



universität  
wien

# MASTERARBEIT

Titel der Masterarbeit

Schätzerreduktion und Konvergenz adaptiver FEM für  
Hindernisprobleme

Verfasser

Marcus Page B.Sc.

angestrebter akademischer Grad  
Master of Science (MSc.)

Wien, im Mai 2010

Studienkennzahl lt. Studienblatt: A 066 821

Studienrichtung lt. Studienblatt: Mathematik

Betreuer: Ao.Univ.Prof. Dipl.Math. Dr.techn. Dirk Praetorius



# Danksagung

Eine Arbeit wie die vorliegende ist ohne die Hilfe weiterer Personen, wenn überhaupt, nur schwer möglich. Aus diesem Grund bin ich sehr dankbar, dass ich stets in den Genuss solcher Hilfe gekommen bin. Allen voran möchte ich mich hierfür bei meinem Betreuer Dirk Praetorius bedanken, der mir jederzeit mit hilfreichen Ratschlägen zur Seite stand und sicher auch selbst viele Stunden mit dieser Arbeit verbracht hat. Außerdem hat er mir, durch die Ressourcen des Institutes für Analysis und Scientific Computing der TU Wien, genügend Rechenleistung zur Verfügung gestellt, um die illustrativen Beispiele zu erzeugen.

Bedanken möchte ich mich auch bei Freunden und Mitstudenten, die ihre Zeit zum Korrekturlesen aufgewendet haben. Aus eigener Erfahrung weiß ich selbst, dass es ziemlich viel Aufwand ist, die Arbeit von jemand anderem Korrektur zu lesen. Dafür also nochmal ein großes Dankeschön.

Zu guter letzt geht ein Riesendank zu meinen Eltern ins schöne Weiterstadt. Danke für die jahrelange Unterstützung in allen möglichen Lebenslagen, die sicherlich nicht immer die einfachste war – und die natürlich auch nach dem Mathematikstudium nicht aufhören soll :-). Ihr seid klasse!



# Zusammenfassung

Diese Arbeit beschäftigt sich mit der Konstruktion adaptiver Algorithmen zur Lösung von Hindernisproblemen durch die Methode der finiten Elemente. Die Grundidee, auf der alle Ausführungen aufbauen, ist hierbei die Verallgemeinerung des Prinzips der *Schätzerreduktion* auf solche (nichtlineare) Problemstellungen. Hierdurch gelingt es, bestehende Resultate zu erweitern und zu verbessern. Eine in Konvergenzbeweisen adaptiver finite-Elemente-Methoden oft verwendete Eigenschaft ist die *diskrete lokale Effizienz* des zu Grunde liegenden Fehlerschätzers. Diese hängt jedoch stark von der *innerer-Punkt-Eigenschaft* ab, welche fordert, dass im Inneren eines verfeinerten Elements mindestens ein neuer Knoten entsteht. Durch diese Forderung sind sämtliche Konvergenzresultate von der Strategie der lokalen Netzverfeinerung abhängig und somit nicht allgemein gültig. Mit der Methode der Schätzerreduktion aus [CKNS, AFP] ist es möglich, die diskrete lokale Effizienz zu umgehen und daher von der Verfeinerung weitestgehend unabhängige Konvergenzbeweise zu erzielen. In der vorliegenden Arbeit wird genau dieser, aus dem linearen Fall bekannte, Ansatz verfolgt, um die Resultate aus [BCH1] und [BCH2] entsprechend zu verallgemeinern. Als Modellproblem dient durchgängig die zweidimensionale *Poisson-Gleichung*. Erweiterungen auf den dreidimensionalen Fall sind jedoch problemlos möglich. Im Einzelnen ist die Arbeit wie folgt gegliedert:

- **Kapitel 1** gibt eine Einführung in die Thematik, in der auch die physikalische Interpretation der betrachteten Hindernisprobleme erläutert wird.
- In **Kapitel 2** wird die Methode der finiten Elemente (FEM) ausführlich erläutert. Hier wird zudem ein Großteil der verwendeten Notation definiert, und es werden einige grundlegende Beweisstrategien bereits im linearen Fall verdeutlicht. Das für den Rest der Arbeit wichtigste Resultat ist hierbei Satz 2.2.6, in dem das Prinzip der Schätzerreduktion für den linearen Fall gezeigt wird.
- Die Verallgemeinerung der Schätzerreduktion auf Hindernisprobleme, bei denen das Hindernis durch eine global affine Funktion gegeben ist, erfolgt in **Kapitel 3**. Wie angekündigt, wird ein adaptiver FEM-Algorithmus und ein entsprechender Fehlerschätzer für dieses Problem angegeben. Die Konvergenz gegen die exakte Lösung (Satz 3.3.9) wird durch Schätzerreduktion (Satz 3.3.3) bewiesen, so dass diese von der Verfeinerungsstrategie unabhängig ist. Außerdem werden durch dieses Vorgehen einige Anforderungen an die Oszillationsterme unnötig. Um die Vorteile dieser Herangehensweise vollends nachzuziehen, sind zum Vergleich auch die ursprünglichen Ideen aus [BCH1] angegeben.
- In **Kapitel 4** findet schließlich eine weitere Verallgemeinerung auf Hindernisprobleme mit Hindernissen aus  $H_0^1(\Omega)$  statt. Wie zuvor werden die Ausführungen aus [BCH2], welche die diskrete lokale Effizienz noch verwenden, vorgestellt. Im Anschluss daran werden die Resultate wieder durch Schätzerreduktion verallgemeinert und ein konvergenter adaptiver Algorithmus angegeben. Für die Konvergenzanalyse (Satz 4.4.6 – 4.4.8) sind jedoch einige Voraussetzungen an die Veränderung der aktiven und inaktiven Teilbereiche notwendig. Hierdurch werden die Grenzen der Schätzerreduktion deutlich, die

---

zuvor noch nicht abzusehen waren. Sowohl Kapitel 3 als auch Kapitel 4 werden zudem durch numerische Illustrationen abgerundet.

- **Kapitel 5** gibt ein Fazit über erreichte Resultate und einen Ausblick auf sinnvolle Erweiterungen dieser Arbeit.
- Eine vollständige Implementierung der gezeigten Beispiele findet sich im **Anhang**. Außerdem ist hier eine Liste der verwendeten Notationen, sowie ein wenig mathematische Hintergrundinformation angegeben.

# Abstract

This thesis deals with the construction of adaptive algorithms that solve elliptic obstacle problems by means of the finite element method. Throughout all chapters, the main idea to do so is the generalization of the so-called *estimator reduction* to such (nonlinear) problems. In this way, it is possible to widen and improve known results. One of the ingredients frequently used in the convergence analysis of adaptive finite element methods is the *discrete local efficiency* of the underlying error estimator. This feature, however, strongly relies on the *interior-node-property* which basically states that at least one new node has to be created within the interior of each refined element. All convergence results exploiting the discrete local efficiency thus depend on the local refinement strategy and are therefore non-universal. By means of the method of estimator reduction from [CKNS, AFP] it is possible though, to circumvent the need for this property and thus achieve convergence results that are independent of the local mesh refinement strategy. In this thesis, this ansatz, which is known from the linear case, is pursued in order to generalize the results from [BCH1] and [BCH2]. Throughout the whole work, the two dimensional *Poisson-equation* serves as model problem. Generalizations to the three-dimensional case are, however, easily possible. In full detail, the work is organized as follows:

- **Chapter 1** gives an introduction to the problem and comments on some physical interpretation of the obstacle problems in question.
- In **Chapter 2**, the finite element method (FEM) is explained in full detail. The better part of the used notation as well as some basic ideas used in the upcoming proofs are also illustrated in this linear setting. The main result within this chapter is Theorem 2.2.6 which states the principle of estimator reduction for the linear case.
- A generalization of the estimator reduction to obstacle problems in which the obstacle is globally affine, is given in **Chapter 3**. As mentioned above, an adaptive FEM-algorithm and a corresponding a posteriori error estimator is given. The convergence towards the exact solution (Theorem 3.3.9) is shown via estimator reduction (Theorem 3.3.3) and is therefore independent of the local mesh refinement strategy. By utilizing this approach, some postulations on the decay of oscillation data from [BCH1] additionally become unnecessary. In order to fully appreciate the advantages of this ansatz, the original ideas from [BCH1] are elaborated as well.
- In **Chapter 4** the analysis is widened to include obstacles from  $H_0^1(\Omega)$ . As before, the results from [BCH2], which make use of the discrete local efficiency, are elaborated. In the following, the results are again generalized by estimator reduction and a convergent adaptive algorithm with a corresponding error estimator is given. During the convergence analysis (Theorem 4.4.6 – 4.4.8), however, the need for some additional assumptions on the change of active and inactive regions, arises. Certain limits of the principle of estimator reduction, that could not be foreseen within the linear setting,

---

can therefore be glimpsed. In addition, Chapter 3, as well as Chapter 4 are completed by some numerical illustrations.

- **Chapter 5** gives a conclusion about achieved results and some outlook on possible future work.
- Finally, a complete MATLAB implementation of the numerical examples is given in the **appendix (Anhang)**. A list of notations as well as some mathematical background information can be found here, too.

# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>1</b>
1.1. Einordnung des Problems . . . . .	2
1.2. Beitrag dieser Arbeit zum Themenbereich . . . . .	4
<b>2. Die P1 Finite-Elemente-Methode</b>	<b>5</b>
2.1. Allgemeines . . . . .	5
2.1.1. Problembeschreibung und schwache Formulierung . . . . .	5
2.1.2. Diskretisierung . . . . .	8
2.1.3. A Priori Fehleranalyse . . . . .	12
2.2. Adaptive Verfeinerung . . . . .	15
2.2.1. Fehlerschätzer . . . . .	15
2.2.2. Adaptiver Algorithmus . . . . .	21
2.2.3. Verfeinerungsstrategien . . . . .	23
2.2.4. Konvergenz der AFEM und Schätzerreduktion . . . . .	27
2.2.5. Optimalität der AFEM . . . . .	30
<b>3. Das affine Hindernisproblem</b>	<b>33</b>
3.1. Modellproblem . . . . .	33
3.2. Ein erster konvergenter Algorithmus . . . . .	37
3.2.1. Zuverlässigkeit . . . . .	40
3.2.2. Diskrete lokale Effizienz . . . . .	43
3.2.3. Fehlerreduktion - Ein Konvergenzsatz . . . . .	49
3.3. Schätzerreduktion für das Hindernisproblem . . . . .	51
3.3.1. Fehlerschätzer und adaptiver Algorithmus . . . . .	51
3.3.2. Schätzerreduktion und Konvergenzanalyse . . . . .	53
3.3.3. Eine Erweiterung . . . . .	72
3.4. Numerische Illustrationen . . . . .	73
<b>4. Das nicht-affine Hindernisproblem</b>	<b>81</b>
4.1. Problembeschreibung und Notationen . . . . .	81
4.2. Fehlerschätzer und adaptiver Algorithmus . . . . .	85
4.3. Konvergenzanalyse . . . . .	87
4.3.1. Zuverlässigkeit . . . . .	88
4.3.2. Diskrete lokale Effizienz . . . . .	93
4.3.3. Gestörte Galerkin-Orthogonalität und Fehlerreduktion . . . . .	99
4.4. Schätzerreduktion . . . . .	101
4.5. Numerische Illustrationen . . . . .	107
<b>5. Zusammenfassung und Ausblick</b>	<b>113</b>

<b>A. Notationen</b>	<b>115</b>
<b>B. Mathematische Wahrheiten</b>	<b>117</b>
B.1. Funktionalanalytische Voraussetzungen . . . . .	117
B.2. Theorem aus [BC] . . . . .	119
<b>C. Implementierung</b>	<b>121</b>
<b>Literaturverzeichnis</b>	<b>137</b>
<b>Lebenslauf des Autors</b>	<b>139</b>

# Kapitel 1.

## Einleitung

Wie können wir die Welt beschreiben, in der wir leben? Ein nicht geringer Anteil der mathematischen Gesellschaft würde diese Frage mit *partiellen Differentialgleichungen* beantworten. Mit Hilfe dieser Gleichungen gelingt es uns, ein breites Spektrum von physikalischen, biologischen oder chemischen Vorgängen zu modellieren. Betrachten wir zum Beispiel eine auf einem quadratischen Gebiet im  $\mathbb{R}^2$  eingespannte, dehnbare Membran, wie etwa das Trommelfell. Wir interessieren uns nun für eine Funktion  $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ , deren Werte  $u(x)$  die Auslenkung selbiger Membran unter Einwirkung einer Kraft  $f$  angeben. Gehen wir vereinfachend davon aus, dass  $f$  nur in vertikaler Richtung wirkt, so ergibt sich durch Ausnutzung einfacher physikalischer Zusammenhänge, dass das *Energiefunktional*, also die potentielle Gesamtenergie der ausgelenkten Membran, durch

$$\mathcal{J}(u) := \frac{1}{2} \|\nabla u\|^2 - (f, u)$$

gegeben ist. Die Größen  $\|\cdot\|$  und  $(\cdot, \cdot)$  bezeichnen hierbei Norm und Skalarprodukt in einem geeigneten Hilbertraum. Wie wir ziemlich bald sehen werden ist die Funktion  $u$ , welche obiges Energiefunktional minimiert, also die Auslenkung der Membran im Ruhezustand angibt, (eindeutige) Lösung einer besonderen partiellen Differentialgleichung, der *Poisson-Gleichung*

$$-\Delta u = f,$$

auf dem Berechnungsgebiet  $\Omega \subseteq \mathbb{R}^2$ . Die Tatsache, dass die Membran am Rand eingespannt ist, wird durch die zusätzliche Forderung  $u|_{\partial\Omega} = 0$  modelliert. Ein Blick auf die Lösung dieser Gleichung (Abbildung 1.1) lässt hoffen, dass dadurch tatsächlich der entsprechende physikalische Vorgang beschrieben wird. Dieselbe Gleichung modelliert übrigens auch elektrostatische und astrophysikalische Phänomene, so dass die Berechnung von Lösungen in vielen Bereichen Gegenstand aktueller Forschung ist.

Problematisch ist jedoch, dass die allermeisten partiellen Differentialgleichungen nicht ohne Weiteres exakt zu lösen sind. Man ist daher gezwungen, auf numerische Methoden zurückzugreifen, die die Werte der Lösung näherungsweise berechnen. Die mathematische Aufgabe besteht nun darin, sinnvolle Verfahren zu entwickeln, deren Ausgabe (die angenäherte Lösung), unabhängig von den Daten, immer gegen die gesuchte Funktion (die exakte Lösung) konvergiert, sofern man nur lange genug rechnet. Diese Verfahren, derer es eine wahre Vielfalt gibt und die unterschiedlicher nicht sein könnten, sind meist auf einen gewissen Typ von Gleichung zugeschnitten. Aufgrund der Vielseitigkeit und Unterschiedlichkeit der natürlich auftretenden Phänomene und daher auch der zu untersuchenden Gleichungen gibt es also kein universell anwendbares Lösungsverfahren. Eine für *elliptische Differentialgleichungen*, in deren Klasse auch die Poisson-Gleichung fällt, gern und oft verwendete Prozedur ist die *Methode der finiten Elemente (FEM)*, die auch in der vorliegenden Arbeit behandelt wird. Dieses Verfahren,

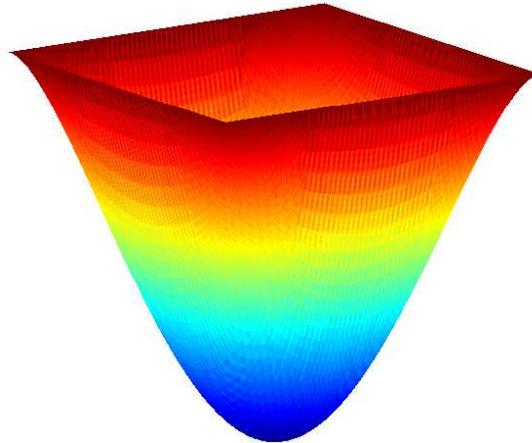


Abbildung 1.1.: Auslenkung einer am Rand eingespannten Membran unter Einwirkung einer vertikal wirkenden Kraft  $f$ .

welches mathematisch in die Klasse der *Galerkin-Methoden* fällt, erfreut sich etwa seit den 1970er Jahren, besonders im Ingenieursbereich, immer stärkerer Beliebtheit. Zur Berechnung der Lösung wird hierbei das Gebiet in ein Gitter aus endlich vielen Teilgebieten zerteilt, und die Lösung an jedem Gitterpunkt approximativ berechnet. Durch kontinuierliche Verfeinerung des zu Grunde liegenden Gitters folgt nun, unter gewissen Voraussetzungen, die Konvergenz gegen die exakte Lösung (vgl. [BRA07, Kapitel 2, §6]).

Da dies jedoch relativ lange dauern kann, ist es erstrebenswert, das Gitter, abhängig von der Lösung, durch Einsatz eines so genannten *a posteriori Fehlerschätzers* an gewissen Stellen stärker zu verfeinern als an anderen. Die Frage, wie dies ohne Kenntnis der exakten Lösung zu bewerkstelligen ist, ist hierbei nur eines der auftretenden Probleme. Im Laufe der Zeit wurden diese *adaptiven Verfahren* intensiv untersucht, so dass sie nun in Forschungs-, wie in Ingenieursbereichen breitgefächert Anwendung finden. Als Pionierleistungen sind hierbei sicherlich die Arbeiten [D] von DÖRFLER und [MNS] von MORIN, NOCHETTO und SIEBERT zu nennen, in denen erstmals Konvergenz adaptiver FEM nachgewiesen wurde. Grundbausteine der Beweise waren damals zwei Eigenschaften des zur lokalen Netzverfeinerung verwendeten Fehlerschätzers, die so genannte *Zuverlässigkeit* und die *diskrete lokale Effizienz*. Den entsprechenden Beweisen lagen jedoch noch starke Einschränkungen an die Netzverfeinerung zu Grunde. Im Jahre 2008 wurde von CASCON, KREUZER, NOCHETTO und SIEBERT in [CKNS] ein allgemeineres Konvergenzresultat erzielt. In derselben Arbeit wurde, unter zusätzlichen Voraussetzungen, sogar gezeigt, dass die adaptive Finite Elemente Methode immer mit optimaler Konvergenzrate konvergiert. Wir werden die genannten Beweise und Einschränkungen im Laufe von Kapitel 2 noch näher untersuchen.

## 1.1. Einordnung des Problems

Was passiert nun aber, wenn wir unsere Aufgabenstellung etwas verkomplizieren? Um bei unserem Beispiel der eingespannten Membran zu bleiben, interessieren wir uns nicht mehr

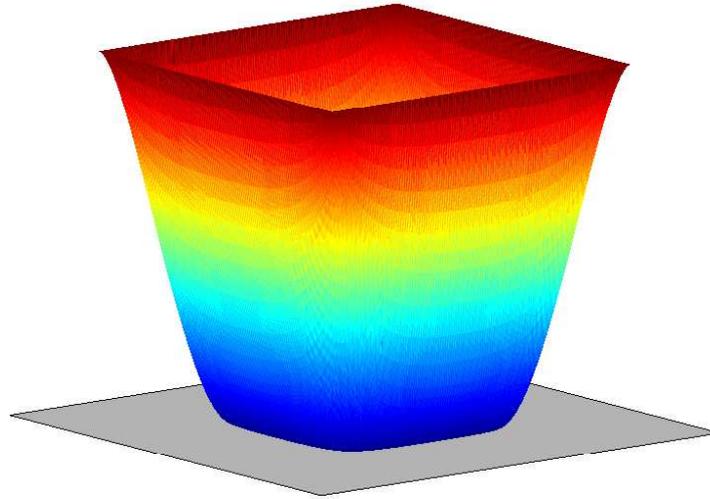


Abbildung 1.2.: Am Rand eingespannte Membran deren Auslenkung durch ein Hindernis begrenzt wird.

nur für die Lösung mit minimaler Energie. Vielmehr verlangen wir, dass die Lösung gewissen Zusatzanforderungen gerecht wird. Eine Möglichkeit wäre zum Beispiel zu fordern – und das ist Gegenstand dieser Arbeit – dass die Lösung größer oder kleiner als eine gegebene Schranke ist. Beispielhaft stellen wir uns wieder die eingespannte Membran vor, die aber nicht frei durchhängen kann, sondern von einem *Hindernis* begrenzt wird (Abbildung 1.2). Mathematisch suchen wir also nicht nach dem globalen Minimum des Energiefunktionals, sondern nach einer Funktion  $u$ , die

$$\mathcal{J}(u) = \min_{v \in K} \mathcal{J}(v)$$

erfüllt. Wir suchen das Minimum also nur innerhalb einer gewissen Menge  $K$  von *zulässigen Funktionen*. Es ist relativ leicht einzusehen ([KS]), dass dieses *elliptische Hindernisproblem*, unter gewissen Voraussetzungen an die Eingabedaten, auch eine eindeutige Lösung hat. Wir sind im Folgenden jedoch an der numerischen Behandlung dieses Problems durch die adaptive Finite-Elemente-Methode interessiert. Mathematisch führen solche Problemstellungen auf (nichtlineare) *Variationsungleichungen*, für welche die adaptive FEM erst seit kurzem angewendet wird. Die Nichtlinearität des Problems fordert hierbei den Einsatz neuer Fehlerschätzer und Beweistechniken, so dass es bislang verhältnismäßig wenig rigorose Konvergenzresultate gibt. Allen voran ist hier die Arbeit von BRAESS, CARSTENSEN und HOPPE ([BCH1]) zu nennen, in der erstmals ein konvergenter FEM-Algorithmus für Hindernisprobleme konstruiert wurde. Die Analyse war in diesem Fall jedoch auf Hindernisse, die durch global affine Funktionen  $\chi$  gegeben sind, beschränkt. Außerdem verwendet der Konvergenzbeweis die *diskrete lokale Effizienz* des Fehlerschätzers, welche an die so genannte *innerer-Punkt-Eigenschaft* gekoppelt ist. Die Verwendung dieser Eigenschaft, die eine starke Anforderung an die Art der Netzverfeinerung darstellt, schränkt somit die Gültigkeit des Konvergenzbeweises auf konkrete Verfeinerungen ein. In der Folgearbeit [BCH2] wird die Analyse schließlich auf den Fall allgemeinerer Hindernisse  $\chi \in H_0^1(\Omega)$  ausgeweitet. Auch in diesem Fall bleibt die

Konvergenz jedoch stark von der Verfeinerungsstrategie abhängig. Außerdem werden noch einige Forderungen an das Abklingen so genannter *Oszillationsterme* gestellt, die zusätzlich garantiert werden müssen.

## 1.2. Beitrag dieser Arbeit zum Themenbereich

Wir orientieren uns in unseren Ausführungen an den Arbeiten [BCH1] und [BCH2], wobei wir ein im linearen Fall verwendetes Beweiskonzept, die so genannte *Schätzerreduktion*, auf elliptische Hindernisprobleme erweitern und so die diskrete lokale Effizienz umgehen. In Kapitel 3 betrachten auch wir zunächst den Fall global affiner Hindernisse, wobei es uns durch Schätzerreduktion gelingt, ein Konvergenzresultat zu erzielen, welches von der Art der lokalen Netzverfeinerung weitestgehend unabhängig ist. Ferner gelingt es uns, durch geschickte Konstruktion des Fehlerschätzers, die zusätzlichen Anforderungen an die Datenoszillationen überflüssig zu machen. Schließlich zeigen wir im Sinne von [CKNS] die Kontraktivität einer *gemischten Fehlergröße*, die Ausgangspunkt für Überlegungen zur Optimalität sein könnte.

Analog zu [BCH2] weiten auch wir unsere Analysen in Kapitel 4 auf den Fall allgemeinerer Hindernisse aus. Wie im affinen Fall ist es unser Ziel, die diskrete lokale Effizienz durch Schätzerreduktion zu umgehen. Dadurch gelingt es uns, unter einigen milden Voraussetzungen, Konvergenzresultate für weitestgehend beliebige Verfeinerungsstrategien zu erzielen. Im Zuge dessen treten jedoch zusätzliche Probleme auf, die auf die Grenzen des Prinzips der Schätzerreduktion hinweisen, im linearen Fall jedoch noch nicht abzusehen waren.

# Kapitel 2.

## Die P1 Finite-Elemente-Methode

### 2.1. Allgemeines

Wie in der Einleitung erwähnt, ist die Methode der Finiten Elemente (FEM) ein numerisches Verfahren zur näherungsweise Berechnung der Lösung elliptischer Randwertprobleme. Obwohl sich diese Methode seit Jahren starker Beliebtheit erfreut, sind rigorose Konvergenzresultate, insbesondere für adaptive Verfahren, vergleichsweise jung. Viele interessante Fragen und Probleme sind sogar mathematisch noch unbehandelt. Wir wollen uns in diesem Kapitel das Prinzip der FEM verdeutlichen und ihre wichtigsten Eigenschaften zeigen. Hierzu beschäftigen wir uns der Einfachheit halber mit dem einfachsten elliptischen Randwertproblem, der *Poissongleichung*, deren physikalische Bedeutung in der Einleitung bereits dargelegt wurde.

#### 2.1.1. Problembeschreibung und schwache Formulierung

In unseren Ausführungen schränken wir uns stets auf ein beschränktes, polygonal berandetes Gebiet  $\Omega \subset \mathbb{R}^2$  ein. Den Rand  $\partial\Omega$  von  $\Omega$  bezeichnen wir mit  $\Gamma$ . Für  $f \in L^2(\Omega)$  suchen wir eine Lösung  $u$  für das *homogene Dirichletproblem*

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega, \\ u &= 0 && \text{auf } \Gamma. \end{aligned} \tag{2.1.1}$$

Da es sich aufgrund der funktionalanalytischen Eigenschaften als sinnvoller herausgestellt hat, gehen wir durch Multiplikation mit einer *Testfunktion*  $v \in H_0^1(\Omega)$  und partieller Integration zur *variationellen Formulierung* über, bei der wir ein  $u \in H_0^1(\Omega)$  suchen, so dass

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx \quad \text{für alle } v \in H_0^1(\Omega)$$

gilt. Wie in der Literatur üblich, bezeichnet  $H_0^1(\Omega)$  hierbei diejenigen Funktionen  $w \in L^2(\Omega)$ , deren schwacher Gradient elementweise wieder in  $L^2(\Omega)$  ist und für die zusätzlich  $w|_{\Gamma} = 0$  im Spürsinne gilt. Den Dualraum von  $H_0^1(\Omega)$  notieren wir mit  $H^{-1}(\Omega)$ . Es ist hinreichend bekannt, dass  $H_0^1(\Omega)$  mit der Norm  $\|w\|_{H_0^1(\Omega)} := (\|w\|_{L^2(\Omega)}^2 + \|\nabla w\|_{L^2(\Omega)}^2)^{1/2}$  ein Hilbertraum ist. Mit der Ungleichung von Poincaré (B.1.1) ist auf  $H_0^1(\Omega)$  außerdem durch  $\|w\|_{(\Omega)} := \|\nabla w\|_{L^2(\Omega)}$  eine äquivalente Norm, die *Energienorm*, gegeben. Wir bezeichnen im Folgenden das  $L^2$ -Skalarprodukt auf  $\Omega$  mit

$$(u, v) := \int_{\Omega} uv \, dx$$

und das *Energieskalarprodukt* auf  $H_0^1(\Omega)$  mit

$$\langle\langle u, v \rangle\rangle := \int_{\Omega} \nabla u \cdot \nabla v \, dx.$$

Betrachten wir diese Größen nicht auf dem gesamten Gebiet  $\Omega$ , sondern nur auf einer Teilmenge  $T$ , so wird dies durch  $(\cdot, \cdot)_T$  bzw.  $\langle\langle \cdot, \cdot \rangle\rangle_T$  notiert. Wie schon aus dem Namen hervorgeht, ist die Energienorm  $\|w\|^2 = \langle\langle w, w \rangle\rangle$  die durch das Energieskalarprodukt induzierte Norm auf  $H_0^1(\Omega)$ . Einsetzen dieser Definitionen liefert die *schwache Formulierung* des Randwertproblems (2.1.1):

**Problem 2.1.1.** *Finde  $u \in H_0^1(\Omega)$ , mit*

$$\langle\langle u, v \rangle\rangle = (f, v) \quad \text{für alle } v \in H_0^1(\Omega). \quad (2.1.2)$$

Dies ist die Problemstellung mit der wir uns im Folgenden beschäftigen wollen. Zunächst leiten wir uns die Existenz einer eindeutigen Lösung her. Hierzu verwenden wir den Satz von Lax-Milgram.

**Satz 2.1.2** (Lax-Milgram). *Sei  $(H, \|\cdot\|_H)$  ein reellwertiger Hilbertraum und  $B : H \times H \rightarrow \mathbb{R}$  eine beschränkte, elliptische Bilinearform, d.h. es existieren Konstanten  $\alpha, \beta > 0$  so, dass*

(i)

$$|B[u, v]| \leq \alpha \|u\|_H \|v\|_H \quad \text{für alle } u, v \in H$$

und

(ii)

$$\beta \|u\|_H^2 \leq B[u, u] \quad \text{für alle } u \in H.$$

Sei ferner  $f : H \rightarrow \mathbb{R}$  ein beschränktes, lineares Funktional auf  $H$ . Dann existiert ein eindeutiges  $u \in H$  mit

$$B[u, v] = \langle f, v \rangle \quad \text{für alle } v \in H,$$

wobei  $\langle \cdot, \cdot \rangle$  die duale Paarung zwischen  $H$  und  $H'$  bezeichnet.

*Beweis.* Siehe EVANS [E, §6.2.1, Theorem 1]. □

**Korollar 2.1.3.** *Für jedes  $f \in H^{-1}(\Omega)$  existiert ein eindeutiges  $u \in H_0^1(\Omega)$  welches Problem 2.1.1 löst.*

*Beweis.* Per Definition ist  $f \in H^{-1}(\Omega)$  ein beschränktes lineares Funktional auf  $H_0^1(\Omega)$ . Ferner ist das Energieskalarprodukt  $\langle\langle \cdot, \cdot \rangle\rangle$  als Skalarprodukt offensichtlich auch eine Bilinearform auf  $H_0^1(\Omega)$ . Es bleibt zu zeigen, dass  $\langle\langle \cdot, \cdot \rangle\rangle$  elliptisch und beschränkt ist.

Beschränktheit:

$$\langle\langle u, v \rangle\rangle = \int_{\Omega} \nabla u \cdot \nabla v \, dx \leq \|u\|_{\Omega} \|v\|_{\Omega} \leq \|u\|_{H_0^1(\Omega)} \|v\|_{H_0^1(\Omega)}.$$

Elliptizität:

$$\langle u, u \rangle = \int_{\Omega} (\nabla u)^2 dx = \|u\|_{\Omega}^2 \geq C_{\text{Poincaré}}^{-1} \|u\|_{H_0^1(\Omega)}^2,$$

wobei wir hier die Ungleichungen von Hölder und Poincaré verwendet haben. Damit folgt die Existenz einer eindeutigen Lösung aus dem Satz von Lax-Milgram.  $\square$

**Bemerkung.** Da der Laplace-Operator in der schwachen Formulierung sogar eine symmetrische Bilinearform, und damit ein äquivalentes Skalarprodukt  $\langle \cdot, \cdot \rangle$ , erzeugt, würden Existenz und Eindeutigkeit der Lösung in diesem Fall sogar direkt aus dem Darstellungssatz von Riesz (B.1.5) folgen. Es ist daher in unserem Fall im Prinzip obsolet, Elliptizität und Beschränktheit zu zeigen. Im Falle eines allgemeineren elliptischen Differentialoperators ist dies allerdings keine Option mehr. Man muss dann zwingend über den Satz von Lax-Milgram argumentieren. Wir haben uns daher entschlossen auch den leichteren Fall durch das allgemeine Vorgehen zu beweisen.

Nach diesen Überlegungen können wir Folgendes festhalten:

- Jede starke Lösung  $u \in C^2(\overline{\Omega})$  des ursprünglichen Problems ist auch eine Lösung der schwachen Formulierung (2.1.2).
- Ist die (eindeutige) Lösung  $u \in H_0^1(\Omega)$  der schwachen Formulierung glatt genug, d.h.  $u \in C^2(\overline{\Omega})$ , so ist sie auch eine starke Lösung von (2.1.1).

Die Lösung der schwachen Formulierung ist somit der einzige sinnvolle Kandidat für eine Lösung von (2.1.1). Das Originalproblem wird im Allgemeinen keine starke Lösung haben. Der Vollständigkeit halber sei hier angemerkt, dass dies mit der Glattheit des Randes  $\Gamma$  zusammenhängt. Der interessierte Leser sei auf [E, §6.3] verwiesen. Aufgrund dieser Tatsache beschäftigt sich die Finite-Elemente-Methode tatsächlich mit einer numerischen Lösung der schwachen Formulierung und zielt nicht in erster Linie darauf ab, das ursprüngliche Problem zu lösen.

Zum Abschluss dieses Abschnitts wollen noch eine zur schwachen Formulierung (2.1.2) äquivalente Darstellung beschreiben, die vor allem in Kapitel 3 Anwendung findet.

**Proposition 2.1.4.** *Das Energiefunktional*

$$\mathcal{J}(v) = \frac{1}{2} \langle v, v \rangle - (f, v) \quad (2.1.3)$$

nimmt sein (eindeutiges) Minimum  $u$  über  $H_0^1(\Omega)$  genau dann an, wenn  $\langle u, v \rangle = (f, v)$  für alle  $v \in H_0^1(\Omega)$  gilt.

*Beweis.* Wir zeigen beide Implikationen einzeln. Sei  $\langle u, v \rangle = (f, v)$  für alle  $v \in H_0^1(\Omega)$ . Für alle  $t \in \mathbb{R}$  gilt

$$\mathcal{J}(u + tv) = \frac{1}{2} \langle u + tv, u + tv \rangle - (f, u + tv) = \mathcal{J}(u) + \underbrace{t(\langle u, v \rangle - (f, v))}_{=0} + \frac{t^2}{2} \langle v, v \rangle.$$

$t = 1$  liefert nun

$$\mathcal{J}(u + v) = \mathcal{J}(u) + \frac{1}{2} \langle v, v \rangle > \mathcal{J}(u) \quad (2.1.4)$$

für alle  $v \neq 0$ , wegen der Positivität von  $\langle \cdot, \cdot \rangle$ . Die Minimierungsaufgabe wird daher durch  $u$  gelöst.

Sei nun  $u$  das Minimum von (2.1.3) über  $H_0^1(\Omega)$ . Dann gilt  $\frac{d}{dt}\mathcal{J}(u + tv)|_{t=0} = 0$ . Nun ist

$$0 = \frac{d}{dt}\mathcal{J}(u + tv)|_{t=0} = [\langle u, v \rangle - (f, v)] + t(f, v)|_{t=0} = [\langle u, v \rangle - (f, v)].$$

Aufgrund der Eindeutigkeit der Lösung von (2.1.2), muss auch das Minimum des Energiefunktionals eindeutig sein. Damit ist die Aussage vollständig bewiesen.  $\square$

**Bemerkung.** *Es ist natürlich möglich, auch andere als nur reine Dirichlet-Randbedingungen, zum Beispiel Neumann-Randbedingungen oder gemischte Randbedingungen, zu betrachten. In diesem Fall erweitert sich die schwache Formulierung um ein Integral über den Neumann-Rand. Da dies allerdings nicht im Fokus dieser Arbeit steht, werden wir solche Fälle im Folgenden nicht behandeln. Gleiches gilt für inhomogene Dirichlet-Randbedingungen. Der interessierte Leser sei diesbezüglich für die Analysis auf [E, Kapitel 6] und für die Numerik auf [BRA07, Kapitel II, §2 und §3] verwiesen.*

### 2.1.2. Diskretisierung

Die Finite-Elemente-Methode besteht im Grunde genommen daraus, den Sobolevraum  $H_0^1(\Omega)$  in der schwachen Formulierung (2.1.2) durch einen endlich dimensionalen Unterraum  $X_\ell \subset H_0^1(\Omega)$  zu ersetzen. Konkret suchen wir also eine Funktion  $U_\ell \in X_\ell$  so, dass

$$\int_{\Omega} \nabla U_\ell \cdot \nabla V_\ell \, dx = \int_{\Omega} f V_\ell \, dx \quad \text{für alle } V_\ell \in X_\ell$$

bzw.

$$\langle U_\ell, V_\ell \rangle = (f, V_\ell) \quad \text{für alle } V_\ell \in X_\ell \tag{2.1.5}$$

gilt. Es handelt sich damit um ein Galerkin-Verfahren. Da  $X_\ell$  als endlich dimensionaler Unterraum eines Hilbertraums abgeschlossen und damit selbst ein Hilbertraum ist, hat auch (2.1.5), analog zu obigen Argumenten, eine eindeutige Lösung  $U_\ell$ , die *Galerkin-Lösung*. Den Parameter  $\ell$  kann man sich als Dimension des Unterraums vorstellen. Man versucht nun  $\ell$ , d.h. die Dimension von  $X_\ell$ , gegen unendlich zu treiben und hofft, dass die Folge der diskreten Lösungen  $(U_\ell)_{\ell \in \mathbb{N}}$  gegen die exakte Lösung  $u \in H_0^1(\Omega)$  konvergiert. Die Konvergenzanalyse hängt hierbei natürlich stark von der Wahl des Unterraums  $X_\ell$  ab. Wir werden später eine gängige Wahl vorstellen. Zuvor wollen wir uns jedoch noch zwei Eigenschaften dieser Diskretisierung anschauen.

Sei  $u \in H_0^1(\Omega)$  die exakte Lösung von Problem 2.1.1 und  $U_\ell \in X_\ell$  die endlich dimensionale Teillösung. Wegen  $V_\ell \in X_\ell \subset H_0^1(\Omega)$  gilt somit

$$\langle u, V_\ell \rangle = (f, V_\ell) \quad \text{für alle } V_\ell \in X_\ell \quad \text{und} \quad \langle U_\ell, V_\ell \rangle = (f, V_\ell) \quad \text{für alle } V_\ell \in X_\ell.$$

Damit ergibt sich die sogenannte *Galerkin-Orthogonalität*

$$\langle u - U_\ell, V_\ell \rangle = 0 \quad \text{für alle } V_\ell \in X_\ell. \tag{2.1.6}$$

Der zweite Punkt betrifft die Implementierung der FEM. Tatsächlich lässt sich das endlich dimensionale Teilproblem, wie wir in folgendem Satz erkennen, bequem durch ein lineares Gleichungssystem lösen.

**Satz 2.1.5.** Sei  $\{\phi_1, \dots, \phi_N\}$  eine Basis von  $X_\ell$ . Wir definieren die Galerkin Matrix  $A \in \mathbb{R}^{N \times N}$  und den Vektor  $b \in \mathbb{R}^N$  durch

$$A_{jk} := \langle \phi_k, \phi_j \rangle \quad \text{und} \quad b_j := (f, \phi_j).$$

Dann gilt:

- $A$  ist symmetrisch und positiv definit (und daher regulär).
- $U_\ell = \sum_{j=1}^N x_j \phi_j$ , wobei  $x = (x_1, \dots, x_N) \in \mathbb{R}^N$  die eindeutige Lösung von  $Ax = b$  ist.

*Beweis.* (i) Die Symmetrie von  $A$  folgt direkt aus der Symmetrie von  $\langle \cdot, \cdot \rangle$ .

(ii) Für einen beliebigen Vektor  $x \in \mathbb{R}^N$  mit  $v_h := \sum_{j=1}^N x_j \phi_j$  gilt

$$0 \leq \|v_h\|^2 = \langle v_h, v_h \rangle = \sum_{j,k=1}^N x_j x_k \langle \phi_j, \phi_k \rangle = x \cdot Ax.$$

Damit folgt die Positivität von  $A$ .

(iii) Sei  $x \in \mathbb{R}^N$  die eindeutige Lösung des Galerkin Systems und  $y \in \mathbb{R}^N$  der Koeffizientenvektor der eindeutigen Galerkin Lösung  $U_\ell \in X_\ell$ , d.h.  $U_\ell = \sum_{j=1}^N y_j \phi_j$ . Wir zeigen  $x = y$ . Es gilt

$$(Ax)_k = b_k = (f, \phi_k) = \langle U_\ell, \phi_k \rangle = \sum_{j=1}^N y_j \langle \phi_j, \phi_k \rangle = (Ay)_k.$$

Da  $k \in \{1, \dots, N\}$  beliebig war, ist der Beweis vollständig. □

**Bemerkung.** In den Ingenieurwissenschaften wird die Galerkin-Matrix  $A$  oft auch als Steifigkeitsmatrix bezeichnet.

**Bemerkung.** Die guten Eigenschaften der Galerkin-Matrix, d.h. Symmetrie und positiv-Definitheit, ermöglichen es, sehr schnelle Lösungsverfahren, wie das cg-Verfahren, zu verwenden. Diese numerische Gutartigkeit ist einer der Gründe für die Beliebtheit der FEM.

Wir wollen uns im Folgenden etwas genauer mit der tatsächlichen Diskretisierung und der Wahl der Unterräume  $X_\ell$  beschäftigen. Hierzu benötigen wir einige Begriffe.

**Definition.** Eine Menge  $T \subset \mathbb{R}^2$  ist ein *nicht-degeneriertes Dreieck*, wenn es die konvexe Hülle von drei Knoten  $\{x_T, y_T, z_T\}$  ist und  $|T| > 0$  gilt. Hierbei bezeichnet  $|\cdot|$  die Fläche von  $T$ . Ferner definieren wir den *Durchmesser* von  $T$  durch

$$h_T := \text{diam}(T) := \sup\{|x - y| \mid x, y \in T\},$$

und die *Höhe* über der längsten Kante durch  $\rho_T$ . Diese Definitionen sind in Abbildung 2.1 visualisiert.

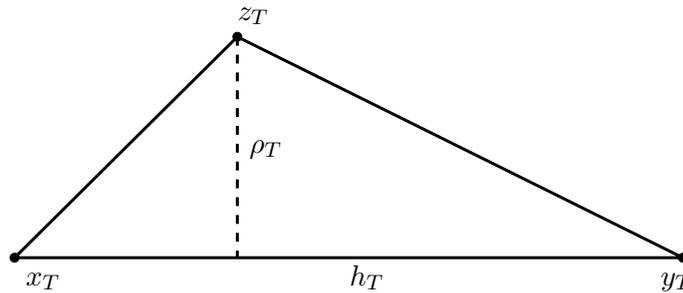


Abbildung 2.1.: Notationen für die verwendeten Dreieckselemente.

Im Zuge der FEM, und auf der Suche nach geeigneten endlich dimensionalen Teilräumen, zerlegen wir nun unser Gebiet  $\Omega$  in eine Menge disjunkter, nicht-degenerierter Dreiecke und erhalten dadurch eine *Triangulierung*, die einigen Eigenschaften gerecht werden muss.

**Definition.** Eine Menge  $\mathcal{T}$  ist genau dann eine *Triangulierung* von  $\Omega$ , wenn

- $\mathcal{T}$  eine endliche Menge nicht-degenerierter Dreiecke ist,
- $\Omega$  durch  $\mathcal{T}$  vollständig bedeckt ist, d.h.  $\bar{\Omega} = \bigcup \mathcal{T}$ ,
- für je zwei Elemente  $T, T' \in \mathcal{T}$  mit  $T \neq T'$  gilt  $|T \cap T'| = 0$ .

Wir nennen die Triangulierung ferner *regulär* (im Sinne von Ciarlet), wenn der Schnitt zweier Elemente  $T, T' \in \mathcal{T}$  mit  $T \neq T'$

- leer,
- ein gemeinsamer Knoten,
- oder eine gemeinsame Kante ist.

Wir bezeichnen die *Menge der Knoten* einer Triangulierung mit  $\mathcal{N}$  und die *Gesamtmenge der Kanten* mit  $\mathcal{E}^*$ . Die Kanten  $E \in \mathcal{E}^*$ , die sich am Rand befinden, d.h.  $E \subset \Gamma$ , bezeichnen wir mit  $\mathcal{E}_\Gamma$ . Aufgrund der Regularität der Triangulierung lässt sich immer unterscheiden, ob eine Kante  $E$  eine Rand- oder eine innere Kante ist. Wir bezeichnen daher die inneren Kanten mit  $\mathcal{E} := \mathcal{E}^* \setminus \mathcal{E}_\Gamma$ .

Wir können nun sinnvolle, endlich dimensionale Unterräume, auf denen wir das diskretisierte Problem 2.1.5 lösen, konstruieren.

**Proposition 2.1.6.** Für eine reguläre Triangulierung  $\mathcal{T}$  von  $\Omega$  definieren wir den diskreten Raum

$$\mathcal{S}_0^1(\mathcal{T}) := \{V_\ell \in C(\Omega) \mid \forall T \in \mathcal{T} \ V_\ell|_T \text{ ist affin}, V_\ell|_\Gamma = 0\} \quad (2.1.7)$$

der stückweise affinen und global stetigen Funktionen mit Nullrand. Es gilt:

- (i)  $\mathcal{S}_0^1(\mathcal{T})$  ist ein endlich dimensionaler Unterraum von  $H_0^1(\Omega)$  mit Dimension  $\#(\mathcal{N} \setminus \Gamma)$ .

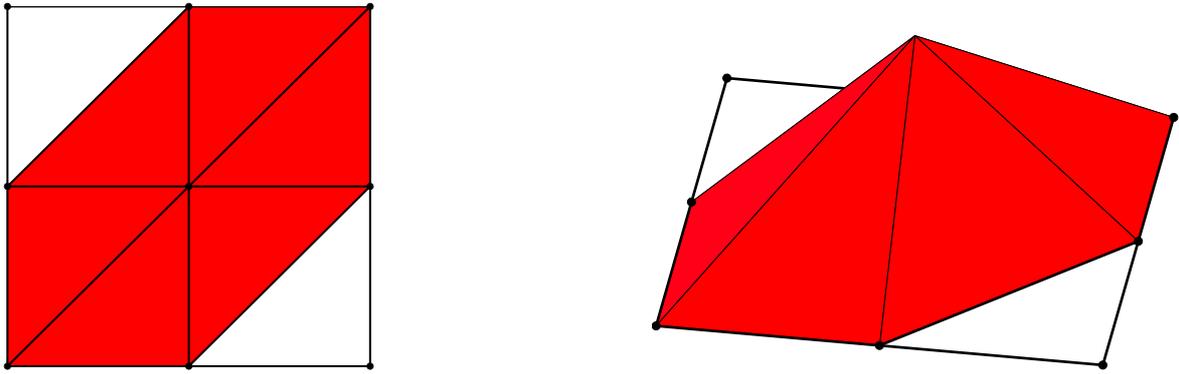


Abbildung 2.2.: Zweidimensionale Hutfunktion  $\eta_z$  über einer regulären Triangulierung eines quadratischen Gebietes  $\Omega$ . Links ist der Träger von  $\eta_z$  visualisiert.

(ii) Für jeden inneren Knoten  $z$  gibt es eine eindeutige Hutfunktion

$$\eta_z \in \mathcal{S}_0^1(\mathcal{T}) \quad \text{mit} \quad \eta_z(z') = \delta_{zz'} \quad \text{für alle } z' \in \mathcal{N} \setminus \Gamma.$$

(iii) Die Menge  $\{\eta_z \mid z \in \mathcal{N} \setminus \Gamma\}$  ist eine Basis von  $\mathcal{S}_0^1(\mathcal{T})$ , die nodale Basis.

*Beweis.* Trivial. □

Ein Element der soeben konstruierten nodalen Basis über einer gegebenen Triangulierung ist in Abbildung 2.2 zu sehen. Deutlich kann man auch die Kompaktheit des Trägers erkennen. Die *P1 Finite-Elemente-Methode* besteht nun darin, die Galerkin Methode mit dem diskreten Raum  $\mathcal{S}_0^1(\mathcal{T})$  zu verwenden, um das homogene Dirichlet Problem zu lösen. Hierbei geht man so vor, dass man auf einer Folge immer feiner werdender Triangulierungen  $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}}$  die Galerkin Lösungen  $U_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)$  berechnet und sich erhofft, dass die Folge der Lösungen  $(U_\ell)_{\ell \in \mathbb{N}}$  in einer geeigneten Norm gegen die exakte Lösung  $u \in H_0^1(\Omega)$  konvergiert. Wir setzen an dieser Stelle voraus, dass die vorkommenden Triangulierungen immer regulär sind. Ausgehend von einer regulären Triangulierung  $\mathcal{T}_\ell$ , erhält man eine feinere Triangulierung  $\mathcal{T}_{\ell+1}$  zum Beispiel durch Aufteilen jedes Elementes  $T \in \mathcal{T}_\ell$  in vier ähnliche Dreiecke (vgl. Abbildung 2.3). Natürlich kann man auch eine beliebige andere Verfeinerungsstrategie aus Abschnitt 2.2.3 verwenden. Abschließend halten wir noch einmal das endlich dimensionale Teilproblem fest, welches durch die Methode der finiten Elemente gelöst wird.

**Problem 2.1.7.** Finde  $U_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)$ , mit

$$\langle\langle U_\ell, V_\ell \rangle\rangle = (f, V_\ell) \quad \text{für alle } V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell). \quad (2.1.8)$$

**Bemerkung.** Per Konstruktion haben die Basisfunktionen  $\eta_z \in \mathcal{S}_0^1(\mathcal{T}_\ell)$  einen sehr begrenzten Träger. Dies führt dazu, dass die Galerkin Matrix im Allgemeinen sehr dünn besetzt sein wird. Bei entsprechender Nummerierung der Knoten und Gleichmäßigkeit der Triangulierung, ist hier sogar eine einfache Bandstruktur möglich. Außerdem ist die Matrix mit einigen weiteren

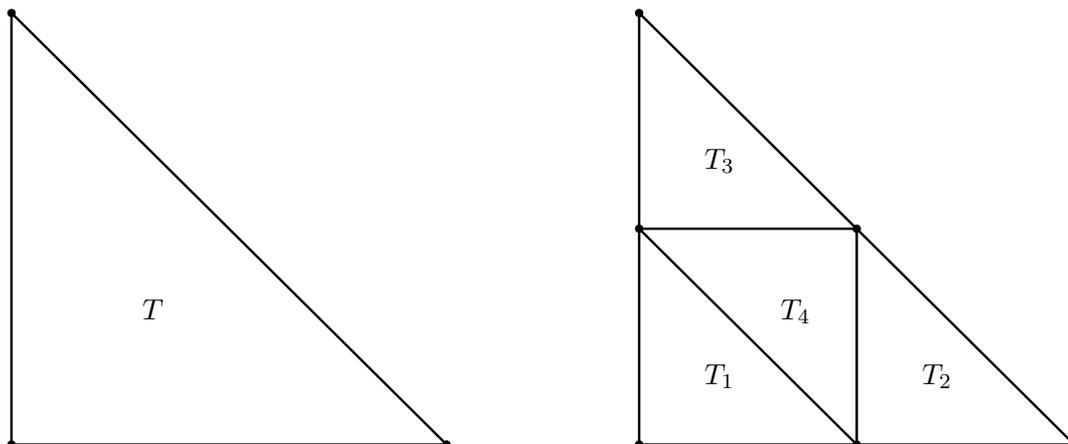


Abbildung 2.3.: Verfeinerung eines Elementes  $T \in \mathcal{T}_\ell$  in vier ähnliche Dreiecke  $T_1, T_2, T_3, T_4 \in \mathcal{T}_{\ell+1}$ . Die neuen Knoten sind gerade die Mittelpunkte der alten Kanten. Das neue Netz  $\mathcal{T}_{\ell+1}$  ist offensichtlich wieder regulär.

Überlegungen relativ günstig exakt zu berechnen. Den Vektor  $b$  der rechten Seite berechnet man normalerweise per Schwerpunktquadratur der Ordnung  $\mathcal{O}(h^2)$ . Für weitere Hinweise zur Implementierung und rechentechnische Betrachtungen sei der interessierte Leser auf [BRA07, FPW] verwiesen.

### 2.1.3. A Priori Fehleranalyse

Wir wollen uns in diesem Abschnitt mit dem Konvergenzverhalten der FEM auseinandersetzen. Bei geeigneter Verfeinerung, d.h. wir verfeinern immer so, dass die Triangulierung in jedem Schritt regulär bleibt, werden wir Aussagen darüber treffen, wann und mit welcher Konvergenzordnung die Folge der Teillösungen gegen die exakte Lösung konvergiert. Da dieser Teil eher als Übersicht anzusehen ist, bleiben die Resultate unbewiesen. Der Vollständigkeit halber werden jedoch immer entsprechende Quellen angegeben. Für die Dauer des gesamten Abschnittes sei  $U_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)$  die eindeutige Lösung von Problem 2.1.7 und  $u \in H_0^1(\Omega)$  die exakte Lösung des ursprünglichen Problems. Wir betrachten mit  $\|\cdot\|_{H_0^1(\Omega)}$  die gewöhnliche Norm auf  $H_0^1(\Omega)$ . Nach der Ungleichung von Poincaré gibt es nun Konstanten  $\alpha$  und  $\beta \in \mathbb{R}$ , so dass

$$\alpha \| \cdot \|_{H_0^1(\Omega)} \leq \| \cdot \| \leq \beta \| \cdot \|_{H_0^1(\Omega)}.$$

Damit lässt sich bereits eine Aussage über die Güte der Approximation treffen.

**Lemma 2.1.8** (Céa). *Der Galerkin Fehler ist quasi-optimal, d.h. es gilt*

$$\|u - U_\ell\|_{H_0^1(\Omega)} \leq \frac{\beta}{\alpha} \min_{V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)} \|u - V_\ell\|_{H_0^1(\Omega)} \quad \text{für alle } u \in H_0^1(\Omega),$$

wobei  $\alpha$  und  $\beta$  die Konstanten der Normäquivalenz bezeichnen. Bezüglich der Energienorm ist die Galerkin Lösung  $U_\ell$  sogar die Bestapproximation, d.h. es gilt

$$\|u - U_\ell\| = \min_{V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)} \|u - V_\ell\| \quad \text{für alle } u \in H_0^1(\Omega). \quad (2.1.9)$$

*Beweis.* Analog zu BRAESS [BRA07, Kapitel II, §4, Satz 4.2] □

Aus dem Céa Lemma ergibt sich sofort eine Konvergenzeigenschaft.

**Proposition 2.1.9.** *Gilt für einen dichten Teilraum  $D \subseteq H_0^1(\Omega)$  die Approximationseigenschaft*

$$\lim_{\ell \rightarrow \infty} \min_{V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)} \|v - V_\ell\|_{H_0^1(\Omega)} = 0 \quad \text{für alle } v \in D,$$

so folgt

$$\lim_{\ell \rightarrow \infty} \|u - U_\ell\|_{H_0^1(\Omega)} = 0 \quad \text{für alle } u \in H_0^1(\Omega).$$

Die Folge der Galerkin Lösungen konvergiert also gegen die exakte Lösung  $u \in H_0^1(\Omega)$ .

*Beweis.* Der Beweis findet sich im Skript von PRAETORIUS [P, Chapter I, Proposition 1.7]. □

**Bemerkung.** *Diese beiden Ergebnisse gelten nicht nur für  $\mathcal{S}_0^1(\mathcal{T}_\ell)$ , sondern lassen sich für einen beliebigen abgeschlossenen Unterraum  $X_\ell \subset H_0^1(\Omega)$  zeigen. Demnach hängt die Konvergenz davon ab, wie gut sich die exakte Lösung  $u$  in  $X_\ell$  approximieren lässt. Da wir stückweise mit Polynomen approximieren ist zu erwarten, dass die Konvergenzgüte von der Glattheit der exakten Lösung abhängen wird.*

Obwohl die Aussage von Proposition 2.1.9 vielversprechend ist, sei angemerkt, dass wir noch keine Aussage über die Konvergenzrate der FEM gewonnen. Die Konvergenz kann bislang also beliebig langsam sein. Für den Rest dieses Abschnittes verbleiben wir bei der so genannten *uniformen Rot-Verfeinerung*, wie sie in Abbildung 2.3 dargestellt ist. Wir verfeinern also gleichmäßig alle Elemente  $T \in \mathcal{T}$ . Ferner definieren wir die *lokale Gitterweite*  $h \in L^\infty(\Omega)$  durch

$$h|_T := h_T \quad \text{für alle } T \in \mathcal{T}$$

und die *Formregularitätskonstante* einer Triangulierung  $\mathcal{T}$  durch

$$\sigma(\mathcal{T}) := \max_{T \in \mathcal{T}} \frac{h_T}{\rho_T} \geq 1. \tag{2.1.10}$$

Wir beobachten die folgenden Eigenschaften, ausgehend von einer regulären Triangulierung  $\mathcal{T}_\ell$ :

- Die neue Triangulierung  $\mathcal{T}_{\ell+1}$  ist auch regulär.
- Es gilt  $h_{T'} = h_T/2$  und  $\rho_{T'} = \rho_T/2$ , wobei  $T' \in \mathcal{T}_{\ell+1}$  einen *Sohn* von  $T \in \mathcal{T}_\ell$ , d.h. ein Element das durch Verfeinerung aus  $T$  hervorgegangen ist, bezeichnet.
- Die Formregularitätskonstante bleibt erhalten, d.h.  $\sigma(\mathcal{T}_{\ell+1}) = \sigma(\mathcal{T}_\ell)$ .

Wir kommen zu einer Aussage über die Konvergenzordnung der Finite-Elemente-Methode.

**Satz 2.1.10** (Approximationssatz). *Für  $u \in H^2(\Omega)$  definieren wir den nodalen Interpolanten durch*

$$I_h u := \sum_{z \in \mathcal{N}} u(z) \eta_z \in \mathcal{S}_0^1(\mathcal{T}_\ell).$$

Für  $\alpha \in \mathbb{R}$  gelten die Abschätzungen

$$\|h^\alpha(u - I_h u)\|_{L^2(\Omega)} \leq C \|h^{2+\alpha} D^2 u\|_{L^2(\Omega)}$$

und

$$\|h^\alpha \nabla(u - I_h u)\|_{L^2(\Omega)} \leq C \sigma(\mathcal{T}) \|h^{1+\alpha} D^2 u\|_{L^2(\Omega)},$$

wobei die Ableitungen  $D^2$  und  $\nabla$  im schwachen Sinne zu verstehen sind. Die Konstante  $C > 0$  ist unabhängig von  $u, \mathcal{T}$  oder  $\Omega$ .

*Beweis.* Analog zu BRAESS [BRA07, Kapitel II, §6, Satz 6.4]. □

Die Definition des nodalen Interpolanten macht an dieser Stelle Sinn, da die Funktion  $u \in H^2(\Omega)$  für  $d = 2, 3$  nach den Soboleveinbettungen automatisch stetig ist. Als Folgerung ergibt sich direkt:

**Korollar 2.1.11.** Für  $u \in H^2(\Omega) \cap H_0^1(\Omega)$  gilt

$$\min_{V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)} \|u - V_\ell\|_{H_0^1(\Omega)} \leq \|u - I_h u\|_{H_0^1(\Omega)} \leq C \sigma(\mathcal{T}_\ell) \|h D^2 u\|_{L^2(\Omega)}. \quad (2.1.11)$$

*Beweis.* Trivial. □

Aus diesem Korollar ergeben sich zwei wichtige Konsequenzen: Erstens ist der Galerkin Fehler nach dem Céa Lemma bis auf Konstanten die Bestapproximation. Demnach hat die P1-FEM für genügend glatte exakte Lösungen  $u \in H^2(\Omega)$  mindestens die Konvergenzordnung  $\mathcal{O}(h)$ . Zweitens haben wir soeben gezeigt, dass die P1-FEM im Besonderen für  $u \in C_c^\infty(\bar{\Omega})$  gegen die exakte Lösung konvergiert. Da  $C_c^\infty(\bar{\Omega})$  in  $H_0^1(\Omega)$  dicht ist (vgl. [E, §5.3.2]), konvergiert die Folge der Galerkin Lösungen somit nach Proposition 2.1.9 immer gegen die exakte Lösung.

Im Rest dieses Abschnittes wollen wir noch zeigen, dass die Konvergenzordnung  $\mathcal{O}(h)$  tatsächlich optimal ist, d.h. dass die P1-FEM unter normalen Umständen exakt mit  $\mathcal{O}(h)$  konvergieren wird. Hierzu benötigen wir eine so genannte *inverse Abschätzung*. Der Raum  $\mathcal{P}^m(T)$  beschreibe im Folgenden die Polynome vom Grad kleiner oder gleich  $m$  auf  $T$ .

**Satz 2.1.12** (inverse Abschätzung). Mit  $\mathcal{P}^m(\mathcal{T}) := \{v_h : \Omega \rightarrow \mathbb{R} \mid \forall T \in \mathcal{T} \ v_h|_T \in \mathcal{P}^m(T)\}$  gilt

$$\|h^\alpha \nabla v_h\|_{L^2(\Omega)} \leq C \sigma(\mathcal{T}) \|h^{\alpha-1} v_h\|_{L^2(\Omega)} \quad \text{für alle } v_h \in \mathcal{P}^m(\mathcal{T}) \text{ und } \alpha \in \mathbb{R}, \quad (2.1.12)$$

und

$$\|h^\alpha D_{\mathcal{T}}^2 v_h\|_{L^2(\Omega)} \leq C \|h^{\alpha-1} v_h\|_{H^1(\mathcal{T})} \quad \text{für alle } v_h \in \mathcal{P}^m(\mathcal{T}) \text{ und } \alpha \in \mathbb{R}, \quad (2.1.13)$$

wobei  $C > 0$  nur von  $m$  und der Formregularitätskonstante  $\sigma(\mathcal{T})$  abhängt.

*Beweis.* Analog zu BRAESS [BRA07, Kapitel II, §6, Satz 6.8] □

Sei nun  $\alpha$  der optimale Exponent für

$$\|v - I_h v\|_{H_0^1(\Omega)} \leq C\sigma(\mathcal{T}) \|h^\alpha D_{\mathcal{T}}^2 v\|_{L^2(\Omega)} \quad \text{für alle } v \in H^2(\Omega) \cap H_0^1(\Omega).$$

Aus (2.1.11) folgt  $\alpha \geq 1$ . Ferner sei  $\beta$  der optimale Exponent für

$$\|h^\beta D_{\mathcal{T}}^2 v\|_{L^2(\Omega)} \leq C \|v\|_{H_0^1(\Omega)} \quad \text{für alle } v \in \mathcal{P}^m(\mathcal{T}).$$

Mit (2.1.13) erhalten wir  $\beta \leq 1$ . Eine kurze Rechnung zeigt

$$\|h^\beta D_{\mathcal{T}}^2 v\|_{L^2(\Omega)} = \|h^\beta D_{\mathcal{T}}^2 (v - I_h v)\|_{L^2(\Omega)} \leq C \|v - I_h v\|_{H_0^1(\Omega)} \leq \tilde{C}\sigma(\mathcal{T}) \|h^\alpha D_{\mathcal{T}}^2 v\|_{L^2(\Omega)}.$$

Hieraus folgt  $\beta \geq \alpha$  und damit  $\alpha = \beta = 1$ . Für  $v \in H^2(\Omega)$  kann die Konvergenzrate  $\mathcal{O}(h)$  also im Allgemeinen nicht mehr verbessert werden. Vielmehr wird die uniforme FEM bei weniger glatten Lösungen sogar langsamer konvergieren. Dies ist auch ein heuristisches Argument dafür, dass es wenig Sinn macht, für die Berechnung der rechten Seite  $b$  Quadraturformeln höherer Ordnung zu verwenden.

## 2.2. Adaptive Verfeinerung

Es ist sehr einfach, die Dreiecke einer gegebenen Triangulierung  $\mathcal{T}_\ell$  wie in Abbildung 2.3 beschrieben, in vier ähnliche Dreiecke zu zerlegen und dadurch das Netz immer weiter zu verfeinern. Oftmals wird das aber nicht das gewünschte Vorgehen widerspiegeln, da die numerische Lösung bei dieser uniformen Verfeinerung überall gleich gut aufgelöst wird. Dies bringt verschiedene Nachteile mit sich. In vielen Anwendungen werden bestimmte Teile der Lösung „interessanter“ sein als andere, dies kann zum Beispiel bei einspringenden Ecken der Fall sein. Zum einen dauert es dann ungleich länger das ganze Netz fein aufzulösen, um eine brauchbare Lösung zu erhalten, als nur die wichtigen Stellen, zum anderen streben wir natürlich danach eben jene „interessanten Stellen“ so gut wie möglich aufzulösen. Es wird daher Sinn machen, an „uninteressanten Stellen“ der Lösung Zeit und Ressourcen zu sparen, um somit schneller zu einer besseren, d.h. für unsere Zwecke geeigneteren, Lösung zu gelangen. Diese Überlegungen führen zu einer ganz neuen Klasse von Finite-Elemente-Methoden, den *adaptiven Verfahren*, bei denen an bestimmten Stellen stärker verfeinert wird, als an anderen. Diese Verfahren können sich sowohl in der Verfeinerungsstrategie, als auch in der Auswahl der „interessanten Stellen“ unterscheiden. Bevor wir uns mit den Eigenschaften adaptiver FEM (AFEM) beschäftigen können, stehen wir jedoch vor dem Problem, herauszufinden an welchen Stellen die Lösung interessant ist, ohne die exakte Lösung  $u$  zu kennen. Dies führt auf den Begriff der *Fehlerschätzer*, den wir im nächsten Abschnitt behandeln wollen.

### 2.2.1. Fehlerschätzer

Wie bereits erwähnt, stehen wir vor der Aufgabe, eine berechenbare Größe  $\eta$  zu finden, welche qualifizierte Aussagen darüber erlaubt, wo das Netz am geschicktesten zu verfeinern ist. Diese darf von der Lösung  $U_\ell$  im  $\ell$ -ten Schritt, der Triangulierung  $\mathcal{T}_\ell$  und den gegebenen Daten  $f$ , **nicht** jedoch von der exakten Lösung  $u$  abhängen. Eine solche Größe nennen wir (*a posteriori*) *Fehlerschätzer*. Wir nennen einen Fehlerschätzer ferner *zuverlässig*, falls

$$\|u - U_\ell\|_{H_0^1(\Omega)} \leq C_{\text{rel}} \eta \tag{2.2.1}$$

gilt, und *effizient*, falls er einer Abschätzung der Form

$$C_{\text{eff}} \eta \leq \|u - U_\ell\|_{H_0^1(\Omega)} \quad (2.2.2)$$

genügt. Für einen zuverlässigen Fehlerschätzer bedeutet die Konvergenz  $\eta \rightarrow 0$  automatisch die Konvergenz des Fehlers gegen 0. Ist der Schätzer zusätzlich effizient, so fällt der Fehler mit der gleichen Konvergenzrate wie der Fehlerschätzer. Mit Hilfe solcher Fehlerschätzer wollen wir zwei Dinge erreichen:

- (i) Wir wollen qualifizierte Aussagen über den Fehler  $\|u - U_\ell\|_{H_0^1(\Omega)}$  in jedem Schritt machen und die Berechnung stoppen, wenn der Fehler ausreichend klein ist.
- (ii) Wir suchen einen Fehlerschätzer mit lokalen Beiträgen, d.h. der Schätzer soll auf jedem Element  $T \in \mathcal{T}_\ell$  einen Beitrag liefern. Auf diese Weise können wir nur an den Stellen verfeinern, wo der Schätzer groß ist, und so unsere adaptive Verfeinerung steuern.

Bevor wir nun zur tatsächlichen Konstruktion solcher Größen kommen, benötigen wir einige Begriffe.

**Definition.** Bezeichne  $\mathcal{N}_T$  die Knoten eines Elements  $T$  und  $\mathcal{N}_E$  die Knoten einer Kante  $E$ . Wir definieren den *Knotenpatch*  $\Omega_z$  durch

$$\tilde{\Omega}_z := \{T \in \mathcal{T} \mid z \in \mathcal{N}_T\} \quad \text{und} \quad \Omega_z := \bigcup \tilde{\Omega}_z := \{x \in \mathbb{R}^2 \mid \exists T \in \tilde{\Omega}_z \ x \in T\}$$

als die Vereinigung aller Elemente die den Knoten  $z$  teilen. Analog definieren wir den *Kantenpatch*  $\Omega_E$  durch

$$\tilde{\Omega}_E := \{T \in \mathcal{T} \mid \mathcal{N}_T \cap \mathcal{N}_E \neq \emptyset\} = \{T \in \tilde{\Omega}_z \mid z \in \mathcal{N}_E\} \quad \text{und} \quad \Omega_E := \bigcup \tilde{\Omega}_E$$

und den *Elementpatch*  $\Omega_T$  durch

$$\tilde{\Omega}_{T'} := \{T \in \mathcal{T} \mid \mathcal{N}_T \cap \mathcal{N}_{T'} \neq \emptyset\} = \{T \in \tilde{\Omega}_z \mid z \in \mathcal{N}_{T'}\} \quad \text{und} \quad \Omega_T := \bigcup \tilde{\Omega}_T.$$

Die Patches  $\Omega_z, \Omega_E$  und  $\Omega_T$  sind in Abbildung 2.4 visualisiert.

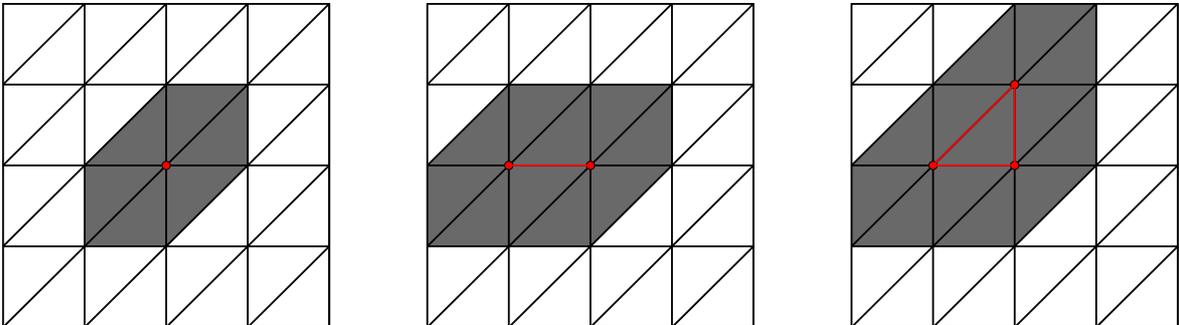


Abbildung 2.4.: Wir unterscheiden die Patches bezüglich Knoten (*links*), Kanten (*mitte*) und Elementen (*rechts*). Der Patch einer Kante oder eines Elements ist hierbei gerade die Vereinigung der entsprechenden Knotenpatches.

Aufgrund der Regularität unseres Netzes folgt, dass jeder Patch nur aus endlich vielen Dreiecken bestehen kann. Die Anzahl der Elemente eines Patches wird also durch eine von der Formregularität abhängende Konstante, der *Patchkonstante*  $C_{\text{patch}}$  beschränkt. Daraus ergibt sich

$$\|v\|_{L^2(\Omega)} \leq \left( \sum_{T \in \mathcal{T}} \|v\|_{L^2(\Omega_T)}^2 \right)^{1/2} \leq C_{\text{patch}} \|v\|_{L^2(\Omega)} \quad \text{für alle } v \in L^2(\Omega). \quad (2.2.3)$$

Für den weiteren Verlauf dieses Kapitels benötigen wir außerdem einen Operator, welcher von  $H_0^1(\Omega)$  nach  $\mathcal{S}_0^1(\mathcal{T})$  abbildet. Der nodale Interpolant ist hierzu nicht ausreichend, da  $H_0^1(\Omega)$  Funktionen im Allgemeinen nicht stetig sind, und daher die Punktauswertung keinen Sinn ergeben würde. Eine Lösung für dieses Problem bietet der *Clément Operator*.

**Satz 2.2.1.** *Seien  $\mathcal{N}_F := \mathcal{N} \setminus \Gamma$  die freien Knoten. Wir definieren den Clément Operator  $J_h : H_0^1(\Omega) \rightarrow \mathcal{S}_0^1(\mathcal{T})$  durch*

$$J_h v := \sum_{z \in \mathcal{N}_F} v_z \eta_z \quad \text{mit} \quad v_z := \frac{1}{|\Omega_z|} \int_{\Omega_z} v dx. \quad (2.2.4)$$

$J_h$  ist damit ein wohldefinierter, linearer Operator. Ferner gilt für jedes  $T \in \mathcal{T}$  die lokale  $H^1$ -Stabilität

$$\|\nabla(v - J_h v)\|_{L^2(T)} \leq C \|\nabla v\|_{L^2(\Omega_T)} \quad \text{für alle } v \in H_0^1(\Omega), \quad (2.2.5)$$

und die lokale Approximationseigenschaft erster Ordnung

$$\|v - J_h v\|_{L^2(T)} \leq Ch_T \|\nabla v\|_{L^2(\Omega_T)} \quad \text{für alle } v \in H_0^1(\Omega). \quad (2.2.6)$$

Außerdem gilt für eine beliebige Kante  $E$  von  $T$

$$\|v - J_h v\|_{L^2(E)} \leq Ch_E^{1/2} \|\nabla v\|_{L^2(\Omega_T)} \quad \text{für alle } v \in H_0^1(\Omega), \quad (2.2.7)$$

wobei  $C > 0$  nur von  $\sigma(\mathcal{T})$  und der Poincaré Konstanten abhängt.

*Beweis.* Siehe VERFÜRTH [V, Lemma 3.1]. □

Wir wollen im Folgenden beispielhaft den *Residualfehlerschätzer*

$$\eta := \left( \sum_{T \in \mathcal{T}} \eta_T^2 \right)^{1/2} \quad (2.2.8)$$

mit

$$\eta_T := \left( h_T^2 \|f\|_{L^2(T)}^2 + h_T \|\partial_n U_\ell\|_{L^2(\partial T \cap \Omega)}^2 \right)^{1/2} \quad (2.2.9)$$

betrachten. Die lokalen Beiträge  $\eta_T$  werden auch *Verfeinerungsindikatoren* genannt. Für eine gemeinsame innere Kante  $E = T_+ \cap T_-$  von zwei Elementen  $T_\pm \in \mathcal{T}$ , ist hierbei der *Sprung der Normalenableitungen* definiert als

$$[\partial_n U_\ell]_E := \frac{\partial U_\ell|_{T_+}}{\partial n_E^+} + \frac{\partial U_\ell|_{T_-}}{\partial n_E^-}, \quad (2.2.10)$$

wobei  $n_E^\pm$  jeweils den äußeren Normalenvektor bezeichnet. Wegen  $n_E^+ = -n_E^-$ , ist die obige Summe eigentlich eine Differenz. Wir kommen nun zu den bereits angesprochenen Abschätzungen.

**Satz 2.2.2** (Zuverlässigkeit). *Der soeben konstruierte Fehlerschätzer  $\eta$  erfüllt die Zuverlässigkeitsabschätzung*

$$\|u - U_\ell\|_{H_0^1(\Omega)} \leq C \eta, \quad (2.2.11)$$

wobei  $C > 0$  nur von der Formregularitätskonstante  $\sigma(\mathcal{T})$  und dem Clément Operator abhängt.

*Beweis.* Für  $w \in H_0^1(\Omega)$  folgt durch elementweise partielle Integration

$$\begin{aligned} R_h(w) &:= (f, w) - \sum_{T \in \mathcal{T}} \langle U_\ell, w \rangle_T \\ &= (f, w) - \sum_{T \in \mathcal{T}} (\partial_n U_\ell, w)_{L^2(\partial T)} \\ &= \sum_{T \in \mathcal{T}} (f, w)_T - \sum_{E \in \mathcal{E}} ([\partial_n U_\ell], w)_E \\ &\leq \sum_{T \in \mathcal{T}} \|f\|_{L^2(T)} \|w\|_{L^2(T)} + \sum_{E \in \mathcal{E}} \|[\partial_n U_\ell]\|_{L^2(E)} \|w\|_{L^2(E)}, \end{aligned}$$

wobei wir hier die Ungleichung von Hölder und  $w|_\Gamma = 0$  verwendet haben. Für ein beliebiges  $v \in H_0^1(\Omega)$  wählen wir nun  $w = v - J_h v$ . Aufgrund der Galerkin-Orthogonalität gilt  $R_h(V_\ell) = 0$  für alle  $V_\ell \in \mathcal{S}_0^1(\mathcal{T})$ , und damit  $R_h(w) = R_h(v)$ . Wir betrachten die beiden Summanden jeweils einzeln. Ausnutzung der Eigenschaften des Clément Operators, der Patchkonstanten und die Anwendung von Cauchy-Schwarz liefern

$$\begin{aligned} \sum_{T \in \mathcal{T}} \|f\|_{L^2(T)} \|v - J_h v\|_{L^2(T)} &\leq C_{\text{Clément}} \left( \sum_{T \in \mathcal{T}} \|h_T f\|_{L^2(T)}^2 \right)^{1/2} \left( \sum_{T \in \mathcal{T}} \|\nabla v\|_{L^2(\Omega_T)}^2 \right)^{1/2} \\ &\leq C_{\text{Clément}} C_{\text{patch}} \left( \sum_{T \in \mathcal{T}} \|h_T f\|_{L^2(T)}^2 \right)^{1/2} \left( \sum_{T \in \mathcal{T}} \|\nabla v\|_{L^2(T)}^2 \right)^{1/2} \\ &= C_{\text{Clément}} C_{\text{patch}} \|h f\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)}. \end{aligned}$$

Wir wählen nun für jede Kante  $E \in \mathcal{E}$  ein Element  $T \in \mathcal{T}$  mit  $T \cap E \neq \emptyset$ . Analoge Argumentation zu vorher zeigt

$$\sum_{E \in \mathcal{E}} \|[\partial_n U_\ell]\|_{L^2(E)} \|v - J_h v\|_{L^2(E)} \leq 3 C_{\text{Clément}} C_{\text{patch}} \|h^{1/2} [\partial_n U_\ell]\|_{L^2(\mathcal{E})} \|\nabla v\|_{L^2(\Omega)},$$

wobei wir hier  $h_T \sim h_E$  ausgenutzt haben. Der zusätzliche Faktor 3 kommt daher, dass wir jedes Element für bis zu 3 Kanten auswählen können. Insgesamt ergibt sich

$$\begin{aligned} R_h(v) &\leq C \|\nabla v\|_{L^2(\Omega)} (\|h f\|_{L^2(\Omega)} + \|h^{1/2} [\partial_n U_\ell]\|_{L^2(\mathcal{E})}) \\ &\leq \sqrt{3} C \|\nabla v\|_{L^2(\Omega)} \eta. \end{aligned}$$

Nach dem Riesz'schen Darstellungssatz (B.1.5) folgt nun  $\|R_h\|_* = \|u - U_\ell\|$ , wobei  $\|\cdot\|_*$  die duale Norm bezeichnet. Nach der Definition der dualen Norm ist dies aber auch durch  $\|R_h\|_* = \sup_{w \in H_0^1(\Omega) \setminus \{0\}} R_h(w) / \|w\|$  gegeben. Nach den obigen Berechnungen gilt

$$R_h(v) \leq C_{\text{rel}} \eta \|v\|$$

für alle  $v \in H_0^1(\Omega)$ . Damit gilt die Aussage auch für das Supremum, so dass der Beweis durch

$$\|u - U_\ell\| \|v\| = \|R_h\|_* \|v\| = \sup_{v \in H_0^1(\Omega) \setminus \{0\}} R_h(v) \leq C_{\text{rel}} \eta \|v\|,$$

vervollständigt wird.  $\square$

Im folgenden Satz stehen die kleineren Elementpatches  $\omega_T \subseteq \Omega_T$  für die Vereinigung der Elemente, die eine gemeinsame Kante mit  $T$  haben. Die kleineren Kantenpatches  $\omega_E := T_E^+ \cup T_E^-$  sind gerade die Vereinigung der Elemente  $T_E^+$  und  $T_E^-$  mit der gemeinsamen Kante  $E$  (siehe Abbildung 2.5). Ferner verwenden wir fortan das Symbol  $A \lesssim B$ , wenn es eine, nur von der Formregularität abhängende Konstante  $c > 0$  gibt, so dass  $A \leq cB$  gilt.

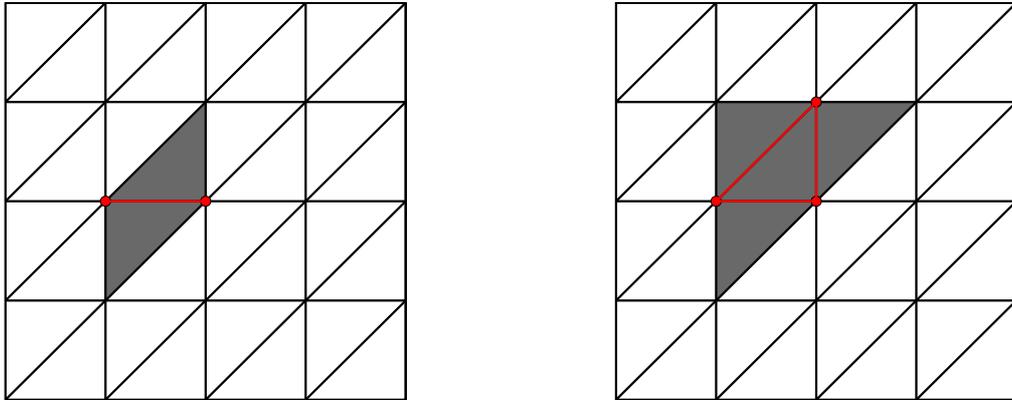


Abbildung 2.5.: Darstellungen der kleineren Patches  $\omega_E \subseteq \Omega_E$  (links) und  $\omega_T \subseteq \Omega_T$  (rechts).

**Satz 2.2.3** (Effizienz). Wir definieren das Integralmittel  $f_{\mathcal{T}} \in \mathcal{P}^0(\mathcal{T})$  durch  $f_{\mathcal{T}|_T} := |T|^{-1} \int_T f \, dx$ . Für jedes  $T \in \mathcal{T}$  erfüllen die Verfeinerungsindikatoren (2.2.9)

$$\eta_T \leq C (\|\nabla(u - U_\ell)\|_{L^2(\omega_T)}^2 + \|h(f - f_{\mathcal{T}})\|_{L^2(\omega_T)}^2)^{1/2}. \quad (2.2.12)$$

Der Schätzer  $\eta$  ist zudem effizient bis auf Oszillationen, d.h.

$$\eta \leq C (\|u - U_\ell\|_{H_0^1(\Omega)} + \|h(f - f_{\mathcal{T}})\|_{L^2(\Omega)}). \quad (2.2.13)$$

Die Konstante  $C > 0$  hängt hierbei nur von der Formregularität der Triangulierung ab.

*Beweis.* Der Beweis ist der Lesbarkeit halber in mehrere Schritte unterteilt.

**Schritt 1:** Mit

$$\eta = \left( \sum_{T \in \mathcal{T}} \eta_T^2 \right)$$

folgt (2.2.13) direkt aus (2.2.12). Im Folgenden betrachten wir die beiden Beiträge von  $\eta_T$  einzeln.

**Schritt 2:** Es gilt

$$\|h_T f\|_{L^2(T)} \lesssim (\|\nabla(u - U_\ell)\|_{L^2(T)} + \|h(f - f_{\mathcal{T}})\|_{L^2(T)}). \quad (2.2.14)$$

Wir definieren die *Element-Bubblefunktion* für  $T \in \mathcal{T}$  durch

$$b_T := \prod_{z \in \mathcal{N}_T} \eta_z \in H_0^1(T) \cap \mathcal{P}^3(T), \quad (2.2.15)$$

und beobachten die Abschätzung

$$\|f_{\mathcal{T}} b_T\|_{L^2(T)} \leq \|f_{\mathcal{T}} b_T^{1/2}\|_{L^2(T)} \leq \|f_{\mathcal{T}}\|_{L^2(T)} \leq C_{\text{ref}} \|f_{\mathcal{T}} b_T\|_{L^2(T)}, \quad (2.2.16)$$

wobei die letzte Abschätzung über affine Transformation auf ein festes *Referenzdreieck* gewonnen wurde. Partielle Integration für  $v = f_{\mathcal{T}} b_T \in H_0^1(T)$  ergibt nun

$$\begin{aligned} C_{\text{ref}}^{-2} \|f_{\mathcal{T}}\|_{L^2(T)}^2 &\leq \|f_{\mathcal{T}} b_T^{1/2}\|_{L^2(T)}^2 = (f_{\mathcal{T}}, v)_{L^2(T)} \\ &= (f_{\mathcal{T}} - f, v)_{L^2(T)} + (f, v)_{L^2(T)} \\ &= (f_{\mathcal{T}} - f, v)_{L^2(T)} + \underbrace{(f - \Delta U_{\ell}, v)}_{=0}{}_{L^2(T)} \\ &= (f_{\mathcal{T}} - f, v)_{L^2(T)} + (\nabla(u - U_{\ell}), \nabla v)_{L^2(T)}, \end{aligned}$$

wobei das Randintegral wegen  $b_T \in H_0^1(T)$  wegfällt. Anwendung der Hölder'schen Ungleichung liefert

$$(f_{\mathcal{T}} - f, v)_{L^2(T)} \leq \|f_{\mathcal{T}} - f\|_{L^2(T)} \|f_{\mathcal{T}} b_T\|_{L^2(T)} \leq \|f_{\mathcal{T}} - f\|_{L^2(T)} \|f_{\mathcal{T}}\|_{L^2(T)},$$

und

$$\begin{aligned} (\nabla(u - U_{\ell}), \nabla v)_{L^2(T)} &\leq \|\nabla(u - U_{\ell})\|_{L^2(T)} \|\nabla(f_{\mathcal{T}} b_T)\|_{L^2(T)} \\ &\leq C_{\text{inv}} h_T^{-1} \|\nabla(u - U_{\ell})\|_{L^2(T)} \|f_{\mathcal{T}} b_T\|_{L^2(T)} \\ &\leq C_{\text{inv}} h_T^{-1} \|\nabla(u - U_{\ell})\|_{L^2(T)} \|f_{\mathcal{T}}\|_{L^2(T)}, \end{aligned}$$

wobei wir hier die inverse Abschätzung (2.1.12) verwendet haben. Umordnen der Terme liefert

$$h_T \|f_{\mathcal{T}}\|_{L^2(T)} \leq C_{\text{ref}}^2 (h_T \|f_{\mathcal{T}} - f\|_{L^2(T)} + C_{\text{inv}} \|\nabla(u - U_{\ell})\|_{L^2(T)}),$$

und mit  $h_T \|f\|_{L^2(T)} = h_T \|f + f_{\mathcal{T}} - f_{\mathcal{T}}\|_{L^2(T)} \leq h_T \|f - f_{\mathcal{T}}\|_{L^2(T)} + h_T \|f_{\mathcal{T}}\|_{L^2(T)}$  schließlich

$$h_T \|f\|_{L^2(T)} \leq (1 + C_{\text{ref}}^2) (h_T \|f_{\mathcal{T}} - f\|_{L^2(T)} + C_{\text{inv}} \|\nabla(u - U_{\ell})\|_{L^2(T)}).$$

**Schritt 3:** Für eine innere Kante  $E \in \mathcal{E}$  gilt

$$h_E^{1/2} \|[\partial_n U_{\ell}]\|_{L^2(E)} \lesssim (\|\nabla(u - U_{\ell})\|_{L^2(\omega_E)} + \|h(f - f_{\mathcal{T}})\|_{L^2(\omega_E)}), \quad (2.2.17)$$

wobei  $\omega_E$  hier die Vereinigung von  $T_+$  und  $T_-$  mit  $E = T_+ \cap T_-$  bezeichnet. Analog zu vorher definieren wir die *Kanten-Bubblefunktion* durch

$$b_E := \prod_{z \in \mathcal{N}_E} \eta_z \in H_0^1(\omega_E) \cap \mathcal{P}^2(\mathcal{T}). \quad (2.2.18)$$

Die essentielle Abschätzung lautet

$$\|b_E\|_{L^2(E)} \leq \|b_E^{1/2}\|_{L^2(E)} \leq h_E^{1/2} \leq C_{\text{ref}} \|b_E\|_{L^2(E)}. \quad (2.2.19)$$

Damit ergibt sich

$$C_{\text{ref}}^{-2} \|[\partial_n U_\ell]\|_{L^2(E)}^2 \leq \|[\partial_n U_\ell] b_E^{1/2}\|_{L^2(E)}^2 = ([\partial_n U_\ell], [\partial_n U_\ell] b_E)_{L^2(E)}.$$

Wegen der geschickten Wahl der Bubblefunktion können wir nun das Skalarprodukt über  $E$  auf den gesamten Rand der Dreiecke  $T_\pm^E$  erweitern. Ferner ergibt sich durch partielle Integration mit  $v := [\partial_n U_\ell] b_E \in \mathcal{P}^2(T_\pm^E)$

$$\begin{aligned} ([\partial_n U_\ell], v)_{L^2(E)} &= (\partial_n U_\ell, v)_{L^2(\partial T_+^E)} + (\partial_n U_\ell)_{L^2(\partial T_-^E)} \\ &= (\nabla U_\ell, \nabla v)_{L^2(\omega_E)} \\ &= (\nabla(U_\ell - u), \nabla v)_{L^2(\omega_E)} + (f, v)_{L^2(\omega_E)} \\ &\leq (C_{\text{inv}} \|\nabla(U_\ell - u)\|_{L^2(\omega_E)} + \|hf\|_{L^2(\omega_E)}) \|h^{-1}v\|_{L^2(\omega_E)}, \end{aligned}$$

wobei die letzte Abschätzung analog zu Schritt 1 folgt. Mit  $|T| \leq \frac{1}{2} h_T h_E$  folgt für  $T \in T_\pm^E$

$$\|v\|_{L^2(T)} = |[\partial_n U_\ell]_E| \|b_E\|_{L^2(T)} \leq |T|^{1/2} |[\partial_n U_\ell]_E| \leq \frac{h_T^{1/2}}{\sqrt{2}} \|[\partial_n U_\ell]\|_{L^2(E)}.$$

Damit erhalten wir

$$h_E^{1/2} \|h^{-1}v\|_{L^2(\omega_E)} \leq \|h^{-1/2}v\|_{L^2(\omega_E)} \leq \|[\partial_n U_\ell]\|_{L^2(\omega_E)}$$

und somit schließlich

$$h_E^{1/2} \|[\partial_n U_\ell]\|_{L^2(E)}^2 \lesssim (C_{\text{inv}} \|\nabla(U_\ell - u)\|_{L^2(\omega_E)} + \|hf\|_{L^2(\omega_E)}) \|[\partial_n U_\ell]\|_{L^2(E)}.$$

Die Verwendung von Schritt 2 zur Gewinnung einer oberen Schranke für  $\|hf\|_{L^2(\omega_E)}$  vollendet den Beweis.  $\square$

**Bemerkung.** *In diesem Beweis haben wir mehrfach Transformationen auf ein so genanntes Referenzdreieck, dessen Maße bekannt sind, durchgeführt. Dort haben wir dann die nötigen Abschätzungen gezeigt und das Ergebnis durch erneute Transformation auf beliebige Dreiecke verallgemeinert. Diese Methode, die garantiert, dass die Konstanten nicht explodieren, wird oft als Skalierungsargument bezeichnet. (vgl. [BRA07, Kapitel II, §6]).*

**Bemerkung.** *In der Literatur sind neben dem von uns verwendete Residualschätzer auch weitere Schätzer in Verwendung. Der Leser sei für Informationen diesbezüglich auf [C] verwiesen.*

### 2.2.2. Adaptiver Algorithmus

Auf den Vorüberlegungen aufbauend, wollen wir in diesem Abschnitt unsere Version eines (konvergenten) adaptiven FEM Algorithmus angeben, welcher die lokalen Beiträge des Residualschätzers (2.2.9) nutzt um, an heuristisch interessanten Stellen, das Netz lokal zu verfeinern. Wir halten dazu die folgenden Notationen fest, die auch in späteren Kapiteln ihre Gültigkeit behalten:

- Der Index  $\ell \in \mathbb{N}$  bezeichnet den  $\ell$ -ten Schritt des Algorithmus'.
- $\mathcal{T}_\ell$  bezeichnet die (reguläre) Triangulierung im  $\ell$ -ten Schritt.

- $\mathcal{N}_\ell$  und  $\mathcal{E}_\ell$  bezeichnen die Knoten bzw. die inneren Kanten der Triangulierung im  $\ell$ -ten Schritt.
- $\mathcal{E}_{\ell,\Gamma}$  und  $\mathcal{E}_\ell^*$  bezeichnen die Randkanten, bzw. alle Kanten von  $\mathcal{T}_\ell$ .
- $\eta_\ell := \left( \sum_{T \in \mathcal{T}_\ell} \eta_\ell(T) \right)^{1/2}$  bezeichnet den Fehlerschätzer mit den Verfeinerungsindikatoren  $\eta_\ell(T)$  aus (2.2.9) im  $\ell$ -ten Schritt.
- Schließlich steht – wie zuvor –  $U_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)$  für die eindeutige Galerkin Lösung im  $\ell$ -ten Schritt.

Ein Schritt unseres adaptiven Verfahrens läuft nun prinzipiell folgendermaßen ab. Wir berechnen die Lösung  $U_\ell$  auf  $\mathcal{T}_\ell$ , die Verfeinerungsindikatoren  $\eta_\ell(T)$  und den Fehlerschätzer  $\eta_\ell$ . Falls die Abbruchbedingung erfüllt ist, beenden wir die Berechnung und geben  $U_\ell$  aus. Ansonsten werden die Elemente, bei denen  $\eta_\ell(T)$  am größten ist zur Verfeinerung ausgewählt. Abschließend erzeugen wir ein neues reguläres Gitter  $\mathcal{T}_{\ell+1}$  bei dem mindestens alle ausgewählten Elemente verfeinert wurden und beginnen von vorne. Formal lässt sich dieser Algorithmus wie folgt niederschreiben:

**Algorithmus 2.2.4** (AFEM - Algorithmus). **Input:** Starttriangulierung  $\mathcal{T}_0$ , maximale Anzahl an Elementen  $n_{E_{max}}$ , Adaptivitätsparameter  $\theta \in (0, 1)$ ,  $\ell := 0$ .

- (i) Berechne die diskrete Teillösung  $U_\ell$ .
- (ii) Berechne die Verfeinerungsindikatoren  $\eta_\ell(T)$ , sowie den Fehlerschätzer  $\eta_\ell$ .
- (iii) STOP, wenn die Anzahl der Elemente größer als  $n_{E_{max}}$  ist.
- (iv) Wähle eine Menge  $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$  minimaler Kardinalität, von markierten Elementen, so dass

$$\theta \eta_\ell^2 = \theta \sum_{T \in \mathcal{T}_\ell} \eta_\ell(T)^2 \leq \sum_{T \in \mathcal{M}_\ell} \eta_\ell(T)^2. \quad (2.2.20)$$

- (v) Erzeuge neue (reguläre) Triangulierung  $\mathcal{T}_{\ell+1}$  so, dass mindestens alle markierten Elemente verfeinert wurden.
- (vi)  $\ell \mapsto \ell + 1$  und beginne wieder bei (i).

**Output:** Endliche Folge von diskreten Lösungen  $U_\ell$  und entsprechende Fehlerschätzer  $\eta_\ell$ .

**Bemerkung.** Die Markierungsstrategie (2.2.20) wurde in [D] von DÖRFLER eingeführt. Markierungsstrategien dieser Form werden daher auch Dörflermarkierung genannt. Sie werden später essentiell in die Konvergenzbeweise eingehen.  $\theta = 1$  kommt hier einer uniformen Verfeinerung gleich, während ein sehr kleines  $\theta$  starke Adaptivität bedeutet.

**Bemerkung.** Das Abbruchkriterium in (iii) ist etwas unkonventionell, da man ein Kriterium der Form  $\|u - U_\ell\| \leq C_{rel} \eta_\ell \leq \tau$  für eine Toleranz  $\tau > 0$  erwarten würde. Dies wäre jedoch nur sinnvoll, wenn man eine qualifizierte Aussage über die Größe der Zuverlässigkeitskonstante  $C_{rel}$  machen könnte. Diese ist im Allgemeinen aber leider unbekannt.

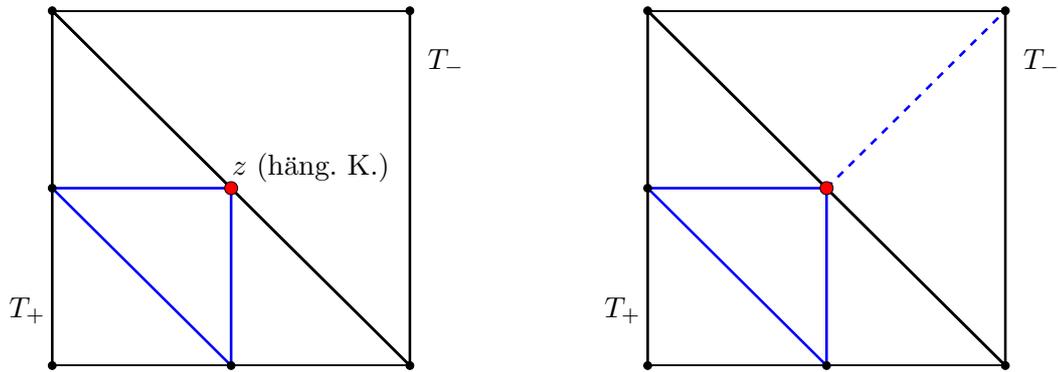


Abbildung 2.6.: Das zur Verfeinerung ausgewählte Element  $T_+$  wird, wie wie in Abschnitt 2.1.2 beschrieben, verfeinert. Im nicht ausgewählten Nachbarelement  $T_-$  entsteht hierdurch der hängende Knoten  $z$  (rot). Da  $\eta_z$  per Definition elementweise affin ist, würde  $\eta_z|_{T_-} \equiv 0$  gelten (*links*). Um dies zu vermeiden wird der *Triangulierungsabschluss* gebildet, bei dem, wenn nötig, auch nicht ausgewählte Elemente verfeinert werden (*rechts*).

Bevor wir uns mit der Konvergenz dieses Algorithmus auseinandersetzen, wollen wir im nächsten Abschnitt die Möglichkeiten und Probleme der Netzverfeinerung aus Schritt (v) genauer beleuchten.

### 2.2.3. Verfeinerungsstrategien

Im Folgenden wollen wir uns verschiedene Strategien zur (lokalen) Netzverfeinerung anschauen. Da wir hierbei nicht einfach alle Elemente verfeinern, treten gewisse zusätzliche Probleme auf. Zum einen hängen alle Abschätzungen aus den vorangegangenen Kapiteln von der Formregularitätskonstante ab und werden damit nichtig, falls  $\sigma(\mathcal{T}_\ell) \xrightarrow{\ell \rightarrow \infty} \infty$ . Zum anderen kann es theoretisch passieren, dass durch Verfeinerung eines Elements  $T_+$  im Nachbarelement  $T_-$  ein so genannter *hängender Knoten*  $z$  entsteht, was dazu führen würde, dass die Hutfunktion  $\eta_z$  auf dem ganzen Element  $T_-$  gleich null ist (Abbildung 2.6). Dies gilt es natürlich zu vermeiden. Ohne Beweis sei an dieser Stelle gesagt, dass alle hier vorgestellten Verfeinerungsregeln die Formregularitätskonstante beschränkt halten, so dass wir uns darum nicht sorgen müssen. Um das zweite Problem in den Griff zu bekommen, ist ein so genannter *Triangulierungsabschluss* nötig, d.h. statt nur der markierten Elemente, werden evtl. auch Nachbarelemente verfeinert um hängende Knoten zu vermeiden. Das Prinzip ist in Abbildung 2.6. dargestellt. Wir kommen nun zu den einzelnen Strategien.

#### Rot-Grün-Blau-Verfeinerung

Dies ist eine kantenbasierte Verfeinerung. Wir müssen also anhand der in Algorithmus 2.2.4 ausgewählten Elemente, Kanten zur Verfeinerung auswählen. Hierbei gehen wir so vor, dass wir einfach die längste Kante  $E \in \mathcal{E}_T$  eines ausgewählten Elementes  $T \in \mathcal{T}_\ell$  (oder alle seine Kanten  $E \in \mathcal{E}_T$ ) markieren.

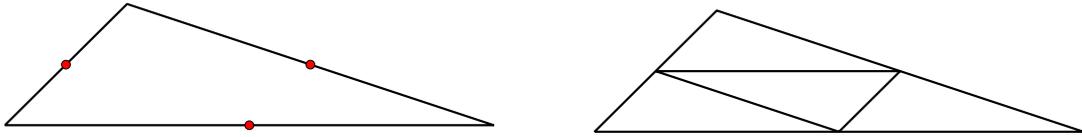


Abbildung 2.7.: *Rot-Verfeinerung*: Falls alle Kanten eines Dreiecks  $T \in \mathcal{T}_\ell$  markiert sind (*links*), so wird es in vier ähnliche Dreiecke zerteilt (*rechts*).

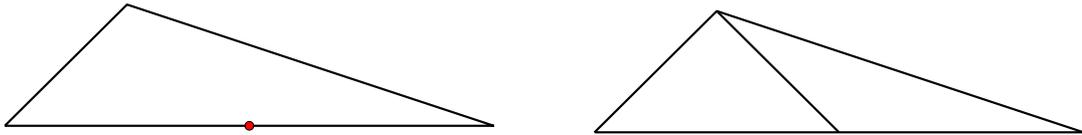


Abbildung 2.8.: *Grün-Verfeinerung*: Falls nur die längste Kante eines Dreiecks  $T \in \mathcal{T}_\ell$  markiert sind (*links*), so wird es in zwei Dreiecke zerteilt (*rechts*).

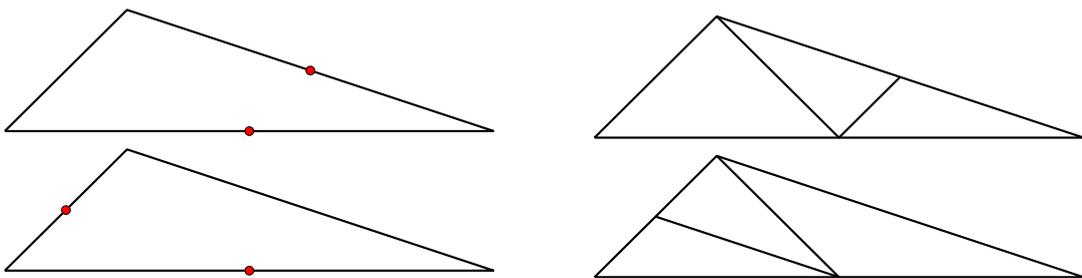


Abbildung 2.9.: *Blau-Verfeinerung*: Sind zwei Kanten von  $T \in \mathcal{T}_\ell$  zur Verfeinerung (*links*), so wird es in insgesamt drei Dreiecke zerteilt (*rechts*).

Den Triangulierungsabschluss definieren wir folgendermaßen rekursiv:

- Enthält ein Element  $T \in \mathcal{T}_\ell$  eine markierte Kante, so markieren wir auch seine längste Kante  $E \in \mathcal{E}_T$  zur Verfeinerung.

Jede markierte Kante wird nun halbiert, so dass ihr Mittelpunkt ein neuer Knoten aus  $\mathcal{N}_{\ell+1}$  ist. Es gelten die folgenden Verfeinerungsregeln:

- Ein Element  $T \in \mathcal{T}_\ell$  wird nur dann verfeinert, wenn mindestens eine seiner Kanten markiert ist, sonst gilt  $T \in \mathcal{T}_{\ell+1}$ .
- Sind alle Kanten von  $T$  markiert, so verwenden wir *rot-Verfeinerung* (Abb. 2.7).
- Falls nur eine Kante von  $T$  (und damit die längste) markiert ist, so verwenden wir *grün-Verfeinerung* (Abb. 2.8).
- Sind schließlich zwei Kanten eines Elements  $T$  markiert, so verwenden wir *blau-Verfeinerung* (Abb. 2.9).

In Abbildung 2.6 wurde somit  $T_+$  rot, und das Nachbarelement  $T_-$  durch den Abschluss grün verfeinert.

### Newest Vertex Bisection

Diese Methode stellt eine beliebte Alternative zur Rot-Grün-Blau-Verfeinerung dar, da sie verhältnismäßig leicht zu implementieren ist. Zur Initialisierung benötigen wir einen Startschritt:

- Wir wählen für jedes Element  $T \in \mathcal{T}_0$  einen beliebigen Knoten aus. Diesen definieren wir als *newest vertex*.
- Die dem newest vertex gegenüberliegende Kante nennen wir *Referenzkante*. Das gilt auch für alle weiteren Verfeinerungen.

Von der Initialisierung ausgehend, gelten diese Verfeinerungsregeln:

- Ein markiertes Element  $T \in \mathcal{T}_\ell$  wird durch Halbieren der Referenzkante in zwei Dreiecke  $T_1, T_2 \in \mathcal{T}_{\ell+1}$  zerteilt.
- Der Mittelpunkt der Referenzkante von  $T$  wird newest vertex der Elemente  $T_1$  und  $T_2$ . Dementsprechend sind die verbleibenden Kanten von  $T$  nun die Referenzkanten von  $T_1$  und  $T_2$ .

Insgesamt ergibt sich analog zu Rot-Grün-Blau die folgende Verfeinerungsstrategie:

- Ist ein Element  $T \in \mathcal{T}_\ell$  zur Verfeinerung markiert, so markieren wir seine Referenzkante  $E \in \mathcal{E}_T$  (oder alle seine Kanten  $E \in \mathcal{E}_T$ ).
- Enthält ein Element  $T$  eine markierte Kante, so markieren wir auch seine Referenzkante  $E \in \mathcal{E}_T$  zur Verfeinerung.

Nun wenden wir die Newest Vertex Bisection (NVB) wiederholt an. Es ergeben sich die folgenden Fälle:

- $T \in \mathcal{T}_\ell$  enthält keine markierte Kante. Es gilt dann  $T \in \mathcal{T}_{\ell+1}$ .
- Sind alle Kanten eines Elements  $T \in \mathcal{T}_\ell$  markiert, so wird es zunächst durch NVB in zwei Söhne  $T_1$  und  $T_2$  unterteilt. Diese werden dann ihrerseits jeweils wieder in zwei Söhne aufgespalten, da Ihre Referenzkanten ja weiterhin markiert sind (Abb. 2.12). Diese Verfeinerung wird auch *bisec<sub>3</sub>* genannt.
- Ist nur eine Kante (und damit die Referenzkante)  $E \in \mathcal{E}_T$  markiert, so wird  $T$  in zwei Söhne  $T_1$  und  $T_2$  aufgespalten (Abb. 2.10).
- Sind schließlich zwei Kanten von  $T \in \mathcal{T}_\ell$  zur Verfeinerung markiert, so verfeinern wir zunächst  $T$  in  $T_1$  und  $T_2$  und schließlich den Sohn, dessen Referenzkante markiert ist noch ein weiteres Mal (Abb. 2.11).

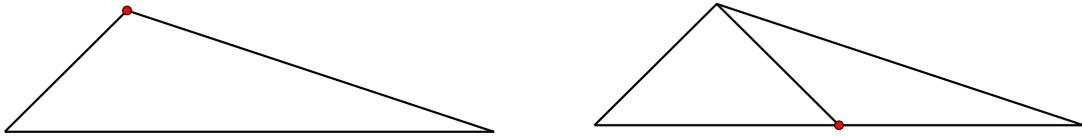


Abbildung 2.10.: *Newset Vertex Bisection*: Für jedes Element  $T \in \mathcal{T}_\ell$  gibt es einen *newset vertex* (rot). Die dem newest vertex gegenüberliegende Kante ist die *Referenzkante* (*links*). Soll ein Element verfeinert werden, so wird die Referenzkante halbiert. Der neue Knoten (rot), der hierbei entsteht, wird newest vertex der beiden Söhne (*rechts*).

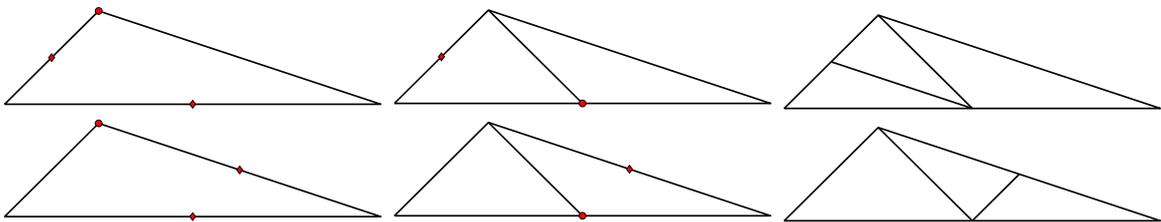


Abbildung 2.11.: *bisec<sub>2</sub>*: Sind nur zwei Kanten von  $T \in \mathcal{T}_\ell$  markiert (*links*), so wird zunächst die Referenzkante zerteilt. Der neue Knoten wird newest vertex der beiden Söhne (*mitte*). Zum Schluss wird noch der eine Sohn mit der markierten Kante durch NVB ein weiteres Mal geteilt (*rechts*).

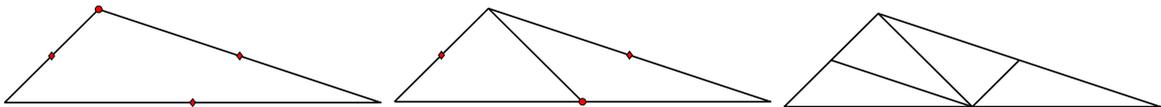


Abbildung 2.12.: *bisec<sub>3</sub>*: Sind alle Kanten eines Elements  $T \in \mathcal{T}_\ell$  markiert (*links*), so halbieren wir zunächst die Referenzkante. Der neue Knoten wird newest vertex der beiden Söhne (*mitte*). Schließlich werden auch die Söhne via NVB noch einmal zerteilt (*rechts*).

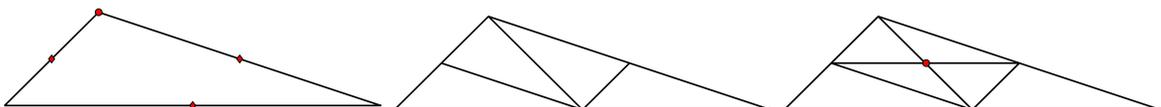


Abbildung 2.13.: *bisec<sub>5</sub>*: Alle Kanten von  $T \in \mathcal{T}_\ell$  werden verfeinert (*mitte*). Zusätzlich wird die in der ersten Verfeinerung neu entstandene Kante erneut verfeinert. Hierdurch entsteht ein neuer Knoten im Inneren des Elements  $T$  (rot) (*rechts*). Die *innerer-Punkt-Eigenschaft* ist somit gewährleistet.

**Bemerkung .** Macht man in der  $\text{bisec}_3$ -Verfeinerung einen weiteren Schritt, indem man die erste neue Kante erneut verfeinert, so erhält man die so genannte  $\text{bisec}_5$ -Verfeinerung (siehe Abb. 2.13). Diese scheint zwar zunächst keinen weiteren Nutzen mit sich zu bringen, ist aber ein sehr wichtiges theoretisches Tool, da hierdurch ein neuer Knoten im Inneren des Elements  $T$  entsteht. Diese Eigenschaft, die in der Literatur innerer-Punkt-Eigenschaft heißt, wird oftmals für Konvergenzbeweise verwendet, schränkt aber dadurch deren Gültigkeit auf die Verfeinerungsstrategie ein.

#### 2.2.4. Konvergenz der AFEM und Schätzerreduktion

Wir wollen den Rest dieses Kapitels nutzen, um zu zeigen, dass auch die adaptive FEM unter den vorgestellten Regeln konvergiert. Besonderes Augenmerk sei hierbei darauf gerichtet, dass die Methode auch, und dies ist nicht trivial, tatsächlich gegen die richtige Lösung  $u \in H_0^1(\Omega)$  strebt.

Der Vollständigkeit halber geben wir ein Resultat an, welches 1996 von DÖRFLER gezeigt, und 2000 von MORIN, NOCHETTO und SIEBERT erweitert wurde. Es wird gezeigt, dass der Diskretisierungsfehler bis auf die Oszillationsterme  $\text{osc}_\ell := \text{osc}_\ell(\mathcal{T}_\ell) = (\sum_{T \in \mathcal{T}_\ell} \text{osc}_\ell(T)^2)^{1/2}$ , mit

$$\text{osc}_\ell(T)^2 := h_T^2 \|f - f_T\|_{L^2(T)}, \quad (2.2.21)$$

eine Kontraktionseigenschaft erfüllt. Hierbei benötigen wir die in der letzten Bemerkung des vorangegangenen Abschnitts angesprochene *innerer-Punkt-Eigenschaft*. Dies liegt daran, dass in den Beweis die so genannte *diskrete lokale Effizienz* eingeht. Es handelt sich dabei um eine Abschätzung, bei der man den Fehlerschätzer  $\eta_\ell$  durch den diskreten Fehler  $\|U_{\ell+1} - U_\ell\|$  dominiert. Wir werden im nächsten Kapitel wieder auf diese Abschätzung, die, soweit wir wissen, nur unter der innerer-Punkt-Eigenschaft zu beweisen ist, zurückkommen (siehe Satz 3.2.12).

**Satz 2.2.5.** [MNS, Theorem 3.1] Die Menge der ausgewählten Elemente  $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$  erfülle die Dörflermarkierung (2.2.20) für  $\theta \in (0, 1)$ . Ferner sei die innerer-Punkt-Eigenschaft gewährleistet. Dann existiert ein  $q \in (0, 1)$  mit

$$\|u - U_{\ell+1}\|_{H_0^1(\Omega)} \leq q \|u - U_\ell\|_{H_0^1(\Omega)} + \text{osc}_\ell \quad \text{für alle } \ell \in \mathbb{N}, \quad (2.2.22)$$

wobei die Kontraktionskonstante nur von  $\Omega, \sigma(\mathcal{T}_\ell)$  und  $\theta$  abhängt. Diese Eigenschaft heißt Fehlerreduktion. Fordern wir zusätzlich, dass die Oszillationsterme verschwinden, so gilt  $\lim_{\ell \rightarrow \infty} U_\ell = u$ .

Wie bereits erwähnt, ist die Gültigkeit obiger Konvergenzaussage quasi auf  $\text{bisec}_5$ -Verfeinerung beschränkt. Um diesen Umstand zu umgehen, wollen wir im Folgenden eine Kontraktionseigenschaft für den Schätzer selbst nachweisen. Mit der Zuverlässigkeit des Fehlerschätzers folgt dann auch die Konvergenz des Fehlers gegen Null. Bevor wir zum eigentlichen Beweis kommen, beobachten wir, dass alle oben vorgestellten Verfeinerungsregeln eine gleichmäßige Verringerung des Durchmessers der markierten Dreiecke, d.h.

$$h_{\ell+1,T} \leq p h_{\ell,T} \quad \text{für ein } p \in (0, 1) \text{ und für alle } T \in \mathcal{M}_\ell$$

gewährleisten, wobei wir hier  $h_{\ell,T} := |T|^{1/2}$  über die Fläche des Elements neu definieren. Mit  $|T| \leq \text{diam}(T)^2 \leq 2\sigma(\mathcal{T}_\ell)|T|$ , sind die beiden Definitionen von  $h$  auf regulären Netzen äquivalent, so dass alle obigen Resultate weiterhin gelten. Das folgende Prinzip der *Schätzerreduktion* stammt aus [CKNS] und [AFP].

**Satz 2.2.6** (Schätzerreduktion). *Sei  $\theta \in (0, 1)$  der Adaptivitätsparameter aus Algorithmus 2.2.4 und  $\eta_\ell$  der oben vorgestellte Residualschätzer (2.2.8). Dann gilt*

$$\eta_{\ell+1} \leq (1 - \theta(1 - p))^{1/2} \eta_\ell + C \|U_{\ell+1} - U_\ell\|, \quad (2.2.23)$$

wobei die Konstante  $C$  nur von der Formregularität  $\sigma(\mathcal{T}_\ell)$  der Triangulierung abhängt und  $p \in (0, 1)$  den Parameter von oben bezeichnet.

*Beweis.* Die Dreiecksungleichung in  $\ell^2$  liefert

$$\begin{aligned} \eta_{\ell+1} &= \left( \sum_{T' \in \mathcal{T}_{\ell+1}} \|h_{T'} f\|_{L^2(T')}^2 + \|h_{T'}^{1/2} [\partial_n U_{\ell+1}]\|_{L^2(\partial T' \cap \Omega)}^2 \right)^{1/2} \\ &\leq \left( \sum_{T' \in \mathcal{T}_{\ell+1}} \|h_{T'} f\|_{L^2(T')}^2 + \|h_{T'}^{1/2} [\partial_n U_\ell]\|_{L^2(\partial T' \cap \Omega)}^2 \right)^{1/2} \\ &\quad + \left( \sum_{T' \in \mathcal{T}_{\ell+1}} \|h_{T'}^{1/2} [\partial_n (U_{\ell+1} - U_\ell)]\|_{L^2(\partial T' \cap \Omega)}^2 \right)^{1/2}. \end{aligned}$$

Sei nun  $E \in \mathcal{E}_\ell = K_+ \cap K_-$  mit  $K_\pm \in \mathcal{T}_\ell$  eine innere Kante. Es gilt

$$\begin{aligned} \|h_T^{1/2} [\partial_n (U_{\ell+1} - U_\ell)]\|_{L^2(E)}^2 &= h_T h_E |\partial_n (U_{\ell+1} - U_\ell)|^2 \\ &= h_T h_E |\nabla(U_{\ell+1} - U_\ell) \cdot n_E - \nabla(U_{\ell+1} - U_\ell) \cdot n_E|^2 \\ &\leq h_T h_E |\nabla(U_{\ell+1} - U_\ell)|_{K_+} - \nabla(U_{\ell+1} - U_\ell)|_{K_-}|^2 \\ &\lesssim h_T h_E |\nabla(U_{\ell+1} - U_\ell)|_{K_+}|^2 + h_T h_E |\nabla(U_{\ell+1} - U_\ell)|_{K_-}|^2 \\ &\sim |K_+| |\nabla(U_{\ell+1} - U_\ell)|_{K_+}|^2 + |K_-| |\nabla(U_{\ell+1} - U_\ell)|_{K_-}|^2 \\ &= \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(K_+)}^2 + \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(K_-)}^2 \\ &= \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\omega_E)}^2. \end{aligned}$$

Damit ergibt sich aufgrund der gleichmäßigen Formregularität von  $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}}$

$$\left( \sum_{T' \in \mathcal{T}_{\ell+1}} \|h_{T'} [\partial_n (U_{\ell+1} - U_\ell)]\|_{L^2(\partial T' \cap \Omega)}^2 \right)^{1/2} \lesssim C \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)} = C \|U_{\ell+1} - U_\ell\|.$$

Sei nun  $\overline{\mathcal{M}}_\ell := \{T' \in \mathcal{T}_{\ell+1} \mid \exists T \in \mathcal{M}_\ell, T' \subseteq T\}$  die Menge der Elemente, die durch Verfeinerung der markierten Elemente entstanden sind. Dann gilt

$$\begin{aligned} &\sum_{T' \in \overline{\mathcal{M}}_\ell} \|h_{T'} f\|_{L^2(T')}^2 + \|h_{T'}^{1/2} [\partial_n U_\ell]\|_{L^2(\partial T' \cap \Omega)}^2 \\ &\leq p \sum_{T \in \mathcal{M}_\ell} \|h_T f\|_{L^2(T)}^2 + \|h_T^{1/2} [\partial_n U_\ell]\|_{L^2(\partial T \cap \Omega)}^2 = p \sum_{T \in \mathcal{M}_\ell} \eta_\ell(T)^2, \end{aligned}$$

wobei wir hier verwendet haben, dass jedes Element  $T \in \mathcal{M}_\ell$  die disjunkte Vereinigung seiner Söhne  $T' \in \overline{\mathcal{M}}_\ell$  ist, und der Sprung  $[\partial_n U_\ell]$  auf neuen inneren Kanten verschwindet. Analog

erhalten wir

$$\begin{aligned} & \sum_{T' \in \mathcal{T}_{\ell+1} \setminus \overline{\mathcal{M}}_\ell} \|h_{T'} f\|_{L^2(T')}^2 + \|h_{T'}^{1/2} [\partial_n U_\ell]\|_{L^2(\partial T' \cap \Omega)}^2 \\ & \leq \sum_{T \in \mathcal{T}_\ell \setminus \mathcal{M}_\ell} \sum_{T \in \mathcal{M}_\ell} \|h_T f\|_{L^2(T)}^2 + \|h_T^{1/2} [\partial_n U_\ell]\|_{L^2(\partial T \cap \Omega)}^2 = \sum_{T \in \mathcal{T}_\ell \setminus \mathcal{M}_\ell} \eta_\ell(T)^2. \end{aligned}$$

Mit der Dörflermarkierung und  $p < 1$  erhalten wir schließlich

$$p \sum_{T \in \mathcal{M}_\ell} \eta_\ell(T)^2 + \sum_{T \in \mathcal{T}_\ell \setminus \mathcal{M}_\ell} \eta_\ell(T)^2 = (p-1) \sum_{T \in \mathcal{M}_\ell} \eta_\ell(T)^2 + \sum_{T \in \mathcal{T}_\ell} \eta_\ell(T)^2 \leq (1 - \theta(1-p)) \eta_\ell^2.$$

Damit ist die Aussage vollständig bewiesen.  $\square$

Wir wollen für den Moment annehmen, dass die Folge  $\|U_{\ell+1} - U_\ell\|$  für  $\ell \rightarrow \infty$  verschwindet. Dann folgt mit der Zuverlässigkeit

$$\|u - U_\ell\| \lesssim \eta_\ell$$

des Schätzers, die Konvergenz der diskreten Lösungen gegen die exakte Lösung  $u \in H_0^1(\Omega)$  aus folgendem Lemma:

**Lemma 2.2.7.** [AFP, Proposition 1.2] *Angenommen die Folge der Fehlerschätzer  $(\eta_\ell)_{\ell \in \mathbb{N}}$  erfülle die Schätzerreduktion*

$$\eta_{\ell+1} \leq q \eta_\ell + \alpha_\ell \quad \text{für alle } \ell \in \mathbb{N}_0,$$

mit  $q \in (0, 1)$  und einer nichtnegativen Folge  $(\alpha_\ell)_{\ell \in \mathbb{N}_0}$  mit  $\lim_{\ell \rightarrow \infty} \alpha_\ell = 0$ , dann gilt auch  $\lim_{\ell \rightarrow \infty} \eta_\ell = 0$ .

Um den Konvergenzbeweis zu vervollständigen, fehlt noch die so genannte *a priori Konvergenz adaptiver Galerkin Verfahren*.

**Lemma 2.2.8** (A priori Konvergenz). *Sei  $\mathcal{H}$  ein Hilbertraum mit der Norm  $\|\cdot\|$  und  $X_\ell$  eine Folge abgeschlossener, konformer Unterräume, d.h.  $X_\ell \subseteq X_{\ell+1}$ . Sei ferner  $U_\ell \in X_\ell$  für ein  $u \in \mathcal{H}$  die Bestapproximation bzgl.  $X_\ell$ , d.h.*

$$\|u - U_\ell\| = \min_{V_\ell \in X_\ell} \|u - V_\ell\|.$$

Dann existiert der Grenzwert  $\lim_{\ell \rightarrow \infty} U_\ell \in \mathcal{H}$  und es gilt  $\lim_{\ell \rightarrow \infty} \|U_{\ell+1} - U_\ell\| = 0$ .

*Beweis.* Sei  $X_\infty$  der Abschluss von  $\bigcup_{\ell=0}^\infty X_\ell \in \mathcal{H}$ . Dann ist  $X_\infty$  ein abgeschlossener Unterraum von  $\mathcal{H}$  und somit existiert auch die Bestapproximation  $U_\infty \in X_\infty$ . Da wir die Bestapproximation in Hilberträumen als Orthogonalprojektion schreiben können, folgt mit dem Satz von Pythagoras

$$\|u - U_\ell\|^2 = \|u - U_\infty\|^2 + \|U_\infty - U_\ell\|^2.$$

Außerdem ist  $U_\ell$  die Bestapproximation von  $U_\infty$  bezüglich  $X_\ell$ . Da  $\bigcup_{\ell=0}^\infty X_\ell$  in  $X_\infty$  dicht ist, finden wir für jedes  $\varepsilon > 0$  einen Index  $\ell_0$  und ein Element  $V_{\ell_0} \in X_{\ell_0}$  so, dass  $\|U_\infty - V_{\ell_0}\| \leq \varepsilon$  gilt. Für  $\ell \geq \ell_0$  folgt nun mit  $X_{\ell_0} \subseteq X_\ell$ :

$$\|U_\infty - U_\ell\| = \min_{V_\ell \in X_\ell} \|U_\infty - V_\ell\| \leq \|U_\infty - V_{\ell_0}\| \leq \varepsilon.$$

$\square$

Da die Galerkin Lösung nach dem Céa-Lemma die Bestapproximation in der Energienorm darstellt, folgt hieraus die (von der Verfeinerungsstrategie unabhängige) Konvergenz von  $U_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)$  gegen die exakte Lösung  $u \in H_0^1(\Omega)$ . Außerdem müssen wir bei diesem Vorgehen keinerlei Anforderungen an die Oszillationsterme stellen.

### 2.2.5. Optimalität der AFEM

Um unsere Analyse der adaptiven Finite Elemente Methode im unbeschränkten Fall abzuschließen, wollen wir noch eine Aussage über quasi-Optimalität machen. Wie wir zuvor gesehen haben, kann man, selbst bei glatten Problemen, keine bessere Konvergenzrate als  $\mathcal{O}(h)$  erwarten. Empirisch ist die tatsächliche Rate bei uniformer Verfeinerung sogar oftmals schlechter. Es stellt sich daher die Frage, ob adaptive Verfahren bezüglich der Konvergenzordnung nun wirklich besser sind als ihr uniformes Pendant oder ob diese Methoden gar langsamer zum Ergebnis führen. Im Jahre 2008 konnte diese Frage von CASCÓN, KREUZER, NOCHETTO und SIEBERT für die Verfeinerung durch Newest Vertex Bisection positiv beantwortet werden, wobei deren Analyse wiederum auf der Vorarbeit von STEVENSON aus dem Jahre 2005 aufbaut. Es konnte sogar nachgewiesen werden, dass durch AFEM in diesem Fall immer die (für die jeweiligen Eingabedaten) optimale Konvergenzrate erreicht wird.

Im Folgenden ordnen wir die exakte Lösung  $u$  und die Daten  $f$  der Approximationsklasse  $\mathbb{A}_\alpha$  zu, wenn es prinzipiell möglich ist, für dieses Problem, eine Konvergenzrate von  $\alpha$  zu erreichen. Das bedeutet also  $(u, f) \in \mathbb{A}_\alpha$  genau dann, wenn es eine geeignete Wahl von diskreten Räumen  $X_\ell$  und entsprechenden Triangulierungen  $\mathcal{T}_\ell$  gibt, so dass  $\|\nabla(U_{\ell+1} - U_\ell)\| = \mathcal{O}(N_\ell^{-\alpha})$ , wobei  $N_\ell$  die Anzahl der Elemente in  $\mathcal{T}_\ell$  bezeichnet. Es gilt nun:

**Satz 2.2.9** (Optimalität von AFEM für Newest Vertex Bisection). *Sei  $\mathcal{T}_0$  eine Starttriangulierung deren newest vertices sinnvoll nummeriert sind. Ferner seien  $C_{\text{eff}}, C_{\text{rel}} > 0$  die Konstanten der Effizienz und der so genannten diskreten lokalen Zuverlässigkeit des Schätzers. Sei außerdem  $\theta \in (0, 1)$  mit*

$$\theta^2 < \theta_*^2 := \frac{1}{(1 + C_{\text{rel}}^2)C_{\text{eff}}^2} > 0.$$

*Zusätzlich erfülle die Menge der markierten Elemente die Dörflermarkierung mit minimaler Kardinalität. Dann gilt*

$$\left(\|\nabla(u - U_\ell)\|_{L^2(\Omega)}^2 + \gamma\eta_\ell^2\right)^{1/2} \leq C(N_\ell - N_0)^{-\alpha}, \quad (2.2.24)$$

für alle  $\alpha > 0$  mit  $(u, f) \in \mathbb{A}_\alpha$ , wobei  $C > 0$  nur von  $\sigma(\mathcal{T})$  abhängt.

*Beweis.* Siehe [CKNS, Theorem 5.11]. □

Die gemischte Fehlergröße  $\Delta_\ell := \|\nabla(u - U_\ell)\|_{L^2(\Omega)}^2 + \gamma\eta_\ell^2$ , und damit auch der Fehler, fällt also mit optimaler Konvergenzordnung. Für den Beweis dieses Satzes benötigt man eine Kontraktionseigenschaft von  $\Delta_\ell$ , welche in der gleichen Arbeit gezeigt wurde, und eine so genannte *diskrete lokale Zuverlässigkeit*, die im Prinzip eine Zuverlässigkeitsaussage für eine kleinere Menge darstellt. Diese Eigenschaft konnte bislang nur für den Residualschätzer gezeigt werden. Außerdem gehen noch einige Abschätzungen ein, die direkt mit der Verfeinerungsstrategie zu tun haben. Dies ist wohl auch der Grund weshalb die Optimalität bislang nur für Newest Vertex Bisection gezeigt werden konnte. Da Optimalitätsaussagen nicht im Mittelpunkt dieser Arbeit stehen, sparen wir weitere Beweise an dieser Stelle aus und verweisen auf

[CKNS] für nähere Informationen. Wir schließen das Kapitel über die P1 Finite-Elemente-Methode mit dem Denkanstoß ab, dass auch im unbeschränkten Fall noch einige interessante Fragen offen sind. Im Folgenden wollen wir uns nun voll und ganz den Hindernisproblemen widmen.



# Kapitel 3.

## Das affine Hindernisproblem

In diesem Abschnitt betrachten wir das affine Hindernisproblem am Beispiel der Poisson-Gleichung. Wie in der Einleitung erwähnt, stelle man sich eine am Rand eingespannte Membran vor, zum Beispiel ein Tuch, die über einem Hindernis hängt, welches durch eine affine Funktion beschrieben wird. Wir interessieren uns nun dafür, an welchen Stellen das Tuch das affine Hindernis berührt und wie es sich an den Nicht-Kontaktstellen verhält.

### 3.1. Modellproblem

Analog zum vorherigen Kapitel sei  $\Omega$  wieder ein beschränktes, polygonal berandetes Gebiet im  $\mathbb{R}^2$  mit Rand  $\Gamma = \partial\Omega$ . Ferner sei durch die affine Funktion  $\chi : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $\chi|_{\Gamma} \leq 0$  auf  $\overline{\Omega}$  ein Hindernis gegeben. Wir beginnen mit dem unbeschränkten Problem: *Finde  $u \in C^2(\overline{\Omega})$ , so dass*

$$\begin{aligned} -\Delta u &= f && \text{auf } \Omega \\ u &= 0 && \text{auf } \Gamma, \end{aligned} \tag{3.1.1}$$

für  $f \in L^2(\Omega)$ . Wie oben beschrieben, verwenden wir für unseren Finite Elemente Ansatz die variationelle Formulierung, bei der wir ein  $u \in H_0^1(\Omega)$  suchen, welches für alle  $v \in H_0^1(\Omega)$  die Gleichheit

$$\langle\langle u, v \rangle\rangle = (f, v) \tag{3.1.2}$$

erfüllt. Hierbei ist das Energieskalarprodukt  $\langle\langle \cdot, \cdot \rangle\rangle$ , wie zuvor, für alle  $v, w \in H_0^1(\Omega)$  durch

$$\langle\langle v, w \rangle\rangle := (\nabla v, \nabla w)$$

gegeben, und  $(\cdot, \cdot) := (\cdot, \cdot)_{L^2(\Omega)}$  bezeichnet, wie immer, das  $L^2$ -Skalarprodukt auf dem Gebiet  $\Omega$ . Nach (2.1.4) ist dieses Problem äquivalent zur Minimierung des Energiefunktionals

$$\mathcal{J}(v) := \frac{1}{2} \langle\langle v, v \rangle\rangle - (f, v), \tag{3.1.3}$$

über  $H_0^1(\Omega)$ . Im Zuge des Hindernisproblems schränken wir nun die Menge der zulässigen Funktionen auf

$$K := \{v \in H_0^1(\Omega) \mid v \geq \chi \text{ f.ü. in } \Omega\}.$$

ein. Unser Hindernisproblem lautet demnach: *Finde  $u \in K$  mit*

$$\mathcal{J}(u) = \min_{v \in K} \mathcal{J}(v). \tag{3.1.4}$$

Zunächst einmal ist nicht klar, ob das Energiefunktional über  $K$  überhaupt ein eindeutiges Minimum annimmt. Dies folgt jedoch nach einer kleinen Vorüberlegung aus einer Fassung des Satzes von Lax-Milgram.

**Lemma 3.1.1.** *Die Menge  $K$  ist abgeschlossen und konvex in  $H^1(\Omega)$ .*

*Beweis.* Seien  $u, v \in K, t \in (0, 1)$ . Punktweise fast überall gilt

$$tu + (1 - t)v \geq t\chi + (1 - t)\chi = \chi.$$

Damit folgt die Konvexität von  $K$ . Für die Abgeschlossenheit betrachten wir nun eine in  $H^1(\Omega)$  konvergente Folge  $(f_n)_n \subset K$  mit  $\lim_{n \rightarrow \infty} f_n = f$ . Da  $H_0^1(\Omega)$  selbst ein abgeschlossener Unterraum von  $H^1(\Omega)$  ist, gilt  $f \in H_0^1(\Omega)$ . Es bleibt zu zeigen, dass  $f \geq \chi$  f.ü. in  $\Omega$ . Aus der Konvergenz in  $H^1(\Omega)$  folgt im Besonderen die Konvergenz von  $(f_n)_n$  gegen  $f$  in  $L^2(\Omega)$  und damit, aus der Umkehrung des Satzes von Lebesgue, die punktweise Konvergenz einer Teilfolge  $(f_{n_k})_{n_k}$  gegen  $f$  fast überall. Da für  $x \in \mathbb{R}$  die Menge  $[\chi(x), \infty)$  in  $\mathbb{R}$  nach unten abgeschlossen ist folgt nun  $f \geq \chi$  fast überall. Damit ist die Aussage bewiesen.  $\square$

**Satz 3.1.2** (Lax-Milgram, Fassung für konvexe Mengen). *Sei  $V$  eine abgeschlossene, konvexe Menge in einem Hilbertraum  $H$  und  $a : H \times H \rightarrow \mathbb{R}$  eine elliptische Bilinearform. Dann hat das Variationsproblem*

$$\mathcal{J}(v) = \frac{1}{2}a(v, v) - \langle \ell, v \rangle \rightarrow \min!$$

*für jedes  $\ell \in H'$  eine eindeutige Lösung in  $V$ . Hierbei bezeichnet  $\langle \cdot, \cdot \rangle$  die duale Paarung zwischen  $H$  und  $H'$ .*

*Beweis.* Der Beweis findet sich bei BRAESS [BRA07, Seite 37, Satz 2.5].  $\square$

**Korollar 3.1.3.** *Das Hindernisproblem (3.1.4) ist über  $K$  eindeutig lösbar.*

*Beweis.* Das Energieskalarprodukt  $\langle \cdot, \cdot \rangle$  ist nach Korollar 2.1.3 eine elliptische Bilinearform. Mit der Konvexität und Abgeschlossenheit von  $K$  (Lemma 3.1.1), folgt die eindeutige Lösbarkeit aus dem Satz von Lax-Milgram.  $\square$

Wir werden im Folgenden eine zu (3.1.4) äquivalente Formulierung herleiten, welche uns in den späteren Ausführungen sehr nützlich sein wird. Sei  $u$  das Minimum des Energiefunktional über  $K$  und  $v \in K$  eine weitere Funktion. Ferner sei  $0 \leq t \leq 1$ . Aufgrund der Konvexität von  $K$  liegt  $(1 - t)u + tv = u + t(v - u)$  wieder in  $K$ . Die Funktion  $\Phi : [0, 1] \rightarrow \mathbb{R}$ , mit

$$\Phi(t) := J(u + t(v - u)) = \frac{1}{2} \int_{\Omega} (\nabla(u + t(v - u)))^2 - f(u + t(v - u)) dx, \quad 0 \leq t \leq 1$$

nimmt somit ihr Minimum bei  $t = 0$  an. Da dies am linken Rand des Definitionsbereiches von  $\Phi$  liegt, folgt aus den Optimalitätsbedingungen erster Ordnung  $\Phi'(0) \geq 0$ . Durch eine kurze Rechnung erhalten wir

$$\begin{aligned} \Phi'(t) &= \int_{\Omega} (\nabla(u + t(v - u))) \cdot (\nabla(u + t(v - u)))' - (f(u + t(v - u)))' dx \\ &= \int_{\Omega} \nabla(u + t(v - u)) \cdot \nabla(v - u) - f(v - u) dx. \end{aligned}$$

Zusammen mit  $\Phi'(0) \geq 0$  ergibt sich:

$$\langle u, v - u \rangle \geq \langle f, v - u \rangle \quad \text{für alle } v \in K. \quad (3.1.5)$$

Um die Äquivalenz dieser beiden Darstellungen einzusehen, zeigen wir noch die andere Richtung. Es gelte (3.1.5). Damit folgt

$$\begin{aligned}\mathcal{J}(v) - \mathcal{J}(u) &= \langle v - u, u \rangle - (f, v - u) - \frac{1}{2} \langle v - u, u \rangle - \frac{1}{2} \langle v, u \rangle + \frac{1}{2} \langle v, v \rangle \\ &= \langle v - u, u \rangle - (f, v - u) + \underbrace{\frac{1}{2} \langle v - u, v - u \rangle}_{\geq 0} \geq \langle v - u, u \rangle - (f, v - u) \geq 0.\end{aligned}$$

Da  $v$  beliebig war, muss das Minimum von  $\mathcal{J}$  über  $K$  demnach durch  $u$  gegeben sein. Wir können nun das fertige, kontinuierliche Hindernisproblem formulieren.

**Problem 3.1.4** (kontinuierlich). *Finde  $u \in K = \{v \in H_0^1(\Omega) \mid v \geq \chi \text{ f.ü. in } \Omega\}$  mit*

$$\langle u, v - u \rangle \geq (f, v - u) \quad \text{für alle } v \in K. \quad (3.1.6)$$

Für die numerische Behandlung von Problem 3.1.4 im Sinne der Finite-Elemente-Methode, wählen wir den endlich dimensionalen Ansatzraum

$$K_\ell = K \cap \mathcal{S}_0^1(\mathcal{T}_\ell) = \{V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell) \mid V_\ell(x) \geq \chi(x) \text{ für alle } x \in \Omega\},$$

wobei  $\mathcal{S}_0^1(\mathcal{T}_\ell)$  wie in Kapitel 2 definiert ist. Offensichtlich ist auch  $K_\ell$  konvex und abgeschlossen, so dass die gleichen Ausführungen wie oben die eindeutige Lösbarkeit und die Äquivalenz zur Variationsungleichung liefern. Unser endlich dimensionales Teilproblem lautet somit:

**Problem 3.1.5** (diskret). *Finde  $U_\ell \in K_\ell = \{V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell) \mid V_\ell(x) \geq \chi(x) \text{ für alle } x \in \Omega\}$  mit*

$$\langle U_\ell, V_\ell - U_\ell \rangle \geq (f, V_\ell - U_\ell) \quad \text{für alle } V_\ell \in K_\ell. \quad (3.1.7)$$

**Bemerkung.** *Wir halten an dieser Stelle fest, dass eine solche Diskretisierung automatisch konform ist, d.h.  $K_\ell \subseteq K_{\ell+1}$ , wobei  $\mathcal{T}_{\ell+1}$  das durch Verfeinerung aus  $\mathcal{T}_\ell$  hervorgegangene Netz bezeichnet. Die Richtigkeit dieser Aussage erhalten wir aus folgender Überlegung: Für ein  $V_\ell \in K_\ell$  gilt nach Definition  $V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)$  und  $V_\ell \geq \chi$ . Da die FEM Räume  $\mathcal{S}_0^1(\mathcal{T}_\ell)$  offensichtlich geschachtelt sind, folgt  $V_\ell \in \mathcal{S}_0^1(\mathcal{T}_{\ell+1})$  und  $V_\ell \geq \chi$ , aber das ist bereits die Definition von  $K_{\ell+1}$ .*

**Definition.** Wir definieren die *aktive Menge*  $\mathcal{A}$  als die größte offene Teilmenge von  $\Omega$  so, dass  $u(x) = \chi(x)$  fast überall in  $\mathcal{A}$  gilt. Ferner definieren wir die *inaktive Menge*  $\mathcal{I}$  durch  $\mathcal{I} := \bigcup_{\varepsilon > 0} B_\varepsilon$ , wobei  $B_\varepsilon$  die größte offene Menge bezeichnet, auf der  $u(x) \geq \chi(x) + \varepsilon$  gilt. Als diskrete Pendanten definieren wir

$$\mathcal{A}_\ell := \{x \in \bar{\Omega} \mid U_\ell(x) = \chi(x)\} \quad (3.1.8)$$

und  $\mathcal{I}_\ell := \bar{\Omega} \setminus \mathcal{A}_\ell$ .

**Bemerkung.** Mit Standardmethoden der Optimierung [IK08, Kapitel 1] kann (3.1.6) unter Einführung eines Lagrange'schen Multiplikators  $\sigma \in H^{-1}(\Omega)$  als unrestringiertes Problem umgeschrieben werden. Es gilt dann

$$\langle\langle u, v \rangle\rangle = (f, v) - \langle \sigma, v \rangle, \quad \text{für alle } v \in H_0^1(\Omega), \quad (3.1.9)$$

wobei  $\langle \cdot, \cdot \rangle$  die duale Paarung zwischen  $H_0^1(\Omega)$  und  $H^{-1}(\Omega)$  bezeichnet. Ferner erhalten wir die Komplementaritätsbedingung

$$\langle \sigma, u - \chi \rangle = 0. \quad (3.1.10)$$

Für das diskrete Problem (3.1.7) ergibt sich analog

$$\langle\langle U_\ell, V_\ell \rangle\rangle = (f, V_\ell) - \langle \sigma_\ell, V_\ell \rangle, \quad \text{für alle } V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell), \quad (3.1.11)$$

wobei  $\langle \cdot, \cdot \rangle$  auch hier die entsprechende duale Paarung bezeichnet. Im diskreten Fall ist nun  $\sigma_\ell \in C(\overline{\Omega})'$  ein Radonmaß, welches sich als Linearkombination von Diracmaßen in der aktiven Menge darstellen lässt (vgl. Lemma 4.1.3). Für Funktionen  $V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)$ , deren Träger kompakt in der inaktiven Menge liegt, gilt daher

$$\langle\langle U_\ell, V_\ell \rangle\rangle = (f, V_\ell). \quad (3.1.12)$$

Innerhalb der inaktiven Menge wird also das unrestringierte Problem gelöst.

Wir halten an dieser Stelle fest, dass es sich, im Gegensatz zum unrestringierten Fall, bei den Problemen 3.1.4 und 3.1.5 um nichtlineare Probleme handelt. Dies kann man zum einen einfach dadurch einsehen, dass der Ansatzraum der zulässigen Funktionen  $K$  bzw.  $K_\ell$  kein Vektorraum ist. Zum anderen macht man sich die Gültigkeit der Aussage leicht an folgender kurzer Rechnung klar. Sei  $u \in K$  die eindeutige Lösung von

$$\langle\langle u, v - u \rangle\rangle \geq (f, v - u) \quad \text{für alle } v \in K,$$

und  $w \in K$  die eindeutige Lösung von

$$\langle\langle w, v - w \rangle\rangle \geq (g, v - w) \quad \text{für alle } v \in K.$$

Dann folgt aber für die Summe  $u + w$

$$\begin{aligned} \langle\langle u + w, v - (u + w) \rangle\rangle &= \langle\langle u, v - (u + w) \rangle\rangle + \langle\langle w, v - (u + w) \rangle\rangle \\ &= \langle\langle u, v - u \rangle\rangle + \langle\langle w, v - w \rangle\rangle - 2\langle\langle u, w \rangle\rangle \\ &\geq (f + g, v - (u + w)), \end{aligned}$$

und somit die Nichtlinearität der betrachteten Probleme.

In den folgenden Abschnitten werden wir einen konvergenten, adaptiven FEM Algorithmus für das Hindernisproblem herleiten. Hierzu verwenden wir zunächst die Methode aus [BCH1], indem wir eine Kontraktion des Fehlers im Energiefunktional, und unter gewissen Voraussetzungen an die Datenoszillationen, die Konvergenz des Fehlers gegen Null in Energienorm zeigen. Der Beweis erfordert die Zuverlässigkeit des Fehlerschätzers und seine diskrete lokale Effizienz. Dieses Vorgehen hat jedoch zwei Nachteile. Erstens muss man zusätzlich zum Fehler dafür sorgen, dass auch die Datenoszillationen kleiner werden und zweitens gilt der Beweis nur

für Verfeinerungsstrategien, die mindestens einen neuen inneren Punkt im zu verfeinernden Dreieck erzeugen.

Im Anschluss daran verfolgen wir ein anderes Vorgehen. Wir wenden die Schätzerreduktion aus [CKNS, AFP] auf das Hindernisproblem an. Mit einigen weiteren Überlegungen können wir so die Notwendigkeit der diskreten lokalen Effizienz umgehen und erhalten einen von der Verfeinerungsstrategie unabhängigen Konvergenzbeweis. Zudem werden wir zeigen, dass es auch möglich ist, die Einschränkung an die Datenoszillationen fallen zu lassen.

Schließlich werden wir auch dieses Resultat noch verbessern indem wir im Sinne von [CKNS] eine Kontraktion der gemischten Größe  $\Delta_\ell = e_\ell + \gamma \eta_\ell^2$  zeigen, wobei  $e_\ell := \mathcal{J}(U_\ell) - \mathcal{J}(u)$  den Fehler im Energiefunktional beschreibt. Hiermit haben wir sogar eine Aussage über die Konvergenzgeschwindigkeit in der Hand. Außerdem ist eine solche Kontraktion einer der Bestandteile für einen möglichen Optimalitätsbeweis für unseren Algorithmus.

Analog zum linearen Fall verwenden wir eine im Sinne von Ciarlet reguläre Triangulierung und verfeinern diese in jedem Schritt adaptiv. Wie oben bezeichnen wir die Menge der inneren Knoten mit  $\mathcal{N}_\ell$  und die Menge der inneren Kanten mit  $\mathcal{E}_\ell$ . Die Menge aller Kanten bezeichnen wir mit  $\mathcal{E}_\ell^*$ . Im Besonderen gilt daher für die Randkanten  $\mathcal{E}_{\ell,\Gamma} = \mathcal{E}_\ell^* \setminus \mathcal{E}_\ell$ . Die zwei elementige Schicht von Randelementen bezeichnen wir mit  $\mathcal{T}_{\ell,\Gamma} := \{T \in \mathcal{T}_\ell \mid T \cap \Gamma \neq \emptyset\}$ , wobei wir mit  $\mathcal{T}_{\ell,c} := \{T \in \mathcal{T}_\ell \mid T \cap \Gamma \neq \emptyset\}$  diejenigen Elemente mit direktem Randkontakt meinen. Schließlich steht  $\Omega_E := T_+ \cup T_-$  für den Patch aus den Dreiecken  $T_+$  und  $T_-$  mit der gemeinsamen Kante  $E = T_+ \cap T_- \in \mathcal{E}_\ell$ .

### 3.2. Ein erster konvergenter Algorithmus

Für die Gesamtdauer dieses Kapitels sei  $u \in K$  die exakte Lösung von Problem 3.1.4 und  $U_\ell \in K_\ell$  die Lösung des diskreten Problems 3.1.5. Analog zum adaptiven Algorithmus 2.2.4 aus Kapitel 2, suchen wir auch hier einen Fehlerschätzer  $\eta_\ell$ , der uns in jedem Schritt angibt, an welchen Stellen das Netz zu verfeinern ist. Etwas konkreter suchen wir eine Größe, die eine *Zuverlässigkeitsabschätzung*

$$\|u - U_\ell\|^2 \lesssim \eta_\ell^2 \tag{3.2.1}$$

und optimalerweise auch eine *Effizienzabschätzung*

$$\eta_\ell^2 \lesssim (\|u - U_\ell\| + \text{osc}_\ell)^2$$

erfüllt. Es stellt sich die Frage, weshalb man nicht einfach den Residualschätzer aus dem linearen Fall verwenden kann. Diese Frage macht zwei tiefgreifende Unterschiede zum unrestringierten Problem deutlich. Zum einen ist die Menge der zulässigen Funktionen  $K$  bzw. ihr diskretes Pendant  $K_\ell$  offensichtlich kein Hilbertraum. Somit findet der Satz von Riesz (B.1.5) keine Anwendung und wir verlieren die wichtige Eigenschaft der *Galerkin-Orthogonalität*. Zum anderen ist die Gleichung (3.1.1) bzw. (3.1.2) nur an den Nicht-Kontaktstellen erfüllt und gilt somit im Allgemeinen nicht mehr. Da diese beiden Aussagen essentielle Bestandteile zu den Beweisen in Kapitel 2 sind, ist die Theorie nicht ohne Weiteres auf den restringierten Fall übertragbar.

Es ist allerdings möglich die gesuchten Aussagen in ähnlicher Form für eine sehr ähnliche

Größe zu beweisen. Wir betrachten im Folgenden den Fehlerschätzer

$$\eta_\ell^2 := \sum_{E \in \mathcal{E}_\ell} \eta_\ell(E)^2, \quad (3.2.2)$$

mit

$$\eta_\ell(E) := h_E^{1/2} \|[\partial_n U_\ell]\|_{L^2(E)} \quad \text{für } E \in \mathcal{E}_\ell. \quad (3.2.3)$$

Hierbei bezeichnet  $h_E$  die Länge der Kante  $E$  und  $[\cdot]$ , wie oben, (2.2.10) den Sprung der Normalenableitung über die Kante  $E$ . Desweiteren betrachten wir die Oszillationsterme

$$\text{osc}_\ell := \left( \sum_{E \in \mathcal{E}_\ell} \text{osc}_\ell(E)^2 \right)^{1/2}, \quad (3.2.4)$$

$$\text{Osc}_\ell := \left( \text{osc}_\ell^2 + \sum_{T \in \mathcal{T}_{\ell,\Gamma}} h_T^2 \|f\|_{L^2(T)}^2 \right)^{1/2}, \quad (3.2.5)$$

mit  $\text{osc}_\ell(E) := |\Omega_E|^{1/2} \|f - f_{\Omega_E}\|_{L^2(\Omega_E)}$ ,  $f_{\Omega_E} := |\Omega_E|^{-1} \int_{\Omega_E} f dx$  der  $L^2$ -Bestapproximation und  $h_T := \text{diam}(T)$  dem Durchmesser des Elements  $T$ . Da unser Fehlerschätzer nur die Kantensprünge berücksichtigt müssen wir für einen konvergenten Algorithmus zusätzlich fordern, dass die Oszillationen beim Übergang von  $\mathcal{T}_\ell$  auf das verfeinerte Gitter  $\mathcal{T}_{\ell+1}$  reduziert werden. Genauer gesagt fordern wir

$$\text{Osc}_{\ell+1} \leq \kappa \text{Osc}_\ell \quad (3.2.6)$$

für ein  $\kappa \in (0, 1)$ . Diese Eigenschaft wird, wenn nötig, durch zusätzliche Verfeinerungen erzwungen. Dass dies möglich ist, sieht man für die Elementoszillationen  $\text{osc}_\ell(T)^2 := h_T^2 \|f\|_{L^2(T)}^2$  an dem folgenden Lemma, welches bereits in [MNS] gezeigt wurde.

**Lemma 3.2.1.** *Sei  $0 < \gamma < 1$  der Faktor um den sich die Elementgröße in jedem Verfeinerungsschritt verkleinert, d.h.  $h_{T'} \leq \gamma h_T$  wobei  $T'$  einer der Söhne von  $T$  ist. Für  $\hat{\theta} \in (0, 1)$  definieren wir  $\hat{\alpha} := (1 - (1 - \gamma^2)\hat{\theta}^2) \in (0, 1)$ . Sei  $\hat{\mathcal{T}}_{\ell,\Gamma}$  eine Teilmenge von  $\mathcal{T}_{\ell,\Gamma}$  von zu verfeinernden Elementen, für die gilt*

$$\sum_{T \in \hat{\mathcal{T}}_{\ell,\Gamma}} \text{osc}_\ell(T)^2 \geq \hat{\theta} \sum_{T \in \mathcal{T}_{\ell,\Gamma}} \text{osc}_\ell(T)^2, \quad (3.2.7)$$

dann folgt

$$\sum_{T' \in \mathcal{T}_{\ell+1,\Gamma}} \text{osc}_{\ell+1}(T')^2 \leq \hat{\alpha} \sum_{T \in \mathcal{T}_{\ell,\Gamma}} \text{osc}_\ell(T)^2. \quad (3.2.8)$$

*Beweis.* Wir betrachten die zu verfeinernden Elemente separat und erhalten

$$\begin{aligned} \sum_{T' \in \mathcal{T}_{\ell+1,\Gamma}} \text{osc}_{\ell+1}(T')^2 &\leq \sum_{T \in \mathcal{T}_{\ell,\Gamma} \setminus \hat{\mathcal{T}}_{\ell,\Gamma}} h_T^2 \|f\|_{L^2(T)}^2 + \gamma^2 \sum_{T \in \hat{\mathcal{T}}_{\ell,\Gamma}} h_T^2 \|f\|_{L^2(T)}^2 \\ &= \sum_{T \in \mathcal{T}_{\ell,\Gamma}} \text{osc}_\ell(T)^2 + (\gamma^2 - 1) \sum_{T \in \hat{\mathcal{T}}_{\ell,\Gamma}} \text{osc}_\ell(T)^2 \\ &\leq (1 - (1 - \gamma^2)\hat{\theta}^2) \sum_{T \in \mathcal{T}_{\ell,\Gamma}} \text{osc}_\ell(T)^2, \end{aligned}$$

wobei wir hier verwendet haben, dass jeder Vater  $T \in \mathcal{T}_{\ell,\Gamma}$  die disjunkte Vereinigung seiner Söhne  $T' \in \mathcal{T}_{\ell+1,\Gamma}$  ist.  $\square$

Da der Beweis für die Kantenoszillationen weitaus komplizierter ist, sei er an dieser Stelle auf später verschoben (vgl. Korollar 3.3.7). Grundlegende Voraussetzung ist jedoch wieder, dass die zu verfeinernde Kantenmenge  $\widehat{\mathcal{E}}_\ell \subseteq \mathcal{E}_\ell$  für die Kantenoszillationen eine Dörflermarkierung der Form

$$\sum_{E \in \widehat{\mathcal{E}}_\ell} \text{osc}_\ell(E)^2 \geq \widehat{\theta} \sum_{E \in \mathcal{E}_\ell} \text{osc}_\ell(E)^2 \quad (3.2.9)$$

erfüllt. Auch für die Kantenoszillationen muss also gegebenenfalls weiter verfeinert werden. Wir können nun die erste Version eines konvergenten, adaptiven FEM Algorithmus für das elliptische Hindernisproblem formulieren.

**Algorithmus 3.2.2** (adaptiver Algorithmus, Version 1). **Input:** Anfangstriangulierung  $\mathcal{T}_0$ , maximale Anzahl Elemente  $n_{E_{max}}$ , Adaptivitätsparameter  $\theta, \theta \in (0, 1), \ell := 0$ .

(i) Berechne diskrete Lösung  $U_\ell \in K_\ell$ .

(ii) Berechne die Verfeinerungsindikatoren  $\eta_\ell(E)$  und den Fehlerschätzer  $\eta_\ell$ .

(iii) STOP, wenn die Anzahl der Elemente größer als  $n_{E_{max}}$  ist.

(iv) Finde eine Menge  $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell$  minimaler Kardinalität von markierten Kanten, so dass

$$\theta \eta_\ell^2 = \theta \sum_{E \in \mathcal{E}_\ell} \eta_\ell(E)^2 \leq \sum_{E \in \mathcal{M}_\ell} \eta_\ell(E)^2.$$

(v) Erweitere die Menge  $\mathcal{M}_\ell \subseteq \mathcal{M}_\ell \cup \mathcal{T}_{\ell,\Gamma}$  ggf. so, dass

$$\widehat{\theta} \sum_{T \in \mathcal{T}_{\ell,\Gamma}} \text{osc}_\ell(T)^2 \leq \sum_{T \in \mathcal{M}_\ell \cap \mathcal{T}_{\ell,\Gamma}} \text{osc}_\ell(T)^2$$

und

$$\widehat{\theta} \sum_{E \in \mathcal{E}_\ell} \text{osc}_\ell(E)^2 \leq \sum_{E \in \mathcal{M}_\ell \cap \mathcal{E}_\ell} \text{osc}_\ell(E)^2.$$

(vi) Erzeuge neues, reguläres Gitter  $\mathcal{T}_{\ell+1}$ , bei dem mindestens alle markierten Elemente und Kanten verfeinert wurden. Hierbei soll in jedem verfeinerten Element mindestens ein neuer, innerer Knoten entstehen.

(vii)  $\ell \mapsto \ell + 1$  und beginne wieder bei (i).

**Output:** Endliche Folge diskreter Lösungen  $U_\ell$  und zugehörige Fehlerschätzer  $\eta_\ell$ .

**Bemerkung.** Ist in (v) eine Kante  $E = T_+ \cap T_-$  zur Verfeinerung ausgewählt worden, so werden  $T_\pm$  nach der in (vi) genannten Regel verfeinert.

Wir werden im Folgenden zeigen, dass die Folge der diskreten Lösungen  $(U_\ell)_{\ell \in \mathbb{N}}$  aus obigem Algorithmus tatsächlich in einer geeigneten Norm gegen die exakte Lösung  $u \in K$  konvergiert. Der Beweis erfordert zwei Zutaten, die Zuverlässigkeit und die diskrete lokale Effizienz des Schätzers  $\eta_\ell$ . Beide Eigenschaften werden in den nächsten Abschnitten gezeigt.

### 3.2.1. Zuverlässigkeit

Um die oben erwähnte, verloren gegangene Galerkin-Orthogonalität zu kompensieren, zeigen wir die Zuverlässigkeit nicht wie in (3.2.1) bezüglich der Energienorm  $\|\cdot\|$ , sondern bezüglich des Fehlers im Energiefunktional  $e_\ell := \mathcal{J}(U_\ell) - \mathcal{J}(u)$ . Der Beweis basiert auf einem Korollar einer allgemeineren Aussage von [BC], die wir hier nicht beweisen wollen. Hierzu betrachten wir kurzzeitig das inhomogene Dirichletproblem mit Randbedingungen  $u|_\Gamma = u_D$ . Ferner sei  $\chi \in H^1(\Omega)$  und  $\chi_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$  eine Approximation von  $\chi$ . Der Elementpatch bezüglich  $T$  wird mit  $\omega_T$  bezeichnet. Es gilt die folgende Aussage:

**Korollar 3.2.3.** *Für alle  $\omega \in H^1(\Omega)$  mit  $\omega|_\Gamma = u_D - U_\ell|_\Gamma$  und  $\chi - U_\ell \leq \omega$  gilt*

$$\begin{aligned} & \|\omega - U_\ell - \omega\| + \|u - U_\ell\| + \langle \sigma, u - U_\ell \rangle^{1/2} \leq C \min_{p_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)^d} (\|\nabla U_\ell - p_\ell\|_{L^2(\Omega)}) \\ & + \min_{q_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)^d} \|\nabla(\chi_\ell - U_\ell) - q_\ell\|_{L^2(\Omega)} \\ & + \left( \sum_{T \in \mathcal{T}_\ell} \|f\|_{L^2(\omega_T)} \|(\chi - \chi_\ell - \omega)_-\|_{L^2(T)} \right)^{1/2} + \|\omega\| \\ & + \left( \sum_{z \in \mathcal{N}_\ell} h_z^2 \min_{q \in \mathbb{R}} \|f - q\|_{L^2(\Omega_z)}^2 \right)^{1/2} \\ & + \left( \sum_{T \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell,\Gamma}} \|h_T^{-1}(\chi - \chi_\ell - \omega)_-\|_{L^2(T)}^2 \right)^{1/2} + \left( \sum_{T \in \mathcal{T}_{\ell,\Gamma}} h_T^2 \|f\|_{L^2(\omega_T)}^2 \right)^{1/2}, \end{aligned}$$

wobei die Konstante  $C > 0$  nur von der Formregularität und  $\text{diam}(\Omega)$  abhängt.

*Beweis.* Der Beweis findet sich bei BARTELS und CARSTENSEN [BC, Theorem 3]. Für das vollständige Theorem, siehe auch Anhang B.2.  $\square$

Bevor wir uns der Zuverlässigkeit von  $\eta_\ell$  widmen benötigen wir noch zwei kleine Lemmata. Das erste beschreibt hierbei die Äquivalenz von  $\eta_\ell$  zu einem anderen Fehlerschätzer, während das zweite eher technischer Natur ist.

**Lemma 3.2.4.** *Sei  $\tilde{\eta}_\ell := \min_{q \in \mathcal{S}^1(\mathcal{T}_\ell)^d} \|\nabla U_\ell - q\|_{L^2(\Omega)}$ . Es gilt*

$$\eta_\ell \sim \tilde{\eta}_\ell, \quad \text{d.h.} \quad \eta_\ell \leq C_1 \tilde{\eta}_\ell \quad \text{und} \quad \tilde{\eta}_\ell \leq C_2 \eta_\ell, \quad (3.2.10)$$

für Konstanten  $C_1, C_2 > 0$  die unabhängig von der Größe der Elemente in  $\mathcal{T}_\ell$  sind (aber von ihrer Form abhängen können).

*Beweis.* Nachzulesen in CARSTENSEN [C, Theorem 3.2 & 3.3].  $\square$

**Lemma 3.2.5.** *Es bezeichne  $\text{osc}_\ell(E) = \sum_{E \in \mathcal{E}_\ell} |\Omega_E| \|f - f_{\Omega_E}\|_{L^2(\Omega_E)}^2$  die Kantenoszillationen und  $\text{osc}_\ell(K) = \sum_{z \in \mathcal{N}_\ell} |\Omega_z| \|f - f_{\Omega_z}\|_{L^2(\Omega_z)}^2$  die Knotenoszillationen von  $f$ . Für stückweise polynomiale  $f$  und Verfeinerung durch newest vertex bisection sind diese beiden Größen äquivalent, d.h.*

$$\text{osc}_\ell(E) \leq C_1 \text{osc}_\ell(K) \leq C_2 \text{osc}_\ell(E), \quad (3.2.11)$$

wobei die Konstanten  $C_1, C_2$  nur von der Formregularität der Triangulierung und dem Polynomgrad von  $f$  abhängen.

*Beweis.* Wir zeigen beide Richtungen einzeln. Die untere Abschätzung gilt immer. Für eine Kante  $E \in \mathcal{E}_\ell$  und einen Knoten  $z \in \mathcal{N}_\ell$  mit  $|\Omega_E| \subseteq |\Omega_z|$  gilt nach der Projektionseigenschaft des Integralmittels

$$|\Omega_E| \|f - f_{\Omega_E}\|_{L^2(\Omega_E)}^2 \leq |\Omega_z| \|f - f_{\Omega_z}\|_{L^2(\Omega_z)}^2 \leq |\Omega_z| \|f - f_{\Omega_z}\|_{L^2(\Omega_z)}^2.$$

Mit der endlichen Überlappung der involvierten Patches erhalten wir

$$\sum_{\substack{E \in \mathcal{E}_\ell \\ \Omega_E \subseteq \Omega_z}} |\Omega_E| \|f - f_E\|_{L^2(\Omega_E)}^2 \leq C_{\text{patch}} |\Omega_z| \|f - f_{\Omega_z}\|_{L^2(\Omega_z)}^2$$

und daher insgesamt

$$\begin{aligned} \sum_{E \in \mathcal{E}_\ell} |\Omega_E| \|f - f_{\Omega_E}\|_{L^2(\Omega_E)}^2 &\leq \sum_{z \in \mathcal{N}_\ell} \sum_{\substack{E \in \mathcal{E}_\ell \\ \Omega_E \subseteq \Omega_z}} |\Omega_E| \|f - f_{\Omega_E}\|_{L^2(\Omega_E)}^2 \\ &\leq C_{\text{patch}} \sum_{z \in \mathcal{N}_\ell} |\Omega_z| \|f - f_{\Omega_z}\|_{L^2(\Omega_z)}^2. \end{aligned}$$

Die obere Abschätzung erhalten wir mittels Skalierungsargument: Wir definieren die Größen

$$|f|_1 := |\Omega_z|^{1/2} \|f - f_{\Omega_z}\|_{L^2(\Omega_z)} \quad \text{und} \quad |f|_2 := \left( \sum_{\substack{E \in \mathcal{E}_\ell \\ \Omega_E \subseteq \Omega_z}} |\Omega_E| \|f - f_{\Omega_E}\|_{L^2(\Omega_E)}^2 \right)^{1/2}.$$

Hierbei handelt es sich offensichtlich um Halbnormen auf  $\mathcal{P}^p(\mathcal{T}_\ell|_{\Omega_z})$ , wobei  $\mathcal{P}^p$  die stückweise polynomiellen Funktionen bezeichnet. Die beiden Halbnormen haben ferner den gleichen Kern, da  $|f|_1 = 0$  direkt impliziert, dass  $f \equiv c$  auf  $\Omega_z$  konstant ist. Analog sieht man unter Zuhilfenahme der Überlappung der einzelnen  $\Omega_E$ , dass auch  $|f|_2 = 0$  nur für konstante Funktionen gilt. Nach Lemma B.1.6 erhalten wir somit

$$|\cdot|_1 \sim |\cdot|_2,$$

wobei die implizit enthaltene Konstante  $C_{\Omega_z} > 0$  vom entsprechenden Polynomgrad und der Form des Patches abhängen kann. Die Unabhängigkeit von  $\text{diam}(\Omega_z)$  erhält man durch Skalierung auf einen Patch bekannter Größe (vgl. *Skalierungsargument*, [BRA07, Kapitel II, §6]). Durch Verfeinerung mit newest vertex bisection entstehen aber, ausgehend von der Starttriangulierung  $\mathcal{T}_0$  nur endlich viele verschiedene Dreiecke und damit auch nur endlich viele

verschiedene Patches  $\Omega_z$  und endlich viele verschiedene Konstanten  $C_{\Omega_z}$ . Insgesamt erhalten wir somit

$$\begin{aligned} \sum_{z \in \mathcal{N}_\ell} |\Omega_z| \|f - f_{\Omega_z}\|_{L^2(\Omega_z)}^2 &\leq C \sum_{z \in \mathcal{N}_\ell} \sum_{\substack{E \in \mathcal{E}_\ell \\ \Omega_E \subseteq \Omega_z}} |\Omega_E| \|f - f_{\Omega_E}\|_{L^2(\Omega_E)}^2 \\ &\leq C C_{\text{patch}} \sum_{E \in \mathcal{E}_\ell} |\Omega_E| \|f - f_{\Omega_E}\|_{L^2(\Omega_E)}^2, \end{aligned}$$

mit  $C = \max\{C_{\Omega_z}\} > 0$ . Damit ist die Aussage vollständig bewiesen.  $\square$

**Satz 3.2.6** (Zuverlässigkeit). *Seien  $\eta_\ell$  und  $Osc_\ell$  wie oben beschrieben, dann gilt*

$$e_\ell^2 \leq C (\eta_\ell^2 + Osc_\ell^2) \quad (3.2.12)$$

für stückweise polynomiale  $f$  und Verfeinerung durch newest vertex bisection. Die Konstante  $C > 0$  ist hierbei unabhängig von der Größe der Elemente in  $\mathcal{T}_\ell$ .

*Beweis.* Wir verwenden den Fehlerschätzer aus Lemma 3.2.4

$$\tilde{\eta}_\ell := \min_{q_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)^d} \|\nabla U_\ell - q_\ell\|_{L^2(\Omega)}$$

und beginnen mit der Abschätzung aus Korollar 3.2.3. Da wir uns in unseren Untersuchungen auf homogene Dirichletrandbedingungen und affine Hindernisse beschränken, gilt  $\chi = \chi_\ell$  und wir können  $\omega \equiv 0$  wählen. Damit fällt die Abschätzung zu

$$\begin{aligned} \|u - U_\ell\|_\Omega + \langle \sigma, u - U_\ell \rangle^{1/2} &\lesssim \tilde{\eta}_\ell + \min_{q_h \in \mathcal{S}^1(\mathcal{T}_\ell)} \|\nabla(\chi_\ell - U_\ell) - q_h\|_{L^2(\Omega)} \\ &\quad + \left( \sum_{E \in \mathcal{E}_\ell} h_E^2 \min_{z \in \mathbb{R}} \|f - z\|_{L^2(\Omega_E)}^2 \right)^{1/2} + \left( \sum_{T \in \mathcal{T}_{\ell, \Gamma}} h_T^2 \|f\|_{L^2(\omega_T)}^2 \right)^{1/2} \end{aligned}$$

zusammen, wobei wir hier außerdem  $h_E^2 \sim h_z^2$  und Lemma 3.2.5 verwendet haben. Aufgrund der gleichmäßigen Formregularität des Netzes  $\mathcal{T}_\ell$  gilt nun

$$\left( \sum_{T \in \mathcal{T}_{\ell, \Gamma}} h_T^2 \|f\|_{L^2(\omega_T)}^2 \right)^{1/2} \lesssim \left( \sum_{T \in \mathcal{T}_{\ell, \Gamma}} h_T^2 \|f\|_{L^2(T)}^2 \right)^{1/2}. \quad (3.2.13)$$

Da  $\chi \equiv \chi_\ell$  global affin ist, folgt  $\nabla \chi_\ell \equiv c$  und daher mit  $q_h = c - q$ , wobei  $q \in \mathcal{S}^1(\mathcal{T}_\ell)^d$  den Schätzer  $\tilde{\eta}_\ell$  minimiert,

$$\min_{q_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)} \|\nabla(\chi_\ell - U_\ell) - q_\ell\|_{L^2(\Omega)} = \tilde{\eta}_\ell.$$

Insgesamt haben wir also

$$\|u - U_\ell\|_\Omega + \langle \sigma, u - U_\ell \rangle^{1/2} \lesssim \tilde{\eta}_\ell + \left( \sum_{E \in \mathcal{E}_\ell} h_E^2 \min_z \|f - z\|_{L^2(\Omega_E)}^2 \right)^{1/2} + \left( \sum_{T \in \mathcal{T}_{\ell, \Gamma}} h_T^2 \|f\|_{L^2(T)}^2 \right)^{1/2}.$$

Wegen  $h_E^2 \sim |\Omega_E|$  und der Cauchy'schen Ungleichung folgt

$$\|u - U_\ell\| + \langle \sigma, u - U_\ell \rangle^{1/2} \lesssim \tilde{\eta}_\ell + Osc_\ell,$$

und damit

$$\|u - U_\ell\|^2 + \langle \sigma, u - U_\ell \rangle \lesssim \tilde{\eta}_\ell^2 + Osc_\ell^2.$$

Nach Definition von  $\mathcal{J}$  erhalten wir für alle  $v \in K$

$$\begin{aligned} \mathcal{J}(v) - \mathcal{J}(u) &= \frac{1}{2} \langle v, v \rangle - (f, v) - \left( \frac{1}{2} \langle u, u \rangle - (f, u) \right) \\ &\quad + (\langle u, u - v \rangle - (f, u - v) - \langle \sigma, u - v \rangle) \\ &= \frac{1}{2} \|v - u\|^2 + \langle \sigma, v - u \rangle. \end{aligned} \quad (3.2.14)$$

Aus obigen Überlegungen erhalten wir für  $v = U_\ell$  eine a posteriori Abschätzung für den Fehler im Energiefunktional

$$\mathcal{J}(U_\ell) - \mathcal{J}(u) \lesssim \tilde{\eta}_\ell^2 + Osc_\ell^2.$$

Schließlich liefert die Äquivalenz (3.2.10) die gesuchte Aussage.  $\square$

**Bemerkung.** Die Originalarbeit [BCH1] kommt an dieser Stelle ohne die zusätzliche Anforderung an  $f$  aus. Ferner besteht die Menge  $\mathcal{T}_{\ell, \Gamma}$  in dieser Arbeit tatsächlich nur aus denjenigen Elementen die den Rand berühren. Der Originalbeweis zur Zuverlässigkeit ist allerdings fragwürdig, da in diesem Fall die Abschätzung 3.2.13 nicht einsichtlich ist. Außerdem ist unklar wie man ohne die Anforderung an stückweise polynomiales  $f$  die Knotenoszillationen aus Korollar 3.2.3 gegen die gewünschten Kantenoszillationen abschätzen kann. In der Originalarbeit finden sich leider keine Hinweise zur Auflösung dieser Problemstellen.

**Bemerkung.** Die zusätzliche Bedingung an  $f$  ist auf den ersten Blick etwas ärgerlich. Da sich jedoch viele Funktionen beliebig gut durch Polynome approximieren lassen, stellt dies in der Praxis keine große Einschränkung dar.

Alle oben genannten Probleme lassen sich übrigens umgehen, wenn man statt  $\eta_\ell$  den aus dem linearen Fall bekannten Fehlerschätzer

$$\hat{\eta}_\ell := \sum_{E \in \mathcal{E}_\ell} h_E \|\partial_n U_\ell\|_{L^2(E)}^2 + \sum_{T \in \mathcal{T}_\ell} |T| \|f\|_{L^2(T)}^2 \quad (3.2.15)$$

verwendet. Für diesen Schätzer lässt sich die Zuverlässigkeit nämlich zeigen, indem man den Term

$$h_z^2 \min_{q \in \mathbb{R}} \|f - q\|_{L^2(\Omega_z)}^2$$

in Korollar 3.2.3 einfach durch  $h_z^2 \|f\|_{L^2(\Omega_z)}^2$  abschätzt. Die Zuverlässigkeit folgt dann ohne weitere Einschränkungen und alle Aussagen aus den nachstehenden Abschnitten lassen sich direkt übernehmen. Da der resultierende Fehlerschätzer allerdings viel gröber wäre, wollen wir weiterhin mit  $\eta_\ell$  arbeiten.

### 3.2.2. Diskrete lokale Effizienz

Der zweite wichtige Bestandteil für den Konvergenzbeweis unseres Algorithmus' ist die diskrete lokale Effizienz. Hierbei suchen wir eine Abschätzung bei der der Fehlerschätzer  $\eta_\ell$  durch den diskreten Term  $\|U_{\ell+1} - U_\ell\|$  dominiert wird. An dieser Stelle sei noch einmal angemerkt, dass hier die in Algorithmus 3.2.2 geforderte *innerer-Punkt-Eigenschaft* massiv eingeht. Konkret fordern wir, dass ein zu verfeinerndes Element  $T$  so verfeinert wird, dass im Inneren

von  $T$  mindestens ein neuer Knoten entsteht. Dies kann zum Beispiel durch wiederholte Bisektion geschehen (vgl. Bemerkung am Ende von Abschnitt 2.2.3). Da die diskrete lokale Effizienz eine Eigenschaft des Fehlerschätzers  $\eta_\ell$  ist, gelten diese Einschränkungen natürlich nur für diejenigen markierten Elemente, die aufgrund der Kantensprünge ausgewählt wurden. Für eine markierte Kante  $E = T_+ \cap T_-$  setzen wir nun  $P := \text{mid}(E) \in \mathcal{N}_{\ell+1} \setminus \mathcal{N}_\ell$  und bezeichnen die neuen inneren Punkte von  $T_\pm$  mit  $P_\pm \in \mathcal{N}_{\ell+1} \setminus \mathcal{N}_\ell$  (siehe Abbildung 3.1). Der Beweis unterscheidet verschiedene Fälle, in Abhängigkeit davon ob die neuen inneren Punkte  $P_\pm$ , bzw. der neue Punkt  $P$  in der Kontaktzone liegt oder nicht. Wir untersuchen also die Differenzenfunktion

$$0 \leq \delta_{\ell+1} := U_{\ell+1} - \chi \in H^1(\Omega).$$

Bevor wir uns die einzelnen Fälle in den Propositionen 3.2.8–3.2.11 anschauen, noch ein vorbereitendes Lemma.

**Lemma 3.2.7.** *Seien  $P_+$  und  $P_-$  die neuen inneren Punkte mit nodalen Basisfunktionen  $\varphi_{P_+}$  und  $\varphi_{P_-}$ . Der Träger dieser Funktionen ist somit eine Teilmenge von  $T_+$  bzw.  $T_-$ . Ferner sei  $\delta_{\ell+1}(P_+) = 0$  oder  $\delta_{\ell+1}(P_-) = 0$ . Für alle  $A \in \mathbb{R}^2$  gilt dann*

$$h_E^{1/2} \|A \cdot n_E\|_{L^2(E)} \leq C \|\nabla \delta_{\ell+1} - A\|_{L^2(T_+)}, \quad (3.2.16)$$

wobei  $n_E$  den äußeren Normalenvektor der Kante  $E$  bezeichnet und  $C > 0$  nur von der Formregularität abhängt.

*Beweis.* Ohne Beschränkung der Allgemeinheit sei  $\delta_{\ell+1}(P_+) = 0$ . Da  $\delta_{\ell+1} \geq 0$ , gilt dies im besonderen auf  $T_+$ . Ferner ist  $\delta_{\ell+1}(P_+) = 0$ . Wir betrachten nun die Menge der verfeinerten Elemente  $\mathcal{T}_{\ell+1}|_{T_+} := \{T \in \mathcal{T}_{\ell+1} \mid T \subseteq T_+\}$ . Da die Differenzenfunktion auf  $T_+$  stückweise affin und bei  $P_+$  gleich Null ist, muss sie dorthin abfallen (oder konstant Null sein). Es finden sich daher Elemente  $T', T'' \in \mathcal{T}_{\ell+1}|_{T_+}$  auf denen  $\nabla \delta_{\ell+1} \cdot n_E$  nicht-positive und nicht-negative Werte annimmt. Sofern diese Werte nicht null sind, haben daher die Größen  $A \cdot n_E$  und  $\nabla \delta_{\ell+1} \cdot n_E$ , für  $A \in \mathbb{R}^2$  auf mindestens einem Element  $T \in \mathcal{T}_{\ell+1}|_{T_+}$  unterschiedliche Vorzeichen. Es ergibt sich

$$|A \cdot n_E| \leq |\nabla(\delta_{\ell+1}|_T - A) \cdot n_E| \leq \|\nabla \delta_{\ell+1}|_T - A\|_2,$$

wobei  $\|\cdot\|_2$  die euklidische Norm auf  $\mathbb{R}^n$  bezeichnet. Für diese Abschätzung haben wir die Ungleichung von Cauchy-Schwarz verwendet. Da  $|T| \sim |T_+| \sim h_E|E|$  und  $\nabla \delta_{\ell+1}$  stückweise konstant ist, folgt

$$\begin{aligned} h_E^{1/2} \|A \cdot n_E\|_{L^2(E)} &= h_E^{1/2} |E|^{1/2} |A \cdot n_E| \sim |T|^{1/2} |A \cdot n_E| \leq |T|^{1/2} |\nabla(\delta_{\ell+1}|_T - A) \cdot n_E| \\ &\leq |T|^{1/2} \|\nabla \delta_{\ell+1}|_T - A\|_2 = \|\nabla \delta_{\ell+1} - A\|_{L^2(T)} \leq \|\nabla \delta_{\ell+1} - A\|_{L^2(T_+)}. \end{aligned}$$

Damit ist die Aussage vollständig bewiesen.  $\square$

Wir wollen uns nun den einzelnen Fällen widmen. Zunächst betrachten wir die Möglichkeit, dass sowohl  $P_+$ , als auch  $P_-$  zur Kontaktmenge gehören. In diesem Fall erhalten wir unter Zuhilfenahme von Lemma 3.2.7 direkt die finale Abschätzung.

**Proposition 3.2.8.** *Sei  $\delta_{\ell+1}(P_+) = \delta_{\ell+1}(P_-) = 0$ , d.h. die numerische Lösung berührt das Hindernis an den neuen inneren Punkten, dann gilt*

$$\eta_\ell(E) \leq C \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega_E)}, \quad (3.2.17)$$

wobei  $C > 0$  nur von der Formregularität von  $\mathcal{T}_\ell$  abhängt.

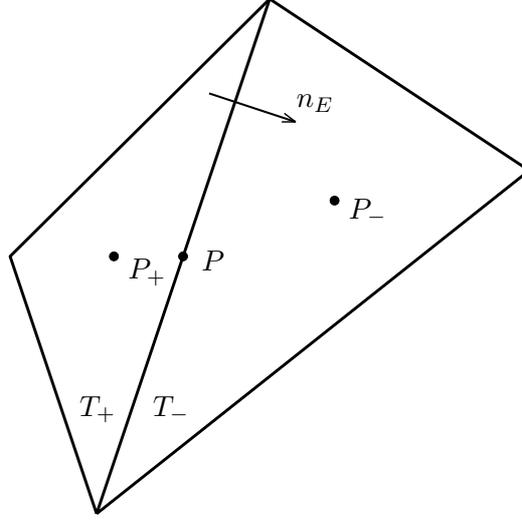


Abbildung 3.1.: Notationen für den gesamten Abschnitt 3.2.2.

*Beweis.* Wir verwenden Lemma 3.2.7 mit  $A = \nabla \delta_\ell|_{T_\pm}$  und erhalten

$$\begin{aligned} \eta_\ell(E) &\leq h_E^{1/2} \|\nabla \delta_\ell|_{T_+} \cdot n_E\|_{L^2(E)} + h_E^{1/2} \|\nabla \delta_\ell|_{T_-} \cdot n_E\|_{L^2(E)} \\ &\stackrel{(3.2.16)}{\lesssim} \|\nabla(\delta_{\ell+1} - \delta_\ell)\|_{L^2(T_+)} + \|\nabla(\delta_{\ell+1} - \delta_\ell)\|_{L^2(T_-)} \\ &\leq \sqrt{2} \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega_E)}, \end{aligned}$$

wobei wir hier die Dreiecksungleichung und die Ungleichung von Cauchy verwendet haben.  $\square$

Wir behandeln nun den Fall in dem mindestens einer der neuen Knoten nicht zur Kontaktmenge gehört.

**Proposition 3.2.9.** *Sei  $\delta_{\ell+1}(P_+) > 0$  oder  $\delta_{\ell+1}(P_-) > 0$ , dann gilt*

$$-\frac{1}{2} |\Omega_E|^{1/2} [\partial_n U_\ell] \leq C (\|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega_E)} + \text{osc}_\ell(E)), \quad (3.2.18)$$

wobei  $C > 0$  nur von der Formregularität abhängt.

*Beweis.* Ohne Beschränkung der Allgemeinheit sei  $\delta_{\ell+1}(P_+) > 0$ . Da sich  $P_+$  in der inaktiven Menge befindet, gilt nach (3.1.12)

$$\langle\langle U_{\ell+1}, \varphi_{P_+} \rangle\rangle = (f, \varphi_{P_+}).$$

Sei  $\varphi_P$ , die nodale Basisfunktion bezüglich des Punktes  $P = \text{mid}(E)$ . Mit  $V_{\ell+1} := \varphi_P + U_{\ell+1} \in K_{\ell+1}$  ergibt sich aus (3.1.7) direkt  $\langle\langle U_{\ell+1}, \varphi_P \rangle\rangle \geq (f, \varphi_P)$ . Wir wählen nun  $\alpha > 0$  so, dass für

$$\varphi_E := \varphi_P - \alpha \varphi_{P_+} \in \mathcal{S}^1(\mathcal{T}_{\ell+1}) \cap H_0^1(\Omega_E)$$

das Integral über den Patch  $\Omega_E$  verschwindet, d.h.

$$\int_{\Omega_E} \varphi_E \, dx = 0.$$

Eine kurze Rechnung zeigt

$$\begin{aligned}
 \langle U_{\ell+1}, \varphi_E \rangle &= \langle U_{\ell+1}, \varphi_P \rangle - \langle U_{\ell+1}, \alpha \varphi_{P_+} \rangle \\
 &= \langle U_{\ell+1}, \varphi_P \rangle - \alpha \langle f, \varphi_{P_+} \rangle \\
 &\geq \langle f, \varphi_P \rangle - \alpha \langle f, \varphi_{P_+} \rangle \\
 &= \langle f, \varphi_E \rangle.
 \end{aligned}$$

Damit erhalten wir

$$\begin{aligned}
 -\langle U_{\ell+1}, \varphi_E \rangle &\leq -\langle f, \varphi_E \rangle + \overbrace{\langle f_{\Omega_E}, \varphi_E \rangle}^{=0} \\
 &= -\langle \varphi_E, f - f_{\Omega_E} \rangle_{L^2(\Omega_E)} \\
 &\leq |\langle \varphi_E, f - f_{\Omega_E} \rangle_{L^2(\Omega_E)}| \\
 &\leq \|\varphi_E\|_{L^2(\Omega_E)} \|f - f_{\Omega_E}\|_{L^2(\Omega_E)} \\
 &\leq \text{osc}_\ell(E),
 \end{aligned} \tag{3.2.19}$$

wobei hier die Hölder'sche Ungleichung verwendet wurde. Durch partielle Integration ergibt sich

$$\begin{aligned}
 \frac{1}{2} \int_E [\partial_n U_\ell] ds &= \frac{1}{2} [\partial_n U_\ell] \int_E 1 ds = [\partial_n U_\ell] \int_E \varphi_E ds = \int_E \varphi_E [\partial_n U_\ell] ds \\
 &= \int_E \varphi_E [\partial_n U_\ell] ds = \int_E \varphi_E \frac{\partial U_\ell|_{T_+}}{\partial_{n_E^+}} ds - \int_E \varphi_E \frac{\partial U_\ell|_{T_-}}{\partial_{n_E^+}} ds \\
 &= \int_{\partial T^+} \varphi_E \frac{\partial U_\ell|_{T_+}}{\partial_{n_E^+}} ds - \int_{\partial T^-} \varphi_E \frac{\partial U_\ell|_{T_-}}{\partial_{n_E^+}} ds \\
 &= \int_{T^+} \nabla U_\ell|_{T^+} \cdot \nabla \varphi_E dx + \int_{T^+} \underbrace{\Delta U_\ell|_{T^+}}_{=0} \varphi_E dx \\
 &\quad + \int_{T^-} \nabla U_\ell|_{T^-} \cdot \nabla \varphi_E dx + \int_{T^-} \underbrace{\Delta U_\ell|_{T^-}}_{=0} \varphi_E dx \\
 &= \langle U_\ell, \varphi_E \rangle.
 \end{aligned} \tag{3.2.20}$$

Hierbei haben wir von der Gleichheit  $n_E^+ = -n_E^-$  Gebrauch gemacht. Die Kombination der erzielten Resultate mit der erneuten Anwendung der Hölder'schen Ungleichung liefert

$$\begin{aligned}
 -\frac{1}{2} \int_E [\partial_n U_\ell] ds &= -\langle U_\ell, \varphi_E \rangle = \langle U_{\ell+1} - U_\ell, \varphi_E \rangle - \langle U_{\ell+1}, \varphi_E \rangle \\
 &\stackrel{(3.2.19), (3.2.20)}{\lesssim} \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega_E)} + \text{osc}_\ell(E).
 \end{aligned}$$

Die Äquivalenz  $h_E \sim |\Omega_E|^{1/2}$  vervollständigt den Beweis, da  $\nabla U_\ell$  stückweise konstant ist.  $\square$

Proposition 3.2.10 behandelt den Fall, bei dem  $P = \text{mid}(E)$  zur inaktiven Menge gehört, während wir uns in Proposition 3.2.11 mit der Fall beschäftigen, dass  $P$  ein aktiver Punkt ist.

**Proposition 3.2.10.** *Sei  $\delta_{\ell+1}(P) > 0$ , dann gilt*

$$\eta_\ell(E) \leq C \left( \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega_E)} + \text{osc}_\ell(E) \right), \quad (3.2.21)$$

wobei die Konstante  $C > 0$  nur von der Formregularität abhängt.

*Beweis.* Analog zu vorher folgt aus  $\delta_{\ell+1}(P) > 0$ , dass

$$\langle\langle U_{\ell+1}, \varphi_P \rangle\rangle = (f, \varphi_P).$$

Die gleichen Argumente wie in (3.2.20) liefern uns ferner

$$\frac{1}{2}\eta_E = \left| \int_E \varphi_P [\partial_n U_\ell] ds \right| = |\langle\langle U_\ell, \varphi_P \rangle\rangle| = |\langle\langle U_{\ell+1} - U_\ell, \varphi_P \rangle\rangle - (f, \varphi_P)|.$$

Den Fall  $\delta_{\ell+1}(P_+) = 0 = \delta_{\ell+1}(P_-)$  haben wir mit Proposition 3.2.8 bereits abgehandelt. Wir beschränken uns also o.B.d.A. darauf, dass  $\delta_{\ell+1}(P_+) > 0$ . Damit ergibt sich sofort

$$\langle\langle U_{\ell+1}, \varphi_{P_+} \rangle\rangle = (f, \varphi_{P_+}).$$

Partielle Integration und  $\varphi_{P_+} \in H_0^1(T_+)$  zeigen

$$\langle\langle U_\ell, \varphi_{P_+} \rangle\rangle = \int_{T_+} \nabla U_\ell \cdot \nabla \varphi_{P_+} dx = - \int_{T_+} \Delta U_\ell \varphi_{P_+} dx + \int_{\partial T_+} \partial_n U_\ell \varphi_{P_+} ds = 0$$

und mit obiger Gleichheit als direkte Konsequenz

$$\langle\langle U_{\ell+1} - U_\ell, \varphi_{P_+} \rangle\rangle = (f, \varphi_{P_+}). \quad (3.2.22)$$

Analog zu Proposition 3.2.9 definieren wir  $\varphi_E := \varphi_P - \alpha \varphi_{P_+}$ , wobei wir  $\alpha > 0$  so wählen, dass.

$$\int_{\Omega_E} \varphi_E dx = 0.$$

Ausnutzen der Gleichheit (3.2.22) zeigt uns

$$\begin{aligned} |\langle\langle U_{\ell+1} - U_\ell, \varphi_E \rangle\rangle - (f, \varphi_E)| &= |\langle\langle U_{\ell+1} - U_\ell, \varphi_P \rangle\rangle - (f, \varphi_P) - \alpha \langle\langle U_{\ell+1} - U_\ell, \varphi_{P_+} \rangle\rangle + \alpha (f, \varphi_{P_+})| \\ &= |\langle\langle U_{\ell+1} - U_\ell, \varphi_P \rangle\rangle - (f, \varphi_P)| \end{aligned}$$

und damit durch zu Proposition 3.2.9 analoger Argumentation

$$\eta_\ell(E) = 2 |\langle\langle U_{\ell+1} - U_\ell, \varphi_E \rangle\rangle - (f, \varphi_E)| \lesssim \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega_E)} + \text{osc}_\ell(E).$$

Die Aussage ist damit bewiesen.  $\square$

**Proposition 3.2.11.** *Sei  $0 \leq [\partial_n U_\ell]|_E$  und  $\delta_{\ell+1}(P) = 0$ , dann gilt*

$$\eta_\ell(E) \leq C \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega_E)}, \quad (3.2.23)$$

wobei  $C > 0$  wieder nur von der Formregularität abhängt.

*Beweis.* Mit  $\delta_{\ell+1} \geq 0$  und  $\delta_{\ell+1}(P) = 0$ , fällt die Differenzenfunktion linear zu  $P = \text{mid}(E)$  ab. Analog zur Argumentation in Lemma 3.2.7 muss es somit  $K_+ \in \mathcal{T}_{\ell+1}|_{T_+}$  und  $K_- \in \mathcal{T}_{\ell+1}|_{T_-}$  geben, so dass

$$\nabla \delta_{\ell+1}|_{K_+} \cdot n_E \leq 0 \leq \nabla \delta_{\ell+1}|_{K_-} \cdot n_E.$$

Damit ergibt sich

$$\begin{aligned} 0 \leq [\partial_n U_\ell] &= \nabla \delta_\ell|_{K_+} \cdot n_E - \nabla \delta_\ell|_{K_-} \cdot n_E \\ &\leq \nabla(\delta_\ell - \delta_{\ell+1})|_{K_+} \cdot n_E - \nabla(\delta_\ell - \delta_{\ell+1})|_{K_-} \cdot n_E \\ &\leq \|\nabla(\delta_\ell - \delta_{\ell+1})|_{K_+}\|_2 + \|\nabla(\delta_\ell - \delta_{\ell+1})|_{K_-}\|_2, \end{aligned}$$

wobei wir hier ausgenutzt haben, dass  $\nabla \chi$  keine Sprünge aufweist. Außerdem haben wir wieder verwendet, dass  $\nabla \delta_\ell$  auf  $T_\pm$  stückweise konstant ist. Daher ist der Sprung bereits durch die Einschränkung auf  $K_\pm$  korrekt angegeben, da  $\nabla \delta_\ell|_{T_\pm} = \nabla \delta_\ell|_{K_\pm}$ . Damit ergibt sich

$$\begin{aligned} \eta_\ell(E) &= |E|^{1/2} h_E^{1/2} [\partial_n U_\ell] \lesssim |K_+|^{1/2} \|\nabla(\delta_\ell - \delta_{\ell+1})|_{K_+}\|_2 + |K_-|^{1/2} \|\nabla(\delta_\ell - \delta_{\ell+1})|_{K_-}\|_2 \\ &= \|\nabla(\delta_{\ell+1} - \delta_\ell)\|_{L^2(K_+)} + \|\nabla(\delta_{\ell+1} - \delta_\ell)\|_{L^2(K_-)} \\ &\leq \sqrt{2} \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega_E)}. \end{aligned}$$

Hierbei haben wir — wie zuvor — verwendet, dass  $\nabla \delta_{\ell+1}$  stückweise konstant ist und  $|\Omega_E| \sim |K_\pm| \sim h_E |E|$  gilt.  $\square$

**Bemerkung.** Bevor wir zum Beweis der diskreten lokalen Effizienz kommen, sei noch Eines angemerkt. Man könnte auf die Idee kommen in den Beweisen der Propositionen 3.2.9 und 3.2.10 die aus der Effizienz des linearen Falles bekannten Bubblefunktionen zu verwenden, anstatt über  $\varphi_E$  zu argumentieren. Dieser Ansatz führt jedoch nicht besonders weit, da, abgesehen von der Eigenschaft, dass  $\varphi_E$  am Rand von  $\Omega_E$  gleich null ist, auch das Integral verschwinden muss. Dies ist eine Eigenschaft welche die Bubblefunktion nicht erfüllt.

Wir können nun den angekündigten Beweis zur diskreten lokalen Effizienz führen.

**Satz 3.2.12** (Diskrete lokale Effizienz). *Für alle Kanten  $E \in \mathcal{M}_{\text{ell},(iv)} \subseteq \mathcal{M}_\ell \cap \mathcal{E}_\ell$ , die von Algorithmus 3.2.2 in Schritt (iv) ausgewählt wurden, gilt*

$$\eta_\ell(E) \leq \sqrt{C} (\|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega_E)} + \text{osc}_\ell(E)), \quad (3.2.24)$$

wobei  $C > 0$  nur von der Formregularität abhängt.

*Beweis.* Für  $\delta_{\ell+1}(P_+) = 0 = \delta_{\ell+1}(P_-)$  ist die Behauptung durch Proposition 3.2.8 gezeigt. Wir beschäftigen uns also nunmehr mit den Fall  $\delta_{\ell+1}(P_+) > 0$  oder  $\delta_{\ell+1}(P_-) > 0$ . Gilt  $\delta_{\ell+1}(P) > 0$ , so erhalten wir das gesuchte Resultat durch Proposition 3.2.10. Es bleibt also lediglich der Fall  $\delta_{\ell+1}(P) = 0$  zu untersuchen. Ist nun  $[\partial_n U_\ell]|_E \geq 0$ , dann folgt die Aussage aus Proposition 3.2.11. Die letzte Möglichkeit ist folglich  $[\partial_n U_\ell]|_E \leq 0$ . In diesem Fall gilt

$$\eta_\ell(E) \sim -|\Omega_E|^{1/2} [\partial_n U_\ell]|_E$$

und damit durch Proposition 3.2.9

$$\eta_\ell(E) \sim |\Omega_E|^{1/2} [\partial_n U_\ell]|_E \lesssim \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega_E)} + \text{osc}_\ell(E).$$

Damit ist die Aussage vollständig bewiesen.  $\square$

**Bemerkung.** Im Beweis zur diskreten lokalen Effizienz, haben wir öfter ausgenutzt, dass die Größen  $\nabla\delta_{\ell+1}$  und  $[\partial_n U_\ell]$  stückweise konstant sind. Dies macht sich an den Stellen bemerkbar, an denen wir die  $L^2$ -Norm lokal als Produkt von Funktionswert und Fläche bzw. Kante geschrieben haben. Dieses Produkt ist jedoch vom Vorzeichen des Integranden abhängig. Könnten wir zum Beispiel sicherstellen, dass der Sprung  $[\partial_n U_\ell]$  immer negativ ist, so würde sich der Beweis stark vereinfachen, da nur noch die Propositionen 3.2.8 und 3.2.9 nötig wären. Es sei weiterhin bemerkt, dass die Tatsache, dass  $\chi$  global affin ist, stark in den Beweis ein- geht. Wäre dies nicht der Fall, so würden sämtliche Vorzeichenargumente nichtig. Außerdem verwenden wir beispielsweise im Beweis von Proposition 3.2.11 die Tatsache, dass  $\chi$  keine Kantensprünge aufweist. Unsere Einschränkungen an das Hindernis gehen also sowohl in die Zuverlässigkeit, als auch in die diskrete lokale Effizienz ein.

### 3.2.3. Fehlerreduktion - Ein Konvergenzsatz

Wir haben nun all unsere Zutaten, die *Zuverlässigkeit* und die *diskrete lokale Effizienz*, beisammen und können unter den genannten Voraussetzungen die Konvergenz der Folge von diskreten Lösungen aus Algorithmus 3.2.2 gegen die exakte Lösung  $u \in K$  zeigen. Wie immer notieren wir den Fehler im Energiefunktional mit  $e_\ell = \mathcal{J}(U_\ell) - \mathcal{J}(u) > 0$ . Es gilt die folgende Aussage:

**Satz 3.2.13** (Fehlerreduktion). *Unter den obigen Voraussetzungen gibt es Konstanten  $0 \leq \rho < 1$  und  $C > 0$ , die nur von dem  $\theta$  aus der Dörflermarkierung und der Formregularität abhängen, so dass*

$$e_{\ell+1} \leq \rho e_\ell + C \text{Osc}_\ell^2 \quad (3.2.25)$$

*gilt.*

Bevor wir diesen Satz beweisen, wollen wir uns schnell klarmachen, dass hieraus die Konvergenz gegen die exakte Lösung in Energienorm folgt. Aus der Annahme (3.2.6) über die Reduktion der Oszillationen in jedem Schritt folgt

$$\text{Osc}_{\ell+1}^2 \leq \kappa^2 \text{Osc}_\ell^2,$$

mit  $\kappa \in (0, 1)$ . Die Oszillationen sind daher eine Nullfolge, so dass wir das Problem als

$$e_{\ell+1} \leq \rho e_\ell + \alpha_\ell \quad \ell \in \mathbb{N}$$

schreiben können, wobei  $\lim_{\ell \rightarrow \infty} \alpha_\ell = 0$  gilt. Durch Induktion über  $\ell$  erhalten wir

$$e_{\ell+1} \leq \rho^{\ell+1} e_0 + \sum_{j=0}^{\ell} \rho^{\ell-j} \alpha_j \leq \rho^{\ell+1} e_0 + \|(\alpha_\ell)_{\ell \in \mathbb{N}}\|_\infty \sum_{k=0}^{\ell} \rho^k,$$

wobei  $\|(\alpha_\ell)_{\ell \in \mathbb{N}}\|_\infty$  die Supremumsnorm der beschränkten Folge  $(\alpha_\ell)_{\ell \in \mathbb{N}}$  beschreibt. Mit der Beschränktheit von  $\sum_{k=0}^{\ell} \rho^k$  als geometrische Reihe, folgt im besonderen die Beschränktheit von  $e_\ell$  und damit die Existenz von  $0 \leq M := \limsup_{\ell \rightarrow \infty} e_\ell < \infty$ . Wir erhalten daher

$$M = \limsup_{\ell \rightarrow \infty} e_{\ell+1} \leq \rho \limsup_{\ell \rightarrow \infty} e_\ell + \limsup_{\ell \rightarrow \infty} \alpha_\ell = \rho M$$

und damit  $M = 0$ . Da  $\rho \in (0, 1)$  folgt nun mit

$$0 \leq \liminf_{\ell \rightarrow \infty} e_\ell \leq \limsup_{\ell \rightarrow \infty} e_\ell = 0$$

die Konvergenz des Fehlers im Energiefunktional gegen Null. Mit  $v = U_\ell \in K$  in (3.2.14) ergibt sich ferner

$$e_\ell \geq \frac{1}{2} \| \|u - U_\ell\| \|^2$$

und damit die Konvergenz der diskreten Lösungen  $U_\ell$  gegen die exakte Lösung  $u \in K$ . Zum Abschluss dieses Abschnittes bleibt noch die Fehlerreduktion selbst zu beweisen.

*Beweis von Theorem 3.2.13 (Fehlerreduktion).* Analog zu (3.2.14) gilt

$$\mathcal{J}(V_{\ell+1}) = \mathcal{J}(U_{\ell+1}) + \frac{1}{2} \| \|V_{\ell+1} - U_{\ell+1}\| \|^2 + \langle \sigma_{\ell+1}, V_{\ell+1} - U_{\ell+1} \rangle$$

für alle  $V_{\ell+1} \in K_{\ell+1}$ , wobei  $\sigma_{\ell+1} \in \mathcal{S}_0^1(\mathcal{T}_{\ell+1})'$  wie im kontinuierlichen Fall den Lagrange'schen Multiplikator bezeichnet. Für  $V_{\ell+1} = U_\ell \in K_{\ell+1}$  ergibt sich damit aufgrund der Positivität des Multiplikators (3.1.7)

$$\mathcal{J}(U_\ell) - \mathcal{J}(U_{\ell+1}) \geq \frac{1}{2} \| \|U_\ell - U_{\ell+1}\| \|^2.$$

Wir bezeichnen nun die in der Zuverlässigkeit und der diskreten lokalen Effizienz enthaltenen Konstanten mit  $C_r$ , bzw.  $C_{dle}$ . Diese beiden Eigenschaften in Verbindung mit der Dörflermarkierung ergeben.

$$\begin{aligned} 2(e_\ell - e_{\ell+1}) &= 2(\mathcal{J}(U_\ell) - \mathcal{J}(U_{\ell+1})) \geq \| \|U_\ell - U_{\ell+1}\| \|^2 \\ &\geq \frac{1}{C_{dle}} \sum_{E \in \mathcal{M}_{\ell, (iv)}} \eta_\ell(E)^2 - osc_\ell^2 \\ &\geq \frac{\theta}{C_{dle}} \eta_\ell^2 - osc_\ell^2 \\ &\geq \frac{\theta}{C_{dle}} \left( \frac{1}{C_r} e_\ell - Osc_\ell^2 \right) - osc_\ell^2. \end{aligned} \tag{3.2.26}$$

Umstellen dieser Ungleichungskette liefert

$$e_{\ell+1} \leq \left( 1 - \frac{\theta}{2C_{dle}C_r} \right) e_\ell + COsc_\ell^2.$$

Die zeigt die Aussage für  $\rho := 1 - \theta/(2C_{dle}C_r) < 1$ . □

**Bemerkung.** Die Kontraktion im letzten Beweis ergibt sich genau dann, wenn  $\frac{\theta}{2C_{dle}C_r} \leq 1$ . Dass dies der Fall ist lässt sich folgendermaßen einsehen: Mit der Zuverlässigkeit und der diskreten lokalen Effizienz des Schätzers ergibt sich

$$\frac{\theta}{C_{dle}} \eta_\ell^2 - osc_\ell^2 \leq \frac{1}{C_{dle}} \sum_{E \in \mathcal{M}_{\ell, (iv)}} \eta_\ell(E)^2 - osc_\ell^2 \leq \| \|U_\ell - U_{\ell+1}\| \|^2 \leq C_r (\eta_\ell^2 + Osc_\ell^2),$$

wobei wir hier die Dörflermarkierung verwendet haben. Da die Oszillationen mit wachsendem  $\ell$  verschwinden, folgt somit

$$\frac{\theta}{C_{dle}C_r} \leq 1.$$

**Bemerkung.** An dieser Stelle wird auch deutlich, warum die Zuverlässigkeit für den Fehler im Energiefunktional  $e_\ell$ , und nicht, wie sonst, für den Fehler in Energienorm  $\|u - U_\ell\|$  gezeigt wurde. Hätten wir letzteren Ansatz verfolgt, so wäre für den Beweis der Fehlerreduktion nach (3.2.26) eine Abschätzung der Form

$$\|U_\ell - U_{\ell+1}\|^2 \leq \|u - U_\ell\|^2 - \|u - U_{\ell+1}\|^2$$

von Nöten. Im linearen Fall folgt dies (sogar mit Gleichheit) aus der Galerkin-Orthogonalität, welche uns hier leider nicht mehr zur Verfügung steht.

### 3.3. Schätzerreduktion für das Hindernisproblem

In diesem Abschnitt wollen wir die bereits erzielten Resultate in zweierlei Hinsicht verbessern. Zum Einen werden wir einen neuen Fehlerschätzer einführen, der es ermöglicht die Datenoszillationen implizit zu kontrollieren. Dadurch wird es obsolet zusätzlich ein Abklingen der Oszillationen (3.2.6) zu fordern, was dazu führt, dass in unserem Algorithmus nur noch eine einzige Dörflermarkierung nötig sein wird. Dies wirkt sich positiv auf Laufzeit und natürlich Implementierungsaufwand aus. Zum Anderen werden wir das Prinzip der Schätzerreduktion (vgl. Satz 2.2.6) auf das Hindernisproblem ausweiten und dadurch die Notwendigkeit der diskreten lokalen Effizienz umgehen. Als Resultat erhalten wir einen Konvergenzbeweis der ohne die *innerer-Punkt-Eigenschaft* auskommt und so, im Gegensatz zu Satz 3.2.13, von der Verfeinerungsstrategie weitestgehend unabhängig ist.

#### 3.3.1. Fehlerschätzer und adaptiver Algorithmus

Wie zuvor notieren wir die exakte Lösung des kontinuierlichen Problems 3.1.4 mit  $u \in K$  und die diskrete Lösung des endlich dimensionalen Teilproblems 3.1.5 mit  $U_\ell \in K_\ell$ . In Anlehnung an unsere vorherige Notation werden wir den folgenden residualbasierten Fehlerschätzer

$$\eta_\ell^2 := \sum_{E \in \mathcal{E}_\ell} (\eta_\ell(E)^2 + \text{osc}_\ell(E)^2) + \sum_{T \in \mathcal{T}_{\ell,\Gamma}} \text{osc}_\ell(T)^2 \quad (3.3.1)$$

mit dem gleichen Symbol wie oben bezeichnen. Da wir fortan den Schätzer (3.2.2) nichtmehr verwenden, sind Verwechslungen ausgeschlossen. Wir zuvor bezeichnet hierbei  $\eta_\ell(E)^2$  die gewichtete  $L^2$ -Norm des Sprungs der Normalenableitungen

$$\eta_\ell(E)^2 := h_E \|\partial_n U_\ell\|_{L^2(E)}^2 \quad \text{für } E \in \mathcal{E}_\ell, \quad (3.3.2)$$

für innere Kanten. Der Term  $\text{osc}_\ell(E)^2$  steht, wie oben, für die Datenoszillationen von  $f$  über  $E$

$$\text{osc}_\ell(E)^2 := |\Omega_E| \|f - f_{\Omega_E}\|_{L^2(\Omega_E)}^2 \quad \text{für } E \in \mathcal{E}_\ell, \quad (3.3.3)$$

wobei  $\Omega_E = T_+ \cup T_-$  wieder den Patch der Kante  $E = T_+ \cap T_-$ , und  $f_{\Omega_E}$  die  $L^2$ -Bestapproximation beschreibt. Schließlich bleiben noch die Elementbeiträge am Rand mit

$$osc_\ell(T)^2 := |T| \|f\|_{L^2(T)}^2 \quad \text{für } T \in \mathcal{T}_\ell. \quad (3.3.4)$$

Bis auf die Tatsache, dass die Elementbeiträge mit  $|T|$  statt mit  $\text{diam}(T)^2$  gewichtet sind, betrachten wir also essentiell den Schätzer des vorherigen Kapitels, dem wir die Oszillationsterme einfach beigefügt haben. Wir gewinnen dadurch die Freiheit nur noch den Term  $\eta_\ell^2$ , und keine zusätzlichen Größen, betrachten zu müssen. Aufgrund der großen Ähnlichkeit der beiden Schätzer leiten wir direkt die folgende Zuverlässigkeitsaussage ab:

**Satz 3.3.1** (Zuverlässigkeit). *Für stückweise polynomiales  $f \in L^2(\Omega)$  und Verfeinerung durch newest vertex bisection ist der Fehlerschätzer  $\eta_\ell$  zuverlässig. Es gibt also eine Konstante  $C > 0$ , die nur von  $\Omega$  und der Formregularität abhängt, so dass*

$$\frac{1}{2} \|u - U_\ell\|^2 \leq \mathcal{J}(U_\ell) - \mathcal{J}(u) \leq C \eta_\ell^2. \quad (3.3.5)$$

gilt.

*Beweis.* Offensichtlich gilt  $|T| \sim \text{diam}(T)^2 = h_T^2$ . Damit folgt die obere Abschätzung aus dem bereits gezeigten Zuverlässigkeitsresultat (Satz 3.2.6). Für die untere Abschätzung verwenden wir die Variationsungleichung (3.1.6) mit  $v = U_\ell$  und erhalten

$$\begin{aligned} \|u - U_\ell\|^2 &= \langle u, u - U_\ell \rangle + \|U_\ell\|^2 - \langle U_\ell, u \rangle \\ &\leq (f, u - U_\ell) + \|U_\ell\|^2 - \langle U_\ell, u \rangle \\ &= (f, u) - (f, U_\ell) + \|U_\ell\|^2 - \langle U_\ell, u \rangle \\ &= \left( \frac{1}{2} \|U_\ell\|^2 - (f, U_\ell) \right) - \left( \frac{1}{2} \|u\|^2 - (f, u) \right) \\ &\quad + \left( \frac{1}{2} \|U_\ell\|^2 - \langle U_\ell, u \rangle + \frac{1}{2} \|u\|^2 \right) \\ &= \mathcal{J}(U_\ell) - \mathcal{J}(u) + \frac{1}{2} \|u - U_\ell\|^2. \end{aligned}$$

Damit ist die Aussage vollständig bewiesen. □

Bevor wir uns der Analyse dieses Schätzers, und damit einer genaueren Untersuchung der Oszillationsterme widmen, halten wir noch die aktualisierte Version unseres adaptiven Algorithmus' fest.

**Algorithmus 3.3.2** (adaptiver Algorithmus, Version 2). **Input:** Anfangstriangulierung  $\mathcal{T}_0$ , maximale Anzahl Elemente  $n_{E_{max}}$ , Adaptivitätsparameter  $\theta \in (0, 1)$ ,  $\ell := 0$ .

- (i) Berechne diskrete Lösung  $U_\ell \in K_\ell$ .
- (ii) Berechne die Verfeinerungsindikatoren  $\eta_\ell(E)$ ,  $osc_\ell(E)$ ,  $osc_\ell(T)$  und den Fehlerschätzer  $\eta_\ell$ .
- (iii) STOP, wenn die Anzahl der Elemente größer als  $n_{E_{max}}$  ist.
- (iv) Finde eine Menge  $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell \cup \mathcal{T}_{\ell, \Gamma}$  minimaler Kardinalität von markierten Elementen und Kanten, so dass
 
$$\theta \eta_\ell^2 \leq \sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} (\eta_\ell(E)^2 + osc_\ell(E)^2) + \sum_{T \in \mathcal{T}_{\ell, \Gamma} \cap \mathcal{M}_\ell} osc_\ell(T)^2.$$
- (v) Erzeuge neues, reguläres Gitter  $\mathcal{T}_{\ell+1}$ , bei dem mindestens alle markierten Elemente und Kanten verfeinert wurden durch beliebige Verfeinerungsstrategie, zum Beispiel durch Newest Vertex Bisection.
- (vi)  $\ell \mapsto \ell + 1$  und beginne wieder bei (i).

**Output:** Endliche Folge diskreter Lösungen  $U_\ell$  und zugehörige Fehlerschätzer  $\eta_\ell$ .

An dieser Stelle kann man deutlich erkennen, dass durch die geschickte Konstruktion des Schätzers  $\eta_\ell$ , alle zur Verfeinerung ausgewählten Kanten und Elemente in einem einzigen Schritt markiert werden. Die Oszillationen müssen daher nichtmehr separat betrachtet werden. Außerdem ist die Verfeinerungsstrategie, wie angekündigt, beliebig. Im nächsten Abschnitt werden wir zeigen, dass auch die Folge der diskreten Lösungen  $(U_\ell)_{\ell \in \mathbb{N}}$ , die durch diesen Algorithmus erzeugt wird, immer gegen die exakte Lösung  $u \in K$  konvergiert.

### 3.3.2. Schätzerreduktion und Konvergenzanalyse

Analog zum linearen Fall und in Anlehnung an [CKNS, AFP], wollen wir die diskrete lokale Zuverlässigkeit, welche uns die *bisec*<sub>5</sub>-Verfeinerung aufgezwungen hat, umgehen. Dies gelingt, indem wir durch *Schätzerreduktion* zeigen, dass der Schätzer selbst, bis auf die Differenz zweier, aufeinander folgender Folgeglieder  $U_\ell$  und  $U_{\ell+1}$ , eine kontraktive Größe ist. Zusammen mit der Zuverlässigkeit wird dann, wie im linearen Fall, die Konvergenz folgen. Um die Konvergenzbeweise zu führen, definieren wir zunächst eine zu  $\eta_\ell$  äquivalente Größe  $\mu_\ell$ , die eine kleine Abwandlung darstellt, durch

$$\mu_\ell^2 := \sum_{E \in \mathcal{E}_\ell} \eta_\ell(E)^2 + \sum_{E \in \mathcal{E}_\ell^*} osc_\ell(E)^2 + \sum_{T \in \mathcal{T}_{\ell, \Gamma}} osc_\ell(T)^2. \quad (3.3.6)$$

Die Oszillationen an den Randkanten  $E \in \mathcal{E}_{\ell, \Gamma} = \mathcal{E}_\ell^* \setminus \mathcal{E}_\ell$  definieren wir hierbei einfach über die Randelementgrößen

$$osc_\ell(E)^2 := osc_\ell(T)^2 \quad \text{für alle } E \in \mathcal{E}_{\ell, \Gamma} := \mathcal{E}_\ell^* \setminus \mathcal{E}_\ell \text{ und } T \in \mathcal{T}_{\ell, \Gamma} \text{ mit } E \subset \partial T. \quad (3.3.7)$$

Da jedes Element  $T \in \mathcal{T}_{\ell,\Gamma}$  höchstens drei Randkanten haben kann, erhalten wir direkt die Äquivalenz der beiden Schätzer

$$\eta_\ell^2 \leq \mu_\ell^2 \leq 4\eta_\ell^2. \quad (3.3.8)$$

Im besonderen gilt daher die im letzten Abschnitt bewiesene Zuverlässigkeit auch für  $\mu_\ell$ . Wir wollen für den Moment annehmen, dass der Schätzer  $\mu_\ell$  anstelle von  $\eta_\ell$ , zur Steuerung der adaptiven Verfeinerung in obigem Algorithmus 3.3.2, verwendet wird. Dies wirkt sich natürlich auf die Dörflermarkierung in Schritt (iv) aus. Zu einem gegebenen Adaptivitätsparameter  $\theta \in (0, 1)$  suchen wir nun eine Menge  $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell^* \cup \mathcal{T}_{\ell,\Gamma}$  von minimaler Kardinalität, so dass

$$\theta \mu_\ell^2 \leq \sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} \eta_\ell(E)^2 + \sum_{E \in \mathcal{E}_\ell^* \cap \mathcal{M}_\ell} \text{osc}_\ell(E)^2 + \sum_{T \in \mathcal{T}_{\ell,\Gamma} \cap \mathcal{M}_\ell} \text{osc}_\ell(T)^2. \quad (3.3.9)$$

Es können also auch Randkanten  $E \in \mathcal{E}_{\ell,\Gamma}$  zur Verfeinerung ausgewählt werden. Im Folgenden werden wir, der Einfachheit halber, Newest Vertex Bisection als Verfeinerung verwenden. Sämtliche Ausführungen lassen sich aber auch auf andere Verfeinerungsstrategien übertragen. Wir werden später noch etwas genauer darauf eingehen. Zunächst halten wir zwei elementare Eigenschaften dieser Strategie fest:

- Markierte Kanten  $E \in \mathcal{M}_\ell$  werden in zwei neue Kanten  $E', E'' \in \mathcal{E}_{\ell+1}^*$  von halber Länge aufgespalten.
- Ist mindestens eine Kante  $E$  eines Elements  $T \in \mathcal{T}_\ell$  markiert, so wird  $T$  in bis zu vier Söhne  $T' \in \mathcal{T}_{\ell+1}$  mit  $|T|/4 \leq |T'| \leq |T|/2$  aufgeteilt.

Wir kommen nun zur angekündigten Schätzerreduktion.

**Satz 3.3.3** (Schätzerreduktion für das affine Hindernisproblem). *Die Menge  $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell^* \cup \mathcal{T}_{\ell,\Gamma}$  erfülle die Dörflermarkierung 3.3.9 für ein gegebenes  $\theta \in (0, 1)$ . Werden die markierten Elemente und Kanten ferner wie oben beschrieben verfeinert, so gilt*

$$\mu_{\ell+1}^2 \leq q \mu_\ell^2 + C \| \|U_{\ell+1} - U_\ell\| \|^2, \quad (3.3.10)$$

für eine nur von  $\theta$  abhängige Konstante  $q \in (0, 1)$ . Die Konstante  $C > 0$  hängt außerdem von der Formregularität ab.

Da der Beweis dieses Satzes sehr lang und aufwendig ist, teilen wir ihn, der Übersichtlichkeit halber, in drei Lemmata auf, welche jeweils einen Teil des Schätzers  $\mu_\ell$  analysieren.

**Lemma 3.3.4.** *Durch die Verfeinerung der Elemente  $T \in \mathcal{T}_{\ell,\Gamma} \cap \mathcal{M}_\ell$  gilt*

$$\sum_{T' \in \mathcal{T}_{\ell+1,\Gamma}} \text{osc}_{\ell+1}(T')^2 \leq \sum_{T \in \mathcal{T}_{\ell,\Gamma}} \text{osc}_\ell(T)^2 - \frac{1}{2} \sum_{T \in \mathcal{T}_{\ell,\Gamma} \cap \mathcal{M}_\ell} \text{osc}_\ell(T)^2. \quad (3.3.11)$$

*Beweis.* Analog zur Schätzerreduktion im linearen Fall, definieren wir die Menge

$$\overline{\mathcal{M}}_{\mathcal{T},\ell} := \{T' \in \mathcal{T}_{\ell+1} \mid \exists T \in \mathcal{T}_{\ell,\Gamma} \cap \mathcal{M}_\ell, \quad T' \subseteq T\}$$

aller durch Verfeinerung markierter Elemente entstandener Elemente  $T' \in \mathcal{T}_{\ell+1}$ . Mit der Reduktion der Elementgröße  $|T|$  eines markierten Elements  $T \in \mathcal{M}_\ell$ , ergibt sich somit

$$\begin{aligned} \sum_{T' \in \mathcal{T}_{\ell+1, \Gamma}} \text{osc}_{\ell+1}(T')^2 &\leq \sum_{T' \in \mathcal{T}_{\ell+1, \Gamma} \setminus \overline{\mathcal{M}}_{\mathcal{T}, \ell}} |T'| \|f\|_{L^2(T')}^2 + \sum_{T' \in \overline{\mathcal{M}}_{\mathcal{T}, \ell}} |T'| \|f\|_{L^2(T')}^2 \\ &\leq \sum_{T \in \mathcal{T}_{\ell, \Gamma} \setminus \mathcal{M}_\ell} |T| \|f\|_{L^2(T)}^2 + \frac{1}{2} \sum_{T \in \mathcal{T}_{\ell, \Gamma} \cap \mathcal{M}_\ell} |T| \|f\|_{L^2(T)}^2 \\ &= \sum_{T \in \mathcal{T}_{\ell, \Gamma} \setminus \mathcal{M}_\ell} \text{osc}_\ell(T)^2 + \frac{1}{2} \sum_{T \in \mathcal{T}_{\ell, \Gamma} \cap \mathcal{M}_\ell} \text{osc}_\ell(T)^2, \end{aligned}$$

wobei das erste Ungleichheitszeichen dadurch entsteht, dass die Elemente in  $\overline{\mathcal{M}}_{\mathcal{T}, \ell}$  nicht unbedingt Randelemente sein müssen. Für das zweite Ungleichheitszeichen haben wir zusätzlich verwendet, dass sich jeder Vater  $T \in \mathcal{T}_\ell$  als disjunkte Vereinigung seiner Söhne  $T' \in \mathcal{T}_{\ell+1}$  schreiben lässt.  $\square$

**Lemma 3.3.5.** *Durch die Verfeinerung markierter Kanten  $E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell$  gilt*

$$\sum_{E' \in \mathcal{E}_{\ell+1}} h_{E'} \|[\partial_n U_\ell]\|_{L^2(E')}^2 \leq \sum_{E \in \mathcal{E}_\ell} \eta_\ell(E)^2 - \frac{1}{2} \sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} \eta_\ell(E)^2. \quad (3.3.12)$$

*Beweis.* Mit der Menge

$$\overline{\mathcal{M}}_{\mathcal{E}, \ell} := \{E' \in \mathcal{E}_{\ell+1} \mid \exists E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell \quad E' \subseteq E\}$$

aller Kanten, die durch Verfeinerung markierter Kanten entstehen, erhalten wir

$$\begin{aligned} \sum_{E' \in \mathcal{E}_{\ell+1}} h_{E'} \|[\partial_n U_\ell]\|_{L^2(E')}^2 &= \sum_{E' \in \mathcal{E}_{\ell+1} \setminus \overline{\mathcal{M}}_{\mathcal{E}, \ell}} h_{E'} \|[\partial_n U_\ell]\|_{L^2(E')}^2 + \sum_{E' \in \overline{\mathcal{M}}_{\mathcal{E}, \ell}} h_{E'} \|[\partial_n U_\ell]\|_{L^2(E')}^2 \\ &\leq \sum_{E \in \mathcal{E}_\ell \setminus \mathcal{M}_\ell} h_E \|[\partial_n U_\ell]\|_{L^2(E)}^2 + \frac{1}{2} \sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} h_E \|[\partial_n U_\ell]\|_{L^2(E)}^2 \\ &= \sum_{E \in \mathcal{E}_\ell \setminus \mathcal{M}_\ell} \eta_\ell(E)^2 + \frac{1}{2} \sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} \eta_\ell(E)^2, \end{aligned}$$

analog zu den Argumenten in Lemma 3.3.4. Im zweiten Schritt haben wir zusätzlich verwendet, dass der Sprung  $[\partial_n U_\ell]$  auf allen neuen Kanten  $E' \in \mathcal{E}_{\ell+1} \setminus \mathcal{E}_\ell$  im Inneren eines Elements  $T \in \mathcal{T}_\ell$  verschwindet. Dies folgt aus der einfachen Tatsache, dass  $U_\ell$  auf  $T \in \mathcal{T}_\ell$  stückweise affin, und damit die Normalenableitung auf allen Dreiecken im  $\ell$ -ten Schritt konstant ist.  $\square$

**Lemma 3.3.6.** *Durch die Verfeinerung markierter Kanten  $E \in \mathcal{E}_\ell^* \cap \mathcal{M}_\ell$  gilt*

$$\sum_{E' \in \mathcal{E}_{\ell+1}^*} \text{osc}_{\ell+1}(E')^2 \leq \sum_{E \in \mathcal{E}_\ell^*} \text{osc}_\ell(E)^2 - \frac{1}{4} \sum_{E \in \mathcal{E}_\ell^* \cap \mathcal{M}_\ell} \text{osc}_\ell(E)^2. \quad (3.3.13)$$

*Beweis.* Im Gegensatz zu den Elementbeiträgen, müssen wir bei den kantenbasierten Größen beachten, dass durch Verfeinerung neue Kanten im Inneren der zerteilten Dreiecke entstehen. Somit entstehen zusätzliche Kantenoszillationen die kontrolliert werden müssen. Daher ist

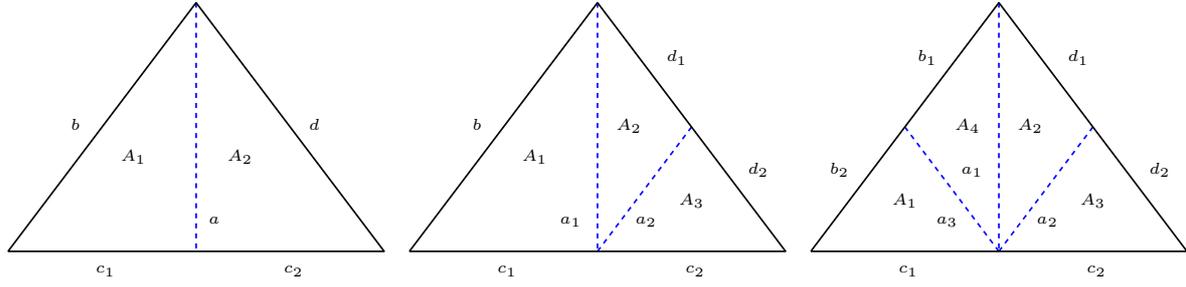


Abbildung 3.2.: Verfeinerung des Elements  $A$  durch eine (*links*), zwei (*mitte*), oder drei (*rechts*) Bisektionen und Notationen für den Beweis von Lemma 3.3.6.

der folgende Beweis ungleich länger als die vorangegangenen. Der Übersichtlichkeit halber führen wir einige Notationen ein. Für eine Kante  $E \in \mathcal{E}_{\ell+1}^*$  und ein Element  $T \in \mathcal{T}_\ell$  mit  $|T \cap \Omega_{\ell+1,E}| > 0$  definieren wir

$$\text{osc}_{\ell+1}(E|T)^2 = |\Omega_{\ell+1,E}| \|f - f_{\Omega_{\ell+1,E}}\|_{L^2(\Omega_{\ell+1,E} \cap T)}^2$$

als den Beitrag der Oszillationen über  $E$  auf dem Element  $T$ . Für Randkanten  $E \in \mathcal{E}_{\ell+1,\Gamma} = \mathcal{E}_{\ell+1}^* \setminus \mathcal{E}_{\ell+1}$  setzen wir  $\Omega_{\ell+1,E} := T'$  und  $f_{\Omega_{\ell+1,E}} := 0$ , wobei  $T' \in \mathcal{T}_{\ell+1,\Gamma}$  das eindeutige Element mit  $E \subset \partial T'$  bezeichnet. Außerdem stellen wir die folgenden Eigenschaften der  $L^2$ -Bestapproximation  $f_\omega = 1/|\omega| \int_\omega f \, dx$  fest:

- $\|f - f_\omega\|_{L^2(\omega)} \leq \|f - \alpha\|_{L^2(\omega)}$  für alle  $\alpha \in \mathbb{R}$ ,
- $\|f - f_\omega\|_{L^2(\omega)} \leq \|f - f_{\hat{\omega}}\|_{L^2(\hat{\omega})}$  für alle messbaren Mengen  $\hat{\omega} \supset \omega$ .

Für ein Element  $A \in \mathcal{T}_\ell$  können nun die folgenden vier Fälle auftreten (siehe Abbildung 3.2):

- (i)  $A$  wird nicht verfeinert, d.h.  $A \in \mathcal{T}_\ell \cap \mathcal{T}_{\ell+1}$ .
- (ii)  $A$  wird in zwei Dreiecke zerteilt.
- (iii)  $A$  wird in drei Dreiecke zerteilt.
- (iv)  $A$  wird in vier Dreiecke zerteilt.

Wir wollen diese Fälle im Folgenden einzeln betrachten und untersuchen als erstes ein Element  $A \in \mathcal{T}_\ell \cap \mathcal{T}_{\ell+1}$  welches nicht verfeinert wird. Seine drei Kanten bezeichnen wir mit  $a, b$  und  $c$ . Das Vorgehen dieses Beweises ist es nun, die Oszillationen auf den Kanten der Elemente in  $\mathcal{T}_{\ell+1}$  auf die Oszillationen auf den Kanten der Elemente in  $\mathcal{T}_\ell$  zurückzuführen. Da in diesem Fall keine neuen Kanten hinzukommen, definieren wir

$$o_\ell(b|A)^2 = \text{osc}_{\ell+1}(b|A)^2, \quad o_\ell(c|A)^2 = \text{osc}_{\ell+1}(c|A)^2, \quad \text{und} \quad o_\ell(d|A)^2 = \text{osc}_{\ell+1}(d|A)^2.$$

Per Definition ergibt sich somit

$$\sum_{\substack{E \in \mathcal{E}_{\ell+1}^* \\ |A \cap \Omega_{\ell+1,E}| > 0}} \text{osc}_{\ell+1}(E|A)^2 \leq \sum_{\substack{E \in \mathcal{E}_\ell^* \\ E \subset \partial A}} o_\ell(E|A)^2, \quad (3.3.14)$$

wobei hier sogar Gleichheit gilt.

Als zweites beschäftigen wir uns mit dem Fall, dass ein Element  $A \in \mathcal{T}_\ell$  mit Kanten  $b, c, d \in \mathcal{E}_\ell^*$  durch eine Bisektion in zwei Dreiecke aufgespalten wird (siehe Abb. 3.2 *links*). Die Kante  $c \in \mathcal{E}_\ell^*$  wird hierbei in die beiden Kanten  $c_1, c_2 \in \mathcal{E}_{\ell+1}^*$  zerteilt, so dass eine zusätzliche Kante  $a \in \mathcal{E}_{\ell+1}^*$  entsteht.  $A$  wird außerdem in die beiden Dreiecke  $A_1$  und  $A_2 \in \mathcal{T}_{\ell+1}$  mit Fläche  $|A_1| = |A_2| = |A|/2$  zerlegt. Die Nachbarn von  $A$  an den entsprechenden Kanten notieren wir mit  $B, C, D \in \mathcal{T}_\ell$ . Sollten wir es mit einer Randkante zu tun haben, so entspricht das Nachbarelement einfach der leeren Menge. Es gilt nun

$$\begin{aligned} & \sum_{\substack{E \in \mathcal{E}_{\ell+1}^* \\ |A \cap \Omega_{\ell+1, E}| > 0}} \text{osc}_{\ell+1}(E|A)^2 \\ &= \text{osc}_{\ell+1}(c_1|A)^2 + \text{osc}_{\ell+1}(c_2|A)^2 + \text{osc}_{\ell+1}(b|A)^2 + \text{osc}_{\ell+1}(d|A)^2 + \text{osc}_{\ell+1}(a|A)^2 \\ &= (|\Omega_{\ell+1, c_1} \cap C| + |A|/2) \|f - f_{\Omega_{\ell+1, c_1}}\|_{L^2(A_1)}^2 + (|\Omega_{\ell+1, c_2} \cap C| + |A|/2) \|f - f_{\Omega_{\ell+1, c_2}}\|_{L^2(A_2)}^2 \\ &\quad + (|\Omega_{\ell+1, b} \cap B| + |A|/2) \|f - f_{\Omega_{\ell+1, b}}\|_{L^2(A_1)}^2 + (|\Omega_{\ell+1, d} \cap D| + |A|/2) \|f - f_{\Omega_{\ell+1, d}}\|_{L^2(A_2)}^2 \\ &\quad + |A| \|f - f_A\|_{L^2(A)}^2, \end{aligned}$$

wobei der letzte Term von der neuen Kante  $a \in \mathcal{E}_{\ell+1}^* \setminus \mathcal{E}_\ell^*$  stammt. Wir wollen nun wieder versuchen diese Oszillationsterme auf die Oszillationen über den Kanten im  $\ell$ -ten Schritt zurückzuführen. Hierzu teilen wir die Oszillationen über der neuen Kante  $a$  einfach auf und addieren sie an passenden Stellen dazu. Wir definieren also

$$\begin{aligned} o_\ell(b|A)^2 &= (|\Omega_{\ell+1, b} \cap B| + |A|/2) \|f - f_{\Omega_{\ell+1, b}}\|_{L^2(A_1)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2, \\ o_\ell(c|A)^2 &= (|\Omega_{\ell+1, c_1} \cap C| + |A|/2) \|f - f_{\Omega_{\ell+1, c_1}}\|_{L^2(A_1)}^2 \\ &\quad + (|\Omega_{\ell+1, c_2} \cap C| + |A|/2) \|f - f_{\Omega_{\ell+1, c_2}}\|_{L^2(A_2)}^2, \\ o_\ell(d|A)^2 &= (|\Omega_{\ell+1, d} \cap D| + |A|/2) \|f - f_{\Omega_{\ell+1, d}}\|_{L^2(A_2)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2 \end{aligned}$$

und erhalten somit wieder (3.3.14) mit Gleichheit.

Als drittes betrachten wir ein Element  $A \in \mathcal{T}_\ell$  mit den Kanten  $b, c, d \in \mathcal{E}_\ell^*$ , welches durch zwei Bisektionen in die drei Dreiecke  $A_1, A_2, A_3 \in \mathcal{T}_{\ell+1}$  zerteilt wird (siehe Abb. 3.2 *mitte*). Die Kanten  $c$  und  $d$  werden hierbei in  $c_1, c_2, d_1, d_2 \in \mathcal{E}_{\ell+1}^*$  aufgespalten, so dass die neuen Kanten  $a_1$  und  $a_2 \in \mathcal{E}_{\ell+1}^*$  entstehen. Ferner gilt  $|A_1| = |A|/2$  und  $|A_2| = |A_3| = |A|/4$ . Die Nachbarelemente entlang der entsprechenden Kanten bezeichnen wir wieder mit  $B, C, D \in \mathcal{T}_\ell$ . Es gilt

$$\begin{aligned} & \sum_{\substack{E \in \mathcal{E}_{\ell+1}^* \\ |A \cap \Omega_{\ell+1, E}| > 0}} \text{osc}_{\ell+1}(E|A)^2 \\ &= \text{osc}_{\ell+1}(c_1|A)^2 + \text{osc}_{\ell+1}(c_2|A)^2 + \text{osc}_{\ell+1}(d_1|A)^2 + \text{osc}_{\ell+1}(d_2|A)^2 + \text{osc}_{\ell+1}(b|A)^2 \\ &\quad + \text{osc}_{\ell+1}(a_1|A)^2 + \text{osc}_{\ell+1}(a_2|A)^2 \\ &= (|\Omega_{\ell+1, c_1} \cap C| + |A|/2) \|f - f_{\Omega_{\ell+1, c_1}}\|_{L^2(A_1)}^2 + (|\Omega_{\ell+1, c_2} \cap C| + |A|/4) \|f - f_{\Omega_{\ell+1, c_2}}\|_{L^2(A_3)}^2 \\ &\quad + (|\Omega_{\ell+1, d_1} \cap D| + |A|/4) \|f - f_{\Omega_{\ell+1, d_1}}\|_{L^2(A_2)}^2 + (|\Omega_{\ell+1, d_2} \cap D| + |A|/4) \|f - f_{\Omega_{\ell+1, d_2}}\|_{L^2(A_3)}^2 \\ &\quad + (|\Omega_{\ell+1, b} \cap B| + |A|/2) \|f - f_{\Omega_{\ell+1, b}}\|_{L^2(A_1)}^2 \\ &\quad + (3|A|/4) \|f - f_{A_1 \cup A_2}\|_{L^2(A_1 \cup A_2)}^2 + (|A|/2) \|f - f_{A_2 \cup A_3}\|_{L^2(A_2 \cup A_3)}^2. \end{aligned}$$

Die letzten beiden Beiträge stammen hierbei wieder von den neuen Kanten  $a_1, a_2 \in \mathcal{E}_{\ell+1}^*$ . Wir können sie abschätzen durch

$$(3|A|/4) \|f - f_{A_1 \cup A_2}\|_{L^2(A_1 \cup A_2)}^2 + (|A|/2) \|f - f_{A_2 \cup A_3}\|_{L^2(A_2 \cup A_3)}^2 \leq (5|A|/4) \|f - f_A\|_{L^2(A)}^2.$$

Mit denselben Ideen wie zuvor definieren wir

$$\begin{aligned} o_\ell(b|A)^2 &= (|\Omega_{\ell+1,b} \cap B| + |A|/2) \|f - f_{\Omega_{\ell+1,b}}\|_{L^2(A_1)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2, \\ o_\ell(c|A)^2 &= (|\Omega_{\ell+1,c_1} \cap C| + |A|/2) \|f - f_{\Omega_{\ell+1,c_1}}\|_{L^2(A_1)}^2 \\ &\quad + (|\Omega_{\ell+1,c_2} \cap C| + |A|/4) \|f - f_{\Omega_{\ell+1,c_2}}\|_{L^2(A_3)}^2 + (|A|/4) \|f - f_A\|_{L^2(A)}^2, \\ o_\ell(d|A)^2 &= (|\Omega_{\ell+1,d_1} \cap D| + |A|/4) \|f - f_{\Omega_{\ell+1,d_1}}\|_{L^2(A_2)}^2 \\ &\quad + (|\Omega_{\ell+1,d_2} \cap D| + |A|/4) \|f - f_{\Omega_{\ell+1,d_2}}\|_{L^2(A_3)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2, \end{aligned}$$

und erhalten wieder die Gültigkeit von (3.3.14).

Als viertes und letztes schauen wir uns den Fall an, bei dem  $A \in \mathcal{T}_\ell$  mit den Kanten  $b, c, d \in \mathcal{E}_\ell^*$  durch drei Bisektionen in die Dreiecke  $A_1, A_2, A_3, A_4 \in \mathcal{T}_{\ell+1}$  zerlegt wird (siehe Abb. 3.2 rechts). Die Kanten  $b, c, d$  werden hierbei zu  $b_1, b_2, c_1, c_2, d_1, d_2 \in \mathcal{E}_{\ell+1}^*$  aufgespalten. Im Inneren entstehen dadurch die neuen Kanten  $a_1, a_2, a_3 \in \mathcal{E}_{\ell+1}^*$ . Ferner hat jedes der neuen Dreiecke ein viertel der Fläche von  $A$ . Wie in den anderen Fällen notieren wir die Nachbarelemente wieder mit  $B, C, D \in \mathcal{T}_{\ell+1}$ , bzw. der leeren Menge, sollte es sich um eine Randkante handeln. Wir erhalten nun

$$\begin{aligned} &\sum_{\substack{E \in \mathcal{E}_{\ell+1}^* \\ |A \cap \Omega_{\ell+1,E}| > 0}} \text{osc}_{\ell+1}(E|A)^2 \\ &= \text{osc}_{\ell+1}(b_1|A)^2 + \text{osc}_{\ell+1}(b_2|A)^2 + \text{osc}_{\ell+1}(c_1|A)^2 + \text{osc}_{\ell+1}(c_2|A)^2 \\ &\quad + \text{osc}_{\ell+1}(d_1|A)^2 + \text{osc}_{\ell+1}(d_2|A)^2 + \text{osc}_{\ell+1}(a_1|A)^2 + \text{osc}_{\ell+1}(a_2|A)^2 + \text{osc}_{\ell+1}(a_3|A)^2 \\ &\leq (|\Omega_{\ell+1,b_1} \cap B| + |A|/4) \|f - f_{\Omega_{\ell+1,b_1}}\|_{L^2(A_4)}^2 + (|\Omega_{\ell+1,b_2} \cap B| + |A|/4) \|f - f_{\Omega_{\ell+1,b_2}}\|_{L^2(A_1)}^2 \\ &\quad + (|\Omega_{\ell+1,c_1} \cap C| + |A|/4) \|f - f_{\Omega_{\ell+1,c_1}}\|_{L^2(A_1)}^2 + (|\Omega_{\ell+1,c_2} \cap C| + |A|/4) \|f - f_{\Omega_{\ell+1,c_2}}\|_{L^2(A_3)}^2 \\ &\quad + (|\Omega_{\ell+1,d_1} \cap D| + |A|/4) \|f - f_{\Omega_{\ell+1,d_1}}\|_{L^2(A_2)}^2 + (|\Omega_{\ell+1,d_2} \cap D| + |A|/4) \|f - f_{\Omega_{\ell+1,d_2}}\|_{L^2(A_3)}^2 \\ &\quad + (3|A|/2) \|f - f_A\|_{L^2(A)}^2 \end{aligned}$$

und definieren

$$\begin{aligned} o_\ell(b|A)^2 &= (|\Omega_{\ell+1,b_1} \cap B| + |A|/4) \|f - f_{\Omega_{\ell+1,b_1}}\|_{L^2(A_4)}^2 \\ &\quad + (|\Omega_{\ell+1,b_2} \cap B| + |A|/4) \|f - f_{\Omega_{\ell+1,b_2}}\|_{L^2(A_1)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2 \\ o_\ell(c|A)^2 &= (|\Omega_{\ell+1,c_1} \cap C| + |A|/4) \|f - f_{\Omega_{\ell+1,c_1}}\|_{L^2(A_1)}^2 \\ &\quad + (|\Omega_{\ell+1,c_2} \cap C| + |A|/4) \|f - f_{\Omega_{\ell+1,c_2}}\|_{L^2(A_3)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2 \\ o_\ell(d|A)^2 &= (|\Omega_{\ell+1,d_1} \cap D| + |A|/4) \|f - f_{\Omega_{\ell+1,d_1}}\|_{L^2(A_2)}^2 \\ &\quad + (|\Omega_{\ell+1,d_2} \cap D| + |A|/4) \|f - f_{\Omega_{\ell+1,d_2}}\|_{L^2(A_3)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2. \end{aligned}$$

Wir zuvor erhalten wir wieder (3.3.14).

Im Folgenden gilt es zu zeigen, dass für die nicht-verfeinerten Kanten

$$\sum_{\substack{T \in \mathcal{T}_\ell \\ T \subset \Omega_{\ell,E}}} o_\ell(E|T)^2 \leq \text{osc}_\ell(E)^2 \quad \text{für alle } E \in \mathcal{E}_\ell^* \cap \mathcal{E}_{\ell+1}^*, \quad (3.3.15)$$

und für die verfeinerten Kanten

$$\sum_{\substack{T \in \mathcal{T}_\ell \\ T \subset \Omega_{\ell,E}}} o_\ell(E|T)^2 \leq \frac{3}{4} \text{osc}_\ell(E)^2 \quad \text{für alle } E \in \mathcal{E}_\ell^* \setminus \mathcal{E}_{\ell+1}^* \quad (3.3.16)$$

gilt. Gelingt es uns diese Aussagen zu zeigen, so führen wir den Beweis folgendermaßen zu Ende: Aus (3.3.14) folgt

$$\begin{aligned} \sum_{E \in \mathcal{E}_{\ell+1}^*} \text{osc}_{\ell+1}(E)^2 &= \sum_{T \in \mathcal{T}_\ell} \sum_{\substack{E \in \mathcal{E}_{\ell+1}^* \\ |T \cap \Omega_{\ell+1,E}| > 0}} \text{osc}_{\ell+1}(E|T)^2 \leq \sum_{T \in \mathcal{T}_\ell} \sum_{\substack{E \in \mathcal{E}_\ell^* \\ E \subset \partial T}} o_\ell(E|T)^2 \\ &= \sum_{E \in \mathcal{E}_\ell^*} \sum_{\substack{T \in \mathcal{T}_\ell \\ T \subset \Omega_{\ell,E}}} o_\ell(E|T)^2. \end{aligned}$$

Die beiden Aussagen (3.3.15) und (3.3.16) liefern dann

$$\begin{aligned} \sum_{E \in \mathcal{E}_{\ell+1}^*} \text{osc}_{\ell+1}(E)^2 &\leq \sum_{E \in \mathcal{E}_\ell^* \cap \mathcal{E}_{\ell+1}^*} \text{osc}_\ell(E)^2 + \frac{3}{4} \sum_{E \in \mathcal{E}_\ell^* \setminus \mathcal{E}_{\ell+1}^*} \text{osc}_\ell(E)^2 \\ &= \sum_{E \in \mathcal{E}_\ell^*} \text{osc}_\ell(E)^2 - \frac{1}{4} \sum_{E \in \mathcal{E}_\ell^* \setminus \mathcal{E}_{\ell+1}^*} \text{osc}_\ell(E)^2, \end{aligned}$$

so dass wir durch die Tatsache  $\mathcal{E}_\ell^* \cap \mathcal{M}_\ell \subseteq \mathcal{E}_\ell^* \setminus \mathcal{E}_{\ell+1}^*$  den Beweis abschließen können. Der gesamte Rest dieses Beweises widmet sich nun der Validierung der Aussagen (3.3.15) und (3.3.16). Hierzu müssen wir alle möglichen Konstellationen von Dreiecken und Verfeinerungen betrachten. Es gilt also eine Menge Fälle zu untersuchen. Wir beschäftigen uns zunächst mit denjenigen, in denen die fragwürdige Kante  $E \in \mathcal{E}_\ell^*$  eine Randkante ist. Die möglichen Fälle sind in Abbildung 3.3 dargestellt. Wir werden sie im Folgenden alle einzeln abarbeiten, wobei die Analysen für die Fälle 4 und 6 analog sind.

Fall 1:

$$o_\ell(E|A)^2 = |A| \|f\|_{L^2(A)}^2 = \text{osc}_\ell(E)^2.$$

Fall 2:

$$o_\ell(E|A)^2 = |A|/2 \|f\|_{L^2(A_1)}^2 + (|A|/2) \|f\|_{L^2(A_2)}^2 = (|A|/2) \|f\|_{L^2(A)}^2 \leq \frac{3}{4} \text{osc}_\ell(E)^2.$$

Fall 3:

$$o_\ell(E|A)^2 = (|A|/2) \|f\|_{L^2(A_1)}^2 + (|A|/4) \|f\|_{L^2(A_3)}^2 + (|A|/4) \|f - f_A\|_{L^2(A)}^2 \leq \frac{3}{4} \text{osc}_\ell(E)^2.$$

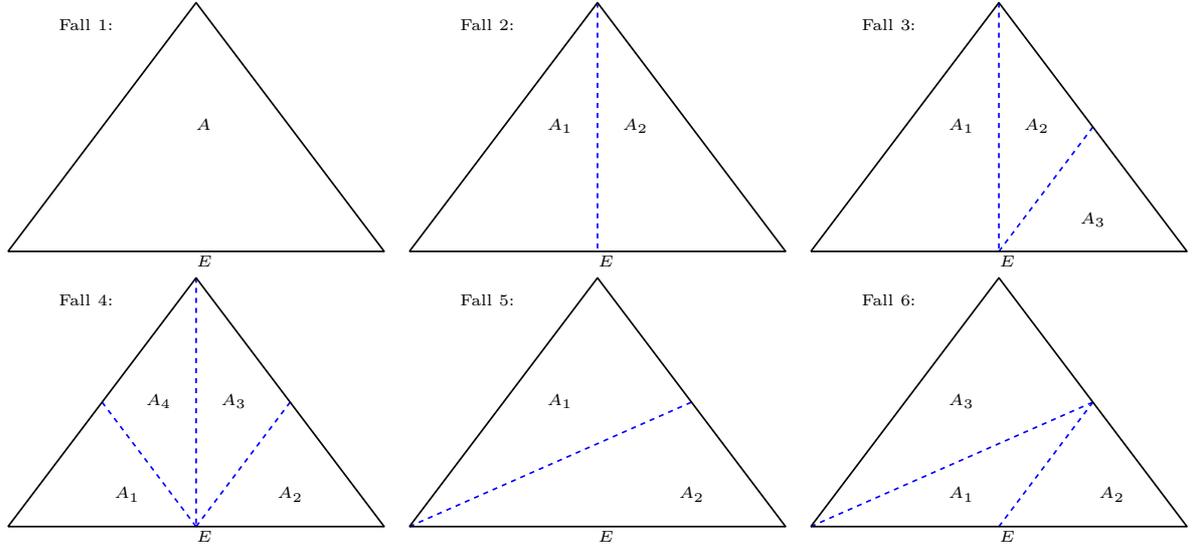


Abbildung 3.3.: Mögliche Fälle bei denen  $E \in \mathcal{E}_\ell^*$  eine Randkante ist.

Fall (4  $\hat{=}$  6):

$$\begin{aligned} o_\ell(E|A)^2 &= (|A|/4)\|f\|_{L^2(A_1)}^2 + (|A|/4)\|f\|_{L^2(A_2)}^2 + (|A|/2)\|f - f_A\|_{L^2(A)}^2 \\ &\leq (|A|/4)\|f\|_{L^2(A)}^2 + (|A|/2)\|f\|_{L^2(A)}^2 \leq \frac{3}{4}osc_\ell(E)^2. \end{aligned}$$

Fall 5:

$$\begin{aligned} o_\ell(E|A)^2 &= (|A|/2)\|f\|_{L^2(A_2)}^2 + (|A|/2)\|f - f_A\|_{L^2(A)}^2 \\ &\leq (|A|/2)\|f\|_{L^2(A)}^2 + (|A|/2)\|f\|_{L^2(A)}^2 = osc_\ell(E)^2. \end{aligned}$$

Es verbleibt die Aussage für innere Kanten  $E \in \mathcal{E}_\ell$  zu zeigen. Die möglichen Fälle sind hierzu in Abbildung 3.4 zusammengefasst. Auch hier führen wieder einige Fälle auf die analoge Analysen.

Fall 7:

$$\begin{aligned} o_\ell(E|A)^2 &= osc_{\ell+1}(E|A)^2 = |A \cup B|\|f - f_{A \cup B}\|_{L^2(A)}^2, \\ o_\ell(E|B)^2 &= osc_{\ell+1}(E|B)^2 = |A \cup B|\|f - f_{A \cup B}\|_{L^2(B)}^2, \end{aligned}$$

und damit

$$o_\ell(E|A)^2 + o_\ell(E|B)^2 = |A \cup B|\|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 = osc_\ell(E).$$

Fall 8:

$$\begin{aligned} o_\ell(E|A)^2 &= (|A|/2 + |B|/2)\|f - f_{A_1 \cup B_1}\|_{L^2(A_1)}^2 + (|B|/2 + |A|/2)\|f - f_{A_2 \cup B_2}\|_{L^2(A_2)}^2, \\ o_\ell(E|B)^2 &= (|A|/2 + |B|/2)\|f - f_{A_1 \cup B_1}\|_{L^2(B_1)}^2 + (|A|/2 + |B|/2)\|f - f_{A_2 \cup B_2}\|_{L^2(B_2)}^2. \end{aligned}$$

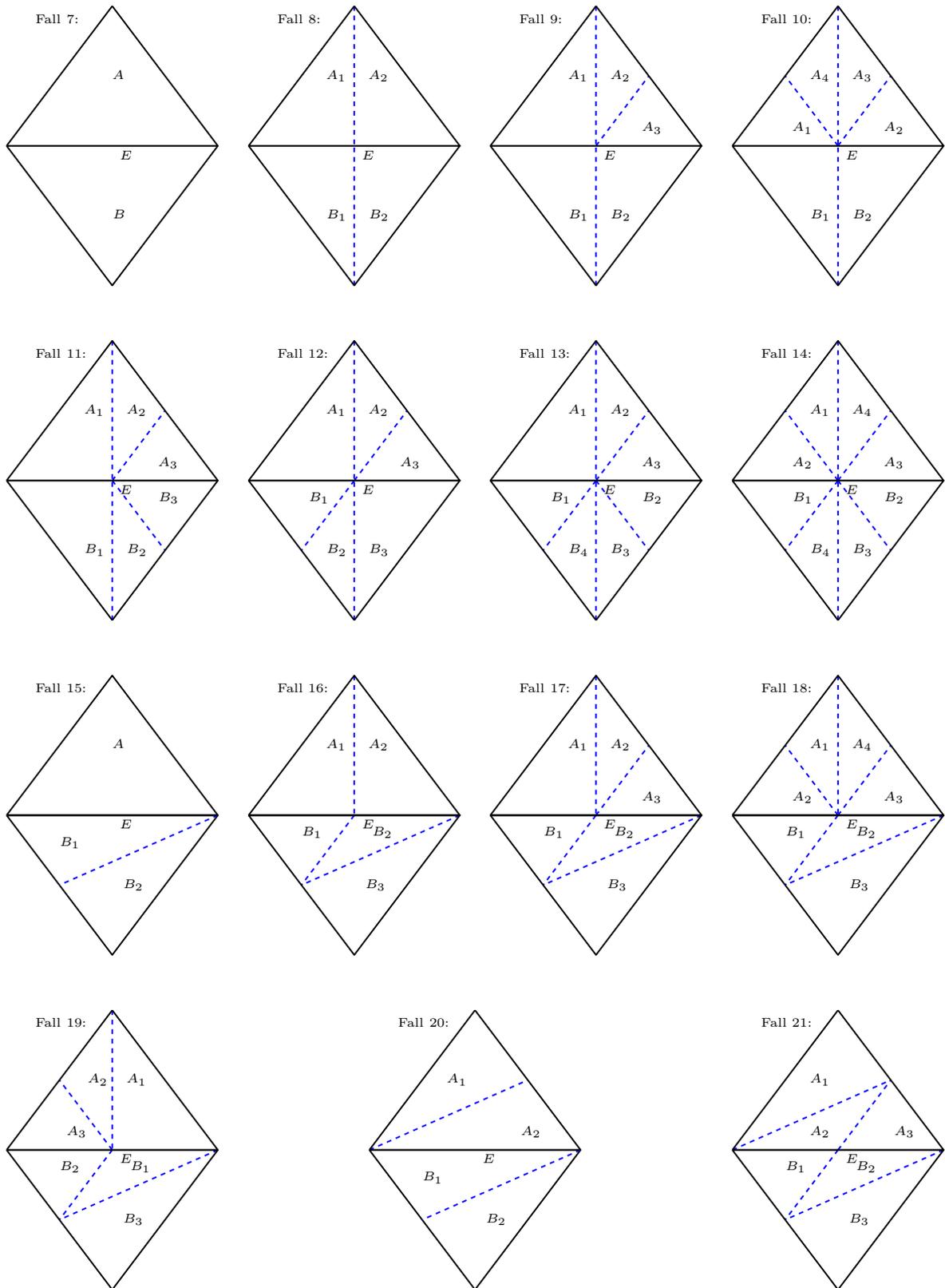


Abbildung 3.4.: Mögliche Fälle bei denen  $E \in \mathcal{E}_\ell$  eine innere Kante ist.

Somit folgt

$$\begin{aligned} o_\ell(E|A)^2 + o_\ell(E|B)^2 &\leq (|A|/2 + |B|/2) \|f - f_{A \cup B}\|_{L^2(A_1 \cup B_1)}^2 \\ &\quad + (|A|/2 + |B|/2) \|f - f_{A \cup B}\|_{L^2(A_2 \cup B_2)}^2 \\ &= (|A|/2 + |B|/2) \|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 \leq \frac{3}{4} \text{osc}_\ell(E)^2. \end{aligned}$$

Fall 9:

$$\begin{aligned} o_\ell(E|A)^2 &= (|A|/2 + |B|/2) \|f - f_{A_1 \cup B_1}\|_{L^2(A_1)}^2 + (|B|/2 + |A|/4) \|f - f_{A_3 \cup B_2}\|_{L^2(A_3)}^2 \\ &\quad + (|A|/4) \|f - f_A\|_{L^2(A)}^2 \\ o_\ell(E|B)^2 &= (|A|/2 + |B|/2) \|f - f_{A_1 \cup B_1}\|_{L^2(B_1)}^2 + (|A|/4 + |B|/2) \|f - f_{A_3 \cup B_2}\|_{L^2(B_2)}^2. \end{aligned}$$

Es ergibt sich daher

$$\begin{aligned} o_\ell(E|A)^2 + o_\ell(E|B)^2 &= (|B|/2 + |A|/2) \|f - f_{A_1 \cup B_1}\|_{L^2(A_1 \cup B_1)}^2 \\ &\quad + (|B|/2 + |A|/4) \|f - f_{A_3 \cup B_2}\|_{L^2(A_3 \cup B_2)}^2 + (|A|/4) \|f - f_A\|_{L^2(A)}^2 \\ &\leq \frac{3}{4} \text{osc}_\ell(E)^2. \end{aligned}$$

Fall (10  $\hat{=}$  16):

$$\begin{aligned} o_\ell(E|A)^2 &= (|B|/2 + |A|/4) \|f - f_{A_1 \cup B_1}\|_{L^2(A_1)}^2 \\ &\quad + (|B|/2 + |A|/4) \|f - f_{A_2 \cup B_2}\|_{L^2(A_2)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2, \\ o_\ell(E|B)^2 &= (|A|/4 + |B|/2) \|f - f_{A_1 \cup B_1}\|_{L^2(B_1)}^2 + (|A|/4 + |B|/2) \|f - f_{A_2 \cup B_2}\|_{L^2(B_2)}^2, \end{aligned}$$

und damit

$$\begin{aligned} o_\ell(E|A)^2 + o_\ell(E|B)^2 &\leq (|B|/2 + |A|/4) \|f - f_{A \cup B}\|_{L^2(A_1 \cup B_1)}^2 \\ &\quad + (|B|/2 + |A|/4) \|f - f_{A \cup B}\|_{L^2(A_2 \cup B_2)}^2 + (|A|/2) \|f - f_{A \cup B}\|_{L^2(A)}^2 \\ &\leq (|B|/2 + |A|/4) \|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 + (|A|/2) \|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 \\ &\leq \frac{3}{4} \text{osc}_\ell(E)^2 \end{aligned}$$

Fall 11:

$$\begin{aligned} o_\ell(E|A)^2 &= (|B|/2 + |A|/2) \|f - f_{A_1 \cup B_1}\|_{L^2(A_1)}^2 \\ &\quad + (|B|/4 + |A|/4) \|f - f_{A_3 \cup B_3}\|_{L^2(A_3)}^2 + (|A|/4) \|f - f_A\|_{L^2(A)}^2, \\ o_\ell(E|B)^2 &= (|B|/2 + |A|/2) \|f - f_{A_1 \cup B_1}\|_{L^2(B_1)}^2 \\ &\quad + (|B|/4 + |A|/4) \|f - f_{A_3 \cup B_3}\|_{L^2(B_3)}^2 + (|B|/4) \|f - f_B\|_{L^2(B)}^2. \end{aligned}$$

Folglich gilt

$$\begin{aligned} o_\ell(E|A)^2 + o_\ell(E|B)^2 &\leq (|B|/2 + |A|/2) \|f - f_{A \cup B}\|_{L^2(A_1 \cup B_1)}^2 \\ &\quad + (|B|/4 + |A|/4) \|f - f_{A \cup B}\|_{L^2(A_3 \cup B_3)}^2 + (|A|/4) \|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 \\ &\quad + (|B|/4) \|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 \leq \frac{3}{4} \text{osc}_\ell(E)^2. \end{aligned}$$

Fall 12:

$$\begin{aligned} o_\ell(E|A)^2 &= (|B|/4 + |A|/2) \|f - f_{A_1 \cup B_1}\|_{L^2(A_1)}^2 \\ &\quad + (|B|/2 + |A|/4) \|f - f_{A_3 \cup B_3}\|_{L^2(A_3)}^2 + (|A|/4) \|f - f_A\|_{L^2(A)}^2, \\ o_\ell(E|B)^2 &= (|A|/2 + |B|/4) \|f - f_{A_1 \cup B_1}\|_{L^2(B_1)}^2 \\ &\quad + (|B|/2 + |A|/4) \|f - f_{A_3 \cup B_3}\|_{L^2(B_3)}^2 + (|B|/4) \|f - f_B\|_{L^2(B)}^2, \end{aligned}$$

und somit

$$\begin{aligned} o_\ell(E|A)^2 + o_\ell(E|B)^2 &\leq (|B|/4 + |A|/2) \|f - f_{A \cup B}\|_{L^2(A_1 \cup B_1)}^2 \\ &\quad + (|B|/2 + |A|/4) \|f - f_{A \cup B}\|_{L^2(A_3 \cup B_3)}^2 + (|A|/4) \|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 \\ &\quad + (|B|/4) \|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 \leq \frac{3}{4} \text{osc}_\ell(E)^2. \end{aligned}$$

Fall (13  $\hat{=}$  17  $\hat{=}$  19):

$$\begin{aligned} o_\ell(E|A)^2 &= (|A|/2 + |B|/4) \|f - f_{A_1 \cup B_1}\|_{L^2(A_1)}^2 \\ &\quad + (|A|/4 + |B|/4) \|f - f_{A_3 \cup B_2}\|_{L^2(A_3)}^2 + (|A|/4) \|f - f_A\|_{L^2(A)}^2, \\ o_\ell(E|B)^2 &= (|A|/2 + |B|/4) \|f - f_{A_1 \cup B_1}\|_{L^2(B_1)}^2 \\ &\quad + (|A|/4 + |B|/4) \|f - f_{A_3 \cup B_2}\|_{L^2(B_2)}^2 + (|B|/2) \|f - f_B\|_{L^2(B)}^2. \end{aligned}$$

Wie zuvor erhalten wir

$$\begin{aligned} o_\ell(E|A)^2 + o_\ell(E|B)^2 &\leq (|A|/2 + |B|/4) \|f - f_{A \cup B}\|_{L^2(A_1 \cup B_1)}^2 \\ &\quad + (|A|/4 + |B|/4) \|f - f_{A \cup B}\|_{L^2(A_3 \cup B_2)}^2 \\ &\quad + (|A|/4) \|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 + (|B|/2) \|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 \\ &\leq \frac{3}{4} \text{osc}_\ell(E)^2. \end{aligned}$$

Fall (14  $\hat{=}$  18  $\hat{=}$  21):

$$\begin{aligned} o_\ell(E|A)^2 &= (|A|/4 + |B|/4) \|f - f_{A_2 \cup B_1}\|_{L^2(A_2)}^2 \\ &\quad + (|A|/4 + |B|/4) \|f - f_{A_3 \cup B_2}\|_{L^2(A_3)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2, \\ o_\ell(E|B)^2 &= (|A|/4 + |B|/4) \|f - f_{A_2 \cup B_1}\|_{L^2(B_1)}^2 \\ &\quad + (|A|/4 + |B|/4) \|f - f_{A_3 \cup B_2}\|_{L^2(B_2)}^2 + (|B|/2) \|f - f_B\|_{L^2(B)}^2, \end{aligned}$$

und daher

$$\begin{aligned} o_\ell(E|A)^2 + o_\ell(E|B)^2 &\leq (|A|/4 + |B|/4) \|f - f_{A \cup B}\|_{L^2(A_2 \cup B_1)}^2 \\ &\quad + (|A|/4 + |B|/4) \|f - f_{A \cup B}\|_{L^2(A_3 \cup B_2)}^2 + (|A|/2) \|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 \\ &\quad + (|B|/2) \|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 \leq \frac{3}{4} \text{osc}_\ell(E)^2. \end{aligned}$$

Fall 15:

$$\begin{aligned} o_\ell(E|A)^2 &= \text{osc}_{\ell+1}(E|A)^2 = (|A| + |B|/2) \|f - f_{A \cup B_1}\|_{L^2(A)}^2, \\ o_\ell(E|B)^2 &= (|A| + |B|/2) \|f - f_{A \cup B_1}\|_{L^2(B_1)}^2 + (|B|/2) \|f - f_B\|_{L^2(B)}^2. \end{aligned}$$

Wir erhalten daher

$$\begin{aligned} o_\ell(E|A)^2 + o_\ell(E|B)^2 &\leq (|A| + |B|/2) \|f - f_{A \cup B}\|_{L^2(A \cup B_1)}^2 \\ &\quad + (|B|/2) \|f - f_{A \cup B}\|_{L^2(A \cup B)}^2 \leq \text{osc}_\ell(E)^2. \end{aligned}$$

Fall 20:

$$\begin{aligned} o_\ell(E|A)^2 &= (|B|/2 + |A|/2) \|f - f_{B_1 \cup A_2}\|_{L^2(A_2)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2, \\ o_\ell(E|B)^2 &= (|B|/2 + |A|/2) \|f - f_{A_2 \cup B_1}\|_{L^2(B_1)}^2 + (|B|/2) \|f - f_B\|_{L^2(B)}^2, \end{aligned}$$

was schließlich zu

$$o_\ell(E|A)^2 + o_\ell(E|B)^2 \leq \text{osc}_\ell(E)^2$$

führt. Damit ist die Aussage vollständig bewiesen.  $\square$

Wir können nun den Beweis der Schätzerreduktion führen.

*Beweis von Satz 3.3.3.* Analog zum linearen Fall (vgl. Satz 2.2.6), erhalten wir durch Anwendung der Dreiecksungleichung im Folgenraum  $\ell^2$

$$\begin{aligned} \mu_{\ell+1} &= \left( \sum_{T' \in \mathcal{T}_{\ell+1, \Gamma}} \text{osc}_{\ell+1}(T')^2 + \sum_{E' \in \mathcal{E}_{\ell+1}^*} \text{osc}_{\ell+1}(E')^2 + \sum_{E' \in \mathcal{E}_{\ell+1}} h_{E'} \|[\partial_n U_{\ell+1}]\|_{L^2(E')}^2 \right)^{1/2} \\ &\leq \left( \sum_{T' \in \mathcal{T}_{\ell+1, \Gamma}} \text{osc}_{\ell+1}(T')^2 + \sum_{E' \in \mathcal{E}_{\ell+1}^*} \text{osc}_{\ell+1}(E')^2 + \sum_{E' \in \mathcal{E}_{\ell+1}} h_{E'} \|[\partial_n U_\ell]\|_{L^2(E')}^2 \right)^{1/2} \\ &\quad + \left( \sum_{E' \in \mathcal{E}_{\ell+1}} h_{E'} \|[\partial_n (U_{\ell+1} - U_\ell)]\|_{L^2(E')}^2 \right)^{1/2} \end{aligned}$$

und durch die Anwendung der Young'schen Ungleichung (B.1.3) für ein beliebiges  $\delta > 0$

$$\begin{aligned} \mu_{\ell+1}^2 &\leq (1 + \delta) \left( \sum_{T' \in \mathcal{T}_{\ell+1, \Gamma}} \text{osc}_{\ell+1}(T')^2 + \sum_{E' \in \mathcal{E}_{\ell+1}^*} \text{osc}_{\ell+1}(E')^2 + \sum_{E' \in \mathcal{E}_{\ell+1}} h_{E'} \|[\partial_n U_\ell]\|_{L^2(E')}^2 \right) \\ &\quad + (1 + \delta^{-1}) \sum_{E' \in \mathcal{E}_{\ell+1}} h_{E'} \|[\partial_n (U_{\ell+1} - U_\ell)]\|_{L^2(E')}^2. \end{aligned} \tag{3.3.17}$$

Analog zum unrestringierten Fall folgt außerdem

$$\sum_{E' \in \mathcal{E}_{\ell+1}} h_{E'} \|[\partial_n (U_{\ell+1} - U_\ell)]\|_{L^2(E')}^2 \lesssim \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}^2 = \|U_{\ell+1} - U_\ell\|^2.$$

Wir wenden nun die soeben mühsam gezeigten Aussagen 3.3.11, 3.3.12 und 3.3.13 an, und

erhalten

$$\begin{aligned}
 & \sum_{T' \in \mathcal{T}_{\ell+1, \Gamma}} \text{osc}_{\ell+1}(T')^2 + \sum_{E' \in \mathcal{E}_{\ell+1}^*} \text{osc}_{\ell+1}(E')^2 + \sum_{E' \in \mathcal{E}_{\ell+1}} h_{E'} \|\partial_n U_\ell\|_{L^2(E')}^2 \\
 & \leq \sum_{T \in \mathcal{T}_{\ell, \Gamma}} \text{osc}_\ell(T)^2 + \sum_{E \in \mathcal{E}_\ell^*} \text{osc}_\ell(E)^2 + \sum_{E' \in \mathcal{E}_\ell} h_E \|\partial_n U_\ell\|_{L^2(E)}^2 \\
 & \quad - \frac{1}{2} \sum_{T \in \mathcal{T}_{\ell, \Gamma} \cap \mathcal{M}_\ell} \text{osc}_\ell(T)^2 - \frac{1}{4} \sum_{E \in \mathcal{E}_\ell^* \cap \mathcal{M}_\ell} \text{osc}_\ell(E)^2 - \frac{1}{2} \sum_{E' \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} h_E \|\partial_n U_\ell\|_{L^2(E)}^2 \\
 & \leq \mu_\ell^2 - \frac{1}{4} \left( \sum_{T \in \mathcal{T}_{\ell, \Gamma} \cap \mathcal{M}_\ell} \text{osc}_\ell(T)^2 + \sum_{E \in \mathcal{E}_\ell^* \cap \mathcal{M}_\ell} \text{osc}_\ell(E)^2 + \sum_{E' \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} h_E \|\partial_n U_\ell\|_{L^2(E)}^2 \right).
 \end{aligned}$$

Schließlich ergibt sich mit der Dörflermarkierung (3.3.9)

$$\begin{aligned}
 & \sum_{T' \in \mathcal{T}_{\ell+1, \Gamma}} \text{osc}_{\ell+1}(T')^2 + \sum_{E' \in \mathcal{E}_{\ell+1}^*} \text{osc}_{\ell+1}(E')^2 + \sum_{E' \in \mathcal{E}_{\ell+1}} h_{E'} \|\partial_n U_\ell\|_{L^2(E')}^2 \\
 & \leq (1 - \theta/4) \mu_\ell^2.
 \end{aligned}$$

Wir wählen nun  $\delta > 0$  so, dass  $q := (1 + \delta)(1 - \theta/4) < 1$ . Einsetzen der bisherigen Resultate in (3.3.17) schließt den Beweis ab.  $\square$

**Korollar 3.3.7.** *Erfüllt die Menge der markierten Kanten und Randelemente eine Dörflermarkierung der Form*

$$\sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} \text{osc}_\ell(E)^2 + \sum_{T \in \mathcal{T}_{\ell, \Gamma} \cap \mathcal{M}_\ell} \text{osc}_\ell(T)^2 \geq \theta \left( \sum_{E \in \mathcal{E}_\ell} \text{osc}_\ell(E)^2 + \sum_{T \in \mathcal{T}_{\ell, \Gamma}} \text{osc}_\ell(T)^2 \right),$$

so gibt es ein  $\kappa \in (0, 1)$  mit

$$\text{Osc}_{\ell+1} \leq \kappa \text{Osc}_\ell.$$

*Beweis.* Obige Dörflermarkierung impliziert mit

$$\begin{aligned}
 \frac{\theta}{4} \left( \sum_{E \in \mathcal{E}_\ell^*} \text{osc}_\ell(E)^2 + \sum_{T \in \mathcal{T}_{\ell, \Gamma}} \text{osc}_\ell(T)^2 \right) & \leq \theta \left( \sum_{E \in \mathcal{E}_\ell} \text{osc}_\ell(E)^2 + \sum_{T \in \mathcal{T}_{\ell, \Gamma}} \text{osc}_\ell(T)^2 \right) \\
 & \leq \sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} \text{osc}_\ell(E)^2 + \sum_{T \in \mathcal{T}_{\ell, \Gamma}} \text{osc}_\ell(T)^2
 \end{aligned}$$

auch eine Dörflermarkierung für den Oszillationsteil von  $\mu_\ell$ . Unter Verwendung der Lemmata 3.3.4 und 3.3.6 folgt die Aussage mit zum Beweis der Schätzerreduktion analoger Argumentation.  $\square$

Bevor wir uns nun mit der tatsächlichen Konvergenz der Folge der diskreten Lösungen  $(U_\ell)_{\ell \in \mathbb{N}}$  gegen die exakte Lösung  $u \in K$  beschäftigen, wollen wir uns kurz klarmachen, dass die bewiesenen Aussagen auch für andere Verfeinerungsstrategien ihre Gültigkeit behalten. Das Prinzip der Schätzerreduktion kann somit von der Verfeinerungsstrategie weitestgehend unabhängig eingesetzt werden.

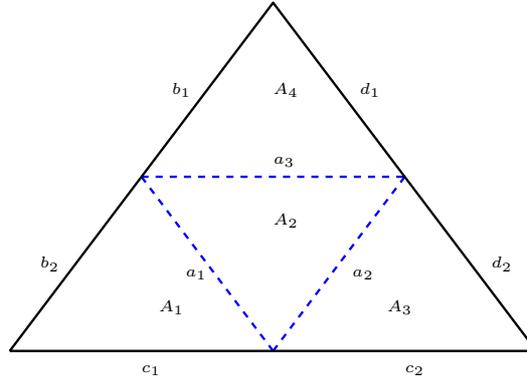


Abbildung 3.5.: Rot-Verfeinerung eines Elementes  $A \in \mathcal{T}_\ell$ .

Werden die markierten Elemente mit  $bisec_5$  bzw. mit Rot-Verfeinerung (siehe Abschnitt 2.2.3) verfeinert, so gelten Lemma 3.3.4 und 3.3.5 offenbar weiterhin. Gleiches gilt auch für Lemma 3.3.6. Dies kann folgendermaßen eingesehen werden: Für den Fall der  $bisec_5$ -Verfeinerung konstruieren wir eine theoretische Zwischentriangulierung  $\tilde{\mathcal{T}}_{\ell+1}$ , bei der die für  $bisec_5$ -Verfeinerung markierten Elemente mit drei Bisektionen verfeinert wurden. Für diesen Zwischenschritt können wir nun Lemma 3.3.6 verwenden und erhalten

$$\tilde{osc}_{\ell+1}^2 \leq \sum_{E \in \mathcal{E}_\ell^*} osc_\ell(E)^2 - \frac{1}{4} \sum_{E \in \mathcal{E}_\ell^* \cap \mathcal{M}_\ell} osc_\ell(E)^2,$$

wobei  $\tilde{osc}_{\ell+1}^2$  die Kantenoszillationen auf  $\tilde{\mathcal{T}}_{\ell+1}$  bezeichnet. Um die  $bisec_5$ -verfeinerte Triangulierung  $\mathcal{T}_{\ell+1}$  zu erhalten, müssen nun einige Elemente  $T' \in \tilde{\mathcal{T}}_{\ell+1}$  durch weitere Bisektionen verfeinert werden. Auf diese Verfeinerung ist jedoch wieder Lemma 3.3.6 anwendbar. Da aus diesem Lemma insbesondere das monotone Abklingen der Kantenoszillationen folgt, erhalten wir mit

$$osc_{\ell+1}^2 \leq \tilde{osc}_{\ell+1}^2 \leq \sum_{E \in \mathcal{E}_\ell^*} osc_\ell(E)^2 - \frac{1}{4} \sum_{E \in \mathcal{E}_\ell^* \cap \mathcal{M}_\ell} osc_\ell(E)^2$$

die gesuchte Aussage. Für den Fall der Rot-Verfeinerung können wir die Aussage von Lemma 3.3.6, wie im Beweis des selbigen, durch einfache Berechnungen verifizieren. Ist ein Element  $A \in \mathcal{T}_\ell$  mit den Kanten  $b, c, d \in \mathcal{E}_\ell^*$  für Rot-Verfeinerung markiert, so wird es, wie in Abbildung 3.5 dargestellt, zerteilt. Die Kanten werden hierbei in  $b_1, b_2, c_1, c_2, d_1, d_2 \in \mathcal{E}_{\ell+1}^*$  aufgespalten, so dass die neuen Kanten  $a_1, a_2, a_3 \in \mathcal{E}_{\ell+1}^*$  entstehen. Für die Söhne  $A_1, A_2, A_3, A_4 \in \mathcal{T}_{\ell+1}$  von  $A$  gilt analog zum NVB-Fall  $|A_j| = |A|/4$ , für  $j \in (1, \dots, 4)$ . Bezeichnen wir die Nachbarelemente entlang der entsprechenden Kanten wieder mit  $B, C, D \in \mathcal{T}_\ell$ , so erhalten wir

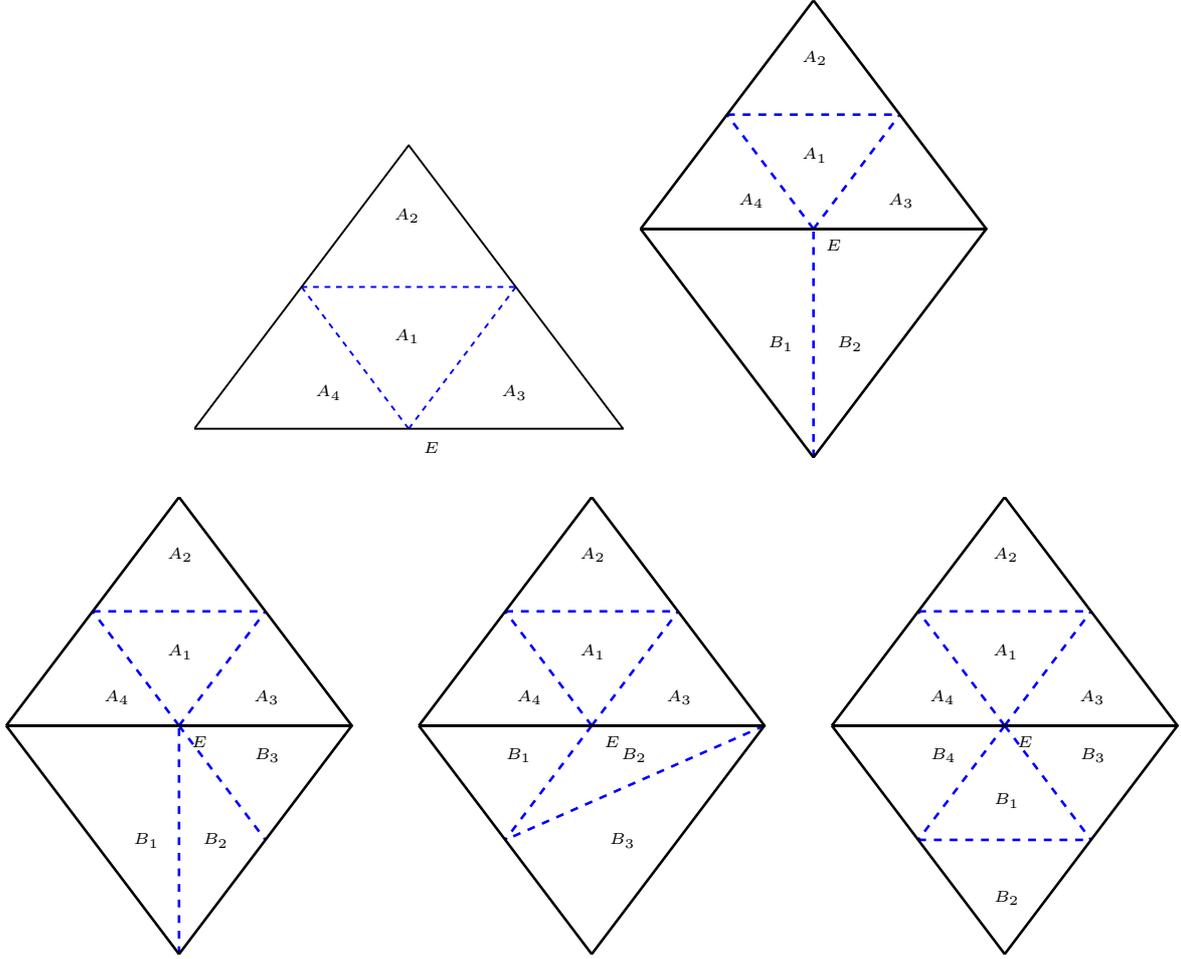


Abbildung 3.6.: Zusätzlich auftretende Fälle bei Rot-Verfeinerung.

$$\begin{aligned}
 & \sum_{\substack{E \in \mathcal{E}_{\ell+1}^* \\ |A \cap \Omega_{\ell+1, E}| > 0}} \text{osc}_{\ell+1}(E|A)^2 \\
 &= \text{osc}_{\ell+1}(b_1|A)^2 + \text{osc}_{\ell+1}(b_2|A)^2 + \text{osc}_{\ell+1}(c_1|A)^2 + \text{osc}_{\ell+1}(c_2|A)^2 \\
 &\quad + \text{osc}_{\ell+1}(d_1|A)^2 + \text{osc}_{\ell+1}(d_2|A)^2 + \text{osc}_{\ell+1}(a_1|A)^2 + \text{osc}_{\ell+1}(a_2|A)^2 + \text{osc}_{\ell+1}(a_3|A)^2 \\
 &\leq (|\Omega_{\ell+1, b_1} \cap B| + |A|/4) \|f - f_{\Omega_{\ell+1, b_1}}\|_{L^2(A_4)}^2 + (|\Omega_{\ell+1, b_2} \cap B| + |A|/4) \|f - f_{\Omega_{\ell+1, b_2}}\|_{L^2(A_1)}^2 \\
 &\quad + (|\Omega_{\ell+1, c_1} \cap C| + |A|/4) \|f - f_{\Omega_{\ell+1, c_1}}\|_{L^2(A_1)}^2 + (|\Omega_{\ell+1, c_2} \cap C| + |A|/4) \|f - f_{\Omega_{\ell+1, c_2}}\|_{L^2(A_3)}^2 \\
 &\quad + (|\Omega_{\ell+1, d_1} \cap D| + |A|/4) \|f - f_{\Omega_{\ell+1, d_1}}\|_{L^2(A_4)}^2 + (|\Omega_{\ell+1, d_2} \cap D| + |A|/4) \|f - f_{\Omega_{\ell+1, d_2}}\|_{L^2(A_3)}^2 \\
 &\quad + (3(|A|/2)) \|f - f_A\|_{L^2(A)}^2,
 \end{aligned}$$

für die Kantenoszillationen im  $(\ell + 1)$ -ten Schritt. Den von den neuen Kanten stammenden Beitrag  $(3(|A|/2)) \|f - f_A\|_{L^2(A)}^2$  können wir nun, analog zum *bisec*<sub>3</sub>-Fall auf die Oszillationen

der ursprünglichen Kanten verteilen, und definieren

$$\begin{aligned}
 o_\ell(b|A)^2 &= (|\Omega_{\ell+1,b_1} \cap B| + |A|/4) \|f - f_{\Omega_{\ell+1,b_1}}\|_{L^2(A_4)}^2 \\
 &\quad + (|\Omega_{\ell+1,b_2} \cap B| + |A|/4) \|f - f_{\Omega_{\ell+1,b_2}}\|_{L^2(A_1)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2, \\
 o_\ell(c|A)^2 &= (|\Omega_{\ell+1,c_1} \cap C| + |A|/4) \|f - f_{\Omega_{\ell+1,c_1}}\|_{L^2(A_1)}^2 \\
 &\quad + (|\Omega_{\ell+1,c_2} \cap C| + |A|/4) \|f - f_{\Omega_{\ell+1,c_2}}\|_{L^2(A_3)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2, \\
 o_\ell(d|A)^2 &= (|\Omega_{\ell+1,d_1} \cap D| + |A|/4) \|f - f_{\Omega_{\ell+1,d_1}}\|_{L^2(A_4)}^2 \\
 &\quad + (|\Omega_{\ell+1,d_2} \cap D| + |A|/4) \|f - f_{\Omega_{\ell+1,d_2}}\|_{L^2(A_3)}^2 + (|A|/2) \|f - f_A\|_{L^2(A)}^2.
 \end{aligned}$$

Da dieser Fall genau die gleichen Größen wie bei der *bisec*<sub>3</sub>-Verfeinerung behandelt, gilt natürlich auch (3.3.14). Es bleibt zu zeigen, dass die Ungleichungen (3.3.15) und (3.3.16) für die in Abbildung 3.6 dargestellten, neuen Fälle auch erfüllt sind. Da die Größen  $o_\ell(b|A)^2$ ,  $o_\ell(c|A)^2$  und  $o_\ell(d|A)^2$  jedoch genau wie bei der *bisec*<sub>3</sub>-Verfeinerung definiert sind, sind die Analysen analog zu den bereits behandelten Fällen 4, 10, 13, 18 und 14.

Wir wollen nun endlich zur angekündigten Konvergenzaussage kommen, deren Beweis noch ein kleines Lemma erfordert.

**Lemma 3.3.8** (a-priori Konvergenz konformer Diskretisierungen). *Die Folge der diskreten Lösungen  $(U_\ell)_{\ell \in \mathbb{N}}$  konvergiert in  $H_0^1(\Omega)$ . Im besonderen gilt  $\lim_{\ell \rightarrow \infty} \|U_{\ell+1} - U_\ell\| = 0$ .*

*Beweis.* Wir definieren die Menge  $K_\infty$  als den Abschluss von  $\bigcup_{\ell=0}^\infty K_\ell$  in  $H_0^1(\Omega)$ . Da die Räume  $K_\ell$  nach der Bemerkung am Anfang dieses Kapitels konform sind, d.h.  $K_\ell \subseteq K_{\ell+1}$ , ist die Menge  $K_\infty$  als Vereinigung geschachtelter konvexer Mengen selbst wieder konvex. Per Definition ist sie außerdem abgeschlossen. Analog zur Argumentation am Beginn dieses Kapitels gibt es somit ein eindeutiges Element  $U_\infty \in K_\infty$ , welches die Variationsungleichung

$$\langle U_\infty, V_\infty - U_\infty \rangle \geq (f, V_\infty - U_\infty) \quad \text{für alle } V_\infty \in K_\infty$$

erfüllt, und somit das Energiefunktional

$$\mathcal{J}(v) = \frac{1}{2} \langle v, v \rangle - (f, v)$$

über  $K_\infty$  minimiert. Da  $\bigcup_{\ell=0}^\infty K_\ell$  in  $K_\infty$  dicht ist, gibt es für jedes  $\varepsilon > 0$  einen Index  $\ell_0$  und eine Funktion  $U_{\ell_0} \in K_{\ell_0}$  mit  $\|U_\infty - U_{\ell_0}\|_{H_0^1(\Omega)} \leq \varepsilon$ . Daher folgt auch für den Fehler in Energie- und die  $L^2$ -Norm  $\|U_\infty - U_{\ell_0}\| \leq \varepsilon$ , bzw.  $\|U_\infty - U_{\ell_0}\| \leq \varepsilon$ . Sei nun  $\ell \geq \ell_0$ . Da die Räume  $K_\ell$  geschachtelt sind folgt  $U_{\ell_0} \in K_\ell$ . Aus der verallgemeinerten Céa-Ungleichung (B.1.4) ergibt sich nun für alle  $V_\ell \in K_\ell$

$$\|U_\infty - U_\ell\|^2 \leq \|U_\infty - V_\ell\|^2 - 2\langle U_\infty, U_\infty - V_\ell \rangle + 2(f, U_\infty - V_\ell).$$

Für  $U_{\ell_0} \in K_\ell$  erhalten wir also im besonderen

$$\begin{aligned}
 \|U_\infty - U_\ell\|^2 &\leq \|U_\infty - U_{\ell_0}\|^2 - 2\langle U_\infty, U_\infty - U_{\ell_0} \rangle + 2(f, U_\infty - U_{\ell_0}) \\
 &\leq \varepsilon^2 - 2(\nabla U_\infty, \nabla(U_\infty - U_{\ell_0})) + 2(f, U_\infty - U_{\ell_0}) \\
 &\leq \varepsilon^2 + 2\|U_\infty\| \|U_\infty - U_{\ell_0}\| + 2\|f\| \|U_\infty - U_{\ell_0}\| \\
 &\leq \varepsilon^2 + 2\|U_\infty\| \varepsilon + 2\|f\| \varepsilon,
 \end{aligned}$$

wobei wir hier die Ungleichung von Hölder verwendet haben. Die Folge  $(U_\ell)_{\ell \in \mathbb{N}}$  ist daher eine Cauchyfolge, so dass die gesuchte Aussage  $\lim_{\ell \rightarrow \infty} \|U_{\ell+1} - U_\ell\| = 0$  folgt. Die Konvergenz in  $H_0^1(\Omega)$  folgt schließlich aus der Abgeschlossenheit dieses Raumes.  $\square$

**Satz 3.3.9** (Konvergenz). *Sei  $\mu_\ell$  der in (3.3.6) vorgestellte Fehlerschätzer. Dennoch werde in Algorithmus 3.3.2 der Residualschätzer  $\eta_\ell$  zur Steuerung der adaptiven Verfeinerung verwendet. Für stückweise polynomiales  $f \in L^2(\Omega)$  und Verfeinerung durch newest vertex bisection konvergiert die Folge der in Algorithmus 3.3.2 berechneten diskreten Lösungen  $(U_\ell)_{\ell \in \mathbb{N}}$  gegen die exakte Lösung  $u \in K$  des kontinuierlichen Problems 3.1.4.*

*Beweis.* Durch die Äquivalenz (3.3.8) der Fehlerschätzer  $\eta_\ell$  und  $\mu_\ell$  und der Dörflermarkierung für  $\eta_\ell$  aus Algorithmus 3.3.2, erhalten wir

$$\frac{\theta}{4} \mu_\ell^2 \leq \theta \eta_\ell^2 \leq \sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} (\eta_\ell(E)^2 + \text{osc}_\ell(E)^2) + \sum_{T \in \mathcal{T}_{\ell, \Gamma} \cap \mathcal{M}_\ell} \text{osc}_\ell(T)^2.$$

Die Menge  $\mathcal{M}_\ell$  der markierten Kanten erfüllt somit auch eine Dörflermarkierung (3.3.9) für  $\mu_\ell$  bezüglich des abgeänderten Parameters  $\tilde{\theta} := \theta/4 \in (0, 1)$ . Konsequenterweise gilt nach Satz 3.3.3

$$\mu_{\ell+1}^2 \leq q \mu_\ell^2 + C \|U_{\ell+1} - U_\ell\|^2$$

für ein  $q \in (0, 1)$  und eine nur von  $\tilde{\theta}$  und der Formregularität abhängende Konstante  $C > 0$ . Mit der eben gezeigten a-priori Konvergenz konformer Diskretisierungen, ist dies ein Ausdruck der Form

$$\mu_{\ell+1}^2 \leq q \mu_\ell^2 + \alpha_\ell^2,$$

wobei  $(\alpha_\ell^2)_{\ell \in \mathbb{N}}$  gegen Null konvergiert. Induktiv erhalten wir daher

$$\mu_{\ell+1}^2 \leq q^{\ell+1} \mu_0^2 + \sum_{j=0}^{\ell} q^{\ell-j} \alpha_j^2 \leq q^{\ell+1} \mu_0^2 + \|(\alpha_\ell^2)\|_\infty \sum_{k=0}^{\ell} q^k,$$

wobei  $\|(\alpha_\ell^2)\|_\infty$  die Supremumsnorm der Nullfolge  $(\alpha_\ell^2)_{\ell \in \mathbb{N}}$  bezeichnet. Da die hintere Summe als geometrische Reihe gegen  $1/(1-q)$  konvergiert, ist die Folge der Fehlerschätzer  $\mu_\ell^2$  beschränkt, so dass  $0 \leq M := \limsup_{\ell \rightarrow \infty} \mu_\ell^2 < \infty$  existiert. Diese Überlegung liefert

$$M = \limsup_{\ell \rightarrow \infty} \mu_{\ell+1}^2 \leq q \limsup_{\ell \rightarrow \infty} \mu_\ell^2 + \limsup_{\ell \rightarrow \infty} \alpha_\ell^2 = q M,$$

und somit  $M = 0$ . Daraus schließen wir

$$0 \leq \liminf_{\ell \rightarrow \infty} \mu_\ell^2 \leq \limsup_{\ell \rightarrow \infty} \mu_\ell^2 = 0.$$

Die Folge der Fehlerschätzer  $\mu_\ell^2$  konvergiert also auch gegen Null. Da  $\eta_\ell^2 \leq \mu_\ell^2$ , folgt diese Tatsache auch für den verwendeten Residualschätzer  $\eta_\ell$ . Die Zuverlässigkeit (3.3.5) von  $\eta_\ell$  schließt den Beweis ab.  $\square$

Um die Vorzüge dieser Aussage gegenüber dem ersten Ansatz vollends einzusehen, wollen wir noch einmal die für den Konvergenzbeweis verwendeten Argumente zusammentragen. Im Zuge dieses Beweises haben wir lediglich die Konformität der Diskretisierung, und damit die Zuverlässigkeit des Schätzers und die a-priori Konvergenz, sowie einige elementare Eigenschaften verschiedener Verfeinerungsstrategien und die Dörflermarkierung selbst verwendet. Wir kommen daher ohne die diskrete lokale Zuverlässigkeit und zusätzliche Annahmen an das Abklingen der Oszillationsterme aus. Nicht nur ist dieses Vorgehen daher auf durch die Verfeinerungsstrategie weit weniger eingeschränkt, man spart sich auch zusätzlichen Programmieraufwand und den länglichen und komplizierten Beweis der diskreten lokalen Zuverlässigkeit aus dem vorangegangenen Abschnitt. Da die Argumente, abgesehen von der Zuverlässigkeit, alle allgemeiner Natur sind, ist daher auch eine Verallgemeinerung auf andere Fehlerschätzer oder Probleme denkbar.

Wir wollen uns an dieser Stelle daran erinnern, dass wir, im Gegensatz zu Abschnitt 3.2 und [BCH1], die Elementoszillationen mit  $|T|$  anstatt  $h_T^2 = \text{diam}(T)^2$  gewichtet haben. Dies hat den Vorteil, dass wir, analog zum linearen Fall, ein Abklingen der Form  $|T'| < |T|$  für einen Sohn  $T' \in \mathcal{T}_{\ell+1}$  eines markierten Elements  $T \in \mathcal{T}_{\ell,\Gamma} \cap \mathcal{M}_\ell$  garantieren können. Die Gewichtung aus [BCH1] garantiert dieses monotone Abklingen nicht. Nichtsdestotrotz führt auch die Verwendung dieses Fehlerschätzers auf einen konvergenten Algorithmus.

**Korollar 3.3.10.** *Wir nehmen an, dass der Schätzer*

$$\tilde{\eta}_\ell^2 := \sum_{E \in \mathcal{E}_\ell} (\eta_\ell(E)^2 + \text{osc}_\ell(E)^2) + \sum_{T \in \mathcal{T}_{\ell,\Gamma}} \tilde{\text{osc}}_\ell(T)^2, \quad (3.3.18)$$

mit

$$\tilde{\text{osc}}_\ell(T)^2 = h_T^2 \|f\|_{L^2(T)}^2 \quad \text{für } T \in \mathcal{T}_{\ell,\Gamma}. \quad (3.3.19)$$

anstelle von  $\eta_\ell$  in Algorithmus 3.3.2 verwendet wird. Dann konvergiert die Folge der diskreten Lösungen  $(U_\ell)_{\ell \in \mathbb{N}}$  unter obigen Voraussetzungen trotzdem gegen die exakte Lösung  $u \in K$ , d.h.  $\lim_{\ell \rightarrow \infty} \|u - U_\ell\| = 0$ . Außerdem gilt  $\lim_{\ell \rightarrow \infty} \tilde{\eta}_\ell = 0$ .

*Beweis.* Mit  $|T| \sim h_T^2$  gilt offensichtlich die Äquivalenz

$$\text{osc}_\ell(T) \leq \tilde{\text{osc}}_\ell(T) \leq C \text{osc}_\ell(T) \quad \text{für alle } T \in \mathcal{T}_{\ell,\Gamma}, \quad (3.3.20)$$

wobei die Konstante  $C > 0$  nur von der Formregularität abhängt. Aus der Dörflermarkierung für den Schätzer  $\tilde{\eta}_\ell$  folgt nun

$$\begin{aligned} \theta \eta_\ell \leq \theta \tilde{\eta}_\ell &\leq \sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} (\eta_\ell(E)^2 + \text{osc}_\ell(E)^2) + \sum_{T \in \mathcal{T}_{\ell,\Gamma} \cap \mathcal{M}_\ell} \tilde{\text{osc}}_\ell(T)^2 \\ &\leq C \left( \sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} (\eta_\ell(E)^2 + \text{osc}_\ell(E)^2) + \sum_{T \in \mathcal{T}_{\ell,\Gamma} \cap \mathcal{M}_\ell} \text{osc}_\ell(T)^2 \right). \end{aligned}$$

Die Menge  $\mathcal{M}_\ell$  erfüllt also auch eine Dörflermarkierung für  $\eta_\ell$  mit abgewandeltem Adaptivitätsparameter  $\tilde{\theta} = \theta/C \in (0, 1)$ . Nach dem Konvergenzsatz 3.3.9 folgt somit die Konvergenz von  $(U_\ell)_{\ell \in \mathbb{N}}$  gegen die exakte Lösung  $u \in K$ . Mit der Äquivalenz der (3.3.20) der beiden Schätzer  $\eta_\ell$  und  $\tilde{\eta}_\ell$  folgt nun auch  $\lim_{\ell \rightarrow \infty} \tilde{\eta}_\ell = 0$ .  $\square$

Zum Abschluss dieses Abschnittes wollen wir noch zeigen, dass man die Schätzerreduktion, unter einigen zusätzlichen Voraussetzungen, auch direkt für den Residualschätzer  $\eta_\ell$  zeigen kann.

**Proposition 3.3.11.** *Die Anfangstriangulierung  $\mathcal{T}_0$  sei so beschaffen, dass ein Element  $T \in \mathcal{T}_0$  höchstens eine Kante am Rand hat. Wird ferner eine Verfeinerungsstrategie verwendet, bei der alle markierten Randelemente durch mindestens drei Bisektionen, oder Rot-Verfeinerung zerteilt werden, dann gilt*

$$\eta_{\ell+1}^2 \leq q \eta_\ell^2 + C \|U_{\ell+1} - U_\ell\|^2.$$

*Beweis.* Für ein gegebenes  $\theta \in (0, 1)$  und eine Menge  $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell \cup \mathcal{T}_{\ell,\Gamma}$  markierter Elemente lautet die Dörflermarkierung

$$\theta \eta_\ell^2 \leq \sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} (\eta_\ell(E)^2 + \text{osc}_\ell(E)^2) + \sum_{T \in \mathcal{T}_{\ell,\Gamma} \cap \mathcal{M}_\ell} \text{osc}_\ell(T)^2. \quad (3.3.21)$$

Sei nun  $\mathcal{T}_{\ell,\Gamma}^* := \{T \in \mathcal{T}_{\ell,\Gamma} \mid |T \cap \Gamma| = 0\}$  die Menge der Elemente der Randschicht, welche den Rand mit keiner Kante berühren. Wegen der angenommenen Regularität der Triangulierungen, ist die Anforderung an  $\mathcal{T}_0$  auch auf allen feineren Triangulierungen  $\mathcal{T}_\ell$  gültig. Wir können daher jeder Randkante  $E \in \mathcal{E}_{\ell,\Gamma} := \mathcal{E}_\ell^* \setminus \mathcal{E}_\ell$  ein eindeutiges Element  $T \in \mathcal{T}_{\ell,\Gamma} \setminus \mathcal{T}_{\ell,\Gamma}^*$  zuordnen und den Schätzer  $\eta_\ell$  somit als

$$\eta_\ell^2 = \sum_{E \in \mathcal{E}_\ell} \eta_\ell(E)^2 + \sum_{E \in \mathcal{E}_\ell^*} \text{osc}_\ell(E)^2 + \sum_{T \in \mathcal{T}_{\ell,\Gamma}^*} \text{osc}_\ell(T)^2$$

umschreiben. Wir interpretieren nun die Menge  $\mathcal{M}_\ell$  als  $\mathcal{M}_\ell^* \subseteq \mathcal{E}_\ell^* \cup \mathcal{T}_{\ell,\Gamma}^*$  und erhalten mit der Dörflermarkierung

$$\theta \eta_\ell^2 \leq \sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell^*} \eta_\ell(E)^2 + \sum_{E \in \mathcal{E}_\ell^* \cap \mathcal{M}_\ell^*} \text{osc}_\ell(E)^2 + \sum_{T \in \mathcal{T}_{\ell,\Gamma}^* \cap \mathcal{M}_\ell^*} \text{osc}_\ell(T)^2.$$

Ist nun ein Element  $T \in \mathcal{T}_{\ell,\Gamma} \cap \mathcal{M}_\ell$  mit  $|T \cap \Gamma| > 0$  zur Verfeinerung ausgewählt, so wird durch die zusätzlichen Anforderungen an die Randverfeinerung sichergestellt, dass auch die entsprechende Randkante  $E = T \cap \Gamma \in \mathcal{E}_\ell^* \setminus \mathcal{E}_\ell$  verfeinert wird (dies wäre nicht der Fall wenn zum Beispiel nur die Referenzkante markierter Elemente verfeinert werden würde). Damit können wir die Lemmata 3.3.5 und 3.3.6 direkt übernehmen. Ersetzen wir in Lemma 3.3.4 die Menge  $\mathcal{T}_{\ell,\Gamma}$  durch  $\mathcal{T}_{\ell,\Gamma}^*$ , so bleiben auch hier sämtliche Ausführungen und damit die Gesamtaussage gültig. Unter Verwendung der Dörflermarkierung erhalten wir nun analog zum Beweis von 3.3.3

$$\eta_{\ell+1}^2 \leq q \eta_\ell^2 + C \|U_{\ell+1} - U_\ell\|^2 \quad (3.3.22)$$

für ein  $q \in (0, 1)$  und eine Konstante  $C > 0$ , die außer von  $\theta$  nur noch von der Formregularität abhängt.  $\square$

**Bemerkung.** *Die zusätzliche Anforderung an die Verfeinerung der Randelemente in der letzten Proposition kann zum Beispiel dadurch erreicht werden, dass von markierten Elementen  $T \in \mathcal{T}_{\ell,\Gamma} \cap \mathcal{M}_\ell$  einfach alle Kanten  $E \in \mathcal{E}_T$  zur Verfeinerung markiert werden. Diese Möglichkeit, die immer gegeben ist, wurde bereits in Abschnitt 2.2.3 kurz angesprochen.*

### 3.3.3. Eine Erweiterung

In diesem kurzen Abschnitt wollen wir das Konvergenzresultat aus dem letzten Teil um eine Erweiterung ergänzen. Im Sinne von [CKNS] zeigen wir, dass die gemischte Größe

$$\Delta_\ell := \mathcal{J}(U_\ell) - \mathcal{J}(u) + \gamma \mu_\ell^2 \quad (3.3.23)$$

für gewisse  $\gamma \in (0, 1)$  eine Kontraktion ist. Mit der bereits oft verwendeten Äquivalenz der beiden Schätzer  $\mu_\ell$  und  $\eta_\ell$  und der Zuverlässigkeit

$$\|u - U_\ell\| \lesssim \eta_\ell$$

folgt auch hieraus die Konvergenz der Folge  $(U_\ell)_{\ell \in \mathbb{N}}$  gegen die exakte Lösung  $u \in K$ . Ein solches Resultat ist jedoch in zweierlei Hinsicht besser als der vorherige Konvergenzsatz. Erstens hat man durch die Kontraktion auch eine Aussage über die Konvergenzgeschwindigkeit der Methode in der Hand, da der Fehler ja mindestens genauso schnell fallen muss wie  $\Delta_\ell$ . Zweitens ist eine Abschätzung der Form (3.3.23) im linearen Fall ein essentieller Bestandteil des Beweises der Optimalität. Eine solche Abschätzung ist daher auch im Hindernisfall ein guter Startpunkt um sich Gedanken über Optimalität zu machen. Mit den gleichen Notationen wie im letzten Teil gilt nun die folgende Aussage:

**Satz 3.3.12.** *Der adaptive Algorithmus 3.3.2 werde mit dem Schätzer  $\eta_\ell$  gesteuert. Dann erfüllt die kombinierte Fehlergröße  $\Delta_\ell$  eine Kontraktionseigenschaft der Form*

$$\Delta_{\ell+1} \leq \kappa \Delta_\ell \quad \text{für alle } \ell \in \mathbb{N}, \quad (3.3.24)$$

wobei die Konstanten  $\gamma, \kappa \in (0, 1)$  nur von  $\theta$  und der Formregularität abhängen. Unter den für die Zuverlässigkeit notwendigen Voraussetzungen gilt im besonderen  $\lim_{\ell \rightarrow \infty} \mathcal{J}(U_\ell) = \mathcal{J}(u)$  und  $\lim_{\ell \rightarrow \infty} \|u - U_\ell\| = 0 = \lim_{\ell \rightarrow \infty} \eta_\ell$ .

*Beweis.* Analog zum Beweis von Satz 3.3.9 erhalten wir mit der Äquivalenz der involvierten Schätzer und der Dörflermarkierung für  $\eta_\ell$

$$\frac{\theta}{4} \mu_\ell^2 \leq \theta \eta_\ell^2 \leq \sum_{E \in \mathcal{E}_\ell \cap \mathcal{M}_\ell} (\eta_\ell(E)^2 + \text{osc}_\ell(E)^2) + \sum_{T \in \mathcal{T}_{\ell, \Gamma} \cap \mathcal{M}_\ell} \text{osc}_\ell(T)^2.$$

Die Menge der markierten Elemente  $\mathcal{M}_\ell$  erfüllt also auch eine Dörflermarkierung für  $\mu_\ell$  für den modifizierten Parameter  $\tilde{\theta} = \theta/4 \in (0, 1)$ . Aus der Schätzerreduktion des vorigen Abschnittes erhalten wir daher

$$\mu_{\ell+1}^2 \leq q \mu_\ell^2 + C \|U_{\ell+1} - U_\ell\|^2.$$

Elementare Umformungen liefern nun

$$\begin{aligned} \Delta_{\ell+1} &= \mathcal{J}(U_\ell) - \mathcal{J}(u) + \gamma \mu_{\ell+1}^2 - (\mathcal{J}(U_\ell) - \mathcal{J}(U_{\ell+1})) \\ &\leq \mathcal{J}(U_\ell) - \mathcal{J}(u) + \gamma q \mu_\ell^2 + \gamma C \|U_{\ell+1} - U_\ell\|^2 - (\mathcal{J}(U_\ell) - \mathcal{J}(U_{\ell+1})). \end{aligned}$$

Analog zum Beweis der Zuverlässigkeit von  $\eta_\ell$  erhalten wir unter Verwendung der Variationsungleichung (3.1.7) im  $(\ell + 1)$ -ten Schritt für  $V_{\ell+1} = U_\ell$

$$\frac{1}{2} \|U_{\ell+1} - U_\ell\|^2 \leq \mathcal{J}(U_\ell) - \mathcal{J}(U_{\ell+1}).$$

Wir wählen nun  $\gamma$  so klein, dass  $(\gamma C - 1/2) \leq 0$  und erhalten

$$\Delta_{\ell+1} \leq \mathcal{J}(U_\ell) - \mathcal{J}(u) + \gamma q \mu_\ell^2 + (\gamma C - 1/2) \|U_{\ell+1} - U_\ell\|^2 \leq \mathcal{J}(U_\ell) - \mathcal{J}(u) + \gamma q \mu_\ell^2.$$

Nach dem Zuverlässigkeitssatz 3.3.1 und der Äquivalenz  $\eta_\ell \sim \mu_\ell$  folgt nun

$$\tilde{C}^{-1} (\mathcal{J}(U_\ell) - \mathcal{J}(u)) \leq \eta_\ell^2 \leq \mu_\ell^2,$$

wobei  $\tilde{C} > 0$  die Konstante aus der Zuverlässigkeit bezeichnet. Sei nun  $\varepsilon > 0$  beliebig, dann gilt

$$\mathcal{J}(U_\ell) - \mathcal{J}(u) + \gamma q \mu_\ell^2 \leq (1 - \gamma \varepsilon \tilde{C}^{-1}) (\mathcal{J}(U_\ell) - \mathcal{J}(u)) + \gamma (q + \varepsilon) \mu_\ell^2 \leq \kappa \Delta_\ell,$$

mit  $\kappa := \max\{1 - \gamma \varepsilon \tilde{C}^{-1}, q + \varepsilon\}$ . Wir wählen nun  $\varepsilon > 0$  so klein, dass  $\kappa \in (0, 1)$ . Hieraus folgt die gesuchte Aussage (3.3.24). Offensichtlich erhalten wir nun  $\lim_{\ell \rightarrow \infty} \Delta_\ell = 0$  und damit

$$\lim_{\ell \rightarrow \infty} \mathcal{J}(U_\ell) = \mathcal{J}(u) \quad \text{und} \quad \lim_{\ell \rightarrow \infty} \mu_\ell = 0.$$

Die Konvergenz gegen die exakte Lösung  $u \in K$  folgt nun, wie üblich, aus der Abschätzung  $\eta_\ell \leq \mu_\ell$  und der Zuverlässigkeit  $\|u - U_\ell\| \lesssim \eta_\ell$ . Damit ist der Beweis abgeschlossen.  $\square$

**Bemerkung.** Analog zur Aussage von Korollar 3.3.10, behält auch dieser Satz seine Gültigkeit, wenn die Elementbeiträge mit  $h_T^2 = \text{diam}(T)^2$  anstelle von  $|T|$  gewichtet werden. Hierzu muss man sich, wie im Beweis von Korollar 3.3.10, nur klarmachen, dass die entsprechende Dörflermarkierung auch eine Dörflermarkierung für den verwendeten Schätzer impliziert.

**Bemerkung.** Unter den in Proposition 3.3.11 beschriebenen zusätzlichen Anforderungen, kann man mit analoger Argumentation auch eine Kontraktion der Größe  $\tilde{\Delta}_\ell := \mathcal{J}(U_\ell) - \mathcal{J}(u) + \gamma \eta_\ell^2$  für entsprechende  $\gamma \in (0, 1)$  zeigen. Man kann in diesem Fall einfach auf die Schätzerreduktion für  $\eta_\ell$  zurückgreifen und den Beweis von Satz 3.3.12 dann analog führen.

**Bemerkung.** Da wir in unseren Ausführungen zu keiner Zeit Gebrauch von der Raumdimension gemacht haben lassen sich sämtliche Resultate analog auf den dreidimensionalen Fall übertragen.

Außerdem haben wir an keiner Stelle die Ungleichung  $u \geq \chi$ , bzw.  $U_\ell \geq \chi$  verwendet. Im Gegensatz zum nicht-affinen Teil, war also die Problemauswahl bislang willkürlich, so dass unsere Ausführungen ihre Gültigkeit auch behalten, wenn wir ein nach oben begrenzendes Hindernis annehmen.

### 3.4. Numerische Illustrationen

Um das Kapitel über affine Hindernisprobleme adäquat zu Ende zu führen, wollen wir im Folgenden die Richtigkeit unserer bisherigen Analysen auch empirisch bestätigen. Hierzu wurde Algorithmus 3.3.2 in MATLAB umgesetzt. Zur Lösung des diskreten Teilproblems in jedem Schritt wurde die *primal-dual active set* Methode aus [HIK] verwendet, von der im Folgenden ein Pseudocode angegeben ist:

---

**Pseudocode 1** PRIMAL-DUAL ACTIVE SET METHODE

---

**Require:**  $(U^{(0)}, \lambda^{(0)}), k = 0, c > 0$ .

- 1: **while** Abbruchkriterium nicht erfüllt **do**
- 2:     Berechne  $\mathcal{I}_k := \{i \in \mathcal{N} : \lambda_i^k + c(U^k - \chi)_i \geq 0\}$  und  $\mathcal{A}_k := \{i \in \mathcal{N} : \lambda_i^k + c(U^k - \chi)_i < 0\}$
- 3:     Löse

$$\begin{aligned} \Delta U^{k+1} + \lambda^{k+1} &= f \\ U^{k+1} &= \chi \quad \text{auf } \mathcal{A}_k, \quad \lambda^{k+1} = 0 \quad \text{auf } \mathcal{I}_k \end{aligned}$$

4: **end while**

---

Zur Initialisierung haben wir  $(U_0^{(0)}, \lambda_0^{(0)}) \equiv (0, 0)$  auf dem Startgitter  $\mathcal{T}_0$  verwendet. Im  $\ell$ -ten Schritt wurden jeweils die erweiterten Lösungen des vorherigen Gitters als Startwerte verwendet, d.h.  $U_\ell^{(0)} := U_{\ell-1}$  und  $\lambda_\ell^{(0)} := \lambda_{\ell-1}$ . Als Abbruchkriterium im  $\ell$ -ten Schritt hat sich

$$\|\nabla(U_\ell^{(j)} - U_\ell^{(j-1)})\|_{L^2(\Omega)} \leq \tau N^{-1/2}, \quad (3.4.1)$$

für eine Toleranz  $\tau > 0$  als sinnvoll herausgestellt.  $N = \#\mathcal{T}_\ell$  bezeichnet hierbei die Anzahl der Elemente der  $\ell$ -ten Triangulierung. Als diskrete Lösungen wurden dann jeweils  $U_\ell := U_\ell^{(j)}$  und  $\lambda_\ell := \lambda_\ell^{(j)}$  verwendet. Eine vollständige Implementierung dieser Methode und des folgenden Beispiels, welches auch in [BCH1] behandelt wurde, findet sich in Anhang C.

Wir betrachten im Folgenden das affine Hindernisproblem mit konstantem Hindernis  $\chi \equiv 0$  auf dem L-förmigen Gebiet  $\Omega := (-2, 2)^2 \setminus [0, 2) \times (-2, 0]$ . Die rechte Seite ist in Polarkoordinaten durch

$$f(r, \varphi) := -r^{2/3} \sin(2\varphi/3) (\gamma_1'(r)/r + \gamma_1''(r)) - \frac{4}{3} r^{-1/3} \gamma_1'(r) \sin(2\varphi/3) - \gamma_2(r), \quad (3.4.2)$$

gegeben. Hierbei bezeichnet  $(\cdot)'$  die radiale Ableitung  $d/dr$ . Ferner sind  $\gamma_1$  und  $\gamma_2$  mit  $\bar{r} := 2(r - 1/4)$  durch

$$\begin{aligned} \gamma_1(r) &= \begin{cases} 1, & \bar{r} < 0, \\ -6\bar{r}^5 + 15\bar{r}^4 - 10\bar{r}^3 + 1, & 0 \leq \bar{r} < 1, \\ 0, & \bar{r} \geq 1, \end{cases} \\ \gamma_2(r) &= \begin{cases} 0, & r \leq 5/4, \\ 1, & \text{sonst.} \end{cases} \end{aligned}$$

definiert. Die exakte Lösung

$$u(r, \varphi) = r^{2/3} \gamma_1(r) \sin(2\varphi/3) \quad (3.4.3)$$

dieses Problems ist in Abbildung 3.7 (*oben*) visualisiert. Sie hat eine Eckensingularität im Ursprung und gehört für jedes  $\varepsilon > 0$  und jeder offenen Umgebung  $D$  um den Ursprung zu  $H^{5/3-\varepsilon}(D)$ . Zum Vergleich ist die numerische Approximation  $U_8$  auf  $\mathcal{T}_8$  in Abbildung 3.7 (*unten*) dargestellt.

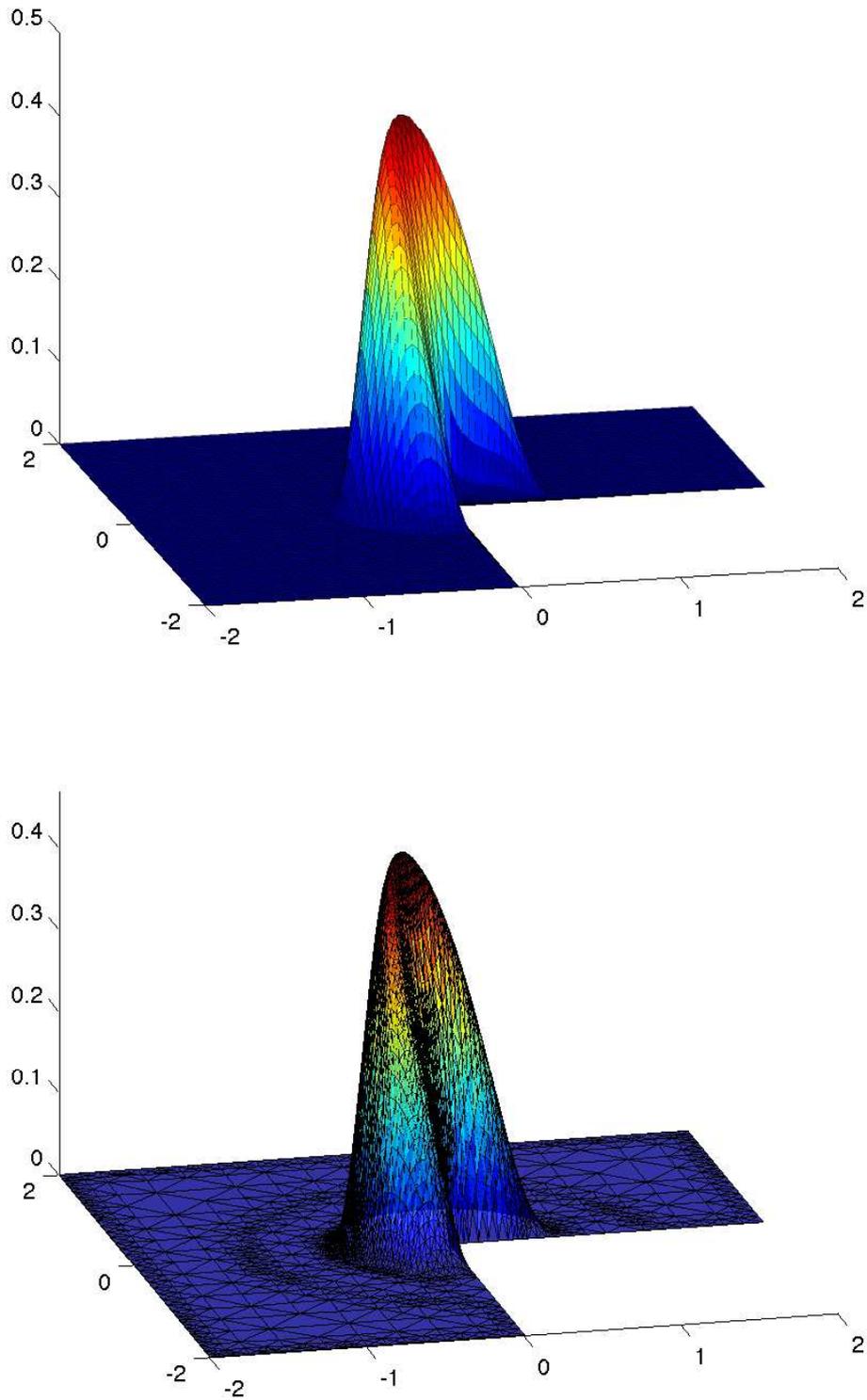


Abbildung 3.7.: Exakte Lösung  $u \in K$  (oben) und diskrete Lösung  $U_8$  auf dem adaptiv generierten Gitter  $\mathcal{T}_8$  mit  $N = 3524$  Elementen für  $\theta = 0.6$  (unten).

$\ell$	$N$	$\sqrt{e_\ell}$	$\eta_\ell$	$osc_\ell$
1	38	9.6584e-01	7.8592e+00	7.6804e+00
2	70	5.7088e-01	4.4866e+00	4.1953e+00
3	156	3.8544e-01	2.6368e+00	2.2612e+00
4	316	3.4255e-01	1.6397e+00	1.1892e+00
5	568	2.1016e-01	1.0907e+00	7.1523e-01
6	1088	1.4429e-01	7.2078e-01	3.9564e-01
7	1966	9.8835e-02	4.9444e-01	2.1653e-01
8	3524	6.9485e-02	3.4353e-01	1.3578e-01
9	6438	4.8982e-02	2.4228e-01	7.6460e-02
10	12270	3.3545e-02	1.6831e-01	4.3110e-02
11	22140	2.3780e-02	1.2136e-01	2.5694e-02
12	42620	1.6327e-02	8.5847e-02	1.4742e-02
13	77918	1.0873e-02	6.2456e-02	9.6058e-03
14	146854	8.6766e-03	4.4498e-02	5.3497e-03
15	268400	6.0689e-03	3.2346e-02	4.4724e-03
16	491250	4.1792e-03	2.3719e-02	2.4099e-03
17	925574	2.8772e-03	1.7260e-02	1.5729e-03

Tabelle 3.1.: Numerische Ergebnisse für adaptive Netzverfeinerung mit  $\theta = 0.6$ , wobei  $N = \#\mathcal{T}_\ell$ ,  $e_\ell = \mathcal{J}(U_\ell) - \mathcal{J}(u)$ , und  $osc_\ell$  wie in (3.4.5) definiert ist.

In den numerischen Berechnungen vergleichen wir uniforme und adaptive Netzverfeinerung, wobei wir den Adaptivitätsparameter  $\theta$  aus  $\{0.2, 0.4, 0.6, 0.8\}$  variieren. Die numerische Konvergenzgeschichte für  $\theta = 0.6$  ist in Tabelle 3.1 dargestellt. Der Fehler ist hierbei im Energiefunktional mit

$$e_\ell := \mathcal{J}(U_\ell) - \mathcal{J}(u) \tag{3.4.4}$$

gegeben. Die Gesamtoszillationen sind, wie im theoretischen Teil angegeben, durch

$$osc_\ell^2 = \sum_{E \in \mathcal{E}_\ell} osc_\ell(E)^2 + \sum_{T \in \mathcal{T}_{\ell,\Gamma}} osc_\ell(T)^2, \tag{3.4.5}$$

definiert. Als Fehlerschätzer wurde der in (3.3.1) definierte Schätzer  $\eta_\ell$ , der aus den Kantsprünge und den Oszillationstermen besteht, verwendet.

In Abbildung 3.8 sind die Konvergenzraten, d.h. der Fehler  $\sqrt{e_\ell}$ , der Schätzer  $\eta_\ell$  und die Oszillationen  $osc_\ell$  im Vergleich mit der Anzahl der Elemente  $N = \#\mathcal{T}_\ell$  für uniforme und adaptive Verfeinerung mit  $\theta = 0.6$  visualisiert. Während uniforme Verfeinerung zu einer empirischen Konvergenzrate von etwa  $\mathcal{O}(N^{-5/12})$  führt, erzielen wir durch adaptive Verfeinerung tatsächlich die optimale Konvergenzrate  $\sqrt{e_\ell} = \mathcal{O}(N^{-1/2})$ . Deutlich ist auch zu erkennen, dass die Oszillationen etwa mit doppelter Geschwindigkeit fallen, d.h.  $osc_\ell = \mathcal{O}(N^{-1})$ . Bemerkenswert ist hierbei auch, dass die Konvergenzraten von  $\eta_\ell$  und  $\sqrt{e_\ell}$  unabhängig von der Verfeinerung parallel sind. Der von uns vorgeschlagene Schätzer  $\eta_\ell$  scheint also auch effizient zu sein. Auf jeden Fall ist dieses Verhalten eine experimentelle Bestätigung der Zuverlässigkeit.

In Abbildung 3.9 ist ein Vergleich des adaptiven Verfahrens für verschiedene Adaptivitätsparameter  $\theta \in \{0.2, 0.4, 0.6, 0.8\}$  zu sehen. Jede adaptive Verfeinerung führt hierbei

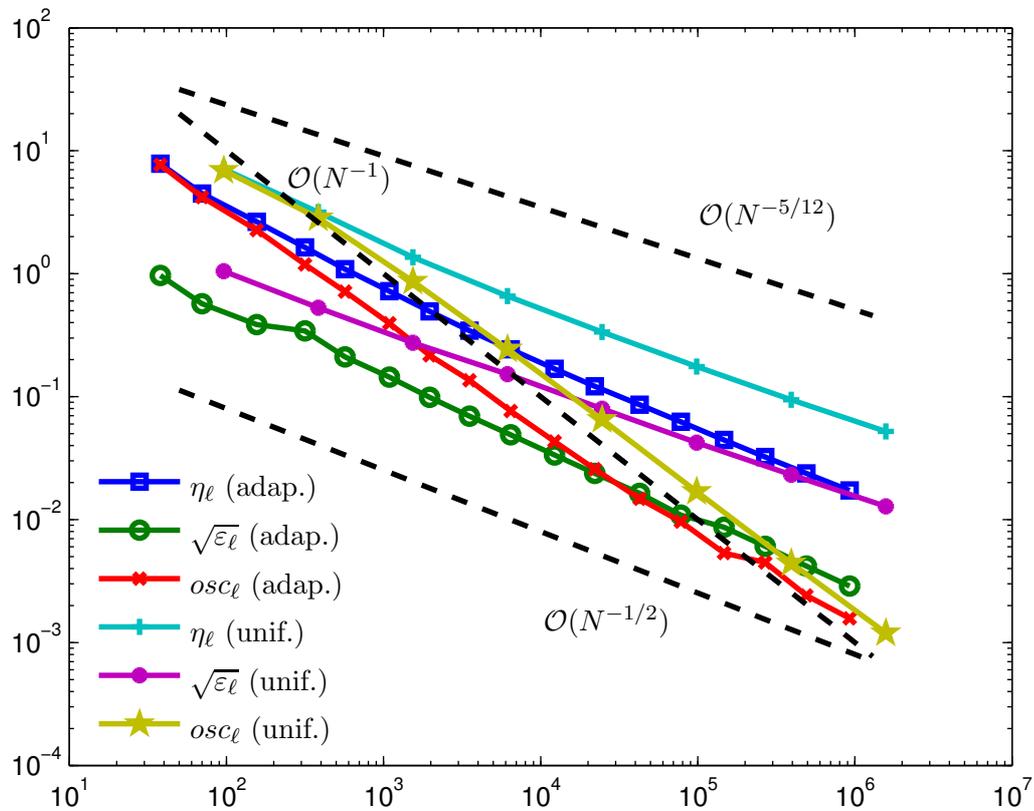


Abbildung 3.8.: Numerische Ergebnisse mit uniformer und adaptiver Verfeinerung mit  $\theta = 0.6$ , wobei  $e_\ell = \mathcal{J}(U_\ell) - \mathcal{J}(u)$ ,  $\eta_\ell$ , und  $osc_\ell$  im Vergleich mit der Anzahl der Elemente  $N = \#\mathcal{T}_\ell$  abgebildet sind.

empirisch zur optimalen Konvergenzrate von  $\mathcal{O}(N^{-1/2})$ . Außerdem sind auch die Vorzüge adaptiver Verfeinerung klar ersichtlich, da hierbei bereits mit viel weniger Elementen kleine Fehler erreicht werden können.

Abbildung 3.10 visualisiert schließlich die adaptiv generierten Triangulierungen nach 5 bzw. 11 Schritten der Verfeinerung für  $\theta = 0.6$ . Wie erwartet bleibt die Verfeinerung hauptsächlich auf die Nicht-Kontaktzone beschränkt. Durch den Einfluss der Elementbeiträge ist jedoch auch am Rand Verfeinerung zu erkennen.

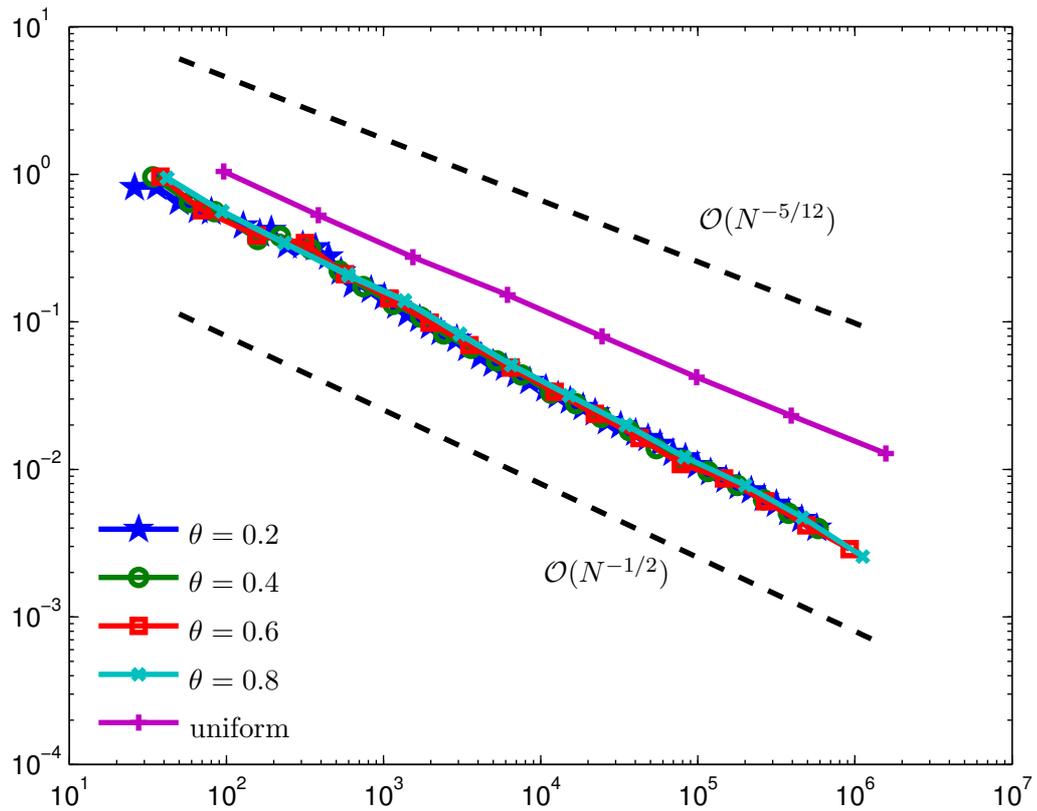


Abbildung 3.9.: Numerische Ergebnisse für  $\sqrt{e_\ell}$  für uniforme und adaptive Netzverfeinerung mit  $\theta \in \{0.2, 0.4, 0.6, 0.8\}$ . Die verschiedenen Größen sind im Vergleich mit der Anzahl der Elemente  $N = \#\mathcal{T}_\ell$  abgebildet.

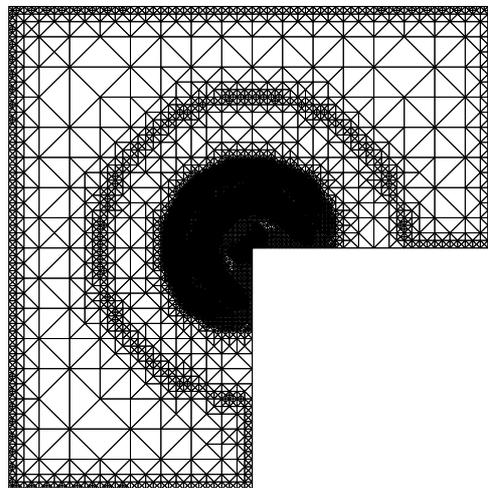
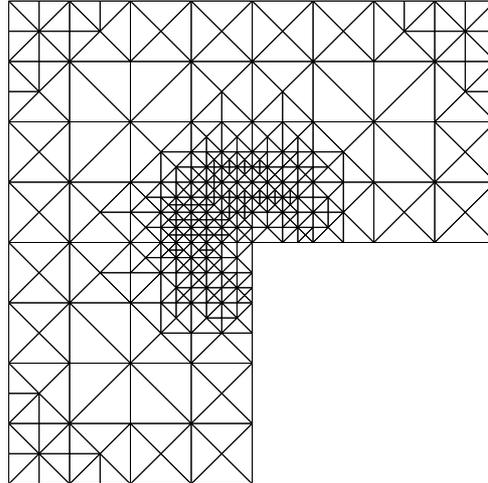


Abbildung 3.10.: Adaptiv generierte Triangulierungen  $\mathcal{T}_5$  (*oben*) und  $\mathcal{T}_{11}$  (*unten*) mit  $N = 568$  bzw.  $N = 22140$  Elementen für  $\theta = 0.6$ .



## Kapitel 4.

### Das nicht-affine Hindernisproblem

In diesem Kapitel untersuchen wir erneut das Hindernisproblem auf einem polygonal berandeten Gebiet  $\Omega$  mit Rand  $\Gamma$ , wobei wir nun die Annahme über die Affinität des Hindernisses fallen lassen. Obwohl sich unsere Ausführungen auf den zweidimensionalen Fall beschränken, lassen sich sämtliche Resultate analog auch auf den dreidimensionalen Fall übertragen, da wir an keiner Stelle Gebrauch von der Raumdimension machen. Die physikalische Interpretation dieses Problems ist dieselbe wie zuvor.

#### 4.1. Problembeschreibung und Notationen

Für die gesamte Dauer dieses Abschnittes sei  $\chi \in H^1(\Omega)$  ein (möglicherweise nicht-affines) Hindernis mit  $\chi|_{\Gamma} \geq 0$ . Analog zum affinen Fall, definieren wir nun die Menge der zulässigen Funktionen, in der wir unsere Lösung suchen, durch

$$K := \{v \in H_0^1(\Omega) \mid v \leq \chi \text{ f.ü. in } \Omega\}.$$

Die Tatsache, dass wir in diesem Teil Funktionen suchen, die kleiner als das Hindernis  $\chi$  sind, wird im Folgenden an mehreren Stellen in die Beweise eingehen, so dass diese Änderung notwendig ist. Da die Menge  $K$  analog zur Argumentation aus Lemma 3.1.1 konvex und abgeschlossen ist, nimmt das Energiefunktional

$$\mathcal{J}(v) := \frac{1}{2} \langle v, v \rangle - (f, v)$$

nach dem Satz von Lax-Milgram (Satz 3.1.2) auch über  $K$  ein eindeutiges Minimum an. Wie zuvor steht  $\langle \cdot, \cdot \rangle = \|\cdot\|^2$  hierbei für das durch die Energienorm induzierte Energieskalarprodukt auf dem gesamten Gebiet  $\Omega$ . Ebenfalls analog zu vorherigen Ausführungen bezeichnet  $(\cdot, \cdot)$  das  $L^2$ -Skalarprodukt auf  $\Omega$ . Wie im affinen Teil können wir nun das kontinuierliche Hindernisproblem formulieren.

**Problem 4.1.1** (kontinuierlich). *Finde  $u \in K = \{v \in H_0^1(\Omega) \mid v \leq \chi \text{ f.ü. in } \Omega\}$  so, dass*

$$\mathcal{J}(u) = \min_{v \in K} \mathcal{J}(v), \tag{4.1.1}$$

*beziehungsweise*

$$\langle u, v - u \rangle \geq (f, v - u) \quad \text{für alle } v \in K \tag{4.1.2}$$

*gilt.*

Die Äquivalenz der beiden Darstellungen (4.1.1) und (4.1.2) ergibt sich auch hier durch zum affinen Fall völlig analoge Berechnungen. Durch die Einführung eines Lagrange'schen Multiplikators  $\sigma \in H^{-1}(\Omega)$ , kann (4.1.2) durch

$$\langle\langle u, v \rangle\rangle = (f, v) - \langle \sigma, v \rangle \quad \text{für alle } v \in H_0^1(\Omega) \quad (4.1.3)$$

als unrestringiertes Problem dargestellt werden. Hierbei bezeichnet  $\langle \cdot, \cdot \rangle$  die duale Paarung zwischen  $H_0^1(\Omega)$  und  $H^{-1}(\Omega)$ . Wir erhalten dabei zusätzlich die Komplementaritätsbedingung

$$\langle \sigma, u - \chi \rangle = 0. \quad (4.1.4)$$

Für die Approximation im endlich dimensionalen Teilraum  $\mathcal{S}_0^1(\mathcal{T}_\ell)$  von  $H_0^1(\Omega)$  durch die Finite-Elemente-Methode, definieren wir  $\chi_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$  als eine Approximation von  $\chi \in H_0^1(\Omega)$ . Falls das Hindernis stetig ist, d.h.  $\chi \in C(\bar{\Omega})$ , könnte dies zum Beispiel durch den nodalen Interpolanten gegeben sein. Wir wählen nun den diskreten Ansatzraum

$$K_\ell := \{V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell) \mid V_\ell(x) \leq \chi_\ell(x) \text{ für alle } x \in \Omega\}.$$

Offensichtlich ist auch  $K_\ell$  nach bekannten Argumenten wieder konvex und abgeschlossen, so dass das Energiefunktional auch über  $K_\ell$  ein eindeutiges Minimum annimmt. Wie gewohnt, definieren wir nun das endlich-dimensionale Teilproblem.

**Problem 4.1.2** (diskret). *Finde  $U_\ell \in K_\ell = \{V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell) \mid V_\ell(x) \leq \chi_\ell(x) \text{ für alle } x \in \Omega\}$  so, dass*

$$\mathcal{J}(U_\ell) = \min_{V_\ell \in K_\ell} \mathcal{J}(V_\ell), \quad (4.1.5)$$

*beziehungsweise*

$$\langle\langle U_\ell, V_\ell - U_\ell \rangle\rangle \geq (f, V_\ell - U_\ell) \quad \text{für alle } V_\ell \in K_\ell \quad (4.1.6)$$

*gilt.*

Die Äquivalenz von (4.1.5) und (4.1.6) folgt wie zuvor. Durch Einführung eines entsprechenden Lagrange'schen Multiplikators  $\sigma_\ell \in C(\bar{\Omega})' = \mathcal{M}_+(\Omega)$  lautet die unrestringierte Version dieses Problems

$$\langle\langle U_\ell, V_\ell \rangle\rangle = (f, V_\ell) - \langle \sigma_\ell, V_\ell \rangle_{\mathcal{M}(\Omega)} \quad \text{für alle } v \in \mathcal{S}_0^1(\mathcal{T}_\ell). \quad (4.1.7)$$

Der Dualraum  $C(\bar{\Omega})'$  von  $C(\bar{\Omega})$  ist hierbei durch den Raum  $\mathcal{M}(\Omega)$  der Radonmaße auf  $\Omega$  gegeben. Ausnutzung der Variationsungleichung (4.1.6) liefert sogar  $\sigma_\ell \in \mathcal{M}_+(\Omega)$ , d.h.

$$\langle \sigma_\ell, V_\ell \rangle_{\mathcal{M}(\Omega)} \geq 0 \quad \text{für alle } V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell) \text{ mit } V_\ell \geq 0. \quad (4.1.8)$$

Hierbei bezeichnet  $\langle \cdot, \cdot \rangle_{\mathcal{M}(\Omega)}$  die entsprechende duale Paarung zwischen  $C(\bar{\Omega})$  und  $\mathcal{M}(\Omega)$ . Analog zum kontinuierlichen Fall erhalten wir die Komplementaritätsbedingung

$$\langle \sigma_\ell, \chi_\ell - U_\ell \rangle_{\mathcal{M}(\Omega)} = 0. \quad (4.1.9)$$

**Bemerkung.** Wir wollen an dieser Stelle festhalten, dass die soeben beschriebene Diskretisierung, abhängig von  $\chi$  und der Wahl von  $\chi_\ell$ , sowohl zu konformen, als auch zu nicht-konformen Diskretisierungen führen kann. Wir beschränken uns in unseren Ausführungen jedoch stets auf den konformen Fall, d.h. wir fordern explizit  $K_\ell \subseteq K_{\ell+1}$ .

**Notationen.** Analog zu vorherigen Ausführungen verwenden wir wieder eine Familie von formregulären Triangulierungen  $(\mathcal{T}_\ell)_\ell \in \mathbb{N}$  des Gebietes  $\Omega$ . Die Menge der Knoten und Kanten einer Teilmenge  $D \subseteq \bar{\Omega}$  bezeichnen wir mit  $\mathcal{N}_\ell(D)$  bzw.  $\mathcal{E}_\ell(D)$ . Die entsprechenden Größen auf dem gesamten Gebiet notieren wir einfach mit  $\mathcal{N}_\ell$  und  $\mathcal{E}_\ell$ . Ferner verwenden wir die Patches

$$\begin{aligned}\omega_\ell^p &:= \bigcup \{T \in \mathcal{T}_\ell \mid p \in \mathcal{N}_\ell(T)\}, \\ \omega_\ell^E &:= \bigcup \{T \in \mathcal{T}_\ell \mid E \in \mathcal{E}_\ell(T)\}, \\ \omega_\ell^T &:= \bigcup \{T' \in \mathcal{T}_\ell \mid \mathcal{N}_\ell(T') \cap \mathcal{N}_\ell(T) \neq \emptyset\}.\end{aligned}$$

Die Menge der Kanten, welche den Punkt  $p$  berühren, bezeichnen wir zusätzlich mit

$$\mathcal{E}_\ell^p := \bigcup \{E \in \mathcal{E}_\ell \mid p \in \mathcal{N}_\ell(E)\}.$$

In Anlehnung an vorher definieren wir die aktive Menge im  $\ell$ -ten Schritt durch

$$\mathcal{A}_\ell := \{x \in \bar{\Omega} \mid U_\ell(x) = \chi_\ell(x)\}$$

und die inaktive Menge durch  $\mathcal{I}_\ell := \bar{\Omega} \setminus \mathcal{A}_\ell$ . Die Menge aller Kanten und Elemente, welche komplett in der aktiven Menge liegen, bezeichnen wir mit

$$\begin{aligned}\mathcal{E}_\ell(\mathcal{A}_\ell) &:= \bigcup \{E \in \mathcal{E}_\ell \mid \mathcal{N}_\ell(E) \subset \mathcal{A}_\ell\}, \\ \mathcal{T}_\ell(\mathcal{A}_\ell) &:= \bigcup \{T \in \mathcal{T}_\ell \mid \mathcal{N}_\ell(T) \subset \mathcal{A}_\ell\}.\end{aligned}$$

Die Komplemente dieser Mengen werden mit  $\mathcal{E}_\ell(\mathcal{I}_\ell)$ , bzw.  $\mathcal{T}_\ell(\mathcal{I}_\ell)$  bezeichnet. Die Mengen der Kanten und Elemente mit genau  $i$  Knoten in der aktiven Menge notieren wir mit

$$\begin{aligned}\mathcal{E}_{\mathcal{A}_\ell}^{(i)} &:= \bigcup \{E \in \mathcal{E}_\ell \mid \#(\mathcal{N}_\ell(E) \cap \mathcal{A}_\ell) = i\}, \quad i \in \{0, 1, 2\}, \\ \mathcal{T}_{\mathcal{A}_\ell}^{(i)} &:= \bigcup \{T \in \mathcal{T}_\ell \mid \#(\mathcal{N}_\ell(T) \cap \mathcal{A}_\ell) = i\}, \quad i \in \{0, 1, 2, 3\}.\end{aligned}$$

Dieselben Größen für die inaktive Menge sind analog durch  $\mathcal{E}_{\mathcal{I}_\ell}^{(i)}$  und  $\mathcal{T}_{\mathcal{I}_\ell}^{(i)}$  gegeben. Schließlich benötigen wir noch diejenigen Elemente und Kanten, die auf der dünnen Schicht zwischen dem gänzlich aktiven und dem gänzlich inaktiven Bereich liegen. Diese definieren wir durch

$$\mathcal{T}_{\mathcal{F}_\ell} := \mathcal{T}_\ell \setminus \left( \mathcal{T}_{\mathcal{A}_\ell}^{(3)} \cup \mathcal{T}_{\mathcal{I}_\ell}^{(3)} \right), \quad \mathcal{E}_{\mathcal{F}_\ell} := \mathcal{E}_\ell \setminus \left( \mathcal{E}_{\mathcal{A}_\ell}^{(2)} \cup \mathcal{E}_{\mathcal{I}_\ell}^{(2)} \right).$$

Für die Dauer des gesamten Kapitels bezeichne  $u \in K$  die eindeutige Lösung von Problem 4.1.1 und  $U_\ell \in K_\ell$  die eindeutige diskrete Lösung von Problem 4.1.2. Um unsere späteren Ausführungen zu erleichtern, wollen wir einen genaueren Blick auf den diskreten Lagrange-Multiplikator  $\sigma_\ell \in \mathcal{M}_+(\Omega)$  werfen. Wie immer bezeichnet  $[\partial_n U_\ell]_E$  hierbei den Sprung der Normalenableitungen über die Kante  $E \in \mathcal{E}_\ell$ .

**Lemma 4.1.3.** *Der Lagrange-Multiplikator  $\sigma_\ell$  kann als Linearkombination von Delta-Distributionen  $\delta_p$  bezüglich  $p \in \mathcal{N}_\ell$  dargestellt werden, d.h. für alle  $p \in \mathcal{N}_\ell$  gibt es ein skalarwertiges  $\alpha_\ell(p) \in \mathbb{R}$  mit*

$$\sigma_\ell = \sum_{p \in \mathcal{N}_\ell} \alpha_\ell(p) \delta_p. \quad (4.1.10)$$

Es gilt ferner

$$\alpha_\ell(p) = \begin{cases} \sum_{T \in \omega_\ell^p} (f, \varphi_p)_T - \sum_{E \in \mathcal{E}_\ell^p} ([\partial_n U_\ell], \varphi_p)_E, & p \in \mathcal{N}_\ell(\mathcal{A}_\ell), \\ 0, & p \in \mathcal{N}_\ell(\mathcal{I}_\ell), \end{cases} \quad (4.1.11)$$

wobei  $\varphi_p$  die nodale Basisfunktion zu  $p \in \mathcal{N}_\ell$  bezeichnet.

*Beweis.* Siehe BRAESS, CARSTENSEN und HOPPE [BCH2, Lemma 2.1].  $\square$

Nach diesem Lemma gilt offensichtlich

$$\langle U_\ell, V_\ell \rangle = (f, V_\ell) \quad (4.1.12)$$

für alle  $V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)$ , deren Träger kompakt in  $\mathcal{I}_\ell$  liegt. Bevor wir uns nun der Aufgabe widmen, die vorgestellten Probleme mittels adaptiver FEM zu lösen, definieren wir vorab noch eine Größe, die eine Erweiterung des Langrange'schen Multiplikators  $\sigma_\ell$  auf allgemeine Funktionen aus  $H_0^1(\Omega)$  darstellt. Wir setzen

$$\langle \tilde{\sigma}_\ell, v \rangle := \sum_{p \in \mathcal{N}_\ell(\mathcal{A}_\ell)} \left( \frac{1}{3} \sum_{T \in \omega_\ell^p} (f, v)_T - \frac{1}{2} \sum_{E \in \mathcal{E}_\ell^p} ([\partial_n U_\ell], v)_E \right), \quad (4.1.13)$$

für alle  $v \in H_0^1(\Omega)$ . Um eine etwas einträglichere Darstellung von (4.1.13) zu erhalten, schauen wir uns die enthaltenen Summen etwas genauer an und analysieren wie oft jedes Element bzw. jede Kante aufsummiert wird. Wir stellen fest, dass das  $L^2$ -Skalarprodukt  $(f, v)_T$  über dem Element  $T \in \mathcal{T}_\ell$  genau dreimal aufsummiert wird, wenn alle drei Knoten von  $\mathcal{N}_\ell(T)$  in der aktiven Menge liegen. In diesem Fall kommt es ja in drei Knotenpatches vor. Sind nur zwei der drei Knoten aktiv, so wird das Skalarprodukt zweimal aufsummiert. Schließlich kommt der Summand  $(f, v)_T$  nur einmal vor, wenn  $T$  nur einen einzigen aktiven Knoten besitzt. Zusammen mit der Gewichtung von  $1/3$  ergibt sich somit ein Gesamtvorkommen von  $1, 2/3$  bzw.  $1/3$  für die jeweiligen Fälle. Ähnliche Überlegungen für das Abzählen der Kantensterme führen auf die einfachere Darstellung

$$\langle \tilde{\sigma}_\ell, v \rangle = \sum_{i=1}^3 \left( \frac{i}{3} \sum_{T \in \mathcal{T}_{\mathcal{A}_\ell}^{(i)}} (f, v)_T \right) - \sum_{i=1}^2 \left( \frac{i}{2} \sum_{E \in \mathcal{E}_{\mathcal{A}_\ell}^{(i)}} ([\partial_n U_\ell], v)_E \right). \quad (4.1.14)$$

Nach [BCH2, (2.19) — (2.20)] ergibt sich hieraus für alle  $V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)$  die Gleichheit

$$\langle \sigma_\ell, V_\ell \rangle_{\mathcal{M}(\Omega)} - \langle \tilde{\sigma}_\ell, V_\ell \rangle = \sum_{T \in \mathcal{T}_{\mathcal{F}_\ell}} \kappa_T (f, V_\ell)_T - \sum_{E \in \mathcal{E}_{\mathcal{F}_\ell}} \kappa_E ([\partial_n U_\ell], V_\ell)_E, \quad (4.1.15)$$

mit

$$\kappa_T := 1 - \frac{i}{3}, T \in \mathcal{T}_{\mathcal{A}_\ell}^{(i)}, \quad \kappa_E := 1 - \frac{i}{2}, E \in \mathcal{E}_{\mathcal{A}_\ell}^{(i)}. \quad (4.1.16)$$

Dieses Resultat ist jedoch leider fragwürdig. Für Elemente und Kanten die gänzlich in der aktiven, oder gänzlich in der inaktiven Menge liegen stimmen wir der Aussage zu. Auf der dünnen Schicht  $\mathcal{E}_{\mathcal{F}_\ell}$  bzw.  $\mathcal{T}_{\mathcal{F}_\ell}$  jedoch, wird  $\tilde{\sigma}_\ell$  zwar auf  $V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell)$  angewendet,  $\sigma_\ell$  aber nicht, da die Koeffizienten der Hutfunktionen auf dem inaktiven Bereich ja mit Null gewichtet werden. Für  $T \in \mathcal{T}_{\mathcal{A}_\ell}^{(1)}$ , wobei  $p \in T$  den Knoten in der aktiven Menge beschreibt, erhalten wir zum Beispiel  $\langle \sigma_\ell, V_\ell(p)\varphi_p \rangle_{\mathcal{M}(\Omega)} \neq \langle \sigma_\ell, V_\ell \rangle_{\mathcal{M}(\Omega)}$ , wobei  $V_\ell(p) \in \mathbb{R}$  der Koeffizient der Basisfunktion zu  $p \in \mathcal{N}_\ell$  ist. Dieser Umstand wirkt sich leider stark auf die Glaubwürdigkeit der Zuverlässigkeit aus, in deren Beweis die genannte Aussage massiv eingeht. In der Originalarbeit [BCH2] findet sich kein Hinweis um das Problem aufzulösen. Da sowohl Intuition als auch numerische Simulation darauf schließen lassen, dass Zuverlässigkeit und daraus resultierende Konvergenz gelten, ist es aber dennoch sinnvoll auf solchen Aussagen aufzubauen, was wir im Laufe dieses Abschnitts auch tun werden. Der Vollständigkeit halber geben wir die Ausführungen aus [BCH2] trotzdem an, kommentieren aber die jeweiligen Problemstellen in den entsprechenden Beweisen. Die ab Abschnitt 4.4 präsentierten Ideen zur Schätzerreduktion für nicht-affine Hindernisprobleme machen aber in jedem Fall Sinn, und tragen, durch bislang unbedachte Komplikationen, sehr zum weiteren Verständnis dieses Prinzips bei.

## 4.2. Fehlerschätzer und adaptiver Algorithmus

Wir wollen nun einen geeigneten Fehlerschätzer und einen entsprechenden adaptiven FEM-Algorithmus zur Lösung der Probleme 4.1.1 bzw. 4.1.2 konstruieren. Hierbei wird, neben Einzelheiten innerhalb der späteren Beweise, ein gravierender Unterschied zum affinen Hindernisproblem deutlich: Die Verwendung eines Fehlerschätzers, welcher, wie im affinen Fall, die Kantensprünge  $[\partial_n U_\ell]_E$  auf Kanten in der aktiven Menge beinhaltet, kann niemals zu einem sinnvollen Verfahren führen. Dies kann man sich leicht klarmachen. Angenommen das Hindernis  $\chi_\ell$  hat einen Knick auf einer Kante  $E$  in der aktiven Menge. Würden wir nun den Sprung der Normalenableitung der diskreten Teillösung  $U_\ell$  auf dieser Kante  $E$  bestimmen, so könnte dieser beliebig groß sein, obwohl der Diskretisierungsfehler selbst schon beliebig klein sein kann. Ein Fehlerschätzer der diese Größe als Verfeinerungsindikator heranzieht, würde somit an Knicken im Hindernis immer weiter verfeinern, obwohl sich die Güte der diskreten Lösung dadurch nicht mehr verbessern würde. Dies liegt ganz einfach daran, dass durch Knicke im Hindernis, Knicke in der Lösung durchaus gewollt sind, und diese nicht durch weitere Verfeinerung aufgelöst werden können bzw. sollen. Wir beschränken uns daher im Folgenden auf einen residualbasierten Fehlerschätzer, welcher Kanten- und Elementterme nur auf der inaktiven Menge und der Schicht zwischen aktiver und inaktiver Menge betrachtet. Wir definieren

$$\eta_\ell := \left( \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} \eta_\ell(T)^2 + \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell)} \eta_\ell(E)^2 \right)^{1/2}, \quad (4.2.1)$$

wobei die einzelnen Beiträge durch

$$\eta_\ell(T) := h_T \|f_T\|_T, \quad \eta_\ell(E) := h_E^{1/2} \|[\partial_n U_\ell]\|_E$$

gegeben sind. Analog zu unserem ersten Ansatz für das affine Problem sind nach [BCH2] auch hier die zwei Zutaten *Zuverlässigkeit* und *diskrete lokale Effizienz* für einen konvergenten Algorithmus vonnöten. Wir verwenden im Folgenden für jeden der Beiträge von  $\eta_\ell$  eine eigene Dörflermarkierung

$$\begin{aligned} \theta \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} \eta_\ell(T)^2 &\leq \sum_{T \in \mathcal{M}_\ell^{(1)}} \eta_\ell(T)^2, \\ \theta \sum_{T \in \mathcal{E}_\ell(\mathcal{I}_\ell)} \eta_\ell(E)^2 &\leq \sum_{T \in \mathcal{M}_\ell^{(2)}} \eta_\ell(E)^2, \end{aligned} \quad (4.2.2)$$

mit  $\theta \in (0, 1)$ ,  $\mathcal{M}_\ell^{(1)} \subseteq \mathcal{T}_\ell(\mathcal{I}_\ell)$  und  $\mathcal{M}_\ell^{(2)} \subseteq \mathcal{E}_\ell(\mathcal{I}_\ell)$ , und verfeinern markierte Elemente so, dass in deren Innerem mindestens ein neuer Punkt entsteht. Wir fordern also erneut die *innerer-Punkt-Eigenschaft*. Bei einer markierten Kante  $E = T_+ \cap T_- \in \mathcal{M}_\ell^{(2)}$  verfeinern wir beide Elemente  $T_\pm \in \mathcal{T}_\ell$ . Wie üblich, sind außerdem einige Oszillationsterme von Bedeutung. Konkret definieren wir

$$osc_\ell^2 := osc_\ell^2(f) + osc_\ell^2(\chi), \quad (4.2.3)$$

wobei die Daten- und Hindernisoszillationen durch

$$osc_\ell^2(f) := \sum_{T \in \mathcal{T}_\ell} osc_T^2(f) + \sum_{E \in \mathcal{E}_\ell} osc_{\omega_\ell^E}^2(f), \quad (4.2.4)$$

$$osc_\ell^2(\chi) := \sum_{T \in \mathcal{T}_\ell} osc_T^2(\chi) + \sum_{E \in \mathcal{E}_\ell} osc_{\omega_\ell^E}^2(\chi), \quad (4.2.5)$$

$$osc_D(f) := \text{diam}(D) \|f - f_D\|_D, \quad (4.2.6)$$

$$osc_D(\chi) := \|\chi - \chi_\ell\|_D, \quad D \in \{T, \omega_\ell^E\} \quad (4.2.7)$$

gegeben sind. Zusätzlich werden auch die Kantensprünge des Hindernisses  $\chi$  von Bedeutung sein. Wir definieren daher

$$\mu_\ell^2 := \sum_{E \in \mathcal{M}_\ell^{(2)}} \mu_E^2(\chi_\ell), \quad \mu_E(\chi_\ell) := h_E^{1/2} \|[\partial_n \chi_\ell]\|_E. \quad (4.2.8)$$

Analog zum affinen Fall fordern wir außerdem ein gleichmäßiges Abklingen dieser Terme, d.h.

$$osc_{\ell+1}^2 \leq \rho_2 osc_\ell^2, \quad \mu_{\ell+1}^2 \leq \rho_3 \mu_\ell^2, \quad (4.2.9)$$

für  $0 < \rho_1, \rho_2 < 1$ . Hierfür kann gegebenenfalls zusätzlich verfeinert werden. Wir können nun den adaptiven Finite-Elemente-Algorithmus formulieren.

**Algorithmus 4.2.1** (adaptiver Algorithmus, Version 1). **Input:** Anfangstriangulierung  $\mathcal{T}_0$ , maximale Anzahl Elemente  $n_{E_{max}}$ , Adaptivitätsparameter  $\theta \in (0, 1)$ ,  $\ell := 0$ .

- (i) Berechne diskrete Lösung  $U_\ell \in K_\ell$ .
- (ii) Berechne die Verfeinerungsindikatoren  $\eta_\ell(E), \eta_\ell(T)$  und den Fehlerschätzer  $\eta_\ell$ .
- (iii) STOP, wenn die Anzahl der Elemente größer als  $n_{E_{max}}$  ist.
- (iv) Finde Mengen  $\mathcal{M}_\ell^{(1)} \subseteq \mathcal{T}_\ell(\mathcal{I}_\ell)$  und  $\mathcal{M}_\ell^{(2)} \subseteq \mathcal{E}_\ell(\mathcal{I}_\ell)$  minimaler Kardinalität von markierten Elementen und Kanten, so dass

$$\theta \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} \eta_\ell(T)^2 \leq \sum_{T \in \mathcal{M}_\ell^{(1)}} \eta_\ell(T)^2,$$

$$\theta \sum_{T \in \mathcal{E}_\ell(\mathcal{I}_\ell)} \eta_\ell(E)^2 \leq \sum_{T \in \mathcal{M}_\ell^{(2)}} \eta_\ell(E)^2,$$

- (v) Erzeuge neues, reguläres Gitter  $\mathcal{T}_{\ell+1}$ , bei dem mindestens alle markierten Elemente und Kanten verfeinert wurden. Hierbei soll die innerer-Punkt-Eigenschaft gewährleistet sein.
- (vi)  $\ell \mapsto \ell + 1$  und beginne wieder bei (i).

**Output:** Endliche Folge diskreter Lösungen  $U_\ell$  und zugehörige Fehlerschätzer  $\eta_\ell$ .

Bevor wir uns im nächsten Abschnitt der Konvergenzanalyse von Algorithmus 4.2.1 widmen, definieren wir noch eine letzte Größe, die im weiteren Verlauf auftauchen wird, den *Konsistenzfehler*

$$con_\ell := con_\ell^{rel} + con_\ell^{ort}, \quad (4.2.10)$$

mit

$$con_\ell^{rel} := |\langle \tilde{\sigma}_\ell, \chi - u \rangle|, \quad con_\ell^{ort} := \langle \sigma, \chi_\ell - U_\ell \rangle. \quad (4.2.11)$$

Die Indizes *rel* und *ort* beziehen sich hierbei auf die Sätze (4.3.1) und (4.3.5), in denen die entsprechenden Größen auftreten. Wir wollen den vorgestellten Algorithmus im Folgenden genauer untersuchen.

### 4.3. Konvergenzanalyse

In diesem Abschnitt wollen wir die Konvergenz der von Algorithmus 4.2.1 erzeugten Folge der diskreten Lösungen  $(U_\ell)_{\ell \in \mathbb{N}}$  gegen die exakte Lösung  $u \in K$  nach [BCH2] zeigen. Konkreter zeigen wir ein Konvergenzresultat, welches bis auf den Konsistenzfehler  $con_\ell$  (4.2.10) gilt. Die Zutaten für den Beweis desselbigen sind, neben der Dörflermarkierung (4.2.2) und dem Abklingen der Oszillationsterme (4.2.9), wie erwähnt, die Zuverlässigkeit und die diskrete lokale Effizienz des Fehlerschätzers  $\eta_\ell$ . Außerdem wird noch eine *gestörte Galerkin-Orthogonalität* in den Beweis einfließen. Wir werden diese Zutaten im Folgenden jeweils einzeln zeigen und

das Konvergenzresultat am Ende aus allen Beiträgen zusammensetzen. Da wir den Beweisen, wie erwähnt, in einigen Schritten nicht zustimmen können, sind die Resultate an dieser Stelle jedoch mit Vorsicht zu genießen. Wir wollen die Originalausführungen dennoch ausarbeiten, da sich hierin wichtige und sinnvolle Ideen finden. Die fragwürdigen Stellen aus der Originalarbeit [BCH2] werden wir jedenfalls herausarbeiten und in kursiver Schrift kommentieren. Die Resultate zur Schätzerreduktion ab Abschnitt 4.4 sind jedoch von den konkreten Beweisen unabhängig und machen daher auch ohne [BCH2] Sinn. Wie immer verwenden wir das Symbol  $A \lesssim B$ , wenn es eine Konstante  $C > 0$  gibt, die nur von der Formregularität der Triangulierung  $\mathcal{T}_\ell$  abhängt, so dass  $A \leq CB$  gilt. Außerdem schreiben wir, wie üblich,  $A \sim B$ , wenn  $A \lesssim B$  und  $B \lesssim A$  gilt.

### 4.3.1. Zuverlässigkeit

In diesem kurzen Abschnitt zeigen wir, dass der Fehlerschätzer (4.2.1) bis auf Oszillationsterme und Konsistenzfehler eine obere Abschätzung für den Fehler liefert. Nach [BCH2, Theorem 4.1] gilt die folgende Aussage:

**Satz 4.3.1** (Zuverlässigkeit). *Seien  $u \in K$  und  $U_\ell \in K_\ell$  die Lösungen des kontinuierlichen, bzw. des diskreten Problems (4.1.1, 4.1.2) und  $\eta_\ell, \text{osc}_\ell$  und  $\text{con}_\ell^{\text{rel}}$  wie oben beschrieben. Dann gilt*

$$\|u - U_\ell\|^2 \leq C \left( \eta_\ell^2 + \text{osc}_\ell^2 + \text{con}_\ell^{\text{rel}} \right), \quad (4.3.1)$$

für eine Konstante  $C > 0$ , die nur von der Formregularität abhängt.

*Beweis nach [BCH2, Theorem 4.1].* Sei  $e_u := u - U_\ell$ , und  $J_h : H_0^1(\Omega) \rightarrow \mathcal{S}_0^1(\mathcal{T}_\ell)$  der Clément Operator (2.2.4), dann gilt

$$\begin{aligned} \|e_u\|^2 &= \langle e_u, e_u \rangle = \langle u - U_\ell, u - U_\ell \rangle \\ &= \langle u, u - U_\ell \rangle - \langle U_\ell, u - U_\ell \rangle \\ &\quad + (f, u - U_\ell) - \langle \sigma, u - U_\ell \rangle - \langle u, u - U_\ell \rangle \\ &\quad - (f, J_h(u - U_\ell)) + \langle \sigma_\ell, J_h(u - U_\ell) \rangle_{\mathcal{M}(\Omega)} + \langle U_\ell, J_h(u - U_\ell) \rangle \\ &\quad + \langle \tilde{\sigma}_\ell, u - U_\ell \rangle - \langle \tilde{\sigma}_\ell, u - U_\ell \rangle \\ &\quad + \langle \tilde{\sigma}_\ell, J_h(u - U_\ell) \rangle - \langle \tilde{\sigma}_\ell, J_h(u - U_\ell) \rangle \\ &\quad + \langle U_\ell, J_h(u - U_\ell) \rangle - \langle U_\ell, J_h(u - U_\ell) \rangle \\ &= (f, u - U_\ell - J_h(u - U_\ell)) - \langle U_\ell, u - U_\ell - J_h(u - U_\ell) \rangle \\ &\quad - \langle \tilde{\sigma}_\ell, u - U_\ell - J_h(u - U_\ell) \rangle \\ &\quad + \langle \sigma_\ell, J_h(u - U_\ell) \rangle_{\mathcal{M}(\Omega)} - \langle \tilde{\sigma}_\ell, J_h(u - U_\ell) \rangle \\ &\quad + \langle \tilde{\sigma}_\ell - \sigma, u - U_\ell \rangle \\ &= r(e_u - J_h(e_u)) + \ell_1(J_h(e_u)) + \ell_2(e_u), \end{aligned}$$

mit

$$\begin{aligned} r(v) &:= (f, v) - \langle U_\ell, v \rangle - \langle \tilde{\sigma}_\ell, v \rangle \quad v \in H_0^1(\Omega), \\ \ell_1(V_\ell) &:= \langle \sigma_\ell, V_\ell \rangle_{\mathcal{M}(\Omega)} - \langle \tilde{\sigma}_\ell, V_\ell \rangle, \quad V_\ell \in \mathcal{S}_0^1(\mathcal{T}_\ell), \\ \ell_2(v) &:= \langle \tilde{\sigma}_\ell - \sigma, v \rangle, \quad v \in H_0^1(\Omega). \end{aligned}$$

Wir betrachten im Folgenden jeden der drei Summanden einzeln. Zunächst erhalten wir durch elementweise partielle Integration

$$\begin{aligned}
 r(v) &= \sum_{T \in \mathcal{T}_\ell} (f, v)_T - \sum_{T \in \mathcal{T}_\ell} ((\Delta U_\ell, v)_T + (\partial_n U_\ell, v)_{\partial T}) - \langle \tilde{\sigma}_\ell, v \rangle \\
 &= \sum_{T \in \mathcal{T}_\ell} (f, v)_T - \sum_{E \in \mathcal{E}_\ell} ([\partial_n U_\ell], v)_E - \langle \tilde{\sigma}_\ell, v \rangle \\
 &= \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} \kappa_T (f_T, v)_T - \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell)} \kappa_E ([\partial_n U_\ell], v)_E \\
 &\quad + \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} \kappa_T (f - f_T, v)_T,
 \end{aligned}$$

wobei die zweite Gleichheit dadurch entsteht, dass  $U_\ell$  stückweise affin ist, und  $\kappa_T$  und  $\kappa_E$  durch (4.1.16) gegeben sind. Für  $v := e_u - J_h(e_u)$  ergibt sich nun

$$\begin{aligned}
 |r(e_u - J_h(e_u))| &= \left| \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} \kappa_T (f_T, u - U_\ell - J_h(u - U_\ell))_T \right. \\
 &\quad - \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell)} \kappa_E ([\partial_n U_\ell], u - U_\ell - J_h(u - U_\ell))_E \\
 &\quad \left. + \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} \kappa_T (f - f_T, u - U_\ell - J_h(u - U_\ell))_T \right|.
 \end{aligned}$$

Anwendung der Ungleichung von Hölder liefert

$$\begin{aligned}
 |r(e_u - J_h(e_u))| &\lesssim \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} \|f_T\|_T \|u - U_\ell - J_h(u - U_\ell)\|_T \\
 &\quad + \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell)} \|[\partial_n U_\ell]\|_E \|u - U_\ell - J_h(u - U_\ell)\|_E \\
 &\quad + \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} \|f - f_T\|_T \|u - U_\ell - J_h(u - U_\ell)\|_T,
 \end{aligned}$$

was mit den Eigenschaften des Clément-Operators (2.2.6)–(2.2.7) durch

$$\begin{aligned}
 |r(e_u - J_h(e_u))| &\lesssim \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} h_T \|f_T\|_T \|u - U_\ell\|_{\omega_\ell^T} \\
 &\quad + \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell)} h_E^{1/2} \|[\partial_n U_\ell]\|_E \|u - U_\ell\|_{\omega_\ell^E} \\
 &\quad + \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} h_T \|f - f_T\|_T \|u - U_\ell\|_{\omega_\ell^T} \\
 &= \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} (\eta_\ell(T) + \text{osc}_T(f)) \|u - U_\ell\|_{\omega_\ell^T} + \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell)} \eta_\ell(E) \|u - U_\ell\|_{\omega_\ell^E} \\
 &\lesssim \frac{1}{\beta} \|u - U_\ell\|^2 + \eta_\ell^2 + \text{osc}_\ell(f)^2
 \end{aligned} \tag{4.3.2}$$

abgeschätzt werden kann. Die Konstante  $\beta > 1$  folgt hierbei aus der Ungleichung von Young. Die Oszillationsterme wurden großzügig abgeschätzt.

Für den Term  $\ell_1(J_h(e_u))$  gilt laut [BCH2, (4.5)] durch analoge Argumentation

$$|\ell_1(J_h(e_u))| \lesssim \frac{1}{\beta} \|u - U_\ell\|^2 + \sum_{T \in \mathcal{T}_{\mathcal{F}_\ell}} (\eta_\ell(T)^2 + \text{osc}_T^2(f)) + \sum_{E \in \mathcal{E}_{\mathcal{F}_\ell}} \eta_\ell(E)^2. \quad (4.3.3)$$

An dieser Stelle wurde jedoch die fragwürdige Gleichung (4.1.15) verwendet. Außerdem ist unklar, woher die für die Verfeinerungsindikatoren notwendigen  $h$ -Potenzen stammen. Die Eigenschaften des Clément-Operators (2.2.5) – (2.2.7) können an dieser Stelle nicht verwendet werden, da sie nur für den Operator  $(1 - J_h)$  Gültigkeit haben. Auch eine Erweiterung zu  $J_h(e_u) = J_h(e_u) - e_u + e_u$  ist hier nicht zielführend, da man den verbleibenden Term  $\|e_u\|_T$  nicht durch  $\|e_u\|_T$  abschätzen kann. Auf den einzelnen Elementen hat dieser Term ja keinen Nullrand, sodass wir die Ungleichung von Poincaré hier nicht verwenden können.

Für den letzten Term  $\ell_2(e_u)$  ergibt sich schließlich

$$\begin{aligned} \ell_2(e_u) &= \langle \tilde{\sigma}_\ell - \sigma, u - U_\ell \rangle \\ &\quad + \langle \tilde{\sigma}_\ell - \sigma, \chi - \chi_\ell \rangle + \langle \tilde{\sigma}_\ell - \sigma, \chi_\ell - \chi \rangle \\ &= \langle \tilde{\sigma}_\ell - \sigma, u - \chi \rangle + \langle \tilde{\sigma}_\ell - \sigma, \chi - \chi_\ell \rangle + \langle \tilde{\sigma}_\ell - \sigma, \chi_\ell - U_\ell \rangle \\ &= \langle \tilde{\sigma}_\ell, u - \chi \rangle + \langle \tilde{\sigma}_\ell - \sigma, \chi - \chi_\ell \rangle + \langle \tilde{\sigma}_\ell - \sigma, \chi_\ell - U_\ell \rangle \\ &\leq \langle \tilde{\sigma}_\ell, u - \chi \rangle + \langle \tilde{\sigma}_\ell - \sigma, \chi - \chi_\ell \rangle + \langle \tilde{\sigma}_\ell, \chi_\ell - U_\ell \rangle, \end{aligned}$$

wobei die letzte Gleichheit aus der Komplementaritätsbedingung (4.1.4) folgt und die letzte Ungleichheit von  $\langle -\sigma, \chi_\ell - U_\ell \rangle \leq 0$  herrührt.

Erneut gilt laut [BCH2, (4.6)]

$$\begin{aligned} \langle \tilde{\sigma}_\ell - \sigma, u - U_\ell \rangle &= \langle \tilde{\sigma}_\ell - \sigma, \chi_\ell - U_\ell + u - \chi \rangle + \langle \tilde{\sigma}_\ell - \sigma, \chi - \chi_\ell \rangle \\ &\sim \sum_{T \in \mathcal{T}_{\mathcal{F}_\ell}} ((f, u - U_\ell)_T + (f, \chi_\ell - \chi)_T) \\ &\quad - \sum_{E \in \mathcal{E}_{\mathcal{F}_\ell}} (([\partial_n U_\ell], u - u_\ell)_E + ([\partial_n U_\ell], \chi_\ell - \chi)_E) \\ &\quad + \langle \tilde{\sigma}_\ell - \sigma, \chi - \chi_\ell \rangle \\ &\lesssim \frac{1}{\beta} \|u - U_\ell\|^2 + \sum_{T \in \mathcal{T}_{\mathcal{F}_\ell}} (\eta_\ell(T)^2 + \text{osc}_T^2(f)) + \sum_{E \in \mathcal{E}_{\mathcal{F}_\ell}} \eta_\ell(E)^2 \\ &\quad + \text{osc}_\ell^2(\chi) + \text{con}_\ell^{\text{rel}} + \langle \tilde{\sigma}_\ell - \sigma, \chi - \chi_\ell \rangle. \end{aligned}$$

An dieser Stelle wurde jedoch anscheinend eine Ungleichung vom Poincaré Typ für  $\chi_\ell - \chi$  verwendet ohne die zusätzlich auftretenden Randterme zu beachten. Außerdem ist unklar wie die Größe auf der Menge  $\mathcal{A}_\ell \cap \mathcal{I}$  abgeschätzt wurde. Dieser Ausdruck kann eigentlich nicht einfach durch  $\text{con}_\ell^{\text{rel}}$  abgeschätzt werden, da man sonst nicht auf den Oszillationsterm  $\text{osc}_\ell^2(\chi)$  kommt.

Der restliche Teil des Beweises beschäftigt sich damit eine entsprechende Abschätzung für  $\langle \tilde{\sigma}_\ell - \sigma, \chi - \chi_\ell \rangle$  zu finden. Lassen wir die Einschränkung an die Randwerte fallen, so erhalten wir aus (4.1.3) und (4.1.7) die Gleichheiten

$$\begin{aligned} \langle u, v \rangle &= (f, v) + (\partial_n u, v)_\Gamma - \langle \sigma, v \rangle, \quad v \in H^1(\Omega), \\ \langle U_\ell, V_\ell \rangle &= (f, V_\ell) + (\partial_n U_\ell, V_\ell)_\Gamma - \langle \sigma_\ell, V_\ell \rangle_{\mathcal{M}(\Omega)}, \quad V_\ell \in \mathcal{S}^1(\mathcal{T}_\ell). \end{aligned} \quad (4.3.4)$$

Wir setzen  $\delta_\chi := \chi - \chi_\ell \in H^1(\Omega)$  und beschreiben mit  $J_h : H^1(\Omega) \rightarrow \mathcal{S}^1(\mathcal{T}_\ell)$  erneut Clément's Operator. Es gilt

$$\langle \tilde{\sigma}_\ell - \sigma, \delta_\chi \rangle = \langle \tilde{\sigma}_\ell - \sigma, J_h(\delta_\chi) \rangle + \langle \tilde{\sigma}_\ell - \sigma, \delta_\chi - J_h(\delta_\chi) \rangle. \quad (4.3.5)$$

Für den ersten Summanden ergibt sich offensichtlich die Gleichheit

$$\langle \tilde{\sigma}_\ell - \sigma, J_h(\delta_\chi) \rangle = (\langle \tilde{\sigma}_\ell, J_h(\delta_\chi) \rangle - \langle \sigma_\ell, J_h(\delta_\chi) \rangle_{\mathcal{M}(\Omega)}) + (\langle \sigma_\ell, J_h(\delta_\chi) \rangle_{\mathcal{M}(\Omega)} - \langle \sigma, J_h(\delta_\chi) \rangle). \quad (4.3.6)$$

Glaubt man den bisherigen Ausführungen, so lässt sich der erste Term mit bekannten Mitteln wieder durch

$$|\langle \tilde{\sigma}_\ell, J_h(\delta_\chi) \rangle - \langle \sigma_\ell, J_h(\delta_\chi) \rangle_{\mathcal{M}(\Omega)}| \lesssim \text{osc}_\ell^2(\chi) + \sum_{T \in \mathcal{T}_{\mathcal{F}_\ell}} (\eta_\ell(T)^2 + \text{osc}_T^2(f)) + \sum_{E \in \mathcal{E}_{\mathcal{F}_\ell}} \eta_\ell(E)^2. \quad (4.3.7)$$

abschätzen.

*Diese Stelle ist Abschätzung [BCH2, (4.11)]. Hier wurde wieder Abschätzung (4.1.15) sowie die Eigenschaften des Clément-Operator direkt für  $J_h$  verwendet.*

Für den zweiten Summanden von (4.3.6) liefert die Young'sche Ungleichung in Verbindung mit (4.3.4) laut [BCH2, (4.13)]

$$\begin{aligned} |\langle \sigma_\ell, J_h(\delta_\chi) \rangle_{\mathcal{M}(\Omega)} - \langle \sigma, J_h(\delta_\chi) \rangle| &= | - \langle U_\ell, J_h(\delta_\chi) \rangle + (f, J_h(\delta_\chi)) + (\partial_n U_\ell, J_h(\delta_\chi))_\Gamma \\ &\quad + \langle u, J_h(\delta_\chi) \rangle - (f, J_h(\delta_\chi)) - (\partial_n u, J_h(\delta_\chi))_\Gamma | \\ &\leq | \langle u - U_\ell, J_h(\delta_\chi) \rangle | + | (\partial_n(u - U_\ell), J_h(\delta_\chi))_\Gamma | \\ &\lesssim \frac{1}{\beta} \|u - U_\ell\|^2 + \text{osc}_\ell^2(\chi), \end{aligned} \quad (4.3.8)$$

wobei für die letzte Abschätzung die *Spur-Ungleichung*

$$\|\partial_n(u - U_\ell)\|_\Gamma \lesssim \|u - U_\ell\| \quad (4.3.9)$$

verwendet wurde.

*An diesem Punkt ist jedoch die grundsätzliche Gültigkeit der Spurungleichung [BCH2, (4.12)] fragwürdig. Unklar ist zudem wie sich der Term  $\text{osc}_\ell(\chi)^2$  aus  $\|J_h(\delta_\chi)\|_\Gamma$  ergibt.*

Es bleibt der zweite Summand von (4.3.5) zu untersuchen. Mit der Darstellung (4.1.13) und (4.3.4) erhalten wir

$$\begin{aligned} \langle \sigma - \tilde{\sigma}_\ell, \delta_\chi - J_h(\delta_\chi) \rangle &= (f, \delta_\chi - J_h(\delta_\chi)) + (\partial_n u, \delta_\chi - J_h(\delta_\chi))_\Gamma - \langle u, \delta_\chi - J_h(\delta_\chi) \rangle \\ &\quad + \sum_{p \in \mathcal{N}_\ell(\mathcal{A}_\ell)} \left( -\frac{1}{3} \sum_{T \in \omega_\ell^p} (f, \delta_\chi - J_h(\delta_\chi))_T + \frac{1}{2} \sum_{E \in \mathcal{E}_\ell^p} ([\partial_n U_\ell], \delta_\chi - J_h(\delta_\chi))_E \right). \end{aligned}$$

Mit

$$\begin{aligned} I_1 &:= (f, \delta_\chi - J_h(\delta_\chi)) - \sum_{p \in \mathcal{N}_\ell(\mathcal{A}_\ell)} \frac{1}{3} \sum_{T \in \omega_\ell^p} (f, \delta_\chi - J_h(\delta_\chi))_T, \\ I_2 &:= (\partial_n u, \delta_\chi - J_h(\delta_\chi))_\Gamma - \langle u, \delta_\chi - J_h(\delta_\chi) \rangle \\ &\quad + \sum_{p \in \mathcal{N}_\ell(\mathcal{A}_\ell)} \frac{1}{2} \sum_{E \in \mathcal{E}_\ell^p} ([\partial_n U_\ell], \delta_\chi - J_h(\delta_\chi))_E, \end{aligned}$$

gilt also

$$\langle \sigma - \tilde{\sigma}_\ell, \delta_\chi - J_h(\delta_\chi) \rangle = I_1 + I_2. \quad (4.3.10)$$

Für den ersten Term ergibt sich nun

$$\begin{aligned} |I_1| &\leq \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} (1 - \kappa_T) (|(f_T, \delta_\chi - J_h(\delta_\chi))_T| + |(f - f_T, \delta_\chi - J_h(\delta_\chi))_T|) \\ &\lesssim \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} (h_T \|f\|_T + h_T \|f - f_T\|_T) \|\delta_\chi\|_{\omega_\ell^T} \\ &\lesssim \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} (\eta_\ell(T)^2 + \text{osc}_T^2(f)) + \text{osc}_\ell^2(\chi). \end{aligned}$$

Für den zweiten Term ergibt sich die Abschätzung

$$\begin{aligned} |I_2| &= |(\partial_n u, \delta_\chi - J_h(\delta_\chi))_\Gamma - \langle u, \delta_\chi - J_h(\delta_\chi) \rangle \\ &\quad + \sum_{p \in \mathcal{N}_\ell(\mathcal{A}_\ell)} \frac{1}{2} \sum_{E \in \mathcal{E}_\ell^p} ([\partial_n U_\ell], \delta_\chi - J_h(\delta_\chi))_E \\ &\quad + \langle U_\ell, \delta_\chi - J_h(\delta_\chi) \rangle - (f, \delta_\chi - J_h(\delta_\chi)) - (\partial_n U_\ell, \delta_\chi - J_h(\delta_\chi))_\Gamma + \langle \sigma_\ell, \delta_\chi - J_h(\delta_\chi) \rangle_{\mathcal{M}(\Omega)} \\ &\leq |\langle u - U_\ell, \delta_\chi - J_h(\delta_\chi) \rangle| + \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell)} (1 - \kappa_E) |([\partial_n U_\ell], \delta_\chi - J_h(\delta_\chi))_E| \\ &\quad + \sum_{E \in \mathcal{E}_\ell(\Gamma)} |(\partial_n(u - U_\ell), \delta_\chi - J_h(\delta_\chi))_E| \\ &\lesssim \frac{1}{\beta} \|u - U_\ell\|^2 + \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell)} \eta_\ell(E)^2 + \text{osc}_\ell^2(\chi), \end{aligned}$$

wobei hier wieder die Spur-Ungleichung (4.3.9), (4.3.4) und

$$\langle U_\ell, \delta_\chi - J_h(\delta_\chi) \rangle = \sum_{E \in \mathcal{E}_\ell} ([\partial_n U_\ell], \delta_\chi - J_h(\delta_\chi))_E + \sum_{E \in \mathcal{E}_\ell(\Gamma)} (\partial_n U_\ell, \delta_\chi - J_h(\delta_\chi))_E$$

verwendet wurde. Aus den beiden letzten Abschätzungen folgt somit

$$|\langle \sigma - \tilde{\sigma}_\ell, \delta_\chi - J_h(\delta_\chi) \rangle| \lesssim \frac{1}{\beta} \|u - U_\ell\|^2 + \eta_\ell^2 + \text{osc}_\ell^2(f) + \text{osc}_\ell^2(\chi). \quad (4.3.11)$$

Schenkt man den fragwürdigen Stellen Glauben, so folgt das gesuchte Resultat als Kombination der Abschätzungen (4.3.2), (4.3.3), (4.3.7), (4.3.8), (4.3.10) und (4.3.11).  $\square$

**Bemerkung.** *Der obige Beweis der Zuverlässigkeit vereinfacht sich sehr stark, wenn wir stückweise affine Hindernisse, d.h.  $\chi_\ell = \chi$  annehmen. In diesem Fall werden sämtliche Berechnungen ab (4.3.4) obsolet, da dann  $\langle \tilde{\sigma}_\ell - \sigma, \chi - \chi_\ell \rangle \equiv 0$  gilt.*

### 4.3.2. Diskrete lokale Effizienz

In diesem Abschnitt zeigen wir die diskrete Effizienz des Fehlerschätzers  $\eta_\ell$ . Konkret zeigen wir, dass die Energienorm der Differenz zweier aufeinander folgender Lösungen  $U_\ell - U_{\ell+1}$  bis auf Oszillations- und Datenterme eine obere Schranke für  $\eta_\ell$  darstellt. Wie im affinen Fall geht auch hier die *innerer-Punkt-Eigenschaft* massiv ein, so dass die Gültigkeit des Resultats im Prinzip auf *bisec<sub>5</sub>*-Verfeinerung beschränkt ist. In den folgenden zwei Lemmata untersuchen wir zunächst die verfeinerten Elemente  $T \in \mathcal{M}_\ell^{(1)}$  und dann die verfeinerten Kanten  $E \in \mathcal{M}_\ell^{(2)}$ . Das zentrale Resultat der diskreten Effizienz (Satz 4.3.4) wird sich dann als Folgerung ergeben. Die Aussagen stammen aus [BCH2, Lemma 5.1, Lemma 5.2 und Theorem 5.1].

**Lemma 4.3.2.** *Sei  $T \in \mathcal{M}_\ell^{(1)}$  ein zur Verfeinerung markiertes Element mit neuem inneren Knoten  $p \in \mathcal{N}_{\ell+1}(T)$ . Wir unterscheiden zwei Fälle.*

(i) *Falls  $p \in \mathcal{N}_{\ell+1}(\mathcal{I}_{\ell+1})$ , dann gilt*

$$\eta_\ell(T)^2 \leq C_1 (\|U_\ell - U_{\ell+1}\|_T^2 + \text{osc}_T^2(f)). \quad (4.3.12)$$

(ii) *Sei  $p \in \mathcal{N}_{\ell+1}(\mathcal{A}_{\ell+1})$ . Wegen  $T \in \mathcal{M}_\ell^{(1)}$  gibt es einen Knoten  $\hat{p} \in \mathcal{N}_\ell(T) \cap \mathcal{N}_\ell(\mathcal{I}_\ell)$ , und es gilt*

$$\eta_\ell(T)^2 \leq C_2 \left( h_T^2 \|f - f_{\omega_\ell^{\hat{p}}}\|_{\omega_\ell^{\hat{p}}}^2 + \sum_{E \in \mathcal{E}_\ell^{\hat{p}}} \eta_\ell(E)^2 \right). \quad (4.3.13)$$

Die Konstanten  $C_1, C_2 > 0$  hängen jeweils nur von der Formregularität und der Geometrie von  $\mathcal{T}_\ell$  ab.

*Beweis.* Sei  $p \in \mathcal{N}_{\ell+1}(T)$  ein innerer Knoten und  $\varphi_p$  seine nodale Basisfunktion. Wir wählen nun  $\psi_p := \kappa \varphi_p$  mit  $\kappa \sim f_T$  so, dass

$$\eta_\ell(T)^2 = h_T^2 \|f_T\|_T^2 \leq h_T^2 (f_T, \psi_p)_T. \quad (4.3.14)$$

Da der Gradient von  $U_\ell$  stückweise konstant ist, erhalten wir durch partielle Integration

$$\begin{aligned} \langle U_\ell, v \rangle &= \langle U_\ell, v \rangle_T \\ &= (\Delta U_\ell, v)_T + (\partial_n U_\ell, v)_{\partial T} = 0, \end{aligned} \quad (4.3.15)$$

für alle  $v \in H_0^1(T)$  deren Träger kompakt in  $T$  liegt. Da  $\psi_p$  diese Anforderungen offensichtlich erfüllt, folgt aus obiger Abschätzung somit

$$h_T^2 \|f_T\|_T^2 \leq h_T^2 ((f_T, \psi_p)_T - \langle U_\ell, \psi_p \rangle_T). \quad (4.3.16)$$

Mit  $\psi_p \in H_0^1(\mathcal{T}_{\ell+1})$  erhalten wir mit (4.1.7) außerdem

$$\langle U_{\ell+1}, \psi_p \rangle - (f, \psi_p) + \langle \sigma_{\ell+1}, \psi_p \rangle_{\mathcal{M}(\Omega)} = 0. \quad (4.3.17)$$

Unsere Abschätzung erweitert sich somit zu

$$h_T^2 \|f_T\|_T^2 \leq h_T^2 ((f_T - f, \psi_p) + \langle U_{\ell+1} - U_\ell, \psi_p \rangle + \langle \sigma_{\ell+1}, \psi_p \rangle_{\mathcal{M}(\Omega)}). \quad (4.3.18)$$

Wir unterscheiden nun die einzelnen Fälle.

**Fall 1** ( $p \in \mathcal{N}_{\ell+1}(\mathcal{I}_{\ell+1})$ ): Anwendung von Lemma 4.1.3 liefert

$$\langle \sigma_{\ell+1}, \psi_p \rangle_{\mathcal{M}(\Omega)} = \kappa \alpha_{\ell+1}(p) = 0,$$

woraus mit (4.3.18) direkt

$$h_T^2 \|f_T\|_T^2 \leq h_T^2 \|U_\ell - U_{\ell+1}\|_T \|\psi_p\|_T + h_T \text{osc}_T(f) \|\psi_p\|_T \quad (4.3.19)$$

folgt. Mit den Abschätzungen

$$\begin{aligned} h_T^2 \|\psi_p\|_T &\sim h_T^2 |\kappa| \sim h_T \|f_T\|_T, \\ h_T \|\psi_p\|_T &\sim h_T |T|^{1/2} |\kappa| \sim h_T \|f_T\|_T, \end{aligned} \quad (4.3.20)$$

folgt die gesuchte Aussage unter Verwendung der Young'schen Ungleichung.

**Fall 2** ( $p \in \mathcal{N}_{\ell+1}(\mathcal{A}_{\ell+1})$ ): Analog zu vorher gilt

$$\begin{aligned} \eta_\ell(T)^2 &= h_T^2 \|f_T\|_T^2 \leq h_T^2 (f_T, \psi_p)_T \\ &= h_T^2 (f_T - f, \psi_p)_T + h_T^2 (f, \psi_p)_T. \end{aligned} \quad (4.3.21)$$

Nach Annahme gibt es ein  $\hat{p} \in \mathcal{N}_\ell(T) \cap \mathcal{N}_\ell(\mathcal{I}_\ell)$  mit nodaler Basisfunktion  $\varphi_{\hat{p}}$ . Wir wählen nun  $\psi_{\hat{p}} := \kappa \varphi_{\hat{p}}$  und  $\alpha > 0$  so, dass

$$\int_{\omega_\ell^{\hat{p}}} (\varphi_p - \alpha \varphi_{\hat{p}}) dx = 0.$$

Da der Träger von  $\psi_{\hat{p}}$  kompakt in der inaktiven Menge  $\mathcal{I}_\ell$  im  $\ell$ -ten Schritt liegt, gilt

$$\langle U_\ell, \psi_{\hat{p}} \rangle = (f, \psi_{\hat{p}})_{\omega_\ell^{\hat{p}}}.$$

Durch partielle Integration erhalten wir andererseits

$$\langle U_\ell, \psi_{\hat{p}} \rangle = \sum_{E \in \mathcal{E}_\ell^{\hat{p}}} ([\partial_n U_\ell], \psi_{\hat{p}})_E,$$

da  $\psi_{\hat{p}} \in H_0^1(\omega_\ell^{\hat{p}})$ . Aus den letzten drei Identitäten folgt nun

$$\begin{aligned} h_T^2 (f, \psi_p)_T &= h_T^2 (f, \psi_p - \alpha \psi_{\hat{p}})_{\omega_\ell^{\hat{p}}} + \alpha h_T^2 (f, \psi_{\hat{p}})_{\omega_\ell^{\hat{p}}} \\ &= h_T^2 (f - f_{\omega_\ell^{\hat{p}}}, \psi_p - \alpha \psi_{\hat{p}})_{\omega_\ell^{\hat{p}}} + \alpha h_T^2 \langle U_\ell, \psi_{\hat{p}} \rangle \\ &= h_T^2 (f - f_{\omega_\ell^{\hat{p}}}, \psi_p - \alpha \psi_{\hat{p}})_{\omega_\ell^{\hat{p}}} + \alpha h_T^2 \sum_{E \in \mathcal{E}_\ell^{\hat{p}}} ([\partial_n U_\ell], \psi_{\hat{p}})_E, \end{aligned} \quad (4.3.22)$$

wobei wir für das zweite Gleichheitszeichen

$$\begin{aligned} (f_{\omega_\ell^{\hat{p}}}, \psi_p - \alpha \psi_{\hat{p}})_{\omega_\ell^{\hat{p}}} &= \int_{\omega_\ell^{\hat{p}}} f_{\omega_\ell^{\hat{p}}} (\psi_p - \psi_{\hat{p}}) dx \\ &= \int_{\omega_\ell^{\hat{p}}} f_{\omega_\ell^{\hat{p}}} \kappa (\varphi_p - \alpha \varphi_{\hat{p}}) dx \\ &= f_{\omega_\ell^{\hat{p}}} \kappa \int_{\omega_\ell^{\hat{p}}} (\varphi_p - \alpha \varphi_{\hat{p}}) dx = 0 \end{aligned}$$

verwendet haben. Die beiden Summanden der rechten Seite von (4.3.22) können durch

$$\begin{aligned} & h_T^2 |(f - f_{\omega_{\hat{p}}}, \psi_p - \alpha\psi_{\hat{p}})_{\omega_{\hat{p}}}| \\ & \lesssim h_T \|f - f_{\omega_{\hat{p}}}\|_{\omega_{\hat{p}}} \left( h_T \|\psi_p\|_T + \alpha |\omega_{\hat{p}}|^{1/2} \|\psi_{\hat{p}}\|_{\omega_{\hat{p}}} \right) \end{aligned} \quad (4.3.23)$$

und

$$h_T^2 |([\partial_n U_\ell], \psi_{\hat{p}})| \lesssim h_E^{1/2} \|[\partial_n U_\ell]\|_E h_E^{3/2} \|\psi_{\hat{p}}\|_E \quad (4.3.24)$$

abgeschätzt werden. Mit

$$\begin{aligned} |\omega_{\hat{p}}|^{1/2} \|\psi_{\hat{p}}\|_{\omega_{\hat{p}}} &= |\omega_{\hat{p}}|^{1/2} |\kappa| \|\varphi_{\hat{p}}\|_{\omega_{\hat{p}}} \lesssim h_T \|f_T\|_T, \\ h_E^{3/2} \|\psi_{\hat{p}}\|_E &= h_E^{3/2} |\kappa| \|\varphi_{\hat{p}}\|_E \lesssim h_T \|f_T\|_T, \end{aligned}$$

(4.3.20), sowie (4.3.23) – (4.3.24), erhalten wir

$$h_T^2 |(f, \psi_p)| \lesssim \left( h_T \|f - f_{\omega_{\hat{p}}}\|_{\omega_{\hat{p}}} + \sum_{E \in \mathcal{E}_{\hat{p}}} h_E^{1/2} \|[\partial_n U_\ell]\|_E \right) h_T \|f_T\|_T. \quad (4.3.25)$$

Die Verwendung von

$$\|f - f_T\|_T \leq \|f - f_{\omega_{\hat{p}}}\|_T \leq \|f - f_{\omega_{\hat{p}}}\|_{\omega_{\hat{p}}}$$

mit (4.3.20) und (4.3.25) in (4.3.21) vollendet schließlich den Beweis.  $\square$

Im nächsten Lemma wollen wir uns mit den zur Verfeinerung markierten Kanten beschäftigen. Analog zum affinen Fall unterscheiden wir hierbei wieder, ob die neuen inneren Punkte  $p, p_\pm$  (siehe Abbildung 4.1) zur aktiven oder zu inaktiven Menge zählen. Zu diesem Zweck betrachten wir wieder die Differenzenfunktion

$$\delta_{\ell+1} := U_{\ell+1} - \chi_{\ell+1} \leq 0 \in \mathcal{S}_0^1(\mathcal{T}_{\ell+1}).$$

Es gilt die folgende Aussage:

**Lemma 4.3.3.** *Sei  $E \in \mathcal{M}_\ell^{(2)}$ ,  $E = T_+ \cap T_-$ ,  $T_\pm \in \mathcal{T}_\ell$  eine zu verfeinernde Ecke mit Mittelpunkt  $p \in \mathcal{N}_{\ell+1}(E)$ . Den Patch von  $E$  bezeichnen wir mit  $\omega_\ell^E := T_+ \cup T_-$ . Es gilt*

$$\eta_\ell(E)^2 \leq C \left( \|U_\ell - U_{\ell+1}\|_{\omega_\ell^E}^2 + \text{osc}_{\omega_\ell^E}^2(f) + \text{osc}_{\omega_\ell^E}^2(\chi) + \mu_E^2(\chi_\ell) \right), \quad (4.3.26)$$

wobei  $\mu_E(\chi_\ell)$  wie in (4.2.8) definiert ist und die Konstante  $C > 0$  nur von der Formregularität und der Geometrie von  $\mathcal{T}_\ell$  abhängt.

*Beweis.* Seien  $p_\pm \in \mathcal{N}_{\ell+1}(T_\pm)$  die neuen inneren Knoten in  $T_\pm$ . Wir unterscheiden im Folgenden zwei Fälle:

- (i)  $\delta_{\ell+1}(p_+) = 0 = \delta_{\ell+1}(p_-)$ ,
- (ii)  $\delta_{\ell+1}(p_+) < 0$  oder  $\delta_{\ell+1}(p_-) < 0$ .

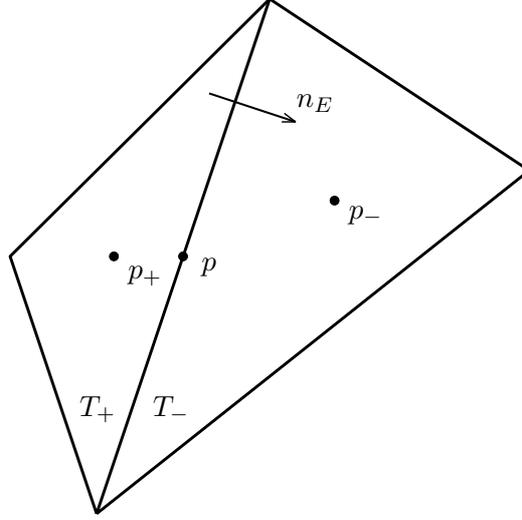


Abbildung 4.1.: Notationen für Lemma 4.3.3.

**Fall (i):** Mit  $\delta_\ell := U_\ell - \chi_\ell$  gilt

$$h_E \|[\partial_n U_\ell]\|_E^2 \lesssim h_E \|[\partial_n \delta_\ell]\|_E^2 + \mu_E^2(\chi_\ell), \quad (4.3.27)$$

wobei wir hier die Ungleichung von Cauchy verwendet haben. Da  $\nabla \delta_\ell|_T$  auf  $T \in \{T_\pm\}$  konstant ist, muss mindestens ein Sohn  $T' \in \mathcal{T}_{\ell+1}(T)$  von  $T \in \mathcal{T}_\ell$  existieren, auf dem  $n_E \cdot \nabla \delta_\ell|_T = n_E \cdot \nabla \delta_\ell|_{T'}$  und  $n_E \cdot \nabla \delta_{\ell+1}|_{T'}$  unterschiedliche Vorzeichen haben (oder Null sind). Mit der Ungleichung von Cauchy-Schwarz ergibt sich nun

$$|n_E \cdot \nabla \delta_\ell|_{T'}| \leq |n_E \cdot \nabla(\delta_\ell - \delta_{\ell+1})|_{T'}| \leq \|\nabla(\delta_\ell - \delta_{\ell+1})|_{T'}\|_2.$$

Hierbei bezeichnet  $\|\cdot\|_2$  die euklidische Norm im  $\mathbb{R}^n$ . Mit  $|T'| \sim |T| \sim h_E|E|$  folgt somit

$$\begin{aligned} h_E \|[\partial_n \delta_\ell]\|_E^2 &= h_E \int_E [\partial_n \delta_\ell]^2 dx \\ &\sim h_E^2 [\partial_n \delta_\ell]^2 \\ &= h_E^2 (\partial_n \delta_\ell|_{T_+} - \partial_n \delta_\ell|_{T_-})^2 \\ &\lesssim h_E^2 (\partial_n \delta_\ell|_{T_+})^2 + h_E^2 (\partial_n \delta_\ell|_{T_-})^2 \\ &= h_E^2 (\partial_n \delta_\ell|_{T'_+})^2 + h_E^2 (\partial_n \delta_\ell|_{T'_-})^2 \\ &\leq h_E^2 |\nabla(\delta_\ell - \delta_{\ell+1})|_{T'_+}|^2 + h_E^2 |\nabla(\delta_\ell - \delta_{\ell+1})|_{T'_-}|^2 \\ &= \|\nabla(\delta_\ell - \delta_{\ell+1})\|_{T'_+}^2 + \|\nabla(\delta_\ell - \delta_{\ell+1})\|_{T'_-}^2 \\ &\leq \|\nabla(\delta_\ell - \delta_{\ell+1})\|_{T_+}^2 + \|\nabla(\delta_\ell - \delta_{\ell+1})\|_{T_-}^2 \\ &= \|\delta_\ell - \delta_{\ell+1}\|_{T_+}^2 + \|\delta_\ell - \delta_{\ell+1}\|_{T_-}^2 \\ &\lesssim \|U_\ell - U_{\ell+1}\|_{\omega_\ell^E}^2 + \text{osc}_{\omega_\ell^E}^2(\chi). \end{aligned} \quad (4.3.28)$$

Die Kombination von (4.3.27) und (4.3.28) liefert schließlich das gewünschte Ergebnis (4.3.26).

**Fall (ii):** Ohne Beschränkung der Allgemeinheit wollen wir den Fall  $\delta_{\ell+1}(p_+) < 0$  betrachten. Der Fall  $\delta_{\ell+1}(p_-) < 0$  folgt dann durch völlig analoge Berechnungen. Wir betrachten die weiteren Fallunterscheidungen

$$(ii)_1 \quad \delta_{\ell+1}(p) < 0,$$

$$(ii)_2 \quad \delta_{\ell+1}(p) = 0.$$

*Fall (ii)<sub>1</sub>:* Mit  $\varphi_p$  und  $\varphi_{p_+}$  bezeichnen wir die nodalen Basisfunktionen bezüglich  $p_+$  und dem Mittelpunkt der Kante  $E$ ,  $p := \text{mid}(E)$ . Da sowohl  $\varphi_p$ , als auch  $\varphi_{p_+}$  nach Anforderung einen kompakten Träger in der inaktiven Menge haben, gilt

$$\langle\langle U_{\ell+1}, \varphi_p \rangle\rangle = (f, \varphi_p) \quad \text{und} \quad \langle\langle U_{\ell+1}, \varphi_{p_+} \rangle\rangle = (f, \varphi_{p_+}). \quad (4.3.29)$$

Unter Verwendung von (4.3.15) folgt hieraus

$$\langle\langle U_{\ell+1} - U_\ell, \varphi_{p_+} \rangle\rangle = (f, \varphi_{p_+}). \quad (4.3.30)$$

Wir setzen nun  $\varphi_E := \varphi_p - \alpha\varphi_{p_+}$  und wählen  $\alpha > 0$  wieder so, dass

$$\int_{\omega_\ell^E} \varphi_E \, dx = 0.$$

Partielle Integration, (4.3.29) – (4.3.30) und Ausnutzen der Tatsache, dass  $\varphi_E|_{\partial T_+} = \varphi_E|_E = \varphi_E|_{\partial T_-}$  führen zu

$$\begin{aligned} \frac{1}{2} \int_E [\partial_n U_\ell] \, ds &= \int_E [\partial_n U_\ell] \varphi_E \, ds \\ &= \langle\langle U_\ell, \varphi_E \rangle\rangle - \langle\langle U_{\ell+1}, \varphi_E \rangle\rangle + (f, \varphi_E) \\ &= \langle\langle U_\ell - U_{\ell+1}, \varphi_E \rangle\rangle + (f, \varphi_E)_{\omega_\ell^E} \\ &= \langle\langle U_\ell - U_{\ell+1}, \varphi_E \rangle\rangle + (f - f_{\omega_\ell^E}, \varphi_E)_{\omega_\ell^E}. \end{aligned}$$

Mit

$$\eta_\ell(E) = h_E^{1/2} \left( \int_E [\partial_n U_\ell]^2 \, ds \right)^{1/2} = h_E^{1/2} |E|^{1/2} ([\partial_n U_\ell]^2)^{1/2} \sim h_E |[\partial_n U_\ell]| = \int_E |[\partial_n U_\ell]| \, ds$$

erhalten wir

$$\eta_\ell(E)^2 \lesssim \|U_\ell - U_{\ell+1}\|_{\omega_\ell^E}^2 + \text{osc}_{\omega_\ell^E}^2(f)$$

und damit die gesuchte Aussage. Den Term  $\|\varphi_E\|$  haben wir hierbei durch eine inverse Abschätzung in den Griff bekommen.

*Fall (ii)<sub>2</sub>:* Für diesen Fall müssen wir erneut eine Fallunterscheidung machen. Wir betrachten

$$(ii)_{2,1} \quad [\partial_n U_\ell] \leq 0,$$

(ii)<sub>2,2</sub>  $[\partial_n U_\ell] > 0$ .

Fall (ii)<sub>2,1</sub>: Mit bekannten Argumenten existiert  $T'_\pm \in \mathcal{T}_{\ell+1}(T_\pm)$  so, dass

$$n_E \cdot \nabla \delta_{\ell+1}|_{T'_+} \geq 0 \geq n_E \cdot \nabla \delta_{\ell+1}|_{T'_-}.$$

Damit ergibt sich direkt

$$\begin{aligned} 0 &\leq -[\partial_n U_\ell] = -\left(n_E \cdot \nabla \delta_\ell|_{T'_+} - n_E \cdot \nabla \delta_\ell|_{T'_-}\right) - [\partial_n \chi_\ell] \\ &\leq -\left(n_E \cdot \nabla (\delta_\ell - \delta_{\ell+1})|_{T'_+} - n_E \nabla (\delta_\ell - \delta_{\ell+1})|_{T'_-}\right) - [\partial_n \chi_\ell] \\ &\leq \|\nabla (\delta_\ell - \delta_{\ell+1})|_{T'_+}\|_2 + \|\nabla (\delta_\ell - \delta_{\ell+1})|_{T'_-}\|_2 + |[\partial_n \chi_\ell]|. \end{aligned}$$

Mit  $|\omega_\ell^E| \sim |T_\pm| \sim h_E^2$  und

$$\eta_\ell(E) \sim -h_E [\partial_n U_\ell]$$

folgt

$$\eta_\ell(E)^2 \lesssim \|U_\ell - U_{\ell+1}\|_{\omega_\ell^E}^2 + \mu_E^2(\chi_\ell)$$

und damit die gesuchte Aussage.

Fall (ii)<sub>2,2</sub>: Als letztes gilt es den Fall  $[\partial_n U_\ell] > 0$  zu untersuchen. Wie zuvor gilt

$$\langle\langle U_{\ell+1}, \varphi_{p_+} \rangle\rangle = (f, \varphi_{p_+}).$$

Mit  $V_\ell := U_{\ell+1} - \varphi_p \in K_{\ell+1}$ , folgt aus der Variationsformulierung (4.1.6) außerdem

$$\langle\langle U_{\ell+1}, \varphi_p \rangle\rangle \leq (f, \varphi_p).$$

Mit  $\varphi_E$  aus Fall (ii)<sub>1</sub> ergibt sich

$$\begin{aligned} 0 &< \frac{1}{2} \int_E [\partial_n U_\ell] ds = \int_E [\partial_n U_\ell] \varphi_E ds \\ &\leq \langle\langle U_\ell, \varphi_E \rangle\rangle - \langle\langle U_{\ell+1}, \varphi_E \rangle\rangle + (f, \varphi_E) \\ &= \langle\langle U_\ell - U_{\ell+1}, \varphi_E \rangle\rangle + (f, \varphi_E) \\ &= \langle\langle U_\ell - U_{\ell+1}, \varphi_E \rangle\rangle + (f - f_{\omega_\ell^E}, \varphi_E)_{\omega_\ell^E}, \end{aligned}$$

woraus das gesuchte Ergebnis analog zu Fall (ii)<sub>1</sub> folgt. Damit ist die Aussage vollständig bewiesen.  $\square$

Wir verwenden nun diese beiden Lemmata um die diskrete Effizienz des Fehlerschätzers nachzuweisen.

**Satz 4.3.4** (Diskrete Effizienz). *Mit den Voraussetzungen und Notationen der Lemmata 4.3.2 und 4.3.3 gilt*

$$\eta_\ell^2 \leq C \left( \|U_\ell - U_{\ell+1}\|^2 + \text{osc}_\ell^2 + \mu_\ell^2 \right), \quad (4.3.31)$$

wobei  $C > 0$ , abgesehen von der Formregularität und der Geometrie der Triangulierung, nur vom Adaptivitätsparameter  $\theta \in (0, 1)$  aus der Dörflermarkierung abhängt.

*Beweis.* Die Aussage folgt aus

$$\begin{aligned} \eta_\ell^2 &= \left( \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} \eta_\ell(T)^2 + \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell)} \eta_\ell(E)^2 \right) \\ &\leq \frac{1}{\theta} \left( \sum_{T \in \mathcal{M}_\ell^{(1)}} \eta_\ell(T)^2 + \sum_{E \in \mathcal{M}_\ell^{(2)}} \eta_\ell(E)^2 \right) \end{aligned}$$

und der Anwendung von Lemma 4.3.2 bzw. 4.3.3 auf die einzelnen Summanden.  $\square$

### 4.3.3. Gestörte Galerkin-Orthogonalität und Fehlerreduktion

Wir haben bisher zwei wichtige Eigenschaften von  $\eta_\ell$ , seine Zuverlässigkeit und seine diskrete Effizienz beisammen. Im Gegensatz zum affinen Fall ist nach [BCH2] jedoch noch eine weitere Zutat für einen Konvergenzbeweis nötig, eine *gestörte Galerkin-Orthogonalität*. Wie bereits in Kapitel 3 erwähnt, sind die Mengen  $K_\ell$  keine Hilberträume, weshalb der Satz von Riesz, und damit auch die Galerkin-Orthogonalität, ihre Gültigkeit verlieren. Es ist allerdings möglich eine schwächere Form dieser Eigenschaft zu zeigen und so  $\|U_\ell - U_{\ell+1}\|$  durch die Fehler in der Energienorm in der  $\ell$ -ten und  $(\ell + 1)$ -ten Iteration abzuschätzen. Der nachfolgende Satz stammt auch aus der Originalarbeit von BRAESS, CARSTENSEN und HOPPE [BCH2, Theorem 5.2]. Auch der Beweis dieses Satzes ist leider fragwürdig. Wie zuvor sind die entsprechenden Stellen deutlich herausgearbeitet.

**Satz 4.3.5** (Gestörte Galerkin-Orthogonalität). *Es seien  $U_\ell \in K_\ell$  und  $U_{\ell+1} \in K_{\ell+1}$  die diskreten Lösungen im  $\ell$ -ten und  $(\ell + 1)$ -ten Schritt. Ferner seien  $osc_\ell$  die Datenoszillationen (4.2.4) und  $con_\ell^{ort}$  der Konsistenzfehler (4.2.10). Ist die Anforderung (4.2.9) and das Abfallen der Oszillationsterme erfüllt, dann gilt für jedes  $\varepsilon > 0$*

$$\begin{aligned} \|U_\ell - U_{\ell+1}\|^2 &\leq \|u - U_\ell\|^2 - (1 - 4\varepsilon) \|u - U_{\ell+1}\|^2 \\ &\quad + \frac{1}{\varepsilon} (1 + \rho_2) osc_\ell^2(\chi) + con_\ell^{ort}. \end{aligned} \quad (4.3.32)$$

*Beweis nach [BCH2, Theorem 5.2].* Durch direktes Ausrechnen ergibt sich

$$\begin{aligned} \|U_\ell - U_{\ell+1}\|^2 &= \langle U_\ell - U_{\ell+1}, U_\ell - U_{\ell+1} \rangle \\ &= \langle U_\ell, U_\ell \rangle - 2\langle U_\ell, U_{\ell+1} \rangle + \langle U_{\ell+1}, U_{\ell+1} \rangle \\ &= \langle U_\ell, U_\ell \rangle - 2\langle U_\ell, U_{\ell+1} \rangle + \langle U_{\ell+1}, U_{\ell+1} \rangle \\ &\quad + \langle u, u \rangle - \langle u, u \rangle + 2\langle u, U_\ell \rangle - 2\langle u, U_\ell \rangle + 2\langle u, U_{\ell+1} \rangle - 2\langle u, U_{\ell+1} \rangle \\ &\quad + \langle U_{\ell+1}, U_{\ell+1} \rangle - \langle U_{\ell+1}, U_{\ell+1} \rangle \\ &= \|u - U_\ell\|^2 - \|u - U_{\ell+1}\|^2 - 2\langle u - U_{\ell+1}, U_\ell - U_{\ell+1} \rangle. \end{aligned} \quad (4.3.33)$$

Aus den Darstellungen mittels Lagrange'schem Multiplikator (4.1.3) und (4.1.7) erhalten wir ferner

$$\begin{aligned} &2\langle u - U_{\ell+1}, U_\ell - U_{\ell+1} \rangle \\ &= \underbrace{2\langle f - f, U_\ell - U_{\ell+1} \rangle}_{=0} + 2(\langle \sigma_{\ell+1}, U_\ell - U_{\ell+1} \rangle_{\mathcal{M}(\Omega)} - \langle \sigma, U_\ell - U_{\ell+1} \rangle). \end{aligned} \quad (4.3.34)$$

Wir nutzen im Folgenden die Komplementaritätsbedingung (4.1.9), sowie  $\sigma_{\ell+1} \in \mathcal{M}_+(\Omega)$  und  $\sigma \in H_+^{-1}(\Omega)$  aus und berechnen

$$\begin{aligned}
& 2(\langle \sigma_{\ell+1}, U_\ell - U_{\ell+1} \rangle_{\mathcal{M}(\Omega)} - \langle \sigma, U_\ell - U_{\ell+1} \rangle) \\
&= 2\langle \sigma_{\ell+1}, U_\ell \rangle_{\mathcal{M}(\Omega)} - 2\langle \sigma_{\ell+1}, U_{\ell+1} \rangle_{\mathcal{M}(\Omega)} - 2\langle \sigma, U_\ell \rangle + 2\langle \sigma, U_{\ell+1} \rangle \\
&= 2\langle \sigma_{\ell+1}, U_\ell \rangle_{\mathcal{M}(\Omega)} - 2\langle \sigma_{\ell+1}, U_{\ell+1} \rangle_{\mathcal{M}(\Omega)} - 2\langle \sigma, U_\ell \rangle + 2\langle \sigma, U_{\ell+1} \rangle \\
&\quad + 2\langle \sigma_{\ell+1}, \chi_\ell \rangle_{\mathcal{M}(\Omega)} - 2\langle \sigma_{\ell+1}, \chi_\ell \rangle_{\mathcal{M}(\Omega)} + 2\langle \sigma, \chi_\ell \rangle - 2\langle \sigma, \chi_\ell \rangle \\
&\quad + 2\langle \sigma_{\ell+1}, \chi_{\ell+1} \rangle_{\mathcal{M}(\Omega)} - 2\langle \sigma_{\ell+1}, \chi_{\ell+1} \rangle_{\mathcal{M}(\Omega)} + 2\langle \sigma, \chi_{\ell+1} \rangle - 2\langle \sigma, \chi_{\ell+1} \rangle \\
&= \underbrace{2\langle \sigma_{\ell+1}, U_\ell - \chi_\ell \rangle_{\mathcal{M}(\Omega)}}_{\leq 0} + 2\langle \sigma_{\ell+1} - \sigma, \chi_\ell - \chi_{\ell+1} \rangle_{\mathcal{M}(\Omega)} \\
&\quad + \underbrace{2\langle \sigma_{\ell+1}, \chi_{\ell+1} - U_{\ell+1} \rangle_{\mathcal{M}(\Omega)}}_{=0} + \underbrace{2\langle \sigma, \chi_\ell - U_\ell \rangle}_{=con_\ell^{ort}} - \underbrace{\langle \sigma, \chi_{\ell+1} - U_{\ell+1} \rangle}_{\leq 0} \\
&\leq 2\langle \sigma_{\ell+1} - \sigma, \chi_\ell - \chi_{\ell+1} \rangle_{\mathcal{M}(\Omega)} + con_\ell^{ort}. \tag{4.3.35}
\end{aligned}$$

Wir setzen erneut  $\delta_{\chi_\ell} := \chi_\ell - \chi_{\ell+1}$ . Laut [BCH2, (5.29)] gilt nun mit (4.3.4)

$$\begin{aligned}
|2\langle \sigma_{\ell+1} - \sigma, \delta_{\chi_\ell} \rangle_{\mathcal{M}(\Omega)}| &= |2\langle \sigma_{\ell+1}, \delta_{\chi_\ell} \rangle_{\mathcal{M}(\Omega)} - 2\langle \sigma, \delta_{\chi_\ell} \rangle| \\
&= | -2\langle U_{\ell+1}, \delta_{\chi_\ell} \rangle + 2(\partial_n U_\ell, \delta_{\chi_\ell})_\Gamma + 2\langle u, \delta_{\chi_\ell} \rangle - 2(\partial_n u, \delta_{\chi_\ell})_\Gamma | \\
&\leq |2\langle u - U_{\ell+1}, \delta_{\chi_\ell} \rangle| + |2(\partial_n(u - U_{\ell+1}), \delta_{\chi_\ell})_\Gamma| \\
&\leq 4\varepsilon \|u - U_{\ell+1}\|^2 + \frac{1}{\varepsilon}(1 + \rho_2)osc_\ell^2(\chi), \tag{4.3.36}
\end{aligned}$$

wobei die letzte Ungleichheit durch Anwendung der Spur-Ungleichung in Verbindung mit der Young'schen Ungleichung und der Abklingbedingung (4.2.9) entstanden ist.

*An dieser Stelle wurde jedoch wieder die fragwürdige Spurungleichung verwendet. Zusätzlich ist außerdem unklar, wie der Term  $osc_\ell(\chi)^2$  aus dem Randterm  $\|\delta_\chi\|_\Gamma$  entsteht.*

Einsetzen von (4.3.34) – (4.3.36) in (4.3.33) vollendet den Beweis.  $\square$

Nun sind alle Vorbereitungen getroffen um das Konvergenzresultat [BCH2, Theorem 3.1] zu zeigen, laut dem der Fehler selbst, bis auf Oszillationsterme und Konsistenzfehler, eine kontraktive Größe darstellt. An dieser Stelle möchten wir erneut darauf aufmerksam machen, dass die Gültigkeit des nachstehenden Satzes nur unter der Prämisse der Gültigkeit der vorangegangenen Aussagen gegeben ist.

**Satz 4.3.6** (Fehlerreduktion). *Seien  $u \in K$  die exakte und  $U_\ell, U_{\ell+1} \in K_\ell, K_{\ell+1}$  die diskreten Lösungen der Probleme 4.1.1 bzw. 4.1.2. Die Datenterme und der Konsistenzfehler  $osc_\ell, \mu_\ell$  und  $con_\ell$  seien wie oben definiert. Außerdem seien die Abklingbedingungen (4.2.9) gegeben. Dann existieren die Konstanten  $\rho_1 \in (0, 1)$  und  $C_1, C_2, C_3 > 0$  die nur von  $\theta$  und der Geometrie der Triangulierung abhängen, so dass*

$$\begin{pmatrix} \|u - U_{\ell+1}\|^2 \\ osc_{\ell+1}^2 \\ \mu_{\ell+1}^2 \end{pmatrix} \leq \begin{pmatrix} \rho_1 & C_1 & C_2 \\ 0 & \rho_2 & 0 \\ 0 & 0 & \rho_3 \end{pmatrix} \begin{pmatrix} \|u - U_\ell\|^2 \\ osc_\ell^2 \\ \mu_\ell^2 \end{pmatrix} + \begin{pmatrix} C_3 con_\ell \\ 0 \\ 0 \end{pmatrix} \tag{4.3.37}$$

*gilt.*

*Beweis.* Die Zuverlässigkeit (4.3.1) liefert in Verbindung mit der diskreten Effizienz (4.3.31) die Existenz einer Konstanten  $C > 0$ , welche nur von  $\theta$  und der Formregularität abhängt, so dass

$$\|u - U_\ell\|^2 \leq C \left( \|U_\ell - U_{\ell+1}\|^2 + \text{osc}_\ell^2 + \text{con}_\ell^{\text{rel}} + \mu_\ell^2 \right)$$

gilt. Die gestörte Galerkin-Orthogonalität (4.3.32) liefert nun

$$\begin{aligned} \|u - U_\ell\|^2 &\leq C \left( \|u - U_\ell\|^2 - (1 - 4\varepsilon) \|u - U_{\ell+1}\|^2 \right. \\ &\quad \left. + \frac{1}{\varepsilon} (1 + \rho_2) \text{osc}_\ell^2(\chi) + \text{con}_\ell^{\text{ort}} + \text{osc}_\ell^2 + \text{con}_\ell^{\text{rel}} + \mu_\ell^2 \right) \end{aligned}$$

und damit

$$\|u - U_{\ell+1}\|^2 \leq \frac{C - 1}{C(1 - 4\varepsilon)} \|u - U_\ell\|^2 + CC_\varepsilon (\text{osc}_\ell^2 + \text{con}_\ell + \mu_\ell^2),$$

mit  $C_\varepsilon := \max(1/C(1 - 4\varepsilon), (1 + \rho_2)/\varepsilon(C(1 - 4\varepsilon)))$ . Damit folgt die gesuchte Aussage für  $\varepsilon < \frac{1}{4C}$ .  $\square$

**Bemerkung.** Analog zum affinen Fall folgt hieraus die Konvergenz der Folge  $(U_\ell)_{\ell \in \mathbb{N}}$  gegen die exakte Lösung  $u \in K$ , sofern der Konsistenzfehler  $\text{con}_\ell$  gegen Null geht.

**Bemerkung.** Für die Gültigkeit der vorangegangenen Aussagen müsste man noch die in den Beweisen derselben markierten Problemstellen beheben. Da die Gültigkeit der Resultate jedoch zu erwarten ist, ist es sinnvoll, sich weiterführende Gedanken zu machen, was im nächsten Abschnitt geschieht.

## 4.4. Schätzerreduktion

In diesem Abschnitt wollen wir, wie in Kapitel 3, das Prinzip der Schätzerreduktion auf unser Hindernisproblem anwenden. Wie zuvor werden unsere Konvergenzresultate dadurch von der Verfeinerungsstrategie weitestgehend unabhängig. Außerdem werden wir die (diskrete) lokale Effizienz des Schätzers vermeiden, so dass wir die Kantensprünge im Hindernis (4.2.8) nicht weiter betrachten müssen. Diese Einschränkung fällt also einfach restlos weg. Zudem können wir die Datenoszillationsterme vernachlässigen, da diese, durch eine Veränderung am Schätzer, nicht mehr benötigt werden. Da die Hindernisoszillationen  $\text{osc}_\ell(\chi)$  (4.2.4) unabhängig von jedwedem Flächenterm, wie z.B.  $|T|$  oder  $h_E$  sind, können wir die Ideen der Schätzerreduktion hierauf nicht anwenden. Das Abklingen der Hindernisoszillationen

$$\text{osc}_{\ell+1}^2(\chi) \leq \rho_\chi \text{osc}_\ell^2(\chi), \quad (4.4.1)$$

mit  $\rho_\chi \in (0, 1)$  müssen wir also weiterhin ausdrücklich voraussetzen. Entgegen (4.2.1) betrachten wir im Folgenden den Fehlerschätzer

$$\eta_\ell := \left( \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} \eta_\ell(T)^2 + \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell)} \eta_\ell(E)^2 \right)^{1/2}, \quad (4.4.2)$$

wobei wir die einzelnen Beiträge durch

$$\begin{aligned}\eta_\ell(T)^2 &:= |T| \|f\|_T^2, \\ \eta_\ell(E)^2 &:= h_E \|[\partial_n U_\ell]\|_E^2,\end{aligned}$$

definieren. Unser adaptiver FEM-Algorithmus lautet nun:

**Algorithmus 4.4.1** (adaptiver Algorithmus, Version 2). **Input:** Anfangstriangulierung  $\mathcal{T}_0$ , maximale Anzahl Elemente  $n_{E_{max}}$ , Adaptivitätsparameter  $\theta \in (0, 1)$ ,  $\ell := 0$ .

- (i) Berechne diskrete Lösung  $U_\ell \in K_\ell$ .
- (ii) Berechne die Verfeinerungsindikatoren  $\eta_\ell(E), \eta_\ell(T)$  und den Fehlerschätzer  $\eta_\ell$ .
- (iii) STOP, wenn die Anzahl der Elemente größer als  $n_{E_{max}}$  ist.
- (iv) Finde eine Menge  $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell(\mathcal{I}_\ell) \cup \mathcal{E}_\ell(\mathcal{I}_\ell)$  minimaler Kardinalität von markierten Elementen und Kanten, so dass

$$\theta \eta_\ell^2 \leq \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell) \cap \mathcal{M}_\ell} \eta_\ell(T)^2 + \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell) \cap \mathcal{M}_\ell} \eta_\ell(E)^2. \quad (4.4.3)$$

- (v) Erzeuge neues, reguläres Gitter  $\mathcal{T}_{\ell+1}$ , bei dem mindestens alle markierten Elemente und Kanten verfeinert wurden durch beliebige Verfeinerungsstrategie, z.B. NVB.
- (vi)  $\ell \mapsto \ell + 1$  und beginne wieder bei (i).

**Output:** Endliche Folge diskreter Lösungen  $U_\ell$  und zugehörige Fehlerschätzer  $\eta_\ell$ .

Wir wollen im Folgenden zeigen, dass die Ausgabe  $(U_\ell)_{\ell \in \mathbb{N}}$  von Algorithmus 4.4.1 tatsächlich eine gegen die exakte Lösung  $u \in K$  konvergente Folge ist. Hierzu machen wir uns zunächst klar, dass die Zuverlässigkeit von  $\eta_\ell$  direkt aus dem vorangegangenen Zuverlässigkeitsresultat (Satz 4.3.1) folgt (und damit natürlich an dessen Gültigkeit gebunden ist).

**Satz 4.4.2** (Zuverlässigkeit). *Der Fehlerschätzer  $\eta_\ell$  (4.4.2) ist zuverlässig, d.h.*

$$\|u - U_\ell\|^2 \leq C \eta_\ell^2 + \text{osc}_\ell^2(\chi) + \text{con}_\ell^{\text{rel}}, \quad (4.4.4)$$

wobei  $\text{con}_\ell^{\text{rel}}$  wie in (4.2.11) definiert ist und die Konstante  $C > 0$  nur von der Formregularität abhängt.

*Beweis.* Der Beweis läuft völlig analog zum Beweis von Satz 4.3.1, wobei wir jedoch an keiner Stelle  $(f, v)_T$  durch  $(f_T, v)_T + (f - f_T, v)_T$  ersetzen. Dadurch verhindern wir das Entstehen der Daten-Oszillationsterme und erhalten Abschätzungen die sich direkt auf unseren neuen Verfeinerungsindikator  $\eta_\ell(T)$  beziehen.  $\square$

Wir wollen im Folgenden zunächst annehmen, dass durch Verfeinerung innerhalb der aktiven Menge keine neuen inaktiven Knoten entstehen können, d.h.

$$\bigcup_{\mathcal{T}_{\ell+1}(\mathcal{I}_{\ell+1})} T \subseteq \bigcup_{\mathcal{T}_\ell(\mathcal{I}_\ell)} T, \quad \bigcup_{\mathcal{E}_{\ell+1}(\mathcal{I}_{\ell+1})} E \subseteq \bigcup_{\mathcal{E}_\ell(\mathcal{I}_\ell)} E \cup \mathcal{E}_{\text{neu}}, \quad (4.4.5)$$

wobei die Menge  $\mathcal{E}_{neu}$  diejenigen Kanten  $E \in \mathcal{E}_{\ell+1}(\mathcal{I}_{\ell+1})$  darstellt, die im Inneren eines Elements  $T \in \mathcal{T}_{\ell}(\mathcal{I}_{\ell})$  liegen. Wir nehmen also an, dass sich die inaktive Menge durch Verfeinerung höchstens verkleinern, aber niemals vergrößern kann. Dass dies keine völlig aus der Luft gegriffene Forderung ist, kann man sich, zumindest im Fall stückweise affiner Hindernisse folgendermaßen klarmachen: In diesem Fall verschwinden die Hindernissoszillationen (4.2.4), so dass Verfeinerung markierter Elemente nurmehr in der inaktiven Menge möglich ist. Die Forderung ist nun gleichbedeutend damit zu verlangen, dass durch Verfeinerung im inaktiven Bereich, im aktiven Bereich keine „Schwingungen“ entstehen, so dass sich die diskrete Lösung hier wieder vom Hindernis entfernt. Mit der Tatsache, dass die diskrete Lösung ein Energiefunktional minimiert, scheint dies eine sinnvolle Annahme zu sein.

**Proposition 4.4.3** (Eingeschränkte Schätzerreduktion). *Die Menge  $\mathcal{M}_{\ell}$  erfülle die Dörflermarkierung (4.4.3). Außerdem sei Voraussetzung (4.4.5) gegeben. Bei Verfeinerung durch NVB oder RGB gilt dann*

$$\eta_{\ell+1}^2 \leq q \eta_{\ell}^2 + C \|U_{\ell+1} - U_{\ell}\|, \quad (4.4.6)$$

wobei  $q \in (0, 1)$  nur von  $\theta$  und  $C > 0$  zusätzlich von der Geometrie der Elemente in  $\mathcal{T}_0$  abhängt.

*Beweis.* Wir definieren

$$\mathcal{T}_{\ell+1}(\mathcal{I}_{alt}) := \{T \in \mathcal{T}_{\ell+1} \mid \exists T' \in \mathcal{T}_{\ell}(\mathcal{I}_{\ell}), \quad T \subseteq T'\}$$

und

$$\mathcal{E}_{\ell+1}(\mathcal{I}_{alt}) := \{E \in \mathcal{E}_{\ell+1} \mid \exists E' \in \mathcal{E}_{\ell}(\mathcal{I}_{\ell}), \quad E \subseteq E'\}$$

als die Mengen der Kanten und Elemente in der  $(\ell + 1)$ -ten Iteration, die aus der aktiven Menge im vorherigen Schritt hervorgegangen sind. Aufgrund der Annahme (4.4.5) gilt

$$\mathcal{T}_{\ell+1}(\mathcal{I}_{\ell+1}) \subseteq \mathcal{T}_{\ell+1}(\mathcal{I}_{alt}), \quad \text{und} \quad \mathcal{E}_{\ell+1}(\mathcal{I}_{\ell+1}) \subseteq \mathcal{E}_{\ell+1}(\mathcal{I}_{alt}) \cup \mathcal{E}_{neu}$$

Wir setzen  $\overline{\mathcal{M}}_{\mathcal{T}, \ell} := \{T \in \mathcal{T}_{\ell+1}(\mathcal{I}_{alt}) \mid \exists T' \in \mathcal{T}_{\ell}(\mathcal{I}_{\ell}) \cap \mathcal{M}_{\ell}, \quad T \subseteq T'\}$  und erhalten mit Standardargumenten der Schätzerreduktion

$$\begin{aligned} \sum_{T' \in \overline{\mathcal{M}}_{\mathcal{T}, \ell}} \eta_{\ell}(T)^2 &\leq \sum_{T \in \mathcal{T}_{\ell+1}(\mathcal{I}_{alt})} \eta_{\ell}(T)^2 \\ &= \sum_{T' \in \overline{\mathcal{M}}_{\mathcal{T}, \ell}} |T'| \|f\|_{T'}^2 + \sum_{T' \in \overline{\mathcal{M}}_{\mathcal{T}, \ell}} |T'| \|f\|_{T'}^2 \\ &\leq \sum_{T \in \mathcal{T}_{\ell}(\mathcal{I}_{\ell}) \setminus \mathcal{M}_{\ell}} |T| \|f\|_T^2 + \frac{1}{2} \sum_{T \in \mathcal{T}_{\ell}(\mathcal{I}_{\ell}) \cap \mathcal{M}_{\ell}} |T| \|f\|_T^2. \end{aligned}$$

Mit  $\overline{\mathcal{M}}_{\mathcal{E}, \ell} := \{E \in \mathcal{E}_{\ell+1}(\mathcal{I}_{alt}) \mid \exists E' \in \mathcal{E}_{\ell}(\mathcal{I}_{\ell}) \cap \mathcal{M}_{\ell}, \quad E \subseteq E'\}$  und der Tatsache, dass der Sprung  $[\partial_n U_{\ell}]_E$  auf Kanten innerhalb von Elementen  $T \in \mathcal{T}_{\ell}(\mathcal{I}_{\ell})$  (also  $E \in \mathcal{E}_{neu}$ ) verschwindet,

erhalten wir analog

$$\begin{aligned}
 \sum_{E' \in \mathcal{E}_{\ell+1}(\mathcal{I}_{\ell+1})} h_{E'} \|\partial_n U_\ell\|_{E'}^2 &\leq \sum_{E' \in \mathcal{E}_{\ell+1}(\mathcal{I}_{\text{alt}})} h_{E'} \|\partial_n U_\ell\|_{E'}^2 \\
 &= \sum_{E' \in \mathcal{E}_{\ell+1}(\mathcal{I}_{\text{alt}}) \setminus \overline{\mathcal{M}}_{\mathcal{E}, \ell}} h_{E'} \|\partial_n U_\ell\|_{E'}^2 + \sum_{E' \in \overline{\mathcal{M}}_{\mathcal{E}, \ell}} h_{E'} \|\partial_n U_\ell\|_{E'}^2 \\
 &\leq \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell) \setminus \mathcal{M}_\ell} h_E \|\partial_n U_\ell\|_E^2 + \frac{1}{2} \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell) \cap \mathcal{M}_\ell} h_E \|\partial_n U_\ell\|_E^2.
 \end{aligned}$$

Die Dörflermarkierung und die Dreiecksungleichung im Folgenraum  $\ell^2$  liefern nun, wie gewohnt, die gesuchte Aussage.  $\square$

Da die Folge  $(U_\ell)_{\ell \in \mathbb{N}}$  aufgrund der Konformität der Diskretisierung a-priori konvergiert (siehe Lemma 3.3.8), erhalten wir, wie zuvor, die Konvergenz gegen die exakte Lösung  $u \in K$  bis auf den Konsistenzfehler.

**Satz 4.4.4** (Eingeschränkte Konvergenz). *Unter der Voraussetzung (4.4.5) konvergiert die Folge  $(U_\ell)_{\ell \in \mathbb{N}}$  der von Algorithmus 4.4.1 erzeugten diskreten Teillösungen gegen die exakte Lösung  $u \in K$ , d.h.  $\lim_{\ell \rightarrow \infty} \|u - U_\ell\|^2 = 0$ , sofern der Konsistenzfehler  $con_\ell^{rel}$  gegen Null geht.*

*Beweis.* Die Aussage folgt aus der Schätzerreduktion (Proposition 4.4.3) in Verbindung mit der Zuverlässigkeit (Satz 4.4.2), der a-priori Konvergenz konformer Diskretisierungen (Lemma 3.3.8) und dem gleichmäßigen Abklingen der Hindernisoszillationen (4.4.1).  $\square$

**Bemerkung.** *Im Fall stückweise affiner Hindernisse, ist auch folgendes, alternatives Vorgehen möglich, bei dem man die Bedingung an das monotone Abnehmen der inaktiven Menge nicht benötigt: Wir definieren den Fehlerschätzer durch*

$$\eta_\ell^2 := \sum_{T \in \mathcal{T}_\ell} \eta_\ell(T)^2 + \sum_{E \in \mathcal{E}_\ell} \eta_\ell(E)^2$$

auf dem gesamten Gebiet  $\Omega$ , wobei wir den Sprung  $[\partial_n U_\ell]_E$  gleich Null setzen, wenn  $[\partial_n U_\ell]_E = [\partial_n \chi_\ell]_E = [\partial_n \chi]_E$  für ein  $E \in \mathcal{E}_\ell(\mathcal{A}_\ell)$  gilt. Auf diese Weise umgehen wir das zu Beginn dieses Kapitels angesprochene Problem, dass wir immer weiter verfeinern, obwohl der Fehler schon beliebig klein ist und der Sprung nur durch das Hindernis gegeben ist. Der Beweis der Schätzerreduktion folgt nun für diejenigen Kanten, auf denen  $[\partial_n U_\ell]$  durch den echten Sprung gegeben ist völlig analog. Für diejenigen Kanten auf denen der Sprung von  $U_\ell$  gleich dem Sprung des Hindernisses  $\chi$ , und damit per Definition  $[\partial_n U_\ell] = 0$  gilt, macht man sich folgende Überlegung zu nutze: Seien  $E', E'' \in \mathcal{E}_{\ell+1}$  die beiden Söhne der Kante  $E \in \mathcal{E}_\ell$ . Der Sprung  $[\partial_n U_\ell]$  ist konstant auf der gesamten Kante  $E$ . Daher gilt  $[\partial_n U_\ell]_{E'} = [\partial_n U_\ell]_{E''} = [\partial_n U_\ell]_E$ . Gilt nun o.B.d.A.  $[\partial_n U_\ell]_{E'} = 0$  wegen obiger Definition, so muss  $U_\ell$  auch schon im  $\ell$ -ten Schritt an der entsprechenden Kante das Hindernis berührt haben. Es gilt daher  $[\partial_n U_\ell]_E = 0$ , so dass die Schätzerreduktion ihre Gültigkeit behält. Problematisch ist jedoch, dass sich der Term

$$\sum_{E' \in \mathcal{E}_{\ell+1}} h_{E'} \|\partial_n (U_{\ell+1} - U_\ell)\|_{E'}^2$$

mit obiger Definition nichtmehr durch  $\|U_{\ell+1} - U_\ell\|^2$  abschätzen lässt. Hierzu müsste man zusätzlich fordern, dass sich das Vorzeichen des Sprunges  $[\cdot]$  über einer Kante  $E' \in \mathcal{E}_{\ell+1}$  im

Übergang von  $U_\ell$  nach  $U_{\ell+1}$  nicht ändert. Dies stellt aber eine recht einsichtliche Annahme dar.

In den folgenden Ausführungen wollen wir versuchen die Voraussetzung an das monotone Abnehmen der inaktiven Menge zu entkräften. Wir lassen zunächst zu, dass sich die inaktive Menge, aufgrund der Verfeinerung im Hindernis, ein einziges Mal vergrößert. Ohne Beschränkung der Allgemeinheit sei dies im  $(\ell + 1)$ -ten Schritt der Fall. Mit

$$\mathcal{T}_{\ell+1}(\mathcal{I}_{\ell+1}^+) := \{T \in \mathcal{T}_{\ell+1}(\mathcal{I}_{\ell+1}) \mid \forall T' \in \mathcal{T}_\ell(\mathcal{I}_\ell), \quad T \not\subseteq T'\}$$

bezeichnen wir die zur inaktiven Menge neu hinzugekommenen Elemente. Die Menge der neu hinzugekommenen Kanten aus diesem Bereich notieren wir mit

$$\mathcal{E}_{\ell+1}(\mathcal{I}_{\ell+1}^+) := \{E \in \mathcal{E}_{\ell+1}(\mathcal{I}_{\ell+1}) \mid E \in \mathcal{E}_{\ell+1}(T) \text{ mit } T \in \mathcal{T}_{\ell+1}(\mathcal{I}_{\ell+1}^+)\}.$$

Wir erhalten eine Aussage über die Verkleinerung des Schätzers  $\eta_\ell$  im  $(\ell + 1)$ -ten Schritt.

**Proposition 4.4.5.** *Die Menge  $\mathcal{M}_\ell$  erfülle die Dörflermarkierung (4.4.3). Mit*

$$\eta_{\ell+1}^2(\mathcal{I}_{\ell+1}^+) := \sum_{T \in \mathcal{T}_{\ell+1}(\mathcal{I}_{\ell+1}^+)} \eta_\ell(T)^2 + \sum_{E \in \mathcal{E}_{\ell+1}(\mathcal{I}_{\ell+1}^+)} \eta_\ell(E)^2 \quad (4.4.7)$$

gilt im  $(\ell + 1)$ -ten Verfeinerungsschritt

$$\eta_{\ell+1}^2 \leq q \eta_\ell^2 + \eta_{\ell+1}^2(\mathcal{I}_{\ell+1}^+) + C \| \|U_{\ell+1} - U_\ell\|^2. \quad (4.4.8)$$

Die Konstante  $q \in (0, 1)$  hängt hierbei nur von  $\theta$  ab, während  $C > 0$  zusätzlich von der Geometrie der Elemente in  $\mathcal{T}_0$  abhängt.

*Beweis.* Wir betrachten  $\eta_{\ell+1}$  separat auf  $\mathcal{I}_{\ell+1} \setminus \mathcal{I}_{\ell+1}^+$  und  $\mathcal{I}_{\ell+1}^+$ . Nach obiger Definition ist die Menge  $\mathcal{T}_{\ell+1}(\mathcal{I}_{\ell+1} \setminus \mathcal{I}_{\ell+1}^+)$  kleiner als  $\mathcal{T}_\ell(\mathcal{I}_\ell)$ . Da selbiges auch für die entsprechenden Kanten gilt, können wir Proposition 4.4.3 verwenden und erhalten

$$\eta_{\ell+1}^2(\mathcal{I}_{\ell+1} \setminus \mathcal{I}_{\ell+1}^+) \leq q \eta_\ell^2 + C \| \|U_{\ell+1} - U_\ell\|^2.$$

Insgesamt ergibt sich also

$$\eta_{\ell+1}^2 = \eta_{\ell+1}^2(\mathcal{I}_{\ell+1} \setminus \mathcal{I}_{\ell+1}^+) + \eta_{\ell+1}^2(\mathcal{I}_{\ell+1}^+) \leq q \eta_\ell^2 + \eta_{\ell+1}^2(\mathcal{I}_{\ell+1}^+) + C \| \|U_{\ell+1} - U_\ell\|^2.$$

Damit ist die Aussage vollständig bewiesen.  $\square$

**Satz 4.4.6** (Erweiterte Konvergenz I). *Es gelte die Voraussetzung (4.4.5) mit der Ausnahme, dass die inaktive Menge im  $(\ell + 1)$ -ten Schritt auch größer werden darf. Dann konvergiert  $(U_\ell)_{\ell \in \mathbb{N}}$  dennoch gegen die exakte Lösung  $u \in K$ , sofern der Konsistenzfehler  $\text{con}_\ell^{\text{rel}}$  gegen Null geht.*

*Beweis.* In jedem Schritt  $k \neq (\ell + 1)$  gilt mit Proposition 4.4.3

$$\eta_{k+1}^2 \leq q \eta_k^2 + C \| \|U_{k+1} - U_k\|^2.$$

Im  $(\ell + 1)$ -ten Schritt hingegen erhalten wir mit Proposition 4.4.5

$$\eta_{\ell+1}^2 \leq q \eta_\ell^2 + \eta_{\ell+1}^2(\mathcal{I}_{\ell+1}^+) + C \| \|U_{\ell+1} - U_\ell\|^2.$$

Insgesamt ergibt sich somit für den  $(\ell + 1 + k)$ -ten Schritt

$$\begin{aligned} \eta_{\ell+1+k}^2 &\leq q^{k+1} \eta_\ell^2 + q^k \eta_{\ell+1}^2(\mathcal{I}_{\ell+1}^+) + C \|U_{\ell+1+k} - U_{\ell+k}\|^2 \\ &\leq q^{\ell+1+k} \eta_0^2 + q^k \eta_{\ell+1}^2(\mathcal{I}_{\ell+1}^+) + C \sum_{j=1}^{\ell+1+k} q^{\ell+1+k-j} \|U_j - U_{j-1}\|^2 + C \|U_{\ell+1+k} - U_{\ell+k}\|^2, \end{aligned}$$

woraus die Konvergenz

$$\lim_{k \rightarrow \infty} \eta_k^2 = \lim_{k \rightarrow \infty} \eta_{\ell+1+k}^2 = 0$$

folgt. Die Zuverlässigkeit (4.4.4) vollendet den Beweis und zeigt die Konvergenz bis auf den Konsistenzfehler.  $\square$

Diese Ideen lassen sich logischerweise auch auf den Fall übertragen, in dem Voraussetzung (4.4.5) in endlich vielen Schritten verletzt wird. Wir lassen also zu, dass sich die inaktive Menge endlich oft vergrößern kann und erhalten:

**Satz 4.4.7** (Erweiterte Konvergenz II). *Es gelte die Voraussetzung (4.4.5) mit endlich vielen Ausnahmen, d.h. die Voraussetzung darf in endlich vielen Schritten verletzt sein. Dann konvergiert  $(U_\ell)_{\ell \in \mathbb{N}}$  dennoch gegen die exakte Lösung  $u \in K$ , sofern der Konsistenzfehler  $con_\ell^{rel}$  gegen Null geht.*

*Beweis.* Da Voraussetzung (4.4.5) nur in endlich vielen Schritten verletzt sein kann, gibt es einen Index  $\ell_0 > 0$ , so dass die Voraussetzung in allen Schritten  $\ell > \ell_0$  gilt. Ab dem  $\ell$ -ten Schritt gilt also die normale Schätzerreduktion. Zusätzlich erhalten wir eine endliche Summe von Schätzern auf Teilmengen von  $\Omega$  die, mit dem Faktor  $q^k$  multipliziert gegen Null konvergieren (siehe Beweis von Satz 4.4.6). Es ergibt sich somit erneut  $\lim_{\ell \rightarrow \infty} \eta_\ell^2 = 0$  und damit die gesuchte Aussage.  $\square$

Schließlich wollen wir noch den Fall betrachten, in dem im Inneren der aktiven Menge unendlich oft neue inaktive Teilmengen entstehen, die zur inaktiven Menge hinzukommenden Mengen allerdings immer kleiner werden.

**Satz 4.4.8** (Erweiterte Konvergenz III). *Voraussetzung (4.4.5) darf in unendlich vielen Schritten verletzt sein; Wir fordern jedoch, dass die neu hinzukommenden inaktiven Teilmengen ab einem gewissen Index  $\ell_0 \in \mathbb{N}$  immer kleiner werden, d.h. sowohl die Größe der hinzukommenden Dreiecke, als auch deren Anzahl nimmt ab. Ferner sei der Sprung  $[\partial_n U_\ell]_E$  mit  $E \in \mathcal{E}_\ell$  für alle  $\ell \in \mathbb{N}$  gleichmäßig beschränkt. Dann konvergiert  $(U_\ell)_{\ell \in \mathbb{N}}$  dennoch gegen die exakte Lösung  $u \in K$ , sofern der Konsistenzfehler  $con_\ell^{rel}$  gegen Null geht.*

*Beweis.* Sei  $\ell > \ell_0$  und  $\varepsilon > 0$  beliebig. Nach Voraussetzung ist die hinzukommende inaktive Menge im  $(\ell + 1)$ -ten Schritt kleiner, als die im  $\ell$ -ten Schritt hinzukommende aktive Menge. Aufgrund der gleichmäßigen Beschränktheit von  $[\partial_n U_\ell]$  gibt es somit einen Index  $\ell_1$  so, dass  $\eta_\ell^2(\mathcal{I}_\ell^+) \leq \varepsilon$  für alle  $\ell > \ell_1$ . Die ab diesem Index hinzukommenden inaktiven Mengen haben somit keinen weiteren Einfluss auf die Konvergenz von  $\eta_\ell$ , da die Werte der Fehlerschätzer hierauf gegen Null konvergieren. Es gibt allerdings nur endlich viele Schritte  $k < \ell_1$ , so dass hier wieder Satz 4.4.7 Anwendung findet. Insgesamt erhalten wir wieder die Konvergenz des Schätzers  $\eta_\ell$  gegen Null und damit die Konvergenz von  $U_\ell$  gegen die exakte Lösung  $u \in K$ .  $\square$

**Bemerkung.** *Eigentlich bliebe hier noch der Fall zu untersuchen, in dem unendlich oft neue inaktive Mengen hinzukommen können, diese jedoch nicht kleiner werden müssen. Das wäre gleichbedeutend mit der Annahme, dass Knoten und ganze Teilbereiche von  $\Omega$  beliebig oft zwischen aktiv und inaktiv hin- und herspringen können. Da die hinzukommenden Mengen  $\eta_\ell(\mathcal{I}_+)$  hierbei keine Nullfolge mehr sein müssen, können wir mit unseren Methoden keine Konvergenz des Schätzers  $\eta_\ell$  gegen Null zeigen. Das Prinzip der Schätzerreduktion versagt daher an dieser Stelle. Die Frage, ob es überhaupt Hindernisse gibt, die ein solches Verhalten hervorrufen, bleibt allerdings offen.*

Bevor wir diesen Abschnitt mit einem numerischen Beispiel beschließen, wollen wir den Unterschied zwischen affinen und nicht-affinen Hindernissen im Bezug auf das Prinzip der Schätzerreduktion kurz Revue passieren lassen. Das Hauptproblem im Falle nicht-affiner Hindernisse liegt darin, dass es nicht sinnvoll sein kann die Kantensprünge der Normalenableitung  $[\cdot]$  im inneren der aktiven Menge als Verfeinerungsindikator heranzuziehen. Dies führt dazu, dass die hier vorgestellten Fehlerschätzer einzig auf der inaktiven Menge arbeiten. Diese kann sich jedoch – zumindest formal – in jedem Schritt ändern. Da das fundamentale Prinzip der Schätzerreduktion darin besteht den Schätzer im  $\ell$ -ten Schritt mit dem Schätzer im Schritt davor zu vergleichen, führt genau diese Einschränkung zu Problemen. Im Prinzip könnte der Schätzer im  $\ell$ -ten Schritt ja völlig andere Elemente und Kanten als derjenige im  $(\ell - 1)$ -ten Schritt betrachten, so dass diese Größen gänzlich unvergleichbar sind. Um dieses Problem anzugehen haben wir Stück für Stück weniger Einschränkungen an die Veränderung der inaktiven Menge gemacht, konnten jedoch den völlig allgemeinen Fall aus genannten Gründen nicht behandeln. Man kann also recht gut sehen, dass das Prinzip der Schätzerreduktion sehr lokal arbeitet und sich daher auch auf Probleme wie sich veränderbare Gebiete nicht ohne Weiteres anwenden lässt.

## 4.5. Numerische Illustrationen

Wir stellen in diesem Abschnitt ein numerisches Beispiel für ein Hindernisproblem mit nicht-affinem Hindernis vor. Zur Berechnung wurde Algorithmus 4.4.1 implementiert. Wie im affinen Fall haben wir zur Lösung des diskreten Teilproblems in jedem Schritt die *primal-dual active set* Methode aus [HIK] verwendet.

Wir betrachten das Hindernisproblem auf dem quadratischen Gebiet  $\Omega := (-1.5, 1.5)^2 \subset \mathbb{R}^2$ . Die rechte Seite ist durch die Konstante Funktion  $f \equiv 2$  gegeben. Als stückweise affines Hindernis verwenden wir

$$\chi = \chi_\ell = |x| + 0.5 \quad x \in (-1.5, 1.5),$$

wobei  $x$  die erste Koordinate darstellt. Durch die Tatsache, dass in jedem Schritt  $\chi = \chi_\ell$  gilt, fallen die Hindernisoszillationen automatisch weg, so dass diese nicht weiter beachtet werden müssen. Die numerische Approximation  $U_9$  auf  $\mathcal{T}_9$  der Lösung ist in Abbildung 4.2 mit und ohne Hindernis dargestellt. Gut ist zu erkennen wie sich die Funktion an das nicht-affine Hindernis  $\chi$  anschmiegt.

In den numerischen Berechnungen vergleichen wir uniforme und adaptive Netzverfeinerung, wobei wir den Adaptivitätsparameter  $\theta$  aus  $\{0.4, 0.6, 0.8\}$  variieren. Die numerische Konvergenzgeschichte für  $\theta = 0.6$  ist in Tabelle 4.1 dargestellt. Analog zum affinen Fall ist auch hier der Fehler wieder im Energiefunktional durch

$$e_\ell := \mathcal{J}(U_\ell) - \mathcal{J}(u)$$

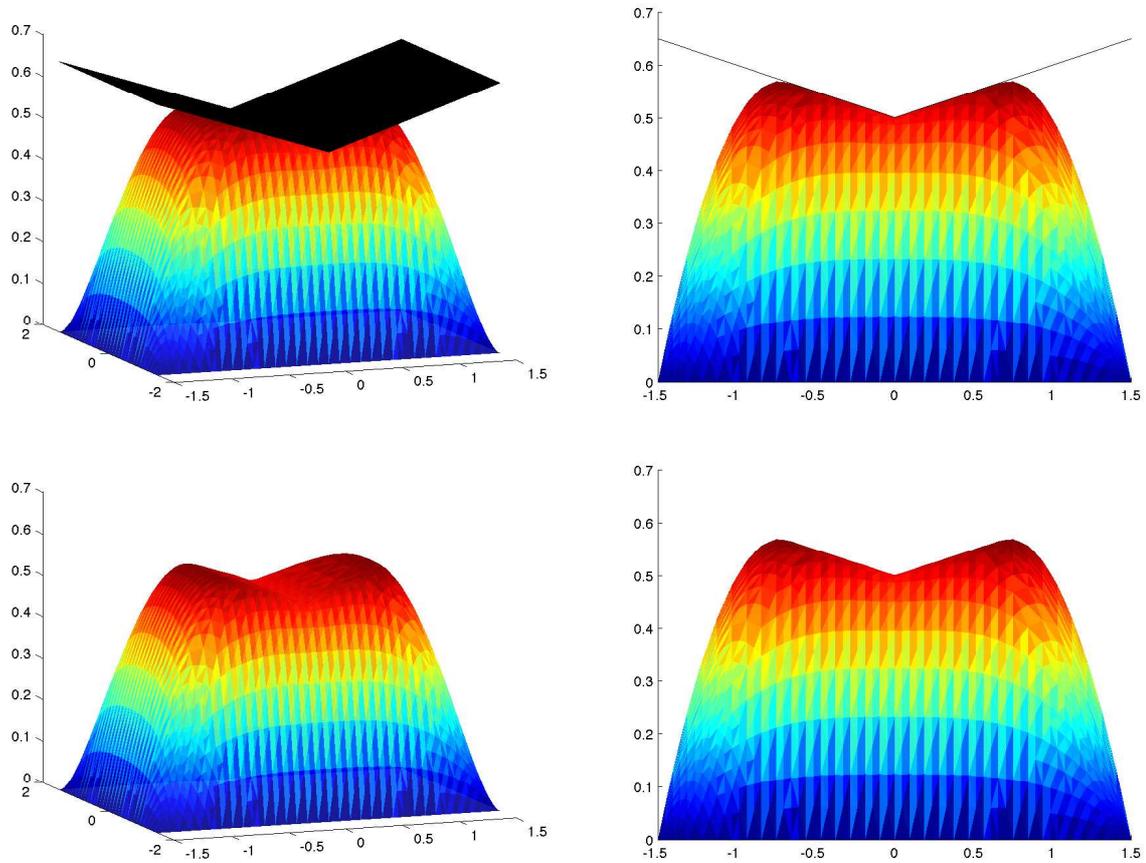


Abbildung 4.2.: Lösung auf adaptiv generiertem Gitter  $\mathcal{T}_9$  mit  $N = 5597$  Elementen und Adaptivitätsparameter  $\theta = 0.6$ . Zur Verdeutlichung ist  $U_9$  mit (oben) und ohne Hindernis (unten) dargestellt.

angegeben. Da wir die exakte Lösung  $u \in K$  nicht kennen, haben wir als Annäherung die numerische Lösung bei uniformer Verfeinerung auf  $N = 524288$  Elementen verwendet. Zur Annäherung der Größe  $\langle u, u \rangle$ , haben wir die Folge  $\langle U_\ell, U_\ell \rangle$  bei uniformer Verfeinerung durch Aitkin's  $\Delta^2$  - Methode extrapoliert. Der Fehlerschätzer  $\eta_\ell$  ist, wie in (4.4.2) beschrieben durch

$$\eta_\ell := \left( \sum_{T \in \mathcal{T}_\ell(\mathcal{I}_\ell)} \eta_\ell(T)^2 + \sum_{E \in \mathcal{E}_\ell(\mathcal{I}_\ell)} \eta_\ell(E)^2 \right)^{1/2},$$

gegeben.

Abbildung 4.3 visualisiert einen Vergleich zwischen adaptivem und uniformem Verfahren für verschiedene Adaptivitätsparameter  $\theta \in \{0.4, 0.6, 0.8\}$ . Dargestellt sind die entsprechenden Fehlerschätzer und der Fehler im Energiefunktional über der Anzahl der Elemente. Aufgrund der hohen Regularität der Lösung ist der Vorteil des adaptiven Verfahrens nur sehr gering. Die Schwankungen im Fehler bei hohen Elementanzahlen sollten sich hierbei noch verbessern, wenn tatsächlich die exakte Lösung  $u \in K$ , oder eine bessere Approximation verwendet wird.

---

$\ell$	$N$	$\sqrt{e_\ell}$	$\eta_\ell$
1	32	9.3545e-01	3.3630e+00
2	63	5.7036e-01	2.3324e+00
3	110	4.9157e-01	1.9705e+00
4	228	3.7624e-01	1.4149e+00
5	416	2.3542e-01	1.0468e+00
6	844	1.9036e-01	7.9144e-01
7	1464	1.5075e-01	5.8077e-01
8	3378	9.3115e-02	3.9343e-01
9	5597	6.7496e-02	2.9935e-01
10	13550	4.7581e-02	1.9593e-01
11	22015	3.3552e-02	1.4882e-01
12	53346	3.4838e-02	1.0125e-01
13	82529	2.3953e-02	7.6044e-02
14	205574	2.3165e-02	4.9509e-02

---

Tabelle 4.1.: Numerische Ergebnisse für adaptive Netzverfeinerung mit  $\theta = 0.6$ , wobei  $N = \#\mathcal{T}_\ell$  und  $e_\ell = \mathcal{J}(U_\ell) - \mathcal{J}(u)$  gilt.

In Abbildung 4.4 sind schließlich die mit  $\theta = 0.6$  adaptiv generierten Triangulierungen  $\mathcal{T}_5$  und  $\mathcal{T}_9$  abgebildet. Wie zu erwarten, beschränkt sich die Verfeinerung hauptsächlich auf den äußeren Bereich. Dass auch im Inneren, entlang des Knicks im Hindernis einige Verfeinerung stattfindet, lässt darauf schließen, dass diese Kanten und Elemente erst relativ spät wirklich zur aktiven Menge hinzukommen.

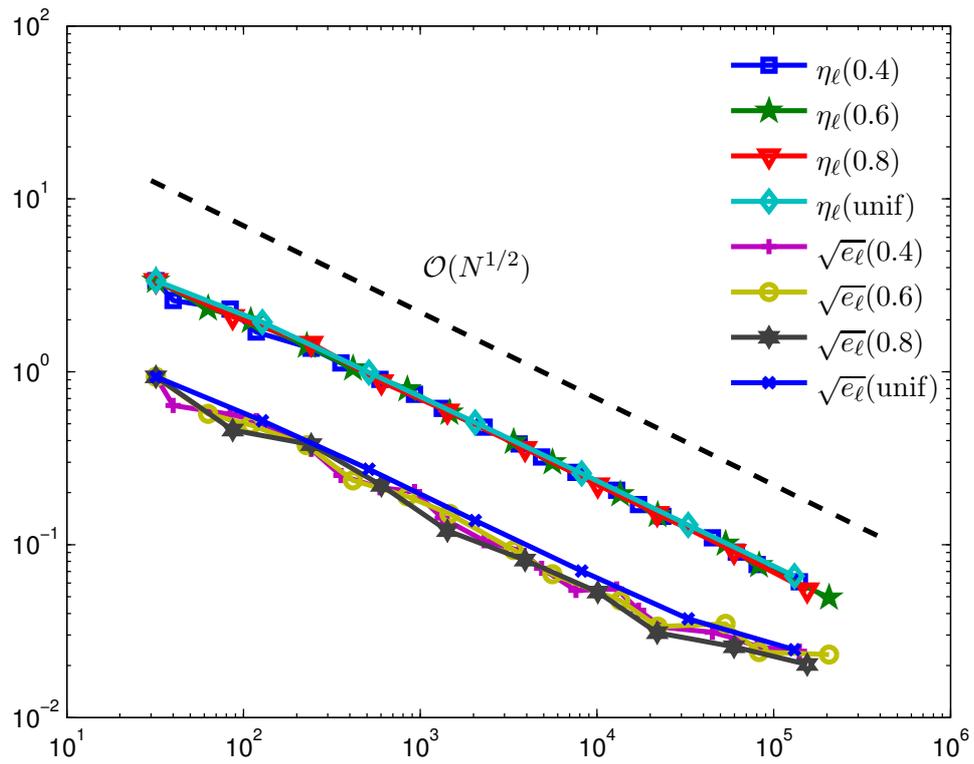


Abbildung 4.3.: Numerische Ergebnisse bei uniformer und adaptiver Verfeinerung wobei  $\theta \in \{0.4, 0.6, 0.8\}$  variiert. Der Fehler  $e_\ell = \mathcal{J}(U_\ell) - \mathcal{J}(u)$  ist im Energiefunktional angegeben. Zu sehen sind die Größen  $\sqrt{e_\ell}$  und  $\eta_\ell$  im Vergleich mit der Anzahl der Elemente  $N = \#\mathcal{T}_\ell$ .

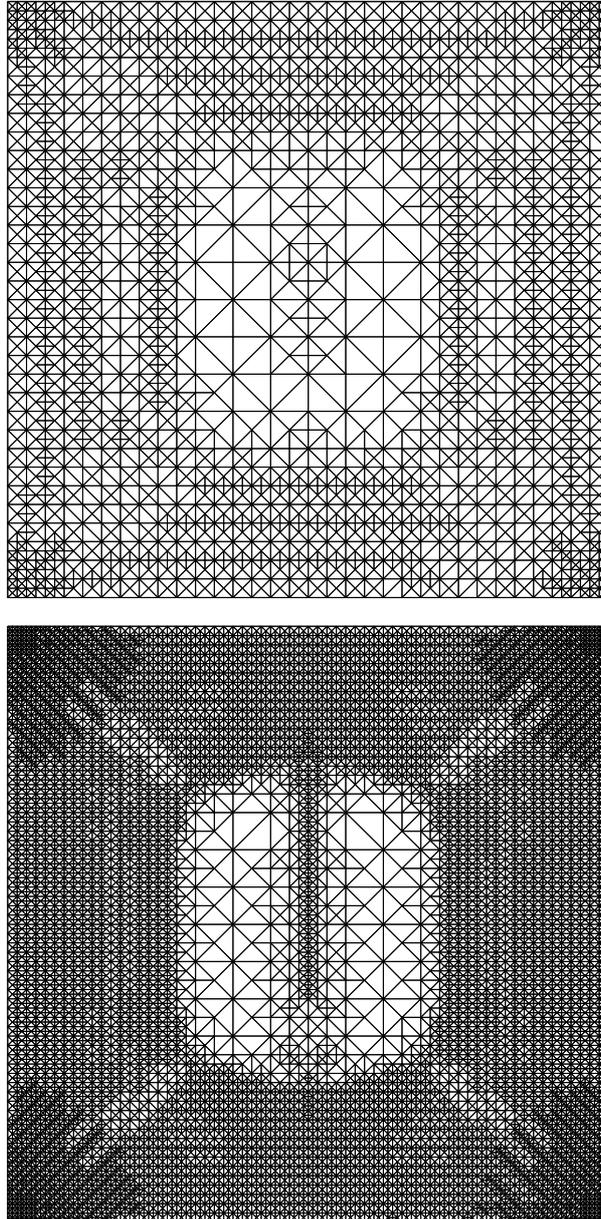


Abbildung 4.4.: Adaptiv generierte Triangulierungen  $\mathcal{T}_5$  (*oben*) und  $\mathcal{T}_8$  (*unten*) mit  $N = 416$  bzw.  $N = 3378$  Elementen für  $\theta = 0.6$ . Verfeinerung ist hauptsächlich auf den äußeren Bereich beschränkt.



## Kapitel 5.

### Zusammenfassung und Ausblick

In dieser Arbeit ist es uns gelungen, konvergente, adaptive finite-Elemente-Algorithmen für *elliptische Hindernisprobleme* zu konstruieren. Hierbei haben wir uns hauptsächlich an den Arbeiten [BCH1] und [BCH2] von BRAESS, CARSTENSEN und HOPPE orientiert. Im Sinne von [BCH1] haben wir zunächst den Fall global affiner Hindernisse betrachtet. Wir haben die Ausführungen von [BCH1] verstanden und nachvollzogen. Im Anschluss daran haben wir das Konzept der *Schätzerreduktion* aus [CKNS, AFP] für Hindernisprobleme verallgemeinert und so einen von der diskreten lokalen Effizienz des Schätzers unabhängigen Konvergenzbeweis erhalten. Dieses Vorgehen wurde in der vorliegenden Arbeit erstmals durchgeführt. Als direkte Konsequenz folgt die Gültigkeit des Konvergenzresultates für weit weniger eingeschränkte Verfeinerungsstrategien. Dies ist auch im Bezug auf Optimalität ein wichtiges Resultat, da ein massiver Bestandteil des Optimalitätsbeweises aus [CKNS] nur für Verfeinerung mit *Newest Vertex Bisection* gilt. Für Untersuchungen in diese Richtung ist es also auch sinnvoll den Hindernisfall mit Verfeinerung durch NVB zu betrachten, was dank unseres Resultates nun legitim ist. Zusätzlich ist es uns außerdem gelungen, die Forderung nach dem Abklingen der Oszillationsterme aus [BCH1] überflüssig zu machen. Dies haben wir erreicht, indem wir den Oszillationsteil direkt mit in den Schätzer aufgenommen und durch Schätzerreduktion gezeigt haben, dass das geforderte Abklingen unter Zuhilfenahme der Dörflermarkierung automatisch eintritt. Es ist daher in diesem Fall obsolet die Oszillationsterme explizit zu kontrollieren. Im Zuge dieses länglichen Beweises wurde somit das Prinzip der Schätzerreduktion auch auf Oszillationsgrößen verallgemeinert. Die größte Schwierigkeit lag hierbei darin, die Oszillationen auf den zusätzlich im Inneren von verfeinerten Elementen entstehenden Kanten zu kontrollieren. Die Gesamtmenge der Kanten wird, im Gegensatz zur Gesamtfläche der Elemente, schließlich in jedem Schritt größer. Zu diesem Zweck wurde der abgewandelte, äquivalente Fehlerschätzer  $\mu_\ell$  eingeführt, welcher zusätzlich Oszillationsterme am Rand beinhaltet, da die Oszillationen von verfeinerten Randkanten ansonsten nicht hätten abgeschätzt werden können. Zum Abschluss des affinen Falles haben wir schließlich die Kontraktivität einer aus Fehler und Fehlerschätzer bestehenden gemischten Fehlergröße gezeigt. Hierdurch erhalten wir nicht nur ein Resultat über die Konvergenzgeschwindigkeit unserer adaptiven Methode, sondern auch einen Ausgangspunkt für einen möglichen Optimalitätsbeweis, da der Beweis aus [CKNS] ein solches Resultat verwendet.

Analog zur Entwicklung von [BCH2] haben auch wir uns im nächsten Schritt allgemeineren (nicht-affinen) Hindernissen zugewandt. Wie zuvor haben wir hierbei zunächst die Resultate aus [BCH2] nachvollzogen und ausgearbeitet. Im Anschluss daran wurden, wieder durch Schätzerreduktion, von der Verfeinerungsstrategie unabhängige Konvergenzresultate erzielt. Das größte Problem hierbei war die Tatsache, dass ein sinnvoller Fehlerschätzer  $\eta_\ell$  im Falle nicht-affiner Hindernisse unmöglich Kantensprünge im inneren der aktiven Menge beinhalten darf. Durch Knicke im Hindernis könnten die Sprünge sonst an eben diesen Stellen groß blei-

ben, obwohl der Fehler schon beliebig klein ist. Dies würde dazu führen, dass an Knicken im Hindernis immer weiter verfeinert wird, obwohl das nichtmehr zur Güte der Lösung beitragen würde, was sicherlich nicht besonders effizient sein kann. Im nicht-affinen Fall ist der Fehlerschätzer also nur auf einem Teilgebiet, der inaktiven Menge, definiert. Die Grundidee der Schätzerreduktion besteht nun darin den Schätzer im  $(\ell + 1)$ -ten Schritt gegen den Schätzer im  $\ell$ -ten Schritt abzuschätzen. Da sich die inaktive Menge allerdings prinzipiell von Schritt zu Schritt ändern kann, mussten wir zusätzliche Einschränkungen an das Entstehen inaktiver Mengen innerhalb vormals aktiver Gebiete treffen. Somit haben wir Probleme und Grenzen des Prinzips der Schätzerreduktion entdeckt, die im linearen Fall nicht abzusehen waren.

Sowohl im affinen als auch im nicht-affinen Fall wurden die theoretischen Ergebnisse schließlich durch numerische Simulationen bestätigt.

Die vorliegende Arbeit konnte zwar einige Fragen aus dem Kontext elliptischer Hindernisprobleme beantworten, hat dafür aber auch neue aufgeworfen. Allen voran steht sicherlich die Frage nach quasi-Optimalität der vorgestellten Algorithmen im Sinne von [CKNS]. Hierzu müsste zusätzlich zur gezeigten Kontraktivität der gemischten Fehlergröße

$$\Delta_\ell = \mathcal{J}(U_\ell) - \mathcal{J}(u) + \gamma\mu_\ell^2$$

die globale Effizienz des involvierten Schätzers, sowie seine *diskrete lokale Zuverlässigkeit* gezeigt werden. Außerdem bräuchte man, aufgrund der verlorenen Galerkin-Orthogonalität, eine Abschätzung der Form

$$\|u - U_\ell\|^2 - \|u - U_{\ell+1}\|^2 \leq \|U_\ell - U_{\ell+1}\|^2,$$

also quasi eine Umkehrung der *gestörten Galerkin-Orthogonalität* aus [BCH2]. Mit diesen Zutaten könnte man die Optimalität der Dörflermarkierung zeigen, und den Beweis analog zu [CKNS] zu Ende führen. Weitere offene Fragen betreffen schließlich Verallgemeinerungen der gezeigten Resultate. So wäre es interessant zu untersuchen, ob man mit den gleichen Argumenten auch einen konvergenten Algorithmus für inhomogene Probleme konstruieren kann. Das numerische Beispiel 1 aus [BCH1] ist ein Indiz dafür, dass eine solche Erweiterung wohl möglich sein sollte. Hierzu wäre sicherlich die Differenz zwischen exakten und diskreten Randdaten zu kontrollieren. Die Erweiterung auf ein gemischtes Problem mit Dirichlet- und Neumann-Randdaten sollte dann relativ direkt möglich sein.

Zu guter letzt stellt sich natürlich noch die Frage, ob es möglich ist die Einschränkungen an das Entstehen inaktiver Mengen im nicht-affinen Fall fallen zu lassen und somit völlig beliebige Hindernisse  $\chi \in H_0^1(\Omega)$  zu betrachten. Dies wäre gleichbedeutend mit der Frage danach, ob der in der Bemerkung am Ende von Abschnitt 4.4 erwähnte Fall überhaupt eintreten kann.

Alles in allem sind wir in ein spannendes Feld eingetaucht bei dem noch viele Probleme mathematisch offen sind. Es stellt sich nur die Frage welches man als nächstes angehen möchte.

# Anhang A.

## Notationen

$\mathbb{N}$	Natürliche Zahlen mit Null, d.h. $0, 1, \dots$
$\Omega$	Gebiet im $\mathbb{R}^2$ auf dem das Problem betrachtet wird.
$\Gamma$	Rand von $\Omega$ .
$C^k(\Omega)$	Raum der $k$ -mal stetig differenzierbaren Funktionen auf $\Omega$ .
$C_c^\infty(\Omega)$	Raum der unendlich oft stetig differenzierbaren Funktionen mit kompaktem Träger auf $\Omega$ .
$L^2(\Omega)$	Sobolevraum der quadratisch integrierbaren Funktionen.
$H_0^1(\Omega)$	Sobolevraum der einmal schwach differenzierbaren $L^2(\Omega)$ -Funktionen, deren Gradient wieder in $L^2(\Omega)$ liegt, mit Nullrand im Spürsinne.
$H^{-1}(\Omega)$	Dualraum von $H_0^1(\Omega)$ .
$H^2(\Omega)$	Sobolevraum der zweimal schwach differenzierbaren $L^2$ -Funktionen auf $(\Omega)$ .
$\mathcal{S}_0^1(\mathcal{T}_\ell)$	Global stetige, stückweise affine Funktionen mit Nullrand.
$\mathcal{P}^m(T)$	Raum der Polynome von Grad kleiner oder gleich $m$ auf $T$ .
$\ \cdot\ $	$L^2$ -Norm auf dem ganzen Gebiet $\Omega$ .
$(\cdot, \cdot)$	$L^2$ -Skalarprodukt auf dem ganzen Gebiet $\Omega$ .
$(\cdot, \cdot)_T$	$L^2$ -Skalarprodukt auf $T \subseteq \Omega$ .
$\ \!\  \cdot \ \!\ $	Energienorm auf $H_0^1(\Omega)$ .
$\langle\!\langle \cdot, \cdot \rangle\!\rangle$	Energieskalarprodukt auf $H_0^1(\Omega)$ .
$u$	Exakte Lösung.
$U_\ell$	Diskrete Teillösung.
$f \in L^2(\Omega)$	Rechte Seite der elliptischen Gleichung.
$T, T', T'', T_\pm, A$	Dreieckselement einer Triangulierung.
$ \cdot $	Fläche (eines Gebietes).
$h_T$	Durchmesser des Elements $T$ .
$\text{diam}(T)$	Durchmesser des Elements $T$ .
$h_E$	Länge der Kante $E$ .
$\rho_T$	Höhe (über der längsten Kante) des Elements $T$ .
$\mathcal{T}(\mathcal{T}_\ell)$	Triangulierung (im $\ell$ -ten Schritt).
$\mathcal{N}(\mathcal{N}_\ell)$	Knoten der Triangulierung $\mathcal{T}(\mathcal{T}_\ell)$ .
$\mathcal{E}^*(\mathcal{E}_\ell^*)$	Gesamtmenge der Kanten von $\mathcal{T}(\mathcal{T}_\ell)$ .
$\mathcal{E}_\Gamma(\mathcal{E}_{\ell, \Gamma})$	Randkanten von $\mathcal{T}(\mathcal{T}_\ell)$ .
$\mathcal{T}_\Gamma(\mathcal{T}_{\ell, \Gamma})$	Randelemente von $\mathcal{T}(\mathcal{T}_\ell)$ .
$\mathcal{T}_{\ell, \Gamma}^*$	Randelemente welche den Rand nur mit einem Knoten berühren.
$\mathcal{E}(\mathcal{E}_\ell)$	Innere Kanten von $\mathcal{T}(\mathcal{T}_\ell)$ .
$\mathcal{N}_T$	Knoten des Elements $T$ .

$\mathcal{N}_E$	Knoten der Kante $E$ .
$\mathcal{E}_T$	Kanten des Elements $T$ .
$x_T, y_T, z_T$	Konkrete Knoten des Elements $T$ .
$\#\mathcal{F}$	Anzahl der Elemente in der Menge $\mathcal{F}$ .
$\eta_z$	Nodale Basisfunktion von Knoten $z \in \mathcal{N}$ .
$\sigma(\mathcal{T})$	Formregularitätskonstante.
$I_h u$	Nodaler Interpolant.
$J_h u$	Clément Operator.
$D^k$	$k$ -te schwache Ableitung in der Multiindex-Notation (siehe EVANS [E, S. 617]).
$\eta(\eta_\ell), \mu(\mu_\ell)$	Fehlerschätzer auf der Triangulierung $\mathcal{T}$ ( $\mathcal{T}_\ell$ ).
$\Omega_z$	Knotenpatch.
$\Omega_E$	Kantenpatch.
$\Omega_T$	Elementpatch.
$\omega_T \subset \Omega_T$	Kleinere Elementpatches.
$f_T$	Integralmittel von $f$ über $T$ .
$\theta, \widehat{\theta}$	Adaptivitätsparameter für die Dörflermarkierung.
$\mathcal{M}_\ell$	Menge der markierten Elemente und Kanten.
$\overline{\mathcal{M}}_{\mathcal{T}, \ell}$	Menge der durch Verfeinerung markierter Elemente entstandenen Elemente $T' \in \mathcal{T}_{\ell+1}$ .
$n_{E_{max}}$	Maximale Anzahl von Elementen als Abbruchkriterium für den adaptiven Algorithmus.
$osc_\ell(T)$	Elementbasierter Oszillationsterm.
$osc_\ell(E)$	Kantenbasierter Oszillationsterm.
$Osc_\ell$	Summe aus Kanten- und Elementoszillationen.
$\eta_\ell(T), \eta_T$	Elementbasierte Verfeinerungsindikatoren.
$\eta_\ell(E)$	Kantenbasierte Verfeinerungsindikatoren.
$\mathbb{A}_\alpha$	Approximationsklasse zur Konvergenzrate $\alpha$ .
$\mathcal{J}(u)$	Energiefunktional, $\mathcal{J}(u) = 1/2 \langle u, u \rangle - (f, u)$ .
$\text{mid}(E)$	Mittelpunkt der Kante $E$ .
$\delta_\ell$	Differenzenfunktion $\delta_\ell := U_\ell - \chi \in H^1(\Omega)$ .
$\langle \cdot, \cdot \rangle_{\mathcal{M}(\Omega)}$	Duale Paarung zwischen $C(\overline{\Omega})$ und $\mathcal{M}(\Omega)$ .
$\delta_p$	Dirac Delta-Funktional bezüglich des Punktes $p$ .
$\omega_\ell^p, \omega_\ell^E, \omega_\ell^T$	Knoten-, Kanten, bzw. Elementpatches in der $\ell$ -ten Triangulierung.
$\mathcal{E}_\ell^p$	Kanten der $\ell$ -ten Triangulierung, welche den Punkt $p$ berühren.
$\mathcal{A}_\ell, \mathcal{I}_\ell$	Aktive bzw. inaktive Menge im $\ell$ -ten Schritt.
$\mathcal{E}_{\mathcal{A}_\ell}^{(i)}, \mathcal{T}_{\mathcal{A}_\ell}^{(i)}$	Menge der Kanten bzw. Elemente mit $i$ Knoten in der aktiven Menge.
$\mathcal{E}_{\mathcal{I}_\ell}^{(i)}, \mathcal{T}_{\mathcal{I}_\ell}^{(i)}$	Menge der Kanten bzw. Elemente mit $i$ Knoten in der inaktiven Menge.
$\mathcal{T}_{\mathcal{F}_\ell}, \mathcal{E}_{\mathcal{F}_\ell}$	Elemente bzw. Kanten die weder komplett in der aktiven, noch komplett in der inaktiven Menge liegen.
$osc_\ell(\chi)$	Oszillationsterme in Abhängigkeit des (nicht-affinen) Hindernisses $\chi$ .
$\mathcal{M}(\Omega)$	Raum der Radonmaße auf $\Omega$ , Dualraum von $C(\overline{\Omega})$ .
$\langle \cdot, \cdot \rangle$	Duale Paarung auf entsprechenden Räumen.
$\delta_\chi$	Differenzenfunktion $\chi - \chi_\ell \in H^1(\Omega)$ .

# Anhang B.

## Mathematische Wahrheiten

### B.1. Funktionalanalytische Voraussetzungen

**Lemma B.1.1** (Ungleichung von Poincaré). Sei  $\Omega \subset \mathbb{R}^d$  ein beschränktes Gebiet im  $\mathbb{R}^d$ , dann gilt für alle  $v \in H_0^1(\Omega)$

$$\|v\|_{L^2(\Omega)} \leq C_{\text{Poincaré}} \|\nabla v\|_{L^2(\Omega)}.$$

Wegen

$$\|\nabla v\|_{L^2(\Omega)} \leq \|v\|_{H_0^1(\Omega)} = \|v\|_{L^2(\Omega)} + \|\nabla v\|_{L^2(\Omega)} \leq (1 + C_{\text{Poincaré}}) \|\nabla v\|_{L^2(\Omega)}$$

ist somit durch  $\|v\| = \|\nabla v\|_{L^2(\Omega)}$  eine äquivalente Norm auf  $H_0^1(\Omega)$  gegeben.

*Beweis.* Siehe BRAESS [BRA07, S. 29] □

**Lemma B.1.2** (Ungleichung von Hölder). Sei  $\Omega \subseteq \mathbb{R}^d$  ein Gebiet im  $\mathbb{R}^d$ . Ferner sei  $1 \leq p, q \leq \infty$  mit  $\frac{1}{p} + \frac{1}{q} = 1$ , dann gilt für  $u \in L^p(\Omega), v \in L^q(\Omega)$

$$\int_{\Omega} |uv| \, dx \leq \|u\|_{L^p(\Omega)} \|v\|_{L^q(\Omega)}.$$

Im Besonderen gilt daher für  $u, v \in L^2(\Omega)$

$$(u, v) \leq \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}.$$

*Beweis.* Siehe EVANS [E, S. 623]. □

**Lemma B.1.3** (Ungleichung von Young/Cauchy). Seien  $a, b > 0, \varepsilon > 0$ , dann gilt

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

*Beweis.* Siehe EVANS [E, S. 622]. □

**Lemma B.1.4** (Verallgemeinerte Céa Ungleichung). Seien  $\mathcal{A}$  und  $\mathcal{A}_\ell$  konvexe, abgeschlossene Teilmengen von  $H_0^1(\Omega)$  mit  $\mathcal{A}_\ell \subseteq \mathcal{A}$ . Ferner seien  $u \in \mathcal{A}$  und  $U_\ell \in \mathcal{A}_\ell$  die eindeutigen Lösungen von

$$\mathcal{J}(u) = \min_{v \in \mathcal{A}} \mathcal{J}(v), \quad \text{bzw.} \quad \mathcal{J}(U_\ell) = \min_{V_\ell \in \mathcal{A}_\ell} \mathcal{J}(V_\ell),$$

dann gilt

$$\|u - U_\ell\|^2 \leq \|u - V_\ell\|^2 - 2\langle u, u - V_\ell \rangle + 2(f, u - V_\ell)$$

für alle  $V_\ell \in \mathcal{A}_\ell$ .

*Beweis.* Wie im theoretischen Teil ausführlich dargelegt, gilt für die  $u \in \mathcal{A}$  die Abschätzung

$$\langle u, v - u \rangle \geq (f, v - u) \Leftrightarrow \langle u, u - v \rangle \leq (f, u - v) \quad \text{für alle } v \in \mathcal{A}.$$

Die Verwendung dieser Variationsungleichung für  $v = U_\ell \in \mathcal{A}_\ell$  und  $V_\ell \in \mathcal{A}_\ell$  liefert

$$\begin{aligned} \|u - U_\ell\|^2 &= \|u\|^2 - 2\langle u, U_\ell \rangle + \|U_\ell\|^2 \\ &\leq (\langle u, U_\ell \rangle + (f, u - U_\ell)) - 2\langle u, U_\ell \rangle + (\langle U_\ell, V_\ell \rangle + (f, U_\ell - V_\ell)) \\ &= -\langle U_\ell, u - V_\ell \rangle + (f, u - V_\ell) \\ &= \langle u - U_\ell, u - V_\ell \rangle - \langle u, u - V_\ell \rangle + (f, u - V_\ell) \\ &\leq (\|u - U_\ell\|^2/2) + (\|u - V_\ell\|^2/2) - \langle u, u - V_\ell \rangle + (f, u - V_\ell), \end{aligned}$$

wobei das letzte Ungleichheitszeichen aus der Anwendung der Ungleichungen von Hölder und Cauchy stammt. Die gesuchte Aussage folgt direkt.  $\square$

**Satz B.1.5** (Riesz'scher Darstellungssatz). *Sei  $\mathcal{H}$  ein Hilbertraum mit innerem Produkt  $(\cdot, \cdot)_\mathcal{H}$ , dann existiert für jedes  $f \in \mathcal{H}'$  ein eindeutiges Element  $u \in \mathcal{H}$  so, dass*

$$(u, v)_\mathcal{H} = \langle f, v \rangle \quad \text{für alle } v \in \mathcal{H},$$

wobei  $\langle \cdot, \cdot \rangle$  die duale Paarung zwischen  $\mathcal{H}$  und  $\mathcal{H}'$  bezeichnet.

*Beweis.* Siehe WERNER [W, S. 197, Theorem V.3.6]  $\square$

**Lemma B.1.6.** *Es seien  $|\cdot|_j, j = \{1, 2\}$  zwei Halbnormen auf einem endlich dimensionalen Raum  $X$ . Für die Kerne  $Y_j := \{x \in X \mid |x|_j = 0\} \subseteq X$  gelte ferner  $Y_1 \subseteq Y_2$ . Dann gilt*

$$|x|_2 \lesssim |x|_1 \quad \text{für alle } x \in X.$$

*Im Fall  $Y_1 = Y_2$  sind die beiden Halbnormen auf  $X$  sogar äquivalent. Die implizit enthaltene Konstante hängt hierbei nur von  $\dim(X)$  ab.*

*Beweis.* Der Beweis gliedert sich in zwei Fälle. Nehmen wir zunächst an, dass durch  $|\cdot|_1$  eine Norm auf  $X$  gegeben ist und betrachten den Quotientenraum  $X/Y_2$ . Dieser ist offensichtlich ebenso endlich dimensional. Ferner sind

$$|x + Y_2|_2 = |x|_2 \quad \text{und} \quad |x + Y_2|_1 \leq |x|_1$$

Normen auf  $X/Y_2$  und wir erhalten durch die Normäquivalenz auf endlich dimensionalen Räumen

$$|x|_2 = |x + Y_2|_2 \lesssim |x + Y_2|_1 \leq |x|_1 \quad \forall x \in X.$$

Für den allgemeineren Fall betrachten wir nun außerdem den Quotientenraum  $X/Y_1$ , welcher ebenso endlich dimensional ist. Hierauf ist durch  $|x + Y_1|_1$  eine Norm gegeben. Die Größe  $|x + Y_1|_2 = |x|_2$  bildet eine Halbnorm. Nun können wir aber auf den ersten Fall zurückgreifen und erhalten somit durch

$$|x|_2 = |x + Y_1|_2 \lesssim |x + Y_1|_1 \leq |x|_1 \quad \forall x \in X$$

die gesuchte Aussage.  $\square$

## B.2. Theorem aus [BC]

Dieser Abschnitt behandelt das Hindernisproblem auf einem beschränkten Lipschitzgebiet  $\Omega \in \mathbb{R}^d$ ,  $d=2,3$  mit rechter Seite  $f \in H^1(\Omega)$ . Zugelassen sind allgemeine Hindernisse  $\chi \in H^1(\Omega)$ , allgemeine Dirichlet Randdaten  $u_D \in H^1(\Gamma_D)$  auf dem Dirichlet Rand  $\Gamma_D$ , sowie allgemeine Neumann Randdaten  $g \in H^1(\Gamma_N)$  auf dem Neumann Rand  $\Gamma_N$ . Wie in der Literatur üblich gilt  $\Gamma_N := \Gamma \setminus \Gamma_D$ . In der  $\ell$ -ten Iteration sei  $\chi_\ell$  jeweils eine global stetige, stückweise affine Approximation von  $\chi$ .

**Definition.** Wir definieren

$$\mathcal{T}_D := \{T \in \mathcal{T}_\ell \mid T \cap \Gamma_D \neq \emptyset\}$$

und die Kontaktzone

$$\mathcal{T}_c := \{T \in \mathcal{T} \setminus \mathcal{T}_D \mid (\chi_\ell - U_\ell)|_T = 0\}.$$

Ferner definieren wir für die freien Knoten  $\mathcal{N}_F := \mathcal{N} \setminus \Gamma_D$

$$\mathcal{T}_i := \{T \in \mathcal{T} \mid T \cap \Gamma_D = \emptyset, \exists x, y \in \mathcal{N}_F \cap \bar{\omega}_T, \chi_\ell(x) = U_\ell(x), \chi_\ell(y) < U_\ell(y)\}$$

und assoziieren mit jedem  $T \in \mathcal{T}_i$  einen Knoten  $z_T \in \mathcal{N}_F \cap \bar{\omega}_T$  so, dass  $\chi_\ell(z_T) = U_\ell(z_T)$ . Wir setzen nun  $\tilde{\Omega}_{z_T} := \Omega_{z_T} \cup \omega_T$  und definieren

$$\mathcal{T}_s := \{K \in \mathcal{T} \mid \exists T \in \mathcal{T}_i, K \subseteq \tilde{\Omega}_{z_T}\}$$

und

$$\Omega_s := \bigcup_{T \in \mathcal{T}_s} T.$$

Hierbei bezeichnet  $\omega_z := \{x \in \Omega \mid 0 < \varphi_z(x)\}$  den Knotenpatch auf  $z \in \mathcal{N}$ . Wir wählen nun für jeden Knoten  $z \in \mathcal{N} \setminus \mathcal{N}_F$  einen Nachbarknoten  $\zeta(z) \in \mathcal{N}_F$  und definieren  $I(z) := \{\tilde{z} \in \mathcal{N} \mid \zeta(\tilde{z}) = z\}$ . Für jeden Knoten  $z \in \mathcal{N}_F$  setzen wir

$$\psi_z := \sum_{\zeta \in I(z)} \varphi_\zeta \quad \text{und} \quad \Omega_z := \{x \in \Omega \mid 0 < \psi_z(x)\}.$$

Es gilt nun die folgende Aussage:

**Satz B.2.1.** *Wir nehmen an, dass für jedes  $T \in \mathcal{T}_i$  mit  $z_T \in \mathcal{N}_F \cap \Gamma_N$  der Schnitt von  $\Omega$  und einem beliebigen offenen Halbraum mit Randpunkt  $z_T$  nicht leer ist. Für jedes  $\omega \in H^1(\Omega)$  mit  $\omega|_{\Gamma_D} = u_D - u_{D,h}$  und  $\chi - U_\ell \leq \omega$ , wobei  $u_{D,h}$  eine stückweise affine Approximation an  $u_D$  ist, gilt nun*

$$\begin{aligned} \|(u - U_\ell) - \omega\| + \|u - U_\ell\| &\lesssim \min_{p_h \in \mathcal{S}^1(\mathcal{T})^d} (\|\nabla U_\ell - p_h\| + \|h_E^{1/2}(g - p_h \cdot n)\|_{L^2(\Gamma_N)}) \\ &+ \min_{q_h \in \mathcal{S}^1(\mathcal{T}_s)^d} \|\nabla(\chi_\ell - U_\ell) - q_h\|_{L^2(\Omega_s)} \\ &+ \left( \sum_{T \in \mathcal{T}_c} \|f\|_{L^2(\omega_T)} \|(\chi - \chi_\ell - \omega)_-\|_{L^2(T)} \right)^{1/2} + \|\omega\| + \|h_T^2 \nabla f\| \\ &+ \left( \sum_{T \in \mathcal{T} \setminus \mathcal{T}_D} \|h_T^{-1}(\chi - \chi_\ell - \omega)_-\|_{L^2(T)} \right)^{1/2} + \left( \sum_{T \in \mathcal{T}_D} h_T^2 \|f\|_{L^2(\omega_T)}^2 \right)^{1/2}. \end{aligned}$$

*Beweis.* Siehe BARTELS und CARSTENSEN [BC, Theorem 3]. □

Offensichtlich gilt  $\mathcal{T}_c \subseteq \mathcal{T}_\ell$ . Weiterhin erhalten wir mit Grundmitteln der Analysis

$$\begin{aligned}
 & \min_{q_h \in \mathcal{S}^1(\mathcal{T}_s)^d} \|\nabla(\chi_\ell - U_\ell) - q_h\|_{L^2(\Omega_s)} \\
 & \leq \min_{q_h \in \mathcal{S}^1(\mathcal{T}_s)^d} \|\nabla(\chi_\ell - U_\ell) - q_h\|_{L^2(\Omega_s)} + \min_{r_h \in \mathcal{S}^1(\mathcal{T}_\ell \setminus \mathcal{T}_s)^d} \|\nabla(\chi_\ell - U_\ell) - r_h\|_{L^2(\Omega \setminus \Omega_s)} \\
 & \leq \min_{q_h \in \mathcal{S}^1(\mathcal{T}_\ell)^d} (\|\nabla(\chi_\ell - U_\ell) - q_h\|_{L^2(\Omega_s)} + \|\nabla(\chi_\ell - U_\ell) - q_h\|_{L^2(\Omega \setminus \Omega_s)}) \\
 & = \min_{q_h \in \mathcal{S}^1(\mathcal{T}_\ell)^d} \|\nabla(\chi_\ell - U_\ell) - q_h\|_{L^2(\Omega)}.
 \end{aligned}$$

Wenn man sich nun den Beweis von Theorem 2 aus [BC] etwas genauer anschaut, so sieht man, dass die Größe  $\langle \sigma, u - U_\ell \rangle^{1/2}$  hier einfach weggelassen wurde. Da der Beweis analog zu obigen Theorem ist, gilt die Aussage also auch weiterhin, wenn wir diese Größe mit einfließen lassen. Außerdem geht aus [BC, Lemma 1] hervor, dass der Term  $\|h_T^2 \nabla f\|$  eigentlich von den Knotenoszillationen abstammt. Die Aussage behält also ihre Gültigkeit auch, wenn wir stattdessen den Term

$$\left( \sum_{z \in \mathcal{N}_\ell} h_z^2 \min_{q \in \mathbb{R}} \|f - q\|_{L^2(\Omega_z)}^2 \right)^{1/2}$$

verwenden. Insgesamt erhalten wir somit unter den gleichen Voraussetzungen wie oben die Abschätzung

$$\begin{aligned}
 & \| (u - U_\ell) - \omega \| + \| u - U_\ell \| + \langle \sigma, u - U_\ell \rangle^{1/2} \lesssim \min_{p_h \in \mathcal{S}^1(\mathcal{T})^d} (\| \nabla U_\ell - p_h \| + \| h_E^{1/2} (g - p_h \cdot n) \|_{L^2(\Gamma_N)}) \\
 & + \min_{q_h \in \mathcal{S}^1(\mathcal{T}_\ell)^d} \|\nabla(\chi_\ell - U_\ell) - q_h\|_{L^2(\Omega)} \\
 & + \left( \sum_{T \in \mathcal{T}_\ell} \|f\|_{L^2(\omega_T)} \|(\chi - \chi_\ell - \omega)_-\|_{L^2(T)} \right)^{1/2} + \| \omega \| \\
 & + \left( \sum_{z \in \mathcal{N}_\ell} h_z^2 \min_{q \in \mathbb{R}} \|f - q\|_{L^2(\Omega_z)}^2 \right)^{1/2} \\
 & + \left( \sum_{T \in \mathcal{T} \setminus \mathcal{T}_D} \|h_T^{-1} (\chi - \chi_\ell - \omega)_-\|_{L^2(T)} \right)^{1/2} + \left( \sum_{T \in \mathcal{T}_D} h_T^2 \|f\|_{L^2(\omega_T)}^2 \right)^{1/2}.
 \end{aligned}$$

Hieraus lässt sich nun leicht das Korollar am Anfang von Abschnitt 3.2.1 ableiten, aus dem wir die Aussage der Zuverlässigkeit gewinnen.

# Anhang C.

## Implementierung

Dieser Abschnitt enthält den kompletten MATLAB Code der nötig ist, um alle numerischen Illustrationen in dieser Arbeit zu reproduzieren. Neben eigenen Implementierungen wurde außerdem das `plafem`-package aus [FPW] verwendet, welches von den Autoren freundlicherweise zur Verfügung gestellt wurde. Für nähere Informationen bezüglich dieses Codes sei auf eben jenes Paper verwiesen. Im Verlauf dieses Abschnittes sind zunächst die allgemeinen Methoden, und später auch die genauen Implementierungen der numerischen Beispiele angegeben.

```
1 function [x,coordinates,elements,indicators] ...
2     = adaptiveAlgorithmPrimalDual6(coordinates,elements,dirichlet,neumann, ...
3         f,chi,uD,nEmax,rho,c)
4
5 %     adaptive finite element algorithm for two-dimensional
6 %     obstacle problem
7 %
8 %     adaptiveAlgorithmPrimalDual6 computes the P1-FEM solution x of the
9 %     obstacle problem
10 %     <<x,v-x>> ≥ (f,v-x)           for all v in admissable set K
11 %
12 %     over a finite element mesh that consists of triangles.
13 %
14 %     adaptiveAlgorithmPrimalDual6 is the implementation of the following
15 %     iterative process for a given initial finite element mesh:
16 %
17 %     1) compute the discrete solution (via solvePrimalDual.m)
18 %     2) compute the elementwise error indicators (via computeEtaRobstacle.m)
19 %     3) mark the elements for refinement (via the Doerfler criterion)
20 %     4) refine the finite element mesh (via refineNVB.m)
21 %     5) return to 1)
22 %
23 %Usage:
24 %
25 %[X,COORDINATES,ELEMENTS,INDICATORS] ...
26 %     = ADAPTIVEALGORITHMPRIMALDUAL6(COORDINATES,ELEMENTS,DIRICHLET,NEUMANN, ...
27 %         F,CHI,UD,NEMAX,RHO;C)
28 %
29 %Comments:
30 %
31 %     adaptiveAlgorithmPrimalDual6 expects as input an initial finite element
32 %     mesh described by the fields COORDINATES, ELEMENTS, DIRICHLET, and NEUMANN.
33 %
34 %     Volume force and boundary data, as well as the affine obstacle are given
35 %     as M-files <f.m>, <chi.m> and <uD.m>. Due to the usage of the plafem-
36 %     package, inhomogeneous dirichlet data ist possible. We stress however,
37 %     that all our theoretical results, as well as the example from the thesis,
```

```

38 %   are only valid with constant zero dirichlet boundary. Either of these
39 %   M-files is assumed to take N evaluation points as (N x 2) matrix and
40 %   to return an (N x 1) column vector.
41 %
42 %   The stopping criterion is realized via the maximal number of elements
43 %   NEMAX of the finite element mesh, and the computation is stopped a
44 %   soon as the current number of elements is larger than NEMAX.
45 %
46 %   The parameter RHO in (0,1) corresponds to the marking of elements by
47 %   use of the Doerfler marking with respect to the residual-based error
48 %   estimator. RHO = 1 means that (essentially) all elements are marked
49 %   for refinement, whereas small RHO leads to highly adapted meshes.
50 %
51 %   The function returns the adaptively generated mesh in terms of
52 %   the matrices COORDINATES and ELEMENTS, the vector X of the nodal
53 %   values of the P1-FEM solution as well the corresponding
54 %   refinement INDICATORS, i.e., the value of the error estimator is
55 %   given by SQRT(SUM(INDICATORS)).
56 %
57 %Remark:
58 %
59 %   This program uses large parts of the plafem-package from [FPW]
60 %
61 %Author:
62 %
63 %   Marcus Page      December 2009
64 %
65
66 %Initialization
67 l = 1;
68 x = zeros(length(coordinates),1);
69 lambda = zeros(length(coordinates),1);
70 x_func = @(k,j) griddata(coordinates(:,1),coordinates(:,2),x,k,j);
71 lambda_func = @(k,j) griddata(coordinates(:,1),coordinates(:,2),lambda,k,j);
72 while 1
73     %*** Compute discrete solution
74     [x,energy, lambda] = solvePrimalDual(coordinates,elements,dirichlet, ...
75         neumann,f,uD,chi,c,l,x_func,lambda_func);
76     x_func = @(k,j) griddata(coordinates(:,1),coordinates(:,2),x,k,j);
77     lambda_func = @(k,j) griddata(coordinates(:,1),coordinates(:,2),lambda,k,j);
78     %*** Compute refinement indicators
79     [indicatorsEdges,indicatorsElements,edgeOsc, jumps,error_partial, ...
80         error_nabla] = computeEtaRObstacle(x,coordinates,elements, ...
81         dirichlet,neumann,f,g);
82     amountEdges = length(indicatorsEdges);
83     amountElements = length(indicatorsElements);
84     %*** Concatenate indicators
85     indicators = [indicatorsElements;indicatorsEdges];
86     eta = sqrt(sum(indicators));
87     eta_total(l) = eta;
88     osc_total(l) = sqrt(sum(edgeOsc)+sum(indicatorsElements));
89     jumps_total(l) = sqrt(sum(jumps));
90     error(l) = sqrt(error_partial + 1/2*energy - 1/2*error_nabla);
91     Ndof(l) = length(elements);
92     fix(clock) %only for time measurement
93     % *** Output (for RHO = 0.6)
94     dlmwrite('temperror_06_sqr.txt', errorsqr(l), '-append', 'precision',16);

```

```

95     dlmwrite('tempNdof_06.txt', Ndof(l), '-append', 'precision', 16);
96     dlmwrite('temperror_06.txt', error(l), '-append', 'precision', 16);
97     dlmwrite('tempOsc_06.txt', osc_total(l), '-append', 'precision', 16);
98     dlmwrite('tempjumps_06.txt', jumps_total(l), '-append', 'precision', 16);
99     dlmwrite('tempeta_06.txt', eta_total(l), '-append', 'precision', 16);
100    l = l+1;
101    %*** Stopping criterion
102    if size(elements,1) ≥ nEmax
103        break
104    end
105    %*** Mark elements for refinement
106    [indicators,idx] = sort(indicators, 'descend');
107    sumeta = cumsum(indicators);
108    ell = find(sumeta ≥ sumeta(end) * rho, 1);
109    marked = idx(1:ell);
110    %*** Refine mesh
111    [coordinates,elements,dirichlet,neumann] = ...
112        refineNVB(coordinates,elements,dirichlet,neumann,marked);
113 end

```

```

1 function [x,energy, lambda] = solvePrimalDual(coordinates,elements, ...
2     dirichlet,neumann,f,uD,chi,c,l,x-temp, lambda-temp)
3
4 % implementation of the primal-dual active set method for the poisson
5 % equation
6 %
7 % returns a vector x of nodal values of the P1 solution of current mesh.
8 %
9 % this function implements the following algorithm:
10 % 1) calculate active and inactive set according to [HIK]
11 % 2) Solve Poisson equation on inactive set, and calculate lambda on
12 % active set
13 % 3) stop or return to 1)
14 %
15 %Author:
16 %
17 % Marcus Page December 2009
18 %
19
20 %***enumerate nodes
21 nodes=1:size((coordinates),1);
22
23 %***Parameters (initial solution and multiplier)
24 x = feval(x-temp, coordinates(:,1),coordinates(:,2));
25 lambda = feval(lambda-temp, coordinates(:,1),coordinates(:,2));
26 active = nodes;
27 k = 1;
28
29 %***Step 1:
30 while 1
31     active = calculateActiveSet(coordinates, c, chi, nodes, lambda, x);
32
33     if active == -1
34         disp('no contact zone');
35         active = [];
36     end

```

```

37     nodes = nodes';
38     inactive = setdiff(nodes,active)';
39
40     %***Step 2:
41     % Compute evaluation of chi at nodal points
42     chi_temp = zeros(1,length(coordinates));
43     chi_temp = feval(chi, coordinates(:,1),coordinates(:,2));
44     % Calculate lambda
45     lambda = min(lambda + c*(x-chi_temp),0);
46     lambda_temp = @(x,y)griddata(coordinates(:,1),coordinates(:,2),lambda,x,y);
47     newf = @(x) feval(f,x)-lambda_temp(x(:,1),x(:,2));
48
49     %Compute discrete solution
50     [x,energy,A]=solveLaplace(coordinates,elements,dirichlet,neumann, ...
51         newf,g,uD);
52     %Adaptive tolerance
53     if (length(nodes) ≤ 1e4)
54         tol = 8e-4;
55     end
56     if(length(nodes) ≥ 1e4 && length(nodes)<5e4)
57         tol = 1e-5;
58     end
59     if(length(nodes)≥5e4)
60         tol = 1e-6;
61     end
62
63     if norm(x-x_old) ≤ tol;
64         out = sprintf('algorithm terminated at step %i:', k);
65         disp(out);
66         break;
67     end
68     k = k+1;
69 end

```

```

1 function active = calculateActiveSet(coordinates, c, chi, nodes, ...
2     lambda_old, x_old)
3
4 %***Returns vector ACTIVE of nodes of current discretization that lie
5 %   within the contact zone, according to [HIK]
6 %
7 %Author:
8 %
9 %   Marcus Page           December 2009
10 %
11
12 active = -1;
13 lam_eval = lambda_old;
14 x_eval = x_old;
15 chi_eval = feval(chi,coordinates(:,1),coordinates(:,2));
16 activezone = ((lam_eval + c*(x_eval - chi_eval)) < 0);
17 active = find(activezone);

```

```

1 function [etaRedges,etaRelements,edgeOsc,jumps,error_partial,error_nabla] ...
2     = computeEtaObstacle(x,coordinates,elements,dirichlet,neumann,f,g)

```

```

3
4 %computeEtaRobstacle: computes residual-based error estimator
5 %           for finite element solution of affine obstacle problem
6 %           with homogeneous dirichlet boundary conditions.
7 %
8 %Usage:
9 %
10 % [ETAREDGES, ETARELEMENTS, EDGEOSC, JUMPS, ERROR_PARTIAL, ERROR_NABLA] ...
11 %   = COMPUTEETAR(X, COORDINATES, ELEMENTS, DIRICHLET, NEUMANN, F, G)
12 %
13 %Comments:
14 %
15 %   The column vector X contains the nodal values of the P1 finite element
16 %   solution. The corresponding finite element mesh is given in terms of
17 %   COORDINATES, ELEMENTS, DIRICHLET, NEUMANN. The volume force <f.m> and
18 %   the Neumann data <g.m> are given as M-files. Either file is assumed
19 %   to take N evaluation points as (N x 2) matrix and to return an (N x 1)
20 %   column vector.
21 %
22 %   The function returns the column vectors ETAREDGES, ETARELEMENTS,
23 %   EDGEOSC, JUMPS, ERROR_PARTIAL and ERROR_NABLA.
24 %   EATREDDGES(J) and ETAREDDGES(j) respectively contain the squared
25 %   contribution of the edge and element terms of ETA. In particular, the
26 %   value of the residual error estimator is given by
27 %   SQRT(SUM([ETAREDDGES, ETARELEMENTS])). EDGEOSC and JUMPS contain
28 %   the corresponding contributions of ETA for further usage (plots, etc.).
29 %   ERROR_PARTIAL and ERROR_NABLA finally return different parts of the
30 %   total error in the energy functional.
31 %
32 %Remark:
33 %
34 %   this is a rewritten method from the plafem-package
35 %
36 %Author:
37 %
38 %   Marcus Page           December 2009
39 %
40
41 [edge2nodes, element2edges, dirichlet2edges, neumann2edges] ...
42   = provideGeometricData(elements, dirichlet, neumann);
43 %*** First vertex of elements and corresponding edge vectors
44 c1 = coordinates(elements(:,1),:);
45 d21 = coordinates(elements(:,2),:) - c1;
46 d31 = coordinates(elements(:,3),:) - c1;
47 %*** Vector of element volumes 2*|T|
48 area2 = d21(:,1).*d31(:,2)-d21(:,2).*d31(:,1);
49 %*** Compute curl(uh) = (-duh/dy, duh/dx)
50 u21 = repmat(x(elements(:,2))-x(elements(:,1)), 1,2);
51 u31 = repmat(x(elements(:,3))-x(elements(:,1)), 1,2);
52 curl = (d31.*u21 - d21.*u31)./repmat(area2,1,2);
53 %*** Compute edge terms hE*(duh/dn) for uh
54 dudn21 = sum(d21.*curl,2);
55 dudn13 = -sum(d31.*curl,2);
56 dudn32 = -(dudn13+dudn21);
57 etaR = accumarray(element2edges(:), [dudn21;dudn32;dudn13], ...
58   [size(edge2nodes,1) 1]);
59 %*** Incorporate Neumann data

```

```

60 if ~isempty(neumann)
61     cn1 = coordinates(neumann(:,1),:);
62     cn2 = coordinates(neumann(:,2),:);
63     gmE = feval(g, (cn1+cn2)/2);
64     etaR(neumann2edges) = etaR(neumann2edges) - sqrt(sum((cn2-cn1).^2,2)).*gmE;
65 end
66 %*** Incorporate Dirichlet data
67 etaR(dirichlet2edges) = 0;
68 %*** Assemble edge contributions of indicators
69 etaRedges = etaR.^2;
70 jumps = etaR.^2;
71 %*** Add volume residual to indicators
72 fsT = feval(f, (c1+(d21+d31)/3));
73 %*** Nullify volume residuals at elements that don't touch the boundary
74 tmp = ismember(element2edges, dirichlet2edges);
75 ind = find(max(tmp, [], 2) == 0);
76 fsT(ind) = 0;
77 etaRelements = (0.5*area2.*fsT).^2;
78
79 %*** Compute patch oscillations of inner edges via quadrature from [BCH]
80 [boundary, neighbours] = findNeighbours(1:size(edge2nodes,1), ...
81     element2edges, edge2nodes, dirichlet2edges);
82 idx = find(boundary == 0);
83 [psi, psi_r, psi_s, kappa, value1] = quad3(7, f, neighbours(idx,1), ...
84     elements, coordinates);
85 [psi, psi_r, psi_s, kappa, value2] = quad3(7, f, neighbours(idx,2), ...
86     elements, coordinates);
87 value1 = value1.*area2(neighbours(idx,1));
88 value2 = value2.*area2(neighbours(idx,2));
89 bestApprox = (1./(0.5*(area2(neighbours(idx,1))+ ...
90     area2(neighbours(idx,2)))).*(value1+value2);
91 %*** Compute L^2 Norm of patch oscillation
92 newf = @(x) (feval(f,x)-bestApprox).^2;
93 [psi, psi_r, psi_s, kappa, value1] = quad3(7, newf, neighbours(idx,1), ...
94     elements, coordinates);
95 [psi, psi_r, psi_s, kappa, value2] = quad3(7, newf, neighbours(idx,2), ...
96     elements, coordinates);
97 value1 = value1.*area2(neighbours(idx,1));
98 value2 = value2.*area2(neighbours(idx,2));
99 final = (0.5*area2(neighbours(idx,1))+ ...
100     0.5*area2(neighbours(idx,2))).*(value1+value2);
101 etaRedges(idx) = etaRedges(idx) + final;
102 edgeOsc = final;
103
104 %*** Compute error terms in energy functional
105 u_ell = @(k,j) griddata(coordinates(:,1), coordinates(:,2), x, k, j);
106 exact = @u;
107 grad_exact = @grad_u;
108
109 % 7-point quadrature from [BCH]
110 % \int f u_ell dx
111 [psi, psi_r, psi_s, kappa, val1] = quad32(7, f, 1:size(elements,1), elements, ...
112     coordinates, u_ell);
113 error_int1 = val1.*area2;
114 % \int f u dx
115 [psi, psi_r, psi_s, kappa, val2] = quad32(7, f, 1:size(elements,1), elements, ...
116     coordinates, exact);

```

---

```

117 error_int2 = val2.*area2;
118 % <<grad u, grad u>>
119 [psi,psi_r,psi_s,kappa, val3] = quad32grad(7,1:size(elements,1),elements, ...
120     coordinates,grad_exact);
121 error_int3 = val3.*area2;
122
123 error_partial = -sum(error_int1) + sum(error_int2);
124 error_nabla = sum(error_int3);

```

```

1 function [boundary, neighbours] = findNeighbours(edge, element2edges, ....
2     edge2nodes, dirichlet2edges)
3
4 %
5 %   finds neighbouring elements to a given edge
6 %
7 %Author:
8 %
9 %   Marcus Page           December 2009
10
11 boundary = zeros(length(edge),1);
12 neighbours = zeros(length(edge),2);
13
14 for k = 1:length(edge)
15     boundary(k) = 0;
16     tmp = find(sum(ismember(element2edges, edge(k)),2));
17     if (size(tmp,1) == 1
18         boundary(k) = 1;
19         neighbours(k) = 0;
20     else
21         neighbours(k,:) = tmp;
22     end
23 end

```

```

1 function [coordinates,newElements,varargout] ...
2     = refineNVB(coordinates,elements,varargin)
3
4 %refineNVB: local refinement of finite element mesh by newest vertex
5 %   bisection, where marked elements are refined by bisec(3)
6 %
7 %Usage:
8 %
9 %[COORDINATES,NEWELEMENTS,DIRICHLET,NEUMANN] ...
10 %   = REFINENVB(COORDINATES,ELEMENTS,DIRICHLET,NEUMANN,MARKED)
11 %
12 %Comments:
13 %
14 %   refineNVB expects as input a finite element mesh described by the
15 %   fields COORDINATES, ELEMENTS, DIRICHLET, and NEUMANN. The vector
16 %   MARKED contains the indices of elements which are refined by bisec(3).
17 %   Further elements will be refined by newest vertex bisection to obtain
18 %   a regular triangulation. MARKED also contains the edges marked
19 %   for refinement. Note that ELEMENTS(J,3) provides the index
20 %   of the newest vertex of the j-th element.
21 %

```

```

22 %   The function returns the refined mesh in terms of the same data as
23 %   for the input.
24 %
25 %Remark:
26 %
27 %   this is a rewritten method from the plafem-package
28 %
29 %Author:
30 %
31 %   Marcus Page           December 2009
32 %
33
34 marked = varargin{end};
35 nE = size(elements,1);
36
37 % *** Distinguish between marked Edges and marked Elements
38 markedElements = marked(marked<nE);
39 markedEdges = setdiff(marked, markedElements)-nE;
40
41 %*** Obtain geometric information on edges
42 [edge2nodes, element2edges, boundary2edges{1:nargin-3}] ...
43   = provideGeometricData(elements, varargin{1:end-1});
44 %*** Mark edges for refinement
45 edge2newNode = zeros(max(max(element2edges)),1);
46 edge2newNode(element2edges(markedElements,:)) = 1;
47 edge2newNode(markedEdges) = 1;      % also mark marked Edges for refinement
48 swap = 1;
49 while ~isempty(swap)
50     markedEdge = edge2newNode(element2edges);
51     swap = find( ~markedEdge(:,1) & (markedEdge(:,2) | markedEdge(:,3)) );
52     edge2newNode(element2edges(swap,1)) = 1;
53 end
54 %*** Generate new nodes
55 edge2newNode(edge2newNode~=0) = size(coordinates,1) + (1:nnz(edge2newNode));
56 idx = find(edge2newNode);
57 coordinates(edge2newNode(idx),:) ...
58   = (coordinates(edge2nodes(idx,1),:)+coordinates(edge2nodes(idx,2),:))/2;
59 %*** Refine boundary conditions
60 for j = 1:nargout-2
61     boundary = varargin{j};
62     if ~isempty(boundary)
63         newNodes = edge2newNode(boundary2edges{j});
64         markedEdges = find(newNodes);
65         if ~isempty(markedEdges)
66             boundary = [boundary(~newNodes,:); ...
67                         boundary(markedEdges,1), newNodes(markedEdges); ...
68                         newNodes(markedEdges), boundary(markedEdges,2)];
69         end
70     end
71     varargout{j} = boundary;
72 end
73 %*** Provide new nodes for refinement of elements
74 newNodes = edge2newNode(element2edges);
75 %*** Determine type of refinement for each element
76 markedEdges = (newNodes~=0);
77 none = ~markedEdges(:,1);
78 bisec1   = ( markedEdges(:,1) & ~markedEdges(:,2) & ~markedEdges(:,3) );

```

```

79 bisec12 = ( markedEdges(:,1) & markedEdges(:,2) & ~markedEdges(:,3) );
80 bisec13 = ( markedEdges(:,1) & ~markedEdges(:,2) & markedEdges(:,3) );
81 bisec123 = ( markedEdges(:,1) & markedEdges(:,2) & markedEdges(:,3) );
82 %*** Generate element numbering for refined mesh
83 idx = ones(nE,1);
84 idx(bisec1) = 2; %*** bisec(1): newest vertex bisection of 1st edge
85 idx(bisec12) = 3; %*** bisec(2): newest vertex bisection of 1st and 2nd edge
86 idx(bisec13) = 3; %*** bisec(2): newest vertex bisection of 1st and 3rd edge
87 idx(bisec123) = 4; %*** bisec(3): newest vertex bisection of all edges
88 idx = [1;1+cumsum(idx)];
89 %*** Generate new elements
90 newElements = zeros(idx(end)-1,3);
91 newElements(idx(none),:) = elements(none,:);
92 newElements([idx(bisec1),1+idx(bisec1)],:) ...
93     = [elements(bisec1,3),elements(bisec1,1),newNodes(bisec1,1); ...
94         elements(bisec1,2),elements(bisec1,3),newNodes(bisec1,1)];
95 newElements([idx(bisec12),1+idx(bisec12),2+idx(bisec12)],:) ...
96     = [elements(bisec12,3),elements(bisec12,1),newNodes(bisec12,1); ...
97         newNodes(bisec12,1),elements(bisec12,2),newNodes(bisec12,2); ...
98         elements(bisec12,3),newNodes(bisec12,1),newNodes(bisec12,2)];
99 newElements([idx(bisec13),1+idx(bisec13),2+idx(bisec13)],:) ...
100     = [newNodes(bisec13,1),elements(bisec13,3),newNodes(bisec13,3); ...
101         elements(bisec13,1),newNodes(bisec13,1),newNodes(bisec13,3); ...
102         elements(bisec13,2),elements(bisec13,3),newNodes(bisec13,1)];
103 newElements([idx(bisec123),1+idx(bisec123),2+idx(bisec123),3+ ...
104     idx(bisec123)],:)= [newNodes(bisec123,1),elements(bisec123,3), ...
105     newNodes(bisec123,3);elements(bisec123,1),newNodes(bisec123,1), ...
106     newNodes(bisec123,3);newNodes(bisec123,1),elements(bisec123,2), ...
107     newNodes(bisec123,2);elements(bisec123,3),newNodes(bisec123,1), ...
108     newNodes(bisec123,2)];

```

```

1 function [psi,psi_r,psi_s,kappa,value] = quad32(K_3,f,T,elements,coordinates,g)
2
3 %
4 % Multiple point quadrature from [BCH]
5 %
6
7 switch K_3
8     case 1
9         r = 1/3;
10        s = 1/3;
11        kappa = 1/2;
12     case 3
13        r = [1,4,1]'/6;
14        s = [1,1,4]'/6;
15        kappa = [1,1,1]/6;
16     otherwise
17        pos = [6-sqrt(15),9+2*sqrt(15),6+sqrt(15),9-2*sqrt(15),7]/21;
18        r = pos([1,2,1,3,3,4,5])';
19        s = pos([1,1,2,4,3,3,5])';
20        wts = [155-sqrt(15),155+sqrt(15),270]/2400;
21        kappa = wts([1,1,1,2,2,2,3]);
22 end
23 one = ones(size(kappa,2),1);
24 psi = [1-r-s,r,s,4*r.*(1-r-s),4*r.*s,4*s.*(1-r-s)];
25 psi_r = [-one,one,0*one,4*(1-2*r-s),4*s,-4*s];

```

```

26 psi_s = [-one,0*one,one,-4*r,4*r,4*(1-r-2*s)];
27
28 value = zeros(size(T,1),1);
29 for i = 1:K_3
30     tmp = Phi_T([r(i) s(i)], T, elements, coordinates);
31     %result is not yet multiplied with functional determinant = 2|T|
32     value = value + kappa(i).*feval(f,tmp).*feval(g,tmp(:,1),tmp(:,2));
33 end

```

```

1 function [psi,psi_r,psi_s,kappa,value] = quad32grad(K_3,T,elements, ...
2     coordinates,g)
3
4 %
5 % Multiple point quadrature from [BCH] for 2-dimensional gradients
6 %
7
8 switch K_3
9     case 1
10        r = 1/3;
11        s = 1/3;
12        kappa = 1/2;
13     case 3
14        r = [1,4,1]'/6;
15        s = [1,1,4]'/6;
16        kappa = [1,1,1]/6;
17     otherwise
18        pos = [6-sqrt(15),9+2*sqrt(15),6+sqrt(15),9-2*sqrt(15),7]/21;
19        r = pos([1,2,1,3,3,4,5])';
20        s = pos([1,1,2,4,3,3,5])';
21        wts = [155-sqrt(15),155+sqrt(15),270]/2400;
22        kappa = wts([1,1,1,2,2,2,3]);
23 end
24 one = ones(size(kappa,2),1);
25 psi = [1-r-s,r,s,4*r.*(1-r-s),4*r.*s,4*s.*(1-r-s)];
26 psi_r = [-one,one,0*one,4*(1-2*r-s),4*s,-4*s];
27 psi_s = [-one,0*one,one,-4*r,4*r,4*(1-r-2*s)];
28
29 value = 0;
30 for i = 1:K_3
31     args = Phi_T([r(i) s(i)], T, elements, coordinates);
32     tmp = feval(g,args(:,1),args(:,2));
33     for k = 1:(size(args,1))
34         oneline = tmp(k,:);
35         tmp2(k) =oneline*oneline';
36     end
37     %result is not yet multiplied with functional determinant = 2|T|
38     value = value + kappa(i).*tmp2';
39 end

```

```

1 function out = Phi_T(x, T, elements, coordinates)
2
3 %
4 % The transformation between T and reference triangle
5 %

```

---

```

6
7 nodes_T = elements(T,:);
8 z1 = coordinates(nodes_T(:,1),:);
9 z2 = coordinates(nodes_T(:,2),:);
10 z3 = coordinates(nodes_T(:,3),:);
11
12 out = (z3 + x(1).*(z1-z3) + x(2).*(z2-z3));

```

```

1 function node = coords2nodes(x, y, coordinates)
2
3 %
4 %returns number of the node at given coordinates
5 %
6
7 for j = 1:length(x)
8     for i = 1:length(coordinates)
9         if isequal(coordinates(i,:), [x y])
10            node(j) = i;
11            break;
12        else
13            node(j) = -1;
14        end
15    end
16 end

```

Im Folgenden wird der Code präsentiert, der sich direkt auf das numerische Beispiel am Ende von Kapitel 3 bezieht.

```

1 %Automated sample file that produces the output of the example of
2 %chapter 3 for adaptivity parameter theta = 0.6. See that chapter for
3 %further details.
4 %
5 %Author:
6 %
7 %     Marcus Page           January 2010
8 %
9
10 warning off
11 addpath('..')
12 warning on
13
14 %*** Initialization
15 load elements.dat
16 load coordinates.dat
17 load dirichlet.dat
18 load neumann.dat
19
20 %*** Parameters
21 nEmax = 5e5;
22 theta = 0.6;
23 c = 1;
24
25 %***Generate initial mesh by bisec_3 refinement
26 for i=1:1

```

```

27     [coordinates,elements,dirichlet,neumann] ...
28         = refineNVB(coordinates,elements,dirichlet,neumann,1:size(elements,1));
29 end
30
31 *** Call adaptive algorithm
32 [x,coordinates,elements,indicators] ...
33     = adaptiveAlgorithmPrimalDual6(coordinates,elements,dirichlet, ...
34         neumann,@f,@chi,@uD,nEmax,theta,c);
35
36 *** Visualization
37 hold on;
38 grid on;
39 trisurf(elements(:,1:3),coordinates(:,1),coordinates(:,2),x','FaceAlpha',1);
40 view(-13,26)

```

```

1 function out = chi(x,y)
2 % (affine) obstacle chi(x,y) = 0
3
4 out = zeros(size(x,1),1);

```

```

1 function out = f(x)
2 % Volume force, right hand side according to (3.4.2)
3 out = zeros(size(x,1),1);
4
5 [phi,r] = cart2pol(x(:,1),x(:,2));
6 phi_minus = find(phi < 0);
7 phi(phi_minus) = 2*pi + phi(phi_minus);
8
9 out = -r.^(2/3).*sin(2*phi./3).*(gamma_1_prime(r)./r + ...
10     gamma_1_doubleprime(r)) - (4/3)*r.^(-1/3).* ...
11     gamma_1_prime(r).*sin(2*phi./3) - gamma_2(r);

```

```

1 function out = gammal(r)
2 % gamma_1 from example
3
4 out = zeros(length(r),1);
5 r_bar = 2*(r-1/4);
6 out(r_bar < 0) = 1;
7 out(r_bar ≥ 1) = 0;
8 out(r_bar ≥ 0 & r_bar < 1) = -6*r_bar(r_bar ≥ 0 & r_bar < 1).^5 + ...
9     15*r_bar(r_bar ≥ 0 & r_bar < 1).^4 - 10*r_bar(r_bar ≥ ...
10     0 & r_bar < 1).^3 + 1;

```

```

1 function out_prime = gamma_1_prime(r)
2 % gamma_1' from example
3
4 r_bar = 2*(r-1/4);
5 out_prime = zeros(length(r),1);
6 out_prime(r_bar ≥ 0 & r_bar < 1) = -960*(r(r_bar ≥ ...
7     0 & r_bar < 1)-1/4).^4 + 960*(r(r_bar ≥ 0 & r_bar < 1)-1/4).^3 ...
8     - 240*(r(r_bar ≥ 0 & r_bar < 1)-1/4).^2;

```

---



---

```

1 function out_dp = gamma_1_doubleprime(r)
2 %gamma_1'' from example
3
4 r_bar = 2*(r - 1/4);
5 out_dp = zeros(length(r),1);
6 out_dp(r_bar ≥ 0 & r_bar < 1) = -3840*(r(r_bar ≥ 0 & r_bar < 1) ...
7     -1/4).^3+2880*(r(r_bar ≥ 0 & r_bar < 1)-1/4).^2 - 480*(r(r_bar > ...
8     = 0 & r_bar < 1)-1/4);

```

```

1 function out = gamma_2(r)
2 %gamma_2 from example
3
4 out = ones(size(r,1),1);
5 out(r ≤ 5/4) = 0;

```

```

1 function out = u(x,y)
2 %exact solution u
3
4 [phi,r] = cart2pol(x,y);
5 phi(phi<0) = 2*pi + phi(phi<0);
6 out = r.^(2/3).*gamma1(r).*sin(2*phi./3);

```

```

1 function out = grad_u(x,y)
2 % gradient of exact solution u
3
4 [phi,r] = cart2pol(x,y);
5 phi(phi<0) = 2*pi + phi(phi<0);
6
7 out(:,1) = sin(2*phi./3).*(2*gamma1(r)+3*r.*gamma_1_prime(r))./(3.*r.^(1/3));
8 out(:,2) = (2*cos(2*phi./3).*gamma1(r))./(3*r.^(1/3));

```

```

1 function out = uD(x)
2 %Dirichlet boundary data: u = u_D = 0 in this example
3
4 n = size(x,1);
5 out = zeros(n,1);

```

```

1 function out = g(x,y)
2 %constant zero neumann data
3
4 out = 0;

```

```

1 function plotExactSolution
2 %plots the exact solution u on a certain uniform mesh

```

```
3
4 exact = @u;
5 x = zeros(100,100);
6 y = zeros(100,100);
7 for i = 1:100
8     x(i,:) = linspace(-2,2,100);
9     y(:,i) = linspace(-2,2,100);
10 end
11
12 for i = 1:100
13     for j = 1:100
14         if (x(i,j) > 0 && y(i,j) < 0)
15             x(i,j) = 0;
16             y(i,j) = 0;
17         end
18         if (x(i,j)>0 && y(i,j) < 0)
19             z(i,j) = 0;
20         else
21             z(i,j) = exact(x(i,j),y(i,j));
22         end
23     end
24 end
25
26 surf(x,y,z, 'FaceAlpha',1,'EdgeAlpha', 0.5);
27 view(-13,26)
28 grid off
```

```
1 %reads the output files from the test runs
2
3 dof04 = dlmread('tempNdof_04.txt')
4 err04 = dlmread('temperror_04.txt')
5 eta04 = dlmread('tempeta_04.txt')
6 jumps04 = dlmread('tempjumps_04.txt')
7 osc04 = dlmread('tempOsc_04.txt')
8
9 dof06 = dlmread('tempNdof_06n.txt')
10 err06 = dlmread('temperror_06n.txt')
11 eta06 = dlmread('tempeta_06n.txt')
12 jumps06 = dlmread('tempjumps_06n.txt')
13 osc06 = dlmread('tempOsc_06n.txt')
14
15 dof08 = dlmread('tempNdof_08.txt')
16 err08 = dlmread('temperror_08.txt')
17 eta08 = dlmread('tempeta_08.txt')
18 jumps08 = dlmread('tempjumps_08.txt')
19 osc08 = dlmread('tempOsc_08.txt')
20
21 dof02 = dlmread('tempNdof_02.txt')
22 err02 = dlmread('temperror_02.txt')
23 eta02 = dlmread('tempeta_02.txt')
24 jumps02 = dlmread('tempjumps_02.txt')
25 osc02 = dlmread('tempOsc_02.txt')
26
27 dofunif = dlmread('tempNdof_univn.txt')
28 errunif = dlmread('temperror_univn.txt')
29 etaunif = dlmread('tempeta_univn.txt')
```

---

```
30 jumpsunif = dlmread('tempjumps_univn.txt')
31 oscunif = dlmread('tempOsc_univn.txt')
```

```
1 %plots the convergence history for uniform and adaptive refinement
2
3 loglog(dof06(2:end), eta06(2:end), '-s', dof06(2:end), err06(2:end), ...
4     '-o', dof06(2:end), osc06(2:end), '-x', dof06(2:end), etaunif(2:end), ...
5     '+', dof06(2:end), errunif(2:end), '-*', dof06(2:end), oscunif(2:end), ...
6     '-p', 'LineWidth', 2);
7 l = legend('eta06', 'err06', 'osc06', 'etaunif', 'errunif', 'oscunif');
8 set(l, 'FontSize', 15, 'Location', 'SouthWest')
9 legend boxoff
10 hold on
11 a = linspace(50, 1300000, 16);
12 b = a.^(-1/2);
13 c = a.^(-1/2.4);
14 d = a.^(-1);
15 loglog(a, 0.8*b, '--k', 'LineWidth', 2)
16 text(1e4, 1.5e-3, '12')
17 loglog(a, c+160*c, '--k', 'LineWidth', 2)
18 text(1e5, 3, '512')
19 loglog(a, d+1000*d, '--k', 'LineWidth', 2)
20 text(2.4e2, 6, 'one')
```

```
1 %plots the comparison of different adaptive refinements and uniform
2 %refinement
3
4 loglog(dof02(2:end), err02(2:end), '-p', dof04(2:end), err04(2:end), ...
5     '-o', dof06(2:end), err06(2:end), '-s', dof08(2:end), err08(2:end), ...
6     '-x', dof06(2:end), errunif(2:end), '-+', 'LineWidth', 2);
7 l = legend('02', '04', '06', '08', 'unif');
8 set(l, 'FontSize', 15, 'Location', 'SouthWest')
9 legend boxoff
10 hold on
11 a = linspace(50, 1300000, 16);
12 b = a.^(-1/2);
13 c = a.^(-1/2.4);
14 loglog(a, 0.8*b, '--k', 'LineWidth', 2)
15 text(1e4, 2e-3, '12')
16 loglog(a, c+30*c, '--k', 'LineWidth', 2)
17 text(1e5, 0.5, '512')
```



## Literaturverzeichnis

- [AFP] M. AURADA, S. FERRAZ-LEITE, D. PRAETORIUS: *Estimator reduction and convergence of adaptive FEM and BEM*, ASC Report **27/2009**, Institute for Analysis and Scientific Computing, Vienna University of Technology, Wien 2009.
- [BRA07] D. BRAESS: *Finite Elemente*, 4. Auflage, Springer Verlag 2007.
- [B] D. BRAESS: *A posteriori error estimators for obstacle problems — another look*, Numer. Math. **101** (2005), 415–421.
- [BC] S. BARTELS, C. CARSTENSEN: *Averaging techniques yield reliable a posteriori finite element error control for obstacle problems*, Numer. Math. **99** (2004), 225–249.
- [BCH] S. BARTELS, C. CARSTENSEN, A. HECHT: *P2QIso2D = 2D isoparametric FEM in Matlab*, J. Comput. Appl. Math. **192** (2006), no. 2, 219–250.
- [BCH1] D. BRAESS, C. CARSTENSEN, R. HOPPE: *Convergence analysis of a conforming adaptive finite element method for an obstacle problem*, Numer. Math. **107** (2007), 455–471.
- [BCH2] D. BRAESS, C. CARSTENSEN, R. HOPPE: *Error reduction in adaptive finite element approximations of elliptic obstacle problems*, J. Comput. Math. **27** (2009), 148–169.
- [C] C. CARSTENSEN: *Reliable and efficient averaging techniques as universal tool for a posteriori finite element error control on unstructured grids*, Int. J. Numer. Anal. Model **3** (2006), 333–347.
- [CKNS] J. CASCON, C. KREUZER, R. NOCHETTO, K. SIEBERT: *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal. **46** (2008), 2524–2550.
- [D] W. DÖRFLER: *A convergent adaptive algorithm for Poisson’s equation*, SIAM J. Numer. Anal. **33** (1996), 1106–1124.
- [E] L. C. EVANS: *Partial Differential Equations*, Graduate Studies in Mathematics, Volume 19. American Mathematical Society, Providence, Rhode Island, 2002.
- [FPW] S. FUNKEN, D. PRAETORIUS, P. WISSGOTT: *Efficient implementation of adaptive P1-FEM in MATLAB*, ASC Report **19/2008**, Institute for Analysis and Scientific Computing, Vienna University of Technology, Wien 2008.
- [HIK] M. HINTERMÜLLER, K. ITO, K. KUNISCH: *The primal-dual active set strategy as a semismooth newton method*, SIAM J. Optim., **13** (2003), 865–888.
- [IK] K. ITO, K. KUNISCH: *Optimal control of elliptic variational inequalities*, Appl. Math. Optim., **41** (2000), 343–364.

- [IK08] K. ITO, K. KUNISCH: *Lagrange Multiplier Approach to Variational Problems and Applications*, Society for Industrial Mathematics, 2008.
- [KS] D. KINDERLEHRER, G. STAMPACCHIA: *An introduction to variational inequalities*, Academic Press, New York, 1980.
- [MNS] P. MORIN, R. NOCHETTO, K. SIEBERT: *Data oscillation and convergence of adaptive FEM*, SIAM. J. Numer. Anal.. **18** (2000), 466–488.
- [P] D. PRAETORIUS: *Finite Element Methods*, Vorlesungsskript, WS 2008/09, Technische Universität Wien.
- [PP] M. PAGE, D. PRAETORIUS: *Convergence of adaptive FEM for some elliptic obstacle problem*, ASC Report **05/2010**, Institute for Analysis and Scientific Computing, Vienna University of Technology, Wien 2010.
- [S] P. STRACK: *Primal-Dual-Active-Set-Methoden zur Lösung restringierter Probleme über partiellen Differentialgleichungen*, Diplomarbeit, Institut für Numerische Simulation, Universität Bonn, November 2008.
- [W] D. WERNER: *Funktionalanalysis*, 3. Auflage, Springer Verlag 2000.
- [V] R. VERFÜRTH: *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Wiley-Teubner, 1996.

# LEBENS LAUF DES AUTORS

MARCUS PAGE B.SC.

Geburtstag: 19.06.1984  
Geburtsort: Groß-Gerau  
Staatsbürgerschaft: deutsch  
eMail-Adresse: Marcus.Page@gmx.de

---

## BILDUNGSWEG

seit 10/2008 **Universität Wien**  
Masterstudium der Mathematik mit Vertiefung in  
numerischer Mathematik und Scientific Computing

10/2006 – 9/2007 **Universität Nihon, Tokyo**  
Auslandsaufenthalt und Studium in Japan

10/2004 – 9/2008 **Technische Universität Darmstadt**  
Bachelorstudium im Studiengang *Mathematics with Computer  
Science*  
Bachelorarbeit: *Mathematical Approaches to Speech Recognition*  
Betreuer: Prof. Dr. Jens Lang, Prof. Dr. Tetsuya Hoya

9/2003 – 6/2004 **Zivildienst** bei der Johanniter-Unfallhilfe

9/1994 – 6/2003 **Edith-Stein-Schule, Darmstadt**  
Abitur mit den Leistungskursen Mathematik und Informatik.

---

## VERÖFFENTLICHUNGEN

*Security of Sanitizable Signatures Revisited*, Public-Key Cryptography (PKC) 2009, Lecture Notes in Computer Science, Volume 5443, pp. 317-336, Springer-Verlag, 2009.  
Erstellt im Rahmen des Seminars „Forschungsorientierte Kryptographie“, TU Darmstadt im SS 2008.