen. Laut Kapitel 4 gilt,

$$\|u' - u'_N\| \leq \|\delta^\Delta\|.$$

Für das gewählte Experiment stimmen die beiden Fehler sogar überein, wie in Abbildung 5.8 zu erkennen ist.

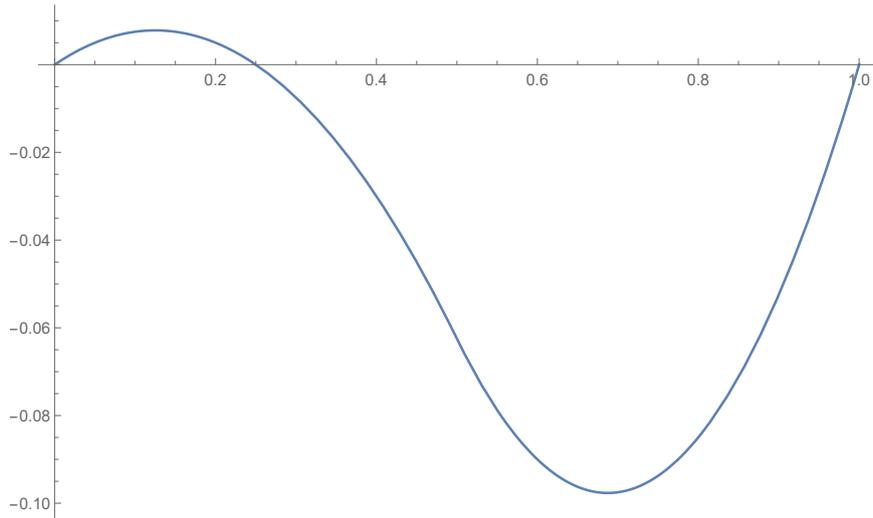


Abbildung 5.1: Exakte Lösung der Poisson-Gleichung (5.1)

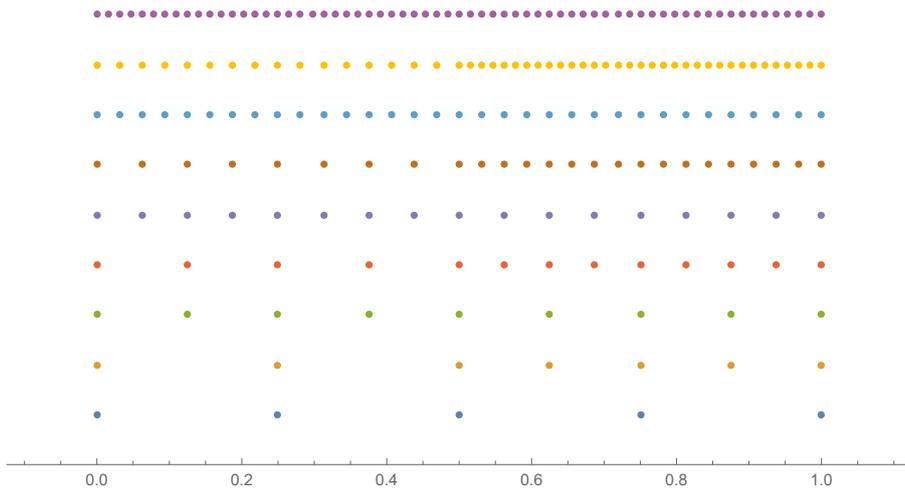


Abbildung 5.2: Verfeinerung des Intervalls nach gewissen Iterationen

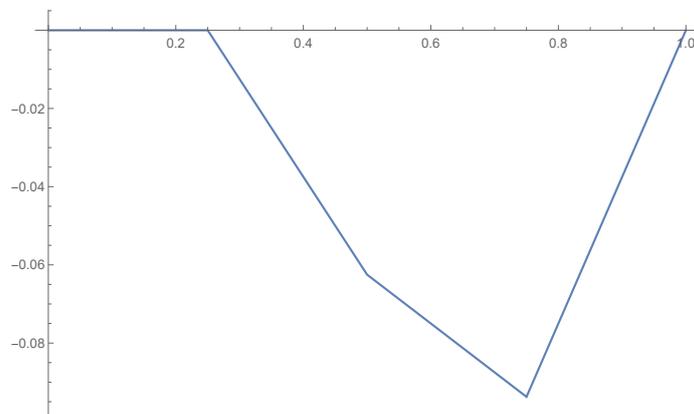


Abbildung 5.3: Näherungslösung bei einer Verfeinerung von 4 Elementen

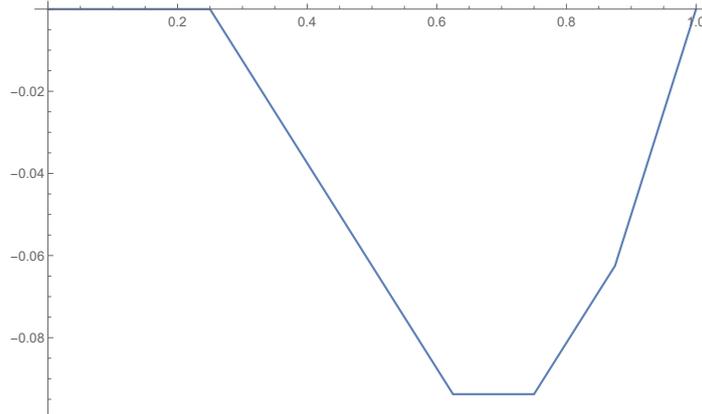


Abbildung 5.4: Näherungslösung bei einer Verfeinerung von 6 Elementen

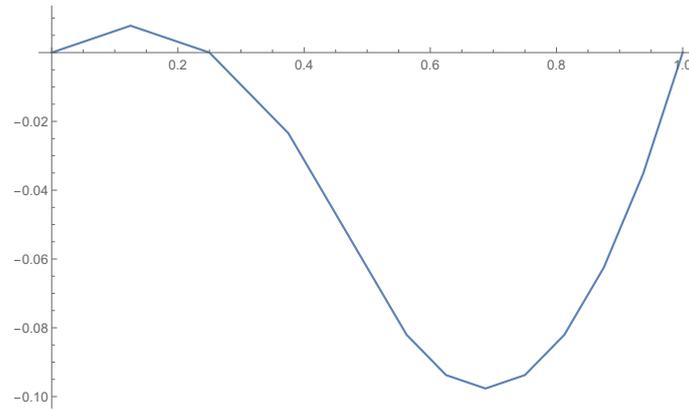


Abbildung 5.5: Näherungslösung bei einer Verfeinerung von 12 Elementen

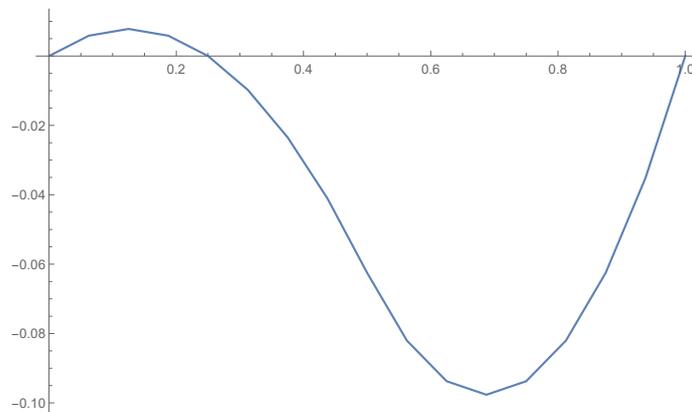


Abbildung 5.6: Näherungslösung bei einer Verfeinerung von 16 Elementen

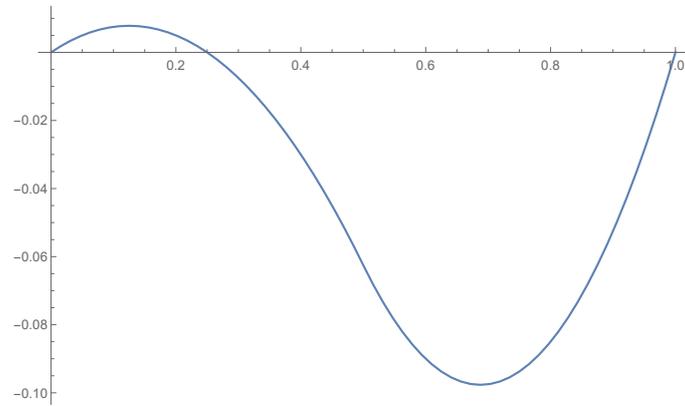


Abbildung 5.7: Näherungslösung bei einer Verfeinerung von 36864 Elementen

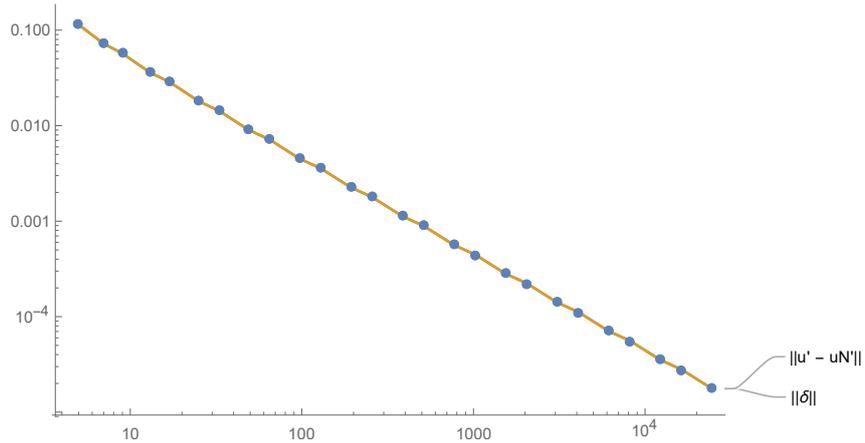


Abbildung 5.8: Vergleich zwischen dem Fehler des äquilibrierten Flusses und dem exakten Fehler.

Kapitel 6

Konklusion

Das Ziel dieser Bachelorarbeit war es, einen äquilibrierten a-posteriori Fehlerschätzer der Poisson-Gleichung zu konstruieren, sodass eine konstant-freie Fehlerabschätzung in der Energienorm erreicht wird. Durch die Implementierung des C++ Programmes, konnte dieser Fehlerschätzer im eindimensionalen Raum realisiert werden.

Beginnend mit dem Kapitel der Grundlagen, wurde die Basis für die gesamte Arbeit aufgebaut. Das klassische Modellproblem der Poisson-Gleichung mit homogenen Dirichlet-Randbedingungen wurde eingeführt und dessen Variationsformulierung wurde hergeleitet. Das Lemma von Lax-Milgram garantierte die Eindeutigkeit und Existenz einer Lösung des abstrakten Variationsproblems der Poisson-Gleichung. Durch die Einführung der Galerkin-Methode, konnte das Problem in ein diskretes Variationsproblem umgeschrieben und dadurch auf ein äquivalentes Gleichungssystem umgeformt werden. Genauso wie für das abstrakte Variationsproblem, konnte eine eindeutige Lösung des diskreten Problems durch das Lemma von Lax-Milgram gewährleistet werden. Die offene Frage dieses Kapitels beschäftigte sich mit der Größe der Fehlerordnung zwischen der exakten Lösung und der Näherungslösung.

Im darauffolgenden Kapitel, wurde der Aufbau eines adaptiven Verfahrens vorgestellt. Dessen Ziel war es, nur jene Stellen des Gebiets, mit großem Fehlerbeitrag zu verfeinern. Dazu eignete sich der residuale Fehlerschätzer als Fehlerindikator, der auf die Berechnung des Residuums beruht. Für diesen Fehlerschätzer gelten die wesentlichen zwei Eigenschaften, die Effizienz und die Zulässigkeit. Die Zuverlässigkeit gab eine obere Schranke der Fehlerabschätzung vor. Im weiteren Verlauf dieses Kapitels, wurden sowohl zwei Markierungsstrategien, als auch zwei Verfeinerungsstrategien vorgestellt, wobei beide Strategien im Algorithmus des adaptiven Verfahrens einsetzbar sind.

Im letzten Kapitel wurde das duale gemischte Variationsproblem eingeführt. Der Satz von Brezzi garantierte eine eindeutige Lösung dieses Problems, welcher die starken Bedingungen, die LBB-Bedingung und die $\ker B$ Bedingung, an den Bilinearformen voraussetzte. Die Einführung des gemischten Problems ermöglichte die Anwendung des zentralen Satzes dieses Kapitels, der sogenannte Satz von Prager-Synge. Mithilfe dieses Satzes konnte eine konstant-freie Fehlerabschätzung gebildet werden. Die Anwendung von Flusskorrekturen ermöglichte eine einfache Berechnung des Fehlerindikators.

Abschließend wurde der äquilibrierte Fehlerschätzer mit einem C++ Programm im eindimensionalen Raum realisiert. Der berechenbare Fehlerschätzer basiert auf die in Kapitel 4 vorgestellten Flusskorrekturen.

Literaturverzeichnis

- [1] Christian Abfalg, Pelin Ekmekci, Christian Kaps. *Adaptive Gitter-Verfeinerung*. Universität Ulm, 2015
- [2] Dietrich Braess. *Finite Elemente: Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*. Springer-Verlag Berlin Heidelberg, 2007
- [3] Lawrence C. Evans. *Partial Differential Equations*. Graduate Studies in Mathematics, American Mathematical Society, 1998
- [4] Robert Gruhlke. *Äquilibrierter Fehlerschätzer für elliptische PDE's*. Bachelorarbeit von Technische Universität Berlin, 2013
- [5] Prof. Dr. R. Herzog. *Numerik partieller Differentialgleichungen - A-posteriori-Fehlerabschätzungen*. Technische Universität Chemnitz, 2014
- [6] Tobias Klas. *Hauptseminar Wissenschaftliches Rechnen - Tetraeder-Gitter-Verfeinerung*. Rheinische Friedrich-Wilhelms-Universität Bonn, 2017
- [7] Martin Neumüller. *Vorlesung Numerik partieller Differentialgleichungen*. Johannes Kepler Universität Linz, 2017
- [8] L. Ridgway Scott and Shangyou Zhang. *Finite element interpolation of nonsmooth functions satisfying boundary conditions*. Mathematics of Computation, 1990
- [9] T. Suttmeier. *Lösung von PDGL mit der Finite Elemente Methode*. Universität Siegen, 2006/07
- [10] Rüdiger Verfürth. *A Posteriori Error Estimation Techniques for Finite Element Methods*. Oxford University Press, 2013

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Bachelorarbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.
Die vorliegende Bachelorarbeit ist mit dem elektronisch übermittelten Textdokument identisch.

Linz, September 2018

Mario Gobrial