



TECHNISCHE  
UNIVERSITÄT  
WIEN  
VIENNA  
UNIVERSITY OF  
TECHNOLOGY

DIPLOMARBEIT

# Adaptive Finite Volumen Methode

ausgeführt am Institut für  
Analysis und Scientific Computing  
der Technischen Universität Wien

unter Anleitung von Ao.Univ.Prof. Dipl.Math. Dr.techn. Dirk Praetorius

durch

**Christoph Erath**

Steinwiesen 4  
A-6824 Schlins

Schlins, am 01. September 2005

Datum

---

Unterschrift



# Kurzfassung

Ziel dieser Arbeit ist die mathematische Rechtfertigung einer adaptiven Netzverfeinerung für die zellenorientierte Finite Volumen Methode zur Lösung eines elliptischen Modellproblems zweiter Ordnung in  $\mathbb{R}^2$ . Als Grundlage dient eine Arbeit von Nicaise [13], in der ein residualbasierter a posteriori Fehlerschätzer eingeführt wird. Auf dessen Grundlage erhalten wir heuristische Verfeinerungsindikatoren, die den lokalen Fehler schätzen und zu einer lokalen Netzverfeinerung herangezogen werden können. Die Analysis in Nicaise [13] lässt aber keine hängenden Knoten zu, so dass die a posteriori Abschätzung de facto nur für uniforme Triangulierungen rigoros ist. In dieser Arbeit wird die Analysis deshalb modifiziert, um auch den Fall hängender Knoten zuzulassen. Wir beweisen (globale) Zuverlässigkeit und (quasi-lokale) Effizienz des Residualschätzers für Triangulierungen mit hängenden Knoten. Die in den Abschätzungen involvierten Konstanten werden in numerischen Beispielen untersucht.

Die Finite Volumen Methode führt für ein lineares konvektiv-diffusives Problem auf eine Bilanzgleichung von konvektivem und diffusivem Fluss. Um auch auf unstrukturierten Triangulierungen rechnen zu können (im Hinblick auf adaptive Verfeinerung), wird der konvektive Fluss mit dem Upwind-Verfahren und der diffusive Fluss mit der Diamond Path Methode approximiert (Coudiere und Villedieu [9]). Leider benötigt die Diamond Path Methode die numerisch aufwendige Berechnung von gewissen Interpolationsgewichten, die von der lokalen Geometrie der Triangulierung abhängen. Es ist daher sinnvoll, a priori nur endlich viele Winkel in der Triangulierung zuzulassen und die Ordnungen der hängenden Knoten nach oben zu beschränken. Für den Fall quadratischer Elemente, auf deren Kanten höchstens ein hängender Knoten liegt, können alle auftretenden Interpolationsgewichte a priori analytisch berechnet werden.

Die Idee in der Arbeit von Nicaise [13] ist es, die stückweise konstante Finite Volumen Lösung  $u_h$  geeignet zu interpolieren, so dass der Interpolant  $I_M u_h$  neben der lokalen Flusserhaltung zusätzliche Orthogonalitätseigenschaften besitzt. Dies erlaubt es dann, die analytischen Techniken aus dem Kontext der Finite Elemente Methode anzuwenden, um einen a posteriori Fehlerschätzer zu konstruieren, der den Fehler  $e := u - I_M u_h$  in Termen des Residuums zu  $I_M u_h$  abschätzt (vergleiche Verfürth [20]). Grundlage für den Interpolationsoperator sind nicht-konforme Elemente vom Morley-Typ. Der Residualschätzer enthält deshalb neben den Sprüngen der Richtungsableitung in Normalrichtung auch die Sprünge der Richtungsableitung in Tangentialrichtung des Interpolanten.

Wir adaptieren die Ideen von Nicaise [13] auf lokal verfeinerte Triangulierungen mit hängenden Knoten. Dazu wird der Begriff der fastregulären Triangulierung eingeführt. Die Definition des Morley-Interpolanten wird so abgeändert, dass die Orthogonalitätseigenschaft des Residuums auch für die fastregulären Triangulierungen erhalten bleibt. Auftretende Spurterme über Kanten werden mit einer Spurgleichung abgeschätzt, die nicht wie in der Literatur üblich den Patch der Kante involviert, sondern nur ein an der Kante gelegenes Element. Dies erlaubt die Behandlung von Teilkanten und damit von hängenden Knoten. Mit diesen Ideen kann der Beweis der (globalen) Zuverlässigkeit des Fehlerschätzers auf fastreguläre Triangulierungen erweitert werden. Der Beweis der (quasi-lokalen) Effizienz basiert wie bei dem entsprechenden Finite Elemente Methode Resultat (Verfürth [20]) auf inversen Abschätzungen für Normen, die auf gewissen Bubble-Funktionen basieren. In der Literatur finden sich diese nur für reguläre Triangulierungen. Mit geeigneter Definition der Bubble-Funktionen für fastreguläre Triangulierungen lässt sich auch die Effizienz unseres Residualschätzers beweisen. Abschließende numerische Experimente bestätigen die theoretischen Ergebnisse.



# Inhaltsverzeichnis

<b>Inhaltsverzeichnis</b>	<b>i</b>
<b>Einleitung</b>	<b>iii</b>
<b>1 Triangulierung</b>	<b>1</b>
1.1 Notationen	1
1.1.1 Knoten- und Kantenmengen, fastreguläre Triangulierung	1
1.1.2 Patches	3
1.1.3 Bezeichnungen anhand einer einfachen Triangulierung	5
1.1.4 Durchmesser $h_T$ und $\varrho_T$	6
1.2 Wahl der Normalvektoren und Tangentialvektoren	7
1.2.1 Normalvektor $\mathbf{n}_E$ und Tangentialvektor $\mathbf{t}_E$	7
1.2.2 Vorzeichen $\sigma_{T,E}$ , $\mathbf{n}_{T,E}$ und $\mathbf{t}_{T,E}$	8
<b>2 Die Finite Volumen Methode</b>	<b>10</b>
2.1 Das Grundprinzip der Diskretisierung	10
2.2 Der konvektive numerische Fluss $F_E^C(u_h)$	12
2.3 Der diffusive numerische Fluss $F_E^D(u_h)$	12
2.3.1 Die Differenzen Methode	13
2.3.2 Diamond Path Methode	14
2.3.3 Herleitung der Diamond Path Methode	15
2.3.4 Berechnung von $\psi_T(a)$	18
<b>3 Residualschätzer</b>	<b>20</b>
3.1 Das Laplace Problem mit reinem Dirichlet-Rand	20
3.2 Das Morley-Element	20
3.3 Der Morley-Interpolant	23
3.4 A posteriori Fehlerschätzer	26
3.4.1 Zuverlässigkeit	27
3.4.2 Effizienz	37
<b>4 Numerische Experimente</b>	<b>52</b>
4.1 Algorithmen und Datenstrukturen	52
4.1.1 Datenstrukturen	52
4.1.2 Adaptiver Algorithmus	54
4.1.3 Netzverfeinerung	55
4.1.4 Berechnung der Finite Volumen Lösung	57
4.1.5 Berechnung des Morley-Interpolanten	57

4.2	Uniforme und adaptive Resultate . . . . .	58
4.2.1	Beispiel mit glatter Lösung und Nullranddaten . . . . .	59
4.2.2	Beispiel mit glatter Lösung und inhomogenen Dirichlet-Daten . . . . .	65
4.2.3	Beispiel mit Eckensingularität und inhomogenen Dirichlet-Daten . . . . .	71
	<b>Symbolverzeichnis</b>	<b>78</b>
	<b>Literaturverzeichnis</b>	<b>81</b>

# Einleitung

Die Finite Volumen Methode wird hauptsächlich für die numerische Lösung von Problemen der Strömungsmechanik eingesetzt. Dort wurde sie zu Beginn der 70-er Jahre von McDonald, MacCormack und Paullay [18] eingeführt. Die Anwendungsmöglichkeiten sind jedoch keineswegs auf Problemstellungen aus diesem Bereich beschränkt. Eine der wichtigsten Eigenschaften von Finite Volumen Methoden ist, dass die Erhaltungsprinzipien, die den mathematischen Modellen kontinuumsmechanischer Problemstellungen zugrunde liegen, per Definition auch für die diskretisierten Gleichungen erfüllt sind (Konservativität).

## Ziel dieser Arbeit

Im Gegensatz zur Finite Differenzen Methode oder zur Finite Elemente Methode verharrete die mathematische Untersuchung der Finite Volumen Methode lange nur in Ansätzen. Erst in den letzten Jahren sind wesentliche Erkenntnisfortschritte zu verzeichnen, auch in der Entwicklung von geeigneten Fehlerschätzern. In der Arbeit von Nicaise [13] wird ein a posteriori Fehlerschätzer für die zellenorientierte Finite Volumen Methode eingeführt, de facto allerdings nur für uniforme Netzverfeinerung. Die Analysis adaptiert Techniken aus dem Kontext der Finite Elemente Methode, wie sie z.B. im Buch von Verfürth [20, Kapitel 1] zu finden sind. Das Ziel der vorliegenden Arbeit ist die Entwicklung eines a posteriori Residualschätzers zur adaptiven Netzverfeinerung für die Finite Volumen Diskretisierung eines elliptischen Modellproblems zweiter Ordnung.

## Aufbau der Arbeit

In dieser Arbeit betrachten wir ausschließlich den zweidimensionalen Fall. Falls nicht anders erwähnt, sind die Elemente  $T \in \mathcal{T}$  der Triangulierung *Dreiecke* oder *Rechtecke*.

In Kapitel 1 wird der Begriff einer *fastregulären* Triangulierung eingeführt. Eine Kante, auf der keine hängenden Knoten liegen, heißt elementare Kante, andernfalls nicht-elementare oder hängende Kante. Eine Triangulierung  $\mathcal{T}$  heißt fastregulär, falls jede hängende Kante die Vereinigung von Elementarkanten ist. Im Gegensatz zu einer regulären Triangulierung lässt also die fastreguläre Triangulierung hängende Knoten zu. Die Notationen für die Verwaltung des Netzes werden eingeführt und neben den üblichen Begriffen Knotenpatch, Kantenpatch und Elementpatch definieren wir auch einen *elementaren* Kantenpatch. Abbildung 1 zeigt den Unterschied zwischen dem klassischen und elementaren Kantenpatch auf einer fastregulären Triangulierung.

Kapitel 2 stellt das Grundprinzip der Finite Volumen Methode anhand der linearen elliptischen

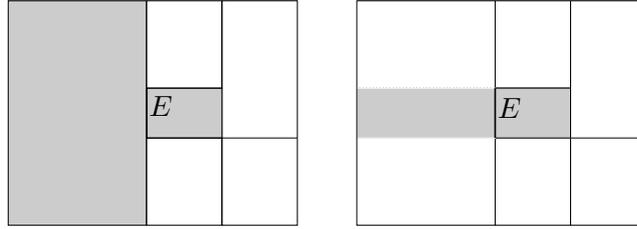


Abbildung 1: Der Kantenpatch links und der neu-eingeführte elementare Kantenpatch rechts auf einer fastregulären Triangulierung.

partiellen Differentialgleichung (gesucht sei  $u \in H^1(\Omega)$ )

$$\begin{aligned} \operatorname{div}(-\nu \nabla u + u \mathbf{v}) &= f && \text{auf } \Omega, \\ u &= u_D && \text{auf } \Gamma = \partial\Omega, \end{aligned} \tag{1}$$

mit  $\nu \in \mathcal{C}^2(\overline{\Omega}, \mathbb{R}_{>0})$ ,  $\mathbf{v} = (v_x, v_y) \in \mathcal{C}^1(\overline{\Omega}, \mathbb{R}^2)$ ,  $f \in L^2(\Omega)$  und  $u_D \in H^{1/2}(\Gamma)$  vor. Das Verfahren der Finite Volumen Methode arbeitet wie die Finite Elemente Methode auf einer Partition  $\mathcal{T}$  (in diesem Kapitel nicht zwingend Dreiecke oder Rechtecke,  $\mathcal{T}$  sei aber eine fastreguläre Triangulierung) des Gebietes  $\Omega$  und führt auf die sogenannte *Bilanzgleichung*

$$\sum_{E \in \mathcal{E}_T} \sigma_{T,E} (\Phi_E^C(u) - \Phi_E^D(u)) = \int_T f(x) dx \quad \forall T \in \mathcal{T}. \tag{2}$$

Wir bezeichnen mit  $\mathcal{E}_T$  die Menge der Kanten  $E$  von  $T \in \mathcal{T}$ . Das Vorzeichen  $\sigma_{T,E}$  ist abhängig von der Richtung des a priori festgelegten Normalvektors auf  $E$ . Der konvektive Fluss  $\Phi_E^C(u)$  und der diffusive Fluss  $\Phi_E^D(u)$  werden durch numerische Flüsse  $F_E^C(u_h)$  und  $F_E^D(u_h)$  ersetzt. Dies führt auf ein lineares Gleichungssystem. Die Finite Volumen Methode projiziert  $u$  auf eine zu den Elementen  $T \in \mathcal{T}$  stückweise konstante Funktion  $u_h$ .

Die Approximation des konvektiven Flusses  $\Phi_E^C(u)$  wurde in den letzten 20 Jahren sogar auf unstrukturierten Netzen weitgehendst untersucht (Coudière und Villedieu [9]). Für den diffusiven Fluss  $\Phi_E^D(u)$  wird hingegen meist einfach eine Differenzen Methode gewählt. Diese Methode hat jedoch den schwerwiegenden Nachteil, dass wir in der Wahl der Triangulierungselemente stark eingeschränkt sind, da die Diskretisierung sonst inkonsistent wird (Handbook of Numerical Analysis [10]). Im Hinblick auf adaptive Verfeinerung wird deshalb die *Diamond Path* Methode für die Approximation von  $\Phi_E^D(u)$  verwendet, wobei der Diamond Path konvex sein muss. Bei dieser Methode wird der ganze Gradient auf einer Kante approximiert. Dies bedarf einer linearen Interpolation der zusätzlichen Unbekannten an den Knoten. Die Berechnung der Interpolationsgewichte, die von der lokalen Geometrie der Triangulierung abhängen, ist numerisch aufwendig. Es ist daher sinnvoll, a priori nur endlich viele Winkel in der Triangulierung zuzulassen und die Ordnungen der hängenden Knoten nach oben zu beschränken. Für den Fall quadratischer Elemente, auf deren Kanten höchstens ein hängender Knoten liegt, können alle auftretenden Interpolationsgewichte a priori analytisch berechnet werden (vergleiche dazu Abbildung 2.5). Das sogenannte *Upwind-Verfahren* für die Approximation von  $\Phi_E^C(u)$  und die *Diamond Path Methode* garantieren, dass das Verfahren für eine hinreichend glatte exakte Lösung  $u \in H^2(\Omega)$  konvergiert. Es gilt dann  $\|u_{\mathcal{T}} - u_h\|_{1,h} = \mathcal{O}(h)$ , wobei  $\|\cdot\|_{1,h}$  eine diskrete  $H^1$ -Norm bezeichnet und  $u_{\mathcal{T}} \in \mathcal{P}_0(\mathcal{T})$  das  $\mathcal{T}$ -stückweise Integralmittel der exakten Lösung  $u$  ist (Coudière und Villedieu [9]).

In Kapitel 3 wird für das Laplace-Problem mit reinem Dirichlet-Rand ein residualbasierter a posteriori Fehlerschätzer hergeleitet. Wie schon in Nicaise [13] werden dabei die analytischen Techniken aus der Finite Elemente Methode adaptiert. Aufgrund zugelassener hängender Knoten müssen die Beweise allerdings teilweise stark modifiziert werden.

Im Ansatz von Nicaise [13] wird die diskrete Finite Volumen Approximation  $u_h \in \mathcal{P}_0(\mathcal{T})$  zunächst durch einen geeigneten Interpolanten  $I_M u_h$  ersetzt. Dieser liegt in einem nicht-konformen Ansatzraum aus Morley-artigen Elementen. Die prinzipielle Eigenschaft des Interpolanten ist die Flusserhaltung auf der fastregulären Triangulierung. Diese kann auf hängenden Kanten sowie Kanten, die nicht Teil einer hängenden Kante sind, erreicht werden, das heißt, der Fluss des Interpolanten ist gleich dem numerischen Fluss der diskreten Lösung. Ferner erzwingen wir die Stetigkeit des Interpolanten in allen Knoten. Aufgrund der Flusseigenschaft lässt sich beweisen, dass das Residuum  $R := f + \Delta I_M u_h$  im  $L^2$ -Sinn *orthogonal* ist zu  $\mathcal{P}_0(\mathcal{T})$ . Diese Orthogonalität ersetzt im Beweis die Galerkin-Orthogonalität einer konformen Finite Elemente Approximation.

Wir zeigen, dass der Fehlerschätzer

$$\eta := \left( \sum_{T \in \mathcal{T}} \eta_T^2 \right)^{1/2} \quad (3)$$

mit Verfeinerungsindikatoren

$$\eta_T^2 := h_T \left( h_T \|f - f_T\|_{L^2(T)}^2 + \sum_{E \in (\mathcal{E}_T \setminus \mathcal{E}_\Gamma)} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)}^2 \right) \quad (4)$$

$$+ \sum_{E \in (\mathcal{E}_T \setminus \mathcal{E}_\Gamma)} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right\|_{L^2(E)}^2 + \sum_{E \in (\mathcal{E}_T \cap \mathcal{E}_\Gamma)} \left\| \frac{\partial (I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}^2 \quad (5)$$

*global zuverlässig* ist (Satz 3.4.9) und (*quasi*-)lokal auch *effizient* bis auf Terme höherer Ordnung (Satz 3.4.16), wobei  $[\cdot]_E$  den Kantensprung und  $\mathbf{n}_E$  bzw.  $\mathbf{t}_E$  den Normal- bzw. Tangentialvektor über eine Kante  $E$ ,  $\mathcal{E}_T$  die Menge der Kanten des Elements  $T \in \mathcal{T}$ ,  $\mathcal{E}_\Gamma$  die Menge der Randkanten und  $f_T \in \mathcal{P}^0(T)$  das  $\mathcal{T}$ -stückweise Integralmittel von  $f$  bezeichne. Das heißt, es gelten für den Fehler  $e := u - I_M u_h \in H^1(\mathcal{T})$  die Abschätzungen

$$\|\nabla_{\mathcal{T}} e\|_{L^2(\Omega)} \leq C_1 \left( \sum_{T \in \mathcal{T}} \eta_T^2 \right)^{1/2}, \quad (6)$$

und

$$\eta_T^2 \leq C_2 \left( \|\nabla_{\mathcal{T}} e\|_{L^2(\omega_T)}^2 + h_T^2 \|f - f_T\|_{L^2(\omega_T)}^2 \right), \quad (7)$$

wobei  $\omega_T$  der Elementpatch ist. Die Konstanten  $C_1$  und  $C_2$  hängen nur von der Form der Elemente in  $\mathcal{T}$ , nicht aber der Größe (Durchmesser  $h_T$ ) oder der Anzahl der Elemente in  $\mathcal{T}$  ab. Insgesamt ist die  $H^1$ -Halbnorm des Fehlers  $e := u - I_M u_h$  also äquivalent zum Fehlerschätzer. Die Idee bei der adaptiven Netzverfeinerung ist es, die Elemente  $T \in \mathcal{T}$  zu verfeinern, bei denen  $\eta_T$  relativ großen Anteil am Gesamtwert des Fehlerschätzers  $\eta$  hat, vergleiche dazu Algorithmus 4.1.4.

Der Morley-Interpolant ist eine nichtkonforme Approximation der exakten Lösung. Zum Beweis der Zuverlässigkeit wird wie bei der nichtkonformen Finite Elemente Methode die *Helmholtz-Zerlegung*  $\nabla_{\mathcal{T}} e = \nabla \rho + \text{curl } \xi$  mit  $\rho \in H_0^1(\Omega)$  und  $\xi \in H^1(\Omega)$ ,  $\int_{\Omega} \xi \, dx = 0$  angewandt. Dies führt

uns im Beweis auf die Berechnung zweier Integrale, wobei dann die Sprünge der Richtungsableitung in Normal- bzw. in Tangentialrichtung des Morley-Interpolanten auftreten. Auftretende Spurterme über Kanten werden mit der *Spurungleichung*

$$\|v - v_E\|_{L^2(E)} \leq Ch_T^{1/2} \|\nabla v\|_{L^2(T)} \quad (8)$$

abgeschätzt. Dabei ist  $v \in H^1(T)$  ( $v_E$  bezeichnet dabei das Integralmittel von  $v$  über die Kante  $E$ ) und die Konstante  $C > 0$  hängt nur von der Form von  $T$  ab. Die Spurungleichung wird mit Hilfe der Spurgleichung in Lemma 3.4.6 für Rechtecke bewiesen. Der Beweis für Dreiecke ist analog. Anders als in der Literatur betrachten wir hier auf der rechten Seite lediglich die Norm über  $T$  und nicht über den Kantenpatch. Dies ist vor allem bei Triangulierungen von Vorteil, die hängende Knoten zulassen.

Der Beweis der Effizienz wird im Wesentlichen aus dem Kontext der Finite Elemente Methode adaptiert (vergleiche dazu Verfürth [20] für einen Beweis auf regulären Gittern im Kontext der Finite Elemente Methode). Wir brauchen dazu inverse Abschätzungen für Normen, die auf gewissen Bubble-Funktionen basieren. In der Literatur finden sich diese nur für reguläre Triangulierungen. Mit geeigneter Definition der Bubble-Funktionen für fastreguläre Triangulierungen lässt sich auch die Effizienz unseres Residualschätzers beweisen.

Kapitel 4 widmet sich numerischen Experimenten. Zunächst erfolgt eine kurze Übersicht über die verwendete Triangulierung, die adaptive Netzverfeinerung und zugehörige Algorithmen. Drei Beispiele zeigen die Zuverlässigkeit und Effizienz des Residualschätzers bei uniformer und adaptiver Verfeinerung.

## Danksagung

Mein Dank gilt Professor Dr. Dirk Praetorius für die Idee und die ausgezeichnete Betreuung dieser Arbeit und allen, die zum Gelingen dieser Arbeit in irgendeiner Weise beigetragen haben. Außerdem möchte ich meiner Familie danken, die mich während meines Studiums der Technischen Mathematik unterstützt hat.

Schlins, im September 2005  
Christoph Erath

# Kapitel 1

## Triangulierung

Sowohl bei der Finite Elemente Methode als auch bei der Finite Volumen Methode wird das gegebene Gebiet  $\Omega$  in *Teilgebiete* zerlegt. Diese Zerlegung wird als *Triangulierung* bezeichnet. Bevor wir uns der Finite Volumen Methode widmen, führen wir in diesem Kapitel einige wichtige Bezeichnungen und Vereinbarungen bezogen auf die Triangulierung  $\mathcal{T}$  ein.

### 1.1 Notationen

Es sei  $\Omega \subseteq \mathbb{R}^2$  ein beschränktes Lipschitz-Gebiet mit dem Rand  $\Gamma = \partial\Omega$ . Die **Triangulierung**  $\mathcal{T}$  ist eine endliche Menge von paarweise disjunkten, einfach-zusammenhängenden, beschränkten Lipschitz-Gebieten  $T \in \mathcal{T}$ , so dass  $\bar{\Omega} = \bigcup\{\bar{T} | T \in \mathcal{T}\}$  gilt. Außerdem sei angenommen, dass die Elemente  $T \in \mathcal{T}$  (und damit auch  $\Omega$ ) stückweise polygonalen Rand haben.

#### 1.1.1 Knoten- und Kantenmengen, fastreguläre Triangulierung

Mit  $\mathcal{N}$  bezeichnen wir die Menge *aller Knoten* von  $\mathcal{T}$ , und

$$\mathcal{N}_\Gamma := \{a \in \mathcal{N} | a \in \Gamma\}$$

sei die Menge aller *Randknoten*, also Knoten die auf  $\Gamma$  liegen.

**Definition 1.1.1 (Hängender Knoten).** Wenn  $a \in T_1 \cap T_2$  mit  $T_1, T_2 \in \mathcal{T}$  eine Ecke von  $T_1$  aber nicht von  $T_2$  ist, dann heißt  $a \in \mathcal{N}$  *hängender Knoten*. Die Menge aller hängenden Knoten bezeichnen wir mit  $\mathcal{N}^h$ .

Dies führt uns zur Definition der Menge aller nichthängenden inneren Knoten

$$\mathcal{N}^* := \mathcal{N} \setminus (\mathcal{N}^h \cup \mathcal{N}_\Gamma).$$

Die Menge *aller Elementkanten* sei  $\mathcal{E}$ . Mit

$$\mathcal{E}_T := \{E \in \mathcal{E} | E \subseteq \partial T \text{ und } \forall E' \in \mathcal{E}, E' \subseteq \partial T : E \subseteq E' \Rightarrow E = E'\}$$

bezeichnen wir die Menge der Kanten von  $T$ .

**Definition 1.1.2 (Elementarkanten).** Eine Kante  $E \in \mathcal{E}$  heißt **elementar**, falls auf ihr keine hängenden Knoten liegen.

**Definition 1.1.3 (Hängende Kante).** Wir bezeichnen eine Kante  $E \in \mathcal{E}$  als **hängend**, wenn  $E$  nicht elementar ist. Die Menge aller hängenden Kanten wird mit  $\mathcal{E}^h$  bezeichnet.

**Definition 1.1.4 (Fastreguläre Triangulierungen).** Es sei  $\mathcal{T}$  eine Triangulierung von  $\Omega$ , und der Rand  $\partial\Omega$  von  $\Omega$  sei in einen abgeschlossenen Teil  $\Gamma_D$  und  $\Gamma_N := \partial\Omega \setminus \Gamma_D$  zerlegt. Dann heißt  $\mathcal{T}$  **fastreguläre Triangulierung** von  $\Omega$  (bzgl.  $\Gamma_D$  und  $\Gamma_N$ ), wenn

1. für alle  $T \in \mathcal{T}$  gilt  $T \cap \Gamma_D = \emptyset$  oder  $T \cap \Gamma_D \in \mathcal{E}$ ,
2. für verschiedene  $T_1, T_2 \in \mathcal{T}$  ist  $T_1 \cap T_2$  entweder leer, ein Knoten  $a \in \mathcal{N}$  oder eine Kante  $E \in \mathcal{E}$ , d.h.  $E \subseteq \partial T_1$  und  $E \subseteq \partial T_2$ ,
3. jede hängende Kante  $E \in \mathcal{E}^h$  sich darstellen lässt als

$$E = \bigcup_{i=1}^n E_i, \quad E_i \in \mathcal{E}, \quad E_i \text{ elementar}, \quad n \in \mathbb{N}, \quad (1.1)$$

4. jedes  $T \in \mathcal{T}$  nichtleeres Inneres hat, d.h.  $\text{int } T \neq \emptyset$ .

$\Gamma_D$  heißt Dirichlet-Rand und  $\Gamma_N$  Neumann-Rand.

**Bemerkung.**

- Der Name *fastreguläre Triangulierung* ist damit begründet, dass sich Definition 1.1.4 nur im Punkt 2. der Definition einer **regulären Triangulierung** nach Ciarlet (siehe z.B. Carstensen [6]) unterscheidet. Bei einer regulären Triangulierung muss  $E = T_1 \cap T_2$  eine gemeinsame Kante sein, d.h.  $E \in \mathcal{E}_{T_1}$  und  $E \in \mathcal{E}_{T_2}$ . Somit sind keine hängenden Kanten zugelassen und Punkt 3. damit überflüssig.
- Bei einer fastregulären Triangulierung ist sichergestellt, dass ein Randknoten  $a \in \mathcal{N}_\Gamma$  niemals ein hängender Knoten sein kann, da dann Punkt 3. von Definition 1.1.4 verletzt ist. Daher ist eine Kante  $E \subseteq \Gamma$  stets elementar. Vergleiche dazu Abbildung 1.1(a) und beachte, dass  $S$  keine Kante aus  $\mathcal{E}$  ist. Falls  $S$  auf  $\Gamma_D$  liegt, ist auch Punkt 1. nicht erfüllt.
- $\mathcal{O}(\mathcal{T})$  sei das kleinste  $n$ , sodass (1.1) für alle hängenden Kanten gilt. Wir bezeichnen  $\mathcal{O}(\mathcal{T})$  auch als **Ordnung** der fastregulären Triangulierung. Die Ordnung gibt also an, dass jede Kante  $E \in \mathcal{E}^h$  die Vereinigung von maximal  $\mathcal{O}(\mathcal{T})$  Elementarkanten ist. Außerdem liegen auf jeder Kante  $E \in \mathcal{E}^h$  höchstens  $\mathcal{O}(\mathcal{T})+1$  hängende Knoten. Insbesondere haben reguläre Triangulierungen Ordnung 1.

Im Folgenden sei  $\mathcal{T}$  immer als fastreguläre Triangulierung vorausgesetzt, wobei die Elemente  $T \in \mathcal{T}$  Dreiecke oder Rechtecke sind. Wir wollen deshalb noch drei weitere Mengen definieren. Mit

$$\mathcal{E}_\Gamma := \{E \in \mathcal{E} \mid E \subseteq \Gamma\}$$

bezeichnen wir die Menge aller Randkanten. Die Menge der *inneren elementaren Kanten* ist

$$\mathcal{E}^* := \mathcal{E} \setminus (\mathcal{E}^h \cup \mathcal{E}_\Gamma).$$

Außerdem sei

$$\mathcal{E}^\circ := \{E \in \mathcal{E} \setminus \mathcal{E}_\Gamma \mid \nexists E' \in \mathcal{E} \setminus \mathcal{E}_\Gamma : E \subset E'\}.$$

In  $\mathcal{E}^\circ$  sind daher alle inneren Elementarkanten, die nicht Teil von hängenden Kanten sind, sowie alle hängenden Kanten enthalten.

Da wir in dieser Arbeit nur Probleme mit **reinem** Dirichlet-Rand betrachten, gilt

$$\Gamma_D = \Gamma.$$

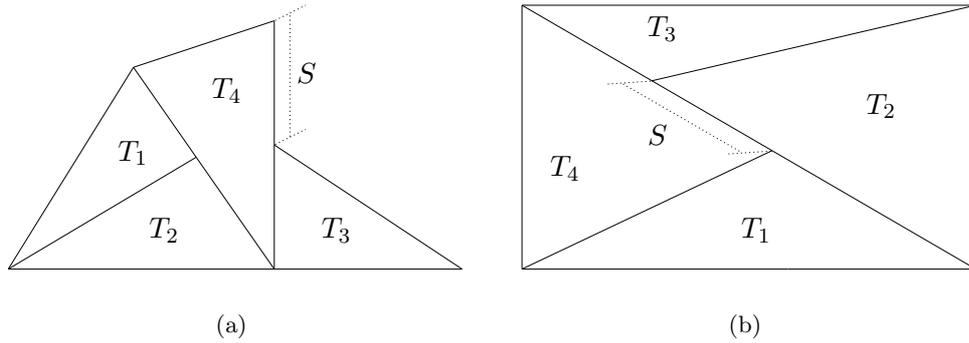


Abbildung 1.1: Zwei einfache Triangulierungen, die beide **nicht** fastregulär sind, da Bedingung 3. von Definition 1.1.4 nicht erfüllt ist (beachte  $S \notin \mathcal{E}$ ).

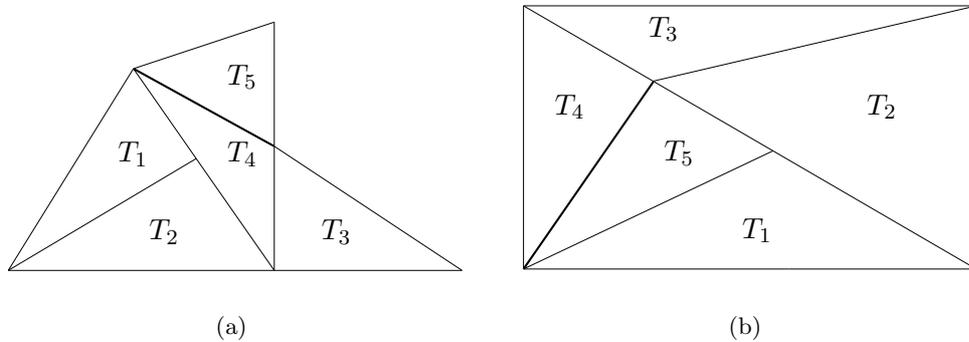


Abbildung 1.2: Durch Hinzunahme einer Kante können aus den nicht fastregulären Triangulierungen aus Abbildung 1.1 fastreguläre Triangulierungen erstellt werden.

### 1.1.2 Patches

#### Knotenpatch

Der Knotenpatch von  $a \in \mathcal{N}$  ist

$$\omega_a := \bigcup \{T \in \mathcal{T} \mid a \subseteq \partial T\},$$

und für die **Elemente** des Knotenpatch von  $a \in \mathcal{N}$  sei

$$\tilde{\omega}_a := \{T \in \mathcal{T} \mid T \subseteq \omega_a\}, \quad \text{d.h.} \quad \omega_a = \bigcup \tilde{\omega}_a.$$

#### Kantenpatch

Der Kantenpatch für  $E \in \mathcal{E} \setminus \mathcal{E}^h$  ist

$$\omega_E := \bigcup \{T \in \mathcal{T} \mid E \subseteq \partial T\}.$$

Für die **Elemente** des Kantenpatch von  $E \in \mathcal{E} \setminus \mathcal{E}^h$  schreiben wir

$$\tilde{\omega}_E := \{T \in \mathcal{T} \mid T \subseteq \omega_E\}.$$

Für  $E \in \mathcal{E}^h$  definiere

$$\omega_E := \bigcup_{i=1}^n \omega_{E_i}, \quad \tilde{\omega}_E = \bigcup_{i=1}^n \tilde{\omega}_{E_i} \quad \text{mit } E = \bigcup_{i=1}^n E_i, \quad E_i \in \mathcal{E}^*.$$

**Elementarer Kantenpatch**

Dazu betrachten wir zunächst eine Kante  $E \in \mathcal{E}^h$ . Dann ist das Element  $T_E$  mit  $E \in \mathcal{E}_{T_E}$  eindeutig bestimmt. Außerdem ist

$$E = \bigcup_{i=1}^n E_i, \quad E_i \in \mathcal{E}^*.$$

$a_{h_j} \in \mathcal{N}^h$  ( $j = 1 \dots n-1$ ) seien die hängenden Knoten, die auf  $E$  liegen,  $a_S$  und  $a_N$  der Anfangs- und Endknoten von  $E$ . O.B.d.A. seien die Anfangs- und Endknoten von  $E_i$  folgendermaßen zugeordnet:

$$a_S, a_{h_1} \text{ von } E_1, \quad a_{h_1}, a_{h_2} \text{ von } E_2, \quad \dots, \quad a_{h_{n-1}}, a_N \text{ von } E_n.$$

Es sei nun  $T_E = \text{conv}\{a_S, a_N, a_G\}$  ein Dreieck, dann gilt (siehe Abbildung 1.3)

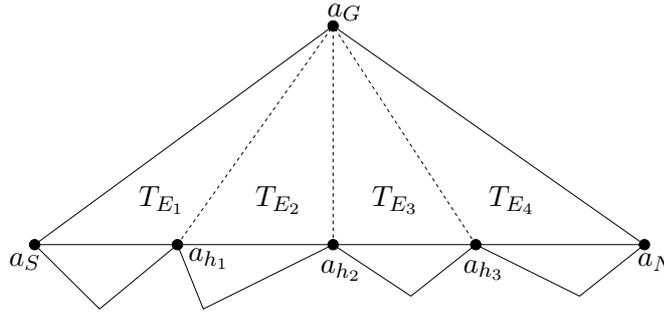


Abbildung 1.3: Beispiel der Dreiecksunterteilung für  $n = 4$ .  $T_{E_i}$  sind keine Elemente aus  $\mathcal{T}$ .

$$T_E = \bigcup_{i=1}^n T_{E_i}, \quad \text{wobei die Dreiecke } T_{E_i} \text{ wie folgt definiert sind:}$$

$$T_{E_1} = \text{conv}\{a_S, a_{h_1}, a_G\}, \quad T_{E_2} = \text{conv}\{a_{h_1}, a_{h_2}, a_G\}, \dots, \quad T_{E_n} = \text{conv}\{a_{h_{n-1}}, a_N, a_G\}.$$

Wenn  $T_E$  ein Rechteck ist, dann gibt es zu jedem hängenden Knoten  $a_{h_j}$  eine Gerade, die orthogonal auf  $E$  steht und  $T_E$  in zwei Rechtecke unterteilt. Es gilt also

$$T_E = \bigcup_{i=1}^n T_{E_i}, \quad T_{E_i} \text{ sind wieder Rechtecke mit der Kante } E_i.$$

**Bemerkung.** Beachte,  $T_{E_i}$  sind keine Elemente aus  $\mathcal{T}$ !

Wir definieren daher den **elementaren Kantenpatch** für die elementaren Kanten  $E_i$ , die Teil der hängenden Kante  $E$  sind, mit

$$\omega_{E_i}^* := T_{E_i} \cup T_W, \quad \tilde{\omega}_{E_i}^* := \{T_{E_i}, T_W\},$$

wobei  $T_W$  mit  $E_i \in \mathcal{E}_{T_W}$  eindeutig bestimmt ist. Für alle anderen Kanten

$$E \in \mathcal{E} \setminus \{E' \mid E' \in \mathcal{E}^*, E' \subset E_h, E_h \in \mathcal{E}^h\} \quad \text{definieren wir:}$$

$$\omega_E^* := \omega_E, \quad \tilde{\omega}_E^* := \tilde{\omega}_E.$$

**Elementpatch**

Der Elementpatch von  $T \in \mathcal{T}$  ist

$$\omega_T := \bigcup \{T' \in \mathcal{T} \mid T \cap T' \in \mathcal{E}\}$$

und analog zu oben

$$\tilde{\omega}_T := \{T \in \mathcal{T} \mid T \subseteq \omega_T\}.$$

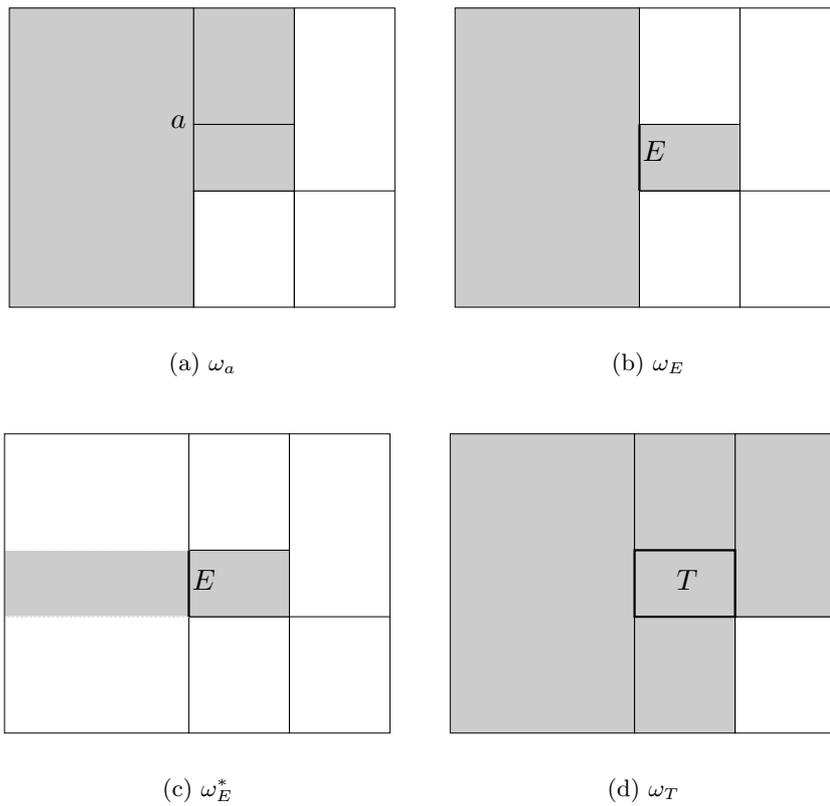


Abbildung 1.4: Die verschiedenen Patches.

**1.1.3 Bezeichnungen anhand einer einfachen Triangulierung**

Anhand der einfachen fastregulären Triangulierung in Abbildung 1.5 zeigen wir, wie sich die oben definierten Mengen zusammensetzen.

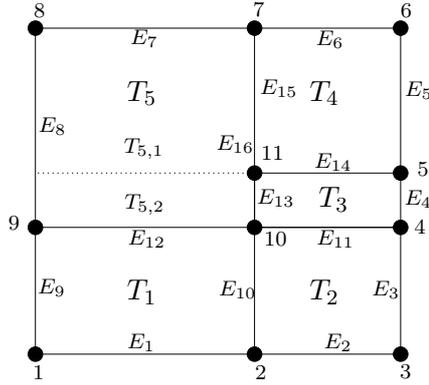


Abbildung 1.5: Beispielt triangulierung mit 5 Elementen.  $T_{5,1}$  und  $T_{5,2}$  sind **keine** Elemente der Triangulierung  $\mathcal{T}$ . Sie dienen lediglich der Veranschaulichung von  $\omega_{E_{13}}^*$  bzw.  $\omega_{E_{15}}^*$ .

$$\begin{aligned} \mathcal{T} &= \{T_1, T_2, T_3, T_4, T_5\}, & T_{5,1} &\subset T_5, & T_{5,2} &\subset T_5, & T_5 &= T_{5,1} \cup T_{5,2}, & T_{5,1}, T_{5,2} &\notin \mathcal{T} \\ \mathcal{N} &= \{1, \dots, 11\}, & \mathcal{N}^h &= \{11\}, & \mathcal{N}_\Gamma &= \{1, \dots, 9\}, & \mathcal{N}^* &= \{10\}, \\ \mathcal{E} &= \{E_1, \dots, E_{16}\}, & \mathcal{E}_\Gamma &= \{E_1, \dots, E_9\}, \\ \mathcal{E}^* &= \{E_{10}, \dots, E_{15}\}, & \mathcal{E}^h &= \{E_{16}\}, & E_{16} &= E_{13} \cup E_{15}, & \mathcal{E}^o &= \{E_{10}, E_{11}, E_{12}, E_{14}, E_{16}\}, \\ \mathcal{E}_{T_1} &= \{E_1, E_9, E_{10}, E_{12}\}, & \mathcal{E}_{T_2} &= \{E_2, E_3, E_{10}, E_{11}\}, & \mathcal{E}_{T_3} &= \{E_4, E_{11}, E_{13}, E_{14}\}, \\ \mathcal{E}_{T_4} &= \{E_5, E_6, E_{14}, E_{15}\}, & \mathcal{E}_{T_5} &= \{E_7, E_8, E_{12}, E_{16}\}. \end{aligned}$$

Die Kante  $E_{16}$  geht also vom Knoten 7 bis zum Knoten 10. Da sie jedoch auch den Knoten 11 enthält, ist sie eine hängende Kante. Für den Kantenpatch geben wir stellvertretend vier an:

$$\omega_{E_1} = T_1, \quad \omega_{E_{10}} = T_1 \cup T_2, \quad \omega_{E_{15}} = T_4 \cup T_5, \quad \omega_{E_{16}} = T_3 \cup T_4 \cup T_5.$$

Der elementare Kantenpatch  $\omega_E^*$  unterscheidet sich nur an den Kanten  $E_{13}$  und  $E_{15}$  vom Kantenpatch  $\omega_E$ :

$$\omega_{E_{13}}^* = T_3 \cup T_{5,2}, \quad \omega_{E_{15}}^* = T_4 \cup T_{5,1}.$$

Der Knotenpatch für den Knoten 10 ist zum Beispiel  $\omega_{10} = T_1 \cup T_2 \cup T_3 \cup T_5$  und der Elementpatch für  $T_3$  ist  $\omega_{T_3} = T_2 \cup T_3 \cup T_4 \cup T_5$ .

#### 1.1.4 Durchmesser $h_T$ und $\varrho_T$

Mit  $h_T := \text{diam}(T)$  bezeichnen wir den Durchmesser von  $T$  und mit  $\varrho_T$  die Höhe von  $T$ . Wenn  $h_E$  die Länge einer Kante von  $T$  ist, gilt  $h_E \leq h_T$  und  $h_T \leq Ch_E$ , wobei die Konstante  $C$  nur von der Form von  $T$  abhängt. Beim Dreieck entspricht  $h_T$  der längsten Seite, dann sei  $\varrho_T$  die Höhe auf  $h_T$ . Es gilt außerdem  $\varrho_T < h_T$ . Annahme  $\varrho_T \geq h_T$ , dann gilt

$$a = \sqrt{\varrho_T^2 + h_{T,1}^2} \geq h_T$$

und damit  $a = h_T := \text{diam}(T)$ , also  $h_{T,1} = 0$ . Dann ist  $h_T = h_{T,2}$ , aber  $b > h_T$ , was zum Widerspruch führt. Beim Rechteck ist  $h_T$  gerade die Diagonale.  $e_2 = \varrho_T$  sei die kürzere Seite.

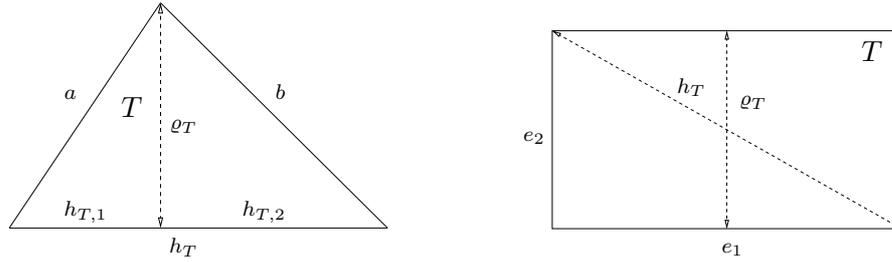


Abbildung 1.6: Der Durchmesser  $h_T$  und die Höhe  $\varrho_T$ .

Dann gelten die Abschätzungen  $\varrho_T < h_T$ ,  $e_2 \leq e_1$ ,  $e_1 \leq h_T$  und  $h_T^2 = e_1^2 + e_2^2 \leq 2e_1^2$  und somit  $h_T \leq \sqrt{2}e_1$ . Die Größe

$$\kappa_T := \frac{h_T}{\varrho_T} > 1$$

beschreibt die *Form* des Elements  $T$ . Wird z.B. bei einem Dreieck ein Winkel immer kleiner, so wird auch  $\varrho_T$  kleiner und somit wächst  $\kappa_T$ . Wenn bei einem Rechteck eine Seite kleiner wird, verkleinert sich ebenfalls  $\varrho_T$  und  $\kappa_T$  wächst. Es sei nun  $h := \max_{T \in \mathcal{T}} h_T$ . Dann heißt eine Folge  $\mathcal{T}_h$  von Triangulierungen **formregulär** (engl. shape regular), wenn  $\kappa_T = h_T/\varrho_T$  für alle  $T \in \mathcal{T}_h$  und alle  $h > 0$  nach oben begrenzt bleibt. In dieser Arbeit werden wir Abschätzungen beweisen, die eine von der Form der Elemente abhängige Konstante enthalten. Diese hängen somit von  $\kappa_T$  ab,  $\kappa_T$  muss also beschränkt bleiben. Zu  $\mathcal{T}$  heißt

$$\kappa_{\mathcal{T}} := \max_{T \in \mathcal{T}_h} \kappa_T$$

„formreguläre“ Konstante.

## 1.2 Wahl der Normalvektoren und Tangentialvektoren

Im Laufe dieser Arbeit benötigen wir Normal- und Tangentialvektoren auf einer Kante. Es ist günstig, diese für die Triangulierung zu fixieren.

### 1.2.1 Normalvektor $\mathbf{n}_E$ und Tangentialvektor $\mathbf{t}_E$

A priori legen wir eine Richtung für den **Normalvektor**  $\mathbf{n}_E$  auf der **Elementarkante**  $E \in \mathcal{E}^* \cup \mathcal{E}_\Gamma$  fest. Dabei haben wir bei inneren Elementarkanten  $E \in \mathcal{E}^*$ , die nicht Teil einer hängenden Kante sind, freie Wahl. Für Randkanten  $E \in \mathcal{E}_\Gamma$  sei  $\mathbf{n}_E$  der nach außen gerichtete Normalvektor. Existiert zu einer Elementarkante  $E \in \mathcal{E}^*$  eine hängende Kante  $E_h \in \mathcal{E}^h$  mit  $E \subseteq E_h$ , dann gibt es eindeutige Elemente  $T_W$  und  $T_E$ , so dass  $E_h$  eine Kante von  $T_E$  und  $E$  eine Kante von  $T_W$  ist. Als Normalvektor  $\mathbf{n}_E$  auf  $E$  sei derjenige gewählt, der vom Element  $T_W$  zum Element  $T_E$  zeigt (**W**est nach **E**ast).

Ist  $E \in \mathcal{E}^h$  eine **hängende Kante** und sind  $E_i \in \mathcal{E}^*$  ( $1 \leq i \leq n$ ) Elementarkanten mit  $E = \bigcup_{i=1}^n E_i$ , so existieren eindeutige Elemente  $T_{W_i}, T_E \in \mathcal{T}$  mit

$$E_i = \overline{T_{W_i}} \cap \overline{T_E} \quad \text{für alle } 1 \leq i \leq n.$$

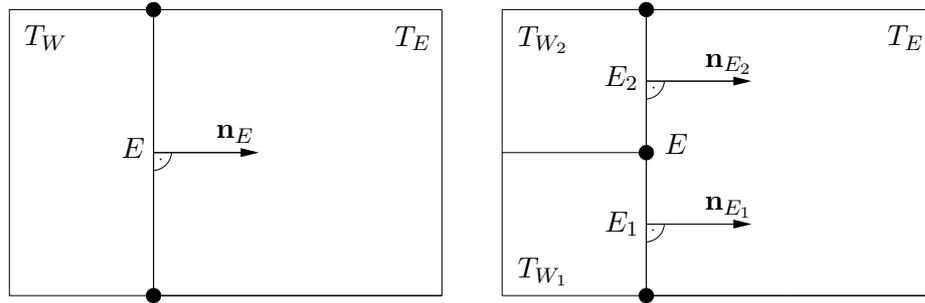


Abbildung 1.7: Wahl der Normalvektoren im Fall nicht hängender Knoten links und einem hängendem Knoten ( $n = 2$ ).

Die gewählten Normalvektoren  $\mathbf{n}_{E_i}$  stimmen dann überein und zeigen alle zum „großen“ Element  $T_E$ . Abbildung 1.7 skizziert die Situation für eine Kante ohne hängenden Knoten und eine Kante mit hängendem Knoten ( $n = 2$ ).

Der Normalvektor  $\mathbf{n}_E$  von  $E \in \mathcal{E}^*$  zeigt also immer von  $T_W$  nach  $T_E$ . Für eine Randkante  $E \in \mathcal{E}_\Gamma$  zeigt der Normalvektor nach Definition immer nach außen, somit bezeichnen wir das dazugehörige Element mit  $T_W$ . Auf diese Vereinbarung werden wir später immer wieder zurückkommen.

Den Tangentialvektor  $\mathbf{t}_E$  einer Kante  $E$  wählen wir orthogonal zu  $\mathbf{n}_E$ , und zwar mathematisch positiv.

### 1.2.2 Vorzeichen $\sigma_{T,E}$ , $\mathbf{n}_{T,E}$ und $\mathbf{t}_{T,E}$

In der Analysis kommt es aber auch vor, dass wir eine Kante  $E \in \mathcal{E}_T$  betrachten und dabei den nach außen gerichteten Normalvektor  $\mathbf{n}_{T,E} := \mathbf{n}_T|_E$  bezüglich  $\partial T$  benötigen. Da wir nun aber den Normalvektor  $\mathbf{n}_E$  zur Kante  $E$  schon fixiert haben, kann dieser unter Umständen nach innen zeigen. In Formeln führen wir deshalb das Vorzeichen  $\sigma_{T,E} = \pm 1$  ein, das so gewählt wird, dass  $\sigma_{T,E}\mathbf{n}_E = \mathbf{n}_{T,E}$  die äußere Normale auf  $E$  bezüglich  $T$  ist.

Wir legen deshalb fest:

$$\sigma_{T,E} = \begin{cases} 1, & \text{falls } \mathbf{n}_E \text{ nach außen zeigt auf der Kante } E \in \mathcal{E}_T \text{ bezogen auf } T, \\ -1, & \text{falls } \mathbf{n}_E \text{ nach innen zeigt auf der Kante } E \in \mathcal{E}_T \text{ bezogen auf } T. \end{cases} \quad (1.2)$$

**Bemerkung.** Wollen wir die Richtungsableitung von  $u \in H^1(\Omega)$  in Richtung des nach außen gerichteten Normalvektors der Kante  $E$  von  $T$  berechnen, so gilt

$$\frac{\partial u}{\partial(\sigma_{T,E}\mathbf{n}_E)} = \sigma_{T,E} \frac{\partial u}{\partial \mathbf{n}_E}.$$

Wird der Tangentialvektor  $\mathbf{t}_{T,E} := \mathbf{t}_T|_E$  auf einer Kante des Elements  $T$  gebraucht, erfolgt die Multiplikation ebenfalls mit  $\sigma_{T,E}$ . Da wir per Definition ein Element  $T$  immer mathematisch positiv umlaufen, zeigt  $\sigma_{T,E}\mathbf{t}_E = \mathbf{t}_{T,E}$  genau die Umlaufrichtung an. Die Endpunkte  $a_S$  und  $a_N$  von  $E \in \mathcal{E}_T$  erfüllen die Eigenschaft

$$(a_N - a_S) \cdot \mathbf{t}_E > 0$$

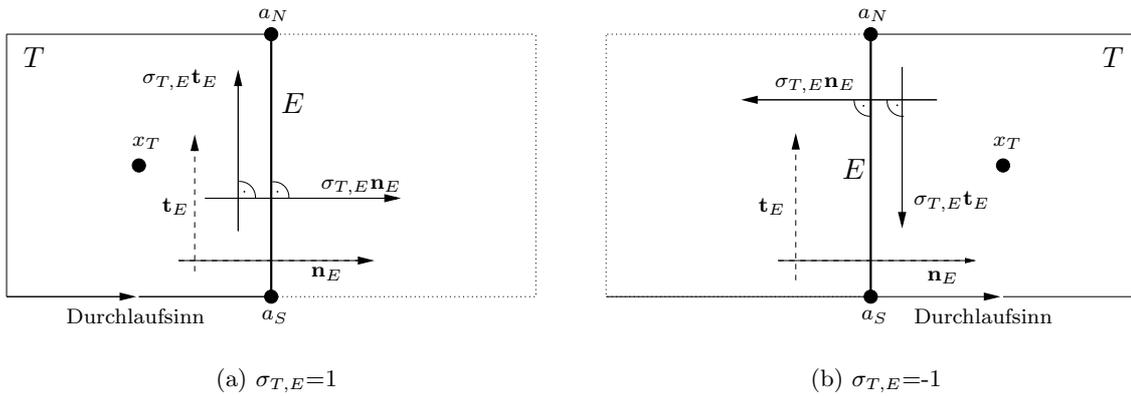


Abbildung 1.8: Die strichlierten Vektoren zeigen die a priori gewählten Normal- bzw. Tangentialvektoren, die eingezeichneten Vektoren geben nur die Richtung an, die Position ist dabei nur symbolisch zu verstehen.

und werden somit ebenfalls a priori für jede Kante  $E \in \mathcal{E}^* \cup \mathcal{E}_\Gamma$  gewählt, d.h.  $\mathbf{t}_E$  zeigt von  $a_S$  nach  $a_N$ . Die Abbildung 1.8 veranschaulicht die Wirkung von  $\sigma_{T,E}$ .

## Kapitel 2

# Die Finite Volumen Methode

Die Finite Volumen Methode wird anhand eines zweidimensionalen konvektiven-diffusiven Problems hergeleitet. Das Schema führt uns zu einer Bilanzgleichung von Flüssen über eine Kante, die die *Konservativitätseigenschaft* erfüllen. Die Konvergenz des Verfahrens hängt entscheidend von der Approximation dieser konvektiven und diffusiven Flüsse ab. Das sogenannte *Upwind-Verfahren* und die *Diamond Path Methode* garantieren nach Coudière und Villedieu [9], dass das Verfahren für eine hinreichend glatte exakte Lösung  $u \in H^2(\Omega)$  konvergiert. Die Elemente  $T \in \mathcal{T}$  seien in diesem Kapitel nicht ausschließlich Dreiecke oder Rechtecke.

### 2.1 Das Grundprinzip der Diskretisierung

Die Methode wird anhand folgender Aufgabe hergeleitet:  
Finde ein  $u \in H^1(\Omega)$ , so dass

$$\begin{aligned} \operatorname{div}(-\nu \nabla u + u \mathbf{v}) &= f & \text{auf } \Omega, \\ u &= u_D & \text{auf } \Gamma = \partial\Omega, \end{aligned} \tag{2.1}$$

mit  $\nu > 0$ ,  $\nu \in C^2(\overline{\Omega})$ ,  $\mathbf{v} = (v_x, v_y) \in C^1(\overline{\Omega}, \mathbb{R}^2)$ ,  $f \in L^2(\Omega)$  und  $u_D \in H^{1/2}(\Gamma)$  gilt.

Für die Approximation durch die Finite Volumen Methode wird das Gebiet  $\Omega$  zunächst in eine endliche Anzahl paarweiser disjunkter *finiter Volumina* (Kontrollvolumina)  $T \in \mathcal{T}$  unterteilt.  $\mathcal{T}$  sei dabei eine fastreguläre Triangulierung.

**Bemerkung.** Es sei darauf hingewiesen, dass die Methode **nicht** ausschließlich auf fastreguläre Triangulierungen angewendet werden kann. Die Diskretisierung bezüglich anderer Triangulierungen kann jedoch anhand dieser Herleitung rasch nachvollzogen werden.

Das Grundprinzip der Finite Volumen Methode besteht nun darin, Gleichung (2.1) über jedes Element  $T \in \mathcal{T}$  zu integrieren und den Gaußschen Integralsatz anzuwenden. Es gilt

$$\begin{aligned} \int_T f(x) dx &= \int_T \operatorname{div}(-\nu \nabla u + u \mathbf{v}) dx = \int_{\partial T} (-\nu \nabla u + u \mathbf{v}) \cdot \mathbf{n}_T ds \\ &= \sum_{E \in \mathcal{E}_T} \int_E (-\nu \nabla u + u \mathbf{v}) \cdot (\sigma_{T,E} \mathbf{n}_E) ds \end{aligned} \tag{2.2}$$

für alle  $T \in \mathcal{T}$ , wobei  $\mathbf{n}_T$  der nach außen gerichtete Normalvektor bezüglich  $\partial T$  ist. Nach (1.2) gilt dann  $\mathbf{n}_T|_E = \sigma_{T,E} \mathbf{n}_E$  für jede Kante  $E \in \mathcal{E}_T$  von  $T$ .

Wir bezeichnen mit

$$\begin{aligned}\Phi_E^C(u) &= \int_E u \mathbf{v} \cdot \mathbf{n}_E ds \quad \text{den konvektiven Fluss und mit} \\ \Phi_E^D(u) &= \int_E (\nu \nabla u) \cdot \mathbf{n}_E ds \quad \text{den diffusiven Fluss}\end{aligned}$$

und schreiben (2.2) in der Form

$$\sum_{E \in \mathcal{E}_T} \sigma_{T,E} (\Phi_E^C(u) - \Phi_E^D(u)) = \int_T f(x) dx \quad \forall T \in \mathcal{T}. \quad (2.3)$$

Diese Gleichung wird auch als **Bilanzgleichung** bezeichnet.

**Konservativität.** Betrachten wir eine Kante  $E = \overline{T_W} \cap \overline{T_E}$ , so ist der Gesamtfluss  $\sigma_{T_W,E}(\Phi_E^C(u) - \Phi_E^D(u))$  bezogen auf  $\mathbf{n}_{T_W,E}$  und  $\sigma_{T_E,E}(\Phi_E^C(u) - \Phi_E^D(u))$  bezogen auf  $\mathbf{n}_{T_E,E}$  durch diese Kante für beide Kontrollvolumina  $T_W$  und  $T_E$  betragsmäßig gleich. Lediglich das Vorzeichen unterscheidet sich, welches durch  $\sigma_{T_W,E}$  bzw.  $\sigma_{T_E,E}$  bestimmt wird. Dieses Vorzeichen gibt jedoch an, ob der Gesamtfluss ins Element reinfließt, oder vom Element wegfließt. Somit gilt auf einer Kante  $E$  immer, dass der Fluss von einem Element ins andere Element fließt. Wir sprechen von der *Konservativität* des Verfahrens.

Es sei darauf hingewiesen, dass wir bis zu dieser Stelle noch keine einzige Approximation durchgeführt haben. Wir suchen nun eine auf jedem Element  $T \in \mathcal{T}$  stückweise konstante Funktion

$$u_h(x)|_T := u_T \quad \forall x \in T, \quad (2.4)$$

wobei  $u_T$  die Approximation des Integralmittels von  $u$  auf  $T$  darstellt, d.h.  $u_T \approx 1/|T| \int_T u dx$ . Der konvektive Fluss  $\Phi_E^C(u)$  über eine Kante  $E \in \mathcal{E}^* \cup \mathcal{E}_\Gamma$  wird dazu durch einen numerischen Fluss  $F_E^C(u_h)$  ersetzt. Den diffusiven Fluss  $\Phi_E^D(u)$  über eine Kante  $E \in \mathcal{E}^* \cup \mathcal{E}_\Gamma$  approximieren wir durch einen numerischen Fluss  $F_E^D(u_h)$ . Der numerische Fluss über eine hängende Kante  $E \in \mathcal{E}^h$  berechnet sich dann als Summe der Teilflüsse der dazugehörigen Elementarkanten:

$$F_E^C(u_h) = \sum_{i=1}^n F_{E_i}^C(u_h), \quad F_E^D(u_h) = \sum_{i=1}^n F_{E_i}^D(u_h) \quad (2.5)$$

$$\text{für } E \in \mathcal{E}^h, \quad E = \bigcup_{i=1}^n E_i, \quad E_1, \dots, E_n \in \mathcal{E}^*, \quad (2.6)$$

was durch die entsprechende Additivität

$$\begin{aligned}\Phi_E^C(u) &= \sum_{i=1}^n \Phi_{E_i}^C(u), \quad \Phi_E^D(u) = \sum_{i=1}^n \Phi_{E_i}^D(u) \\ \text{für } E \in \mathcal{E}^h, \quad E &= \bigcup_{i=1}^n E_i, \quad E_1, \dots, E_n \in \mathcal{E}^*\end{aligned}$$

motiviert wird. Es genügt daher, die eigentliche Approximation der Flüsse über die Kanten aus  $\mathcal{E}^* \cup \mathcal{E}_\Gamma$  zu betrachten. Beide numerischen Flüsse werden dabei so gewählt, dass die Konservativität erhalten bleibt. Somit erhalten wir die folgende Finite Volumen Diskretisierung:

Finde die Lösung  $u_h \in \mathcal{P}_0(\mathcal{T})$  für  $T \in \mathcal{T}$  von

$$\sum_{E \in \mathcal{E}_T} \sigma_{T,E} (F_E^C(u_h) - F_E^D(u_h)) = \int_T f dx \quad \forall T \in \mathcal{T}. \quad (2.7)$$

Bezeichnet  $N := \#\mathcal{T}$  die Anzahl der Elemente, so führt (2.7) auf ein lineares Gleichungssystem mit der Systemmatrix  $A \in \mathbb{R}^{N \times N}$ .

**Bemerkung.** Wir betrachten hier ein reines Dirichlet-Problem. Für Probleme mit zusätzlichem Neumann-Rand müssen nur die Flussapproximationen, die in den nächsten zwei Abschnitten vorgestellt werden, entsprechend modifiziert werden. Die eigentliche Diskretisierung ändert sich nicht. Siehe dazu Coudière und Villedieu [9].

**Fazit.** Die Finite Volumen Methode projiziert  $u$  auf eine zu den Elementen  $T \in \mathcal{T}$  stückweise konstante Funktion  $u_h$ .

## 2.2 Der konvektive numerische Fluss $F_E^C(u_h)$

Für die Approximation von  $\Phi_E^C(u) = \int_E u \mathbf{v} \cdot \mathbf{n}_E ds$  entlang  $E$  wenden wir das sogenannte **Upwind-Verfahren** an. Die prinzipielle Idee ist es, die Approximation abhängig von der Richtung des Geschwindigkeitsvektors  $\mathbf{v}$  zu machen. Damit nützen wir die Transporteigenschaft von Konvektions-Diffusions-Problemen aus, die besagt, dass der konvektive Transport von  $u$  nur „stromabwärts“ erfolgt.

Auf einer Kante  $E$  soll der bereits festgelegte Normalvektor  $\mathbf{n}_E$  von  $T_W$  nach  $T_E$  zeigen. Es ist  $\mathbf{v} \cdot \mathbf{n}_E$  der Anteil des Geschwindigkeitsvektors  $\mathbf{v}$  in  $\mathbf{n}_E$  Richtung. Wir berechnen den Mittelwert dieses Anteils auf jeder Kante  $E$

$$v_E = \frac{1}{h_E} \int_E \mathbf{v} \cdot \mathbf{n}_E ds. \quad (2.8)$$

Beim **Upwind-Verfahren erster Ordnung** ordnen wir  $u$  entlang der Kante  $E$  einen Wert unserer Finite Volumen Lösung  $u_h$  in Abhängigkeit der Richtung von  $\mathbf{v}$  (genauer: in Abhängigkeit von  $v_E$ ) zu und können somit das Integral lösen. Somit ergibt sich

$$\forall E \in \mathcal{E}^* \quad F_E^C(u_h) = \begin{cases} h_E \mathbf{v} \cdot \mathbf{n}_E u_W & \text{falls } v_E \geq 0, \\ h_E \mathbf{v} \cdot \mathbf{n}_E u_E & \text{sonst,} \end{cases} \quad (2.9)$$

$$\forall E \in \mathcal{E}_\Gamma \quad F_E^C(u_h) = \begin{cases} h_E \mathbf{v} \cdot \mathbf{n}_E u_W & \text{falls } v_E \geq 0, \\ h_E \mathbf{v} \cdot \mathbf{n}_E g(x_E) & \text{sonst.} \end{cases} \quad (2.10)$$

Im zweiten Fall von  $E \in \mathcal{E}_\Gamma$  reduziert sich  $T_E$  auf die Kante  $E$ , deshalb nehmen wir  $u = g(x_E)$ , wobei  $x_E$  der Mittelpunkt von  $E$  ist. Beachte, dass  $\mathbf{n}_E$  auf Randkanten nach Definition immer nach außen zeigt und somit  $T_W$  existiert.

## 2.3 Der diffusive numerische Fluss $F_E^D(u_h)$

Für die Approximation beim diffusiven Fluss wird  $u_T$  als diskrete Variable betrachtet, welche  $u$  in einem Punkt des Kontrollvolumens  $T$  approximiert. Da wir genau eine diskrete Unbekannte  $u_T$  pro Kontrollvolumen betrachten, die innerhalb von  $T$  liegt, sprechen wir auch von der *zellenorientierten* (engl. cell-centered) Finite Volumen Methode<sup>1</sup>.

Für die Berechnung von  $F_E^D(u_h)$  müssen wir zunächst die Richtungsableitung  $\nabla u \cdot \mathbf{n}_E = \partial u / \partial \mathbf{n}_E$  auf  $E$  mit Hilfe der diskreten Unbekannten  $u_T$  annähern.

<sup>1</sup>Im Unterschied dazu sei auf die eckenorientierte (engl. cell vertex) Methode verwiesen. Hier wird auf zwei verschiedenen Triangulierungen gerechnet. In dieser Arbeit beschäftigen wir uns allerdings ausschließlich mit der zellenorientierten Methode, weshalb wir auf die eckenorientierte Methode nicht näher eingehen wollen.

### 2.3.1 Die Differenzen Methode

Eine einfache Approximation der Richtungsableitung ist

$$\frac{\partial u}{\partial \mathbf{n}_E} \approx \frac{u_E - u_W}{d(x_W, x_E)}, \quad (2.11)$$

wobei  $E = \overline{T_W} \cap \overline{T_E}$ , und  $d(x_W, x_E)$  der Abstand zwischen den Mittelpunkten  $x_W$  und  $x_E$  der Volumina  $T_W$  und  $T_E$  ist.  $u_W$  bzw.  $u_E$  ist dabei die Approximation von  $u(x_W)$  bzw.  $u(x_E)$ . Handelt es sich um eine Randkante, so wird anstelle von  $u_E$  ein Wert von  $g(x)$  mit  $x \subseteq E$  genommen.  $d(x_W, \Gamma)$  ist entsprechend der Abstand von  $x_W$  zum Rand. Nach dem Handbook of Numerical Analysis [10] ist diese Flussapproximation für allgemeine Triangulierungen **nicht konsistent**. Eine konsistente Approximation kann nur dann erreicht werden, wenn die Unbekannten  $u_W$  und  $u_E$  auf einer orthogonalen Linie zur Kante  $E$  liegen. Wollen wir für den diffusiven Fluss Approximation (2.11) verwenden, muss die Triangulierung zulässig im folgenden Sinn sein:

**Definition 2.3.1 (Zulässige Triangulierung).** Eine zulässige Triangulierung  $\mathcal{T}$  von  $\Omega$  mit den Punkten  $\mathcal{M}$ , wobei jedes Kontrollvolumen  $T$  genau einen Punkt  $x_T \in \mathcal{M}$  enthält, erfüllt folgende Eigenschaften:

1.  $\bigcup_{T \in \mathcal{T}} \overline{T} = \overline{\Omega}$
2. Für alle Kontrollvolumen  $T_1$  und  $T_2$  ist  $\overline{T_1} \cap \overline{T_2}$  entweder leer, ein Punkt oder eine gemeinsame Kante  $E \in \mathcal{E}$ .
3. Es seien  $x_W, x_E \in \mathcal{M}$  mit  $x_W \in T_W$  und  $x_E \in T_E$ ,  $T_W, T_E \in \mathcal{T}$ . Falls  $\overline{T_W} \cap \overline{T_E} := E \in \mathcal{E}$  eine Kante ist, so ist die Verbindung  $x_{T_W}$  und  $x_{T_E}$  orthogonal zu  $E$  (Abbildung 2.1).
4. Ist  $E \in \mathcal{E}_\Gamma \cap \mathcal{E}_T$  eine Kante auf dem Rand, so gilt  $\delta_E \cap E \neq \emptyset$ . Dabei sei  $\delta_E$  die Linie zwischen  $x_T$  und  $E$ , wobei  $\delta_E$  orthogonal auf  $E$  steht (Abbildung 2.1).

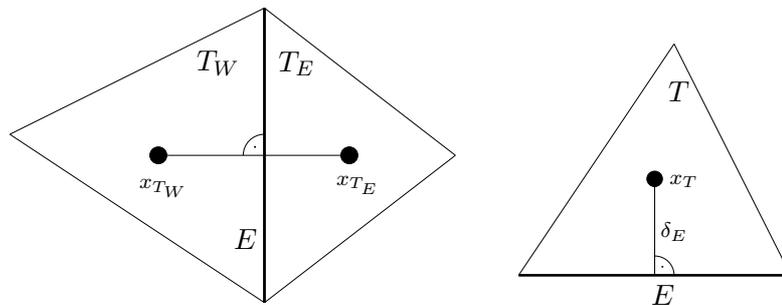


Abbildung 2.1: Veranschaulichung der Bedingungen 3. und 4. von Definition 2.3.1.

Die ersten beiden Bedingungen werden auch als Regularität der Triangulierung nach Ciarlet bezeichnet. Diese lässt keine hängenden Knoten zu. Auch die Bedingungen 3. und 4. stellen eine massive Einschränkung in der Wahl der Triangulierungselemente dar. Eine Triangulierung aus lauter Dreiecke erfüllt die beiden ersten Bedingungen. Damit auch die Bedingungen 3. und 4. erfüllt sind, müssen die Innenwinkel aller Dreiecke echt kleiner  $\pi/2$  bleiben. Denn als  $x_T$  kommt nur der Umkreismittelpunkt in Frage, der durch den Schnittpunkt der Seitensymmetralen gebildet wird. Dieser liegt jedoch nur innerhalb von  $T$ , wenn alle Winkel von  $T$  echt kleiner  $\pi/2$  sind. Wählen wir für  $x_T$  einen anderen Punkt im Dreieck, so ist es unmöglich, zumindest Bedingung 3. zu erfüllen.

### 2.3.2 Diamond Path Methode

Im Hinblick auf adaptive Verfeinerung sind die Einschränkungen aus Definition 2.3.1 sehr unangenehm. Es ist z.B. schwer, eine Dreieckstriangulierung lokal zu verfeinern, ohne dass ein Winkel eines Dreiecks größer gleich  $\pi/2$  wird. Auch die lokale Verfeinerung von Rechteckstriangulierungen ist nahezu unmöglich. Verfeinern wir von einer zulässigen Rechteckstriangulierung ein Rechteck, so entstehen hängende Knoten. Somit ist die diffusive Flussapproximation nach der Differenzen Methode für uns nicht geeignet.

Wollen wir daher den diffusiven Fluss über eine Kante für allgemeine Triangulierungen approximieren, so müssen wir den ganzen Gradienten auf jeder Kante  $E \in \mathcal{E}^*$  betrachten. In der Literatur werden dazu die Konstruktionsmethoden prinzipiell in zwei Klassen unterteilt, die eine ist bekannt als **Green-Gauss-Klasse**, die andere als **polynomiale Lagrange-Interpolations-Klasse**. Die Methoden in beiden Klassen verwenden für die Approximation mehr Kontrollvolumen als nur die zwei, die eine gemeinsame Kante enthalten. In diesem Abschnitt beschränken wir uns auf die **Diamond Path Methode**. Sie ist vom Green-Gauss-Typ.

Der Diamond Path wird dabei durch die zwei Mittelpunkte  $x_W$  und  $x_E$  der beiden Volumen  $T_W$  und  $T_E$  ( $E = \overline{T_W} \cap \overline{T_E}$ ) und der beiden Endknoten ( $x_S := a_S$  und  $x_N := a_N$ ) der Kante  $E$  gebildet. Ist  $E \in \mathcal{E}_\Gamma$  eine Randkante, so ist  $x_E = m_E$ , wobei  $m_E$  der Mittelpunkt der Kante  $E$  ist. Abbildung 2.2 veranschaulicht die Wahl.

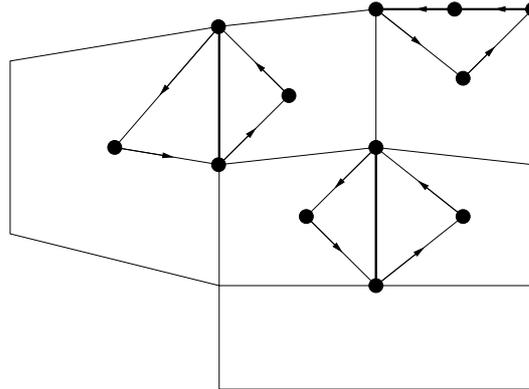


Abbildung 2.2: Drei verschiedene Diamond Pathes.

**Einschränkung.**  $\mathcal{T}$  sei von nun an wieder eine fastreguläre Triangulierung. Außerdem fordern wir, dass der Diamond Path  $\chi_E := \text{conv}\{x_E, x_S, x_W, x_N\}$  **konvex** ist.

Wir führen folgende Approximationen von  $u$  ein:

$$\begin{aligned} u_W \approx u(x_W) \text{ und } u_E \approx u(x_E) \text{ mit } u_W &= u_h|_W, & u_E &= u_h|_E, \\ u_N \approx u(x_N), & u_S \approx u(x_S). \end{aligned}$$

Falls  $E \in \mathcal{E}_\Gamma$  eine Randkante ist, gilt  $u_E = g(m_E)$ . Der Diamond Path selbst arbeitet lokal (in Bezug auf  $E$ ). Es gilt aber zu beachten, dass  $u_N$  und  $u_S$  ebenfalls unbekannte Werte sind. Diese werden durch lineare Interpolation der Werte  $u_T = u_h|_T$  mit  $T \in \tilde{\omega}_{x_N}$  bzw.  $T \in \tilde{\omega}_{x_S}$  dargestellt. Wir schreiben daher allgemein

$$\begin{aligned} \forall a \in \mathcal{N}^* \cup \mathcal{N}^h & \quad u_a = \sum_{T \in \tilde{\omega}_a} \psi_T(a) u_T, \\ \forall a \in \mathcal{N}_\Gamma & \quad u_a = g(a). \end{aligned}$$

Zur Berechnung der Gewichte  $\psi_T(a)$  vergleiche Abschnitt 2.3.4. Bei der Approximation sind somit sämtliche Elemente  $T \in \mathcal{T}$  mit  $E \cap T \neq \emptyset$  für  $T \in \mathcal{T}$  einbezogen.

In Abschnitt 2.3.3 wird hergeleitet, dass die Näherung des Gradienten  $\nabla u$  auf  $E \in (\mathcal{E}^* \cup \mathcal{E}_\Gamma)$  mittels der Diamond Path Methode

$$\begin{aligned} \nabla u_E &= \left[ \frac{u_E - u_W}{d_E} - \alpha_E \frac{u_N - u_S}{h_E} \right] \mathbf{n}_E + \frac{u_N - u_S}{h_E} \mathbf{t}_E, \\ \alpha_E &= \frac{(x_E - x_W) \cdot \mathbf{t}_E}{(x_E - x_W) \cdot \mathbf{n}_E}, \quad d_E = (x_E - x_W) \cdot \mathbf{n}_E \end{aligned} \quad (2.12)$$

ergibt.  $\nabla u_E$  ist dabei in den Basisvektoren  $(\mathbf{n}_E, \mathbf{t}_E)$  dargestellt. Für die Richtungsableitung erhalten wir

$$\frac{\partial u_E}{\partial \mathbf{n}_E} = \nabla u_E \cdot \mathbf{n}_E = \frac{u_E - u_W}{d_E} - \alpha_E \frac{u_N - u_S}{h_E}.$$

**Bemerkung.** Liegt die Verbindung der Punkte  $x_W$  und  $x_E$  orthogonal auf  $E$ , so ist  $\alpha_E = 0$ . Das führt zur exakt gleichen Näherung für  $\partial u / \partial \mathbf{n}_E$  wie bei (2.11).

Für die Approximation von  $\Phi_E^D(u)$  erhalten wir somit den numerischen diffusiven Fluss

$$\begin{aligned} F_E^D(u_h) &= h_E \nu_E \left( \frac{u_E - u_W}{d_E} - \alpha_E \frac{u_N - u_S}{h_E} \right) \\ \text{mit } \nu_E &= \frac{1}{h_E} \int_E \nu \, ds, \quad \alpha_E = \frac{(x_E - x_W) \cdot \mathbf{t}_E}{(x_E - x_W) \cdot \mathbf{n}_E}, \\ d_E &= (x_E - x_W) \cdot \mathbf{n}_E. \end{aligned} \quad (2.13)$$

### 2.3.3 Herleitung der Diamond Path Methode

In diesem Abschnitt wird gezeigt, wie  $\nabla u$  auf einer Kante  $E \in (\mathcal{E}^* \cup \mathcal{E}_\Gamma)$  mit der Diamond Path Methode approximiert wird, d.h.

$$\nabla u|_E \approx \nabla u_E \in \mathbb{R}^2 \quad \text{für } E \in \mathcal{E}^*.$$

Die Bezeichnungen werden aus Abbildung 2.3 entnommen. Wie oben schon erwähnt, fordern wir von unserer fastregulären Triangulierung  $\mathcal{T}$ , dass das sogenannten **Co-Volumen**  $\chi_E := \text{conv}\{x_E, x_N, x_W, x_S\}$  *konvex* ist. Außerdem gilt für  $E \in \mathcal{E}_\Gamma$   $x_E = m_E$  und  $u_E = g(m_E)$  ( $m_E$  ist Mittelpunkt der Kante  $E$ ). Das Integralmittel des Gradienten  $\nabla u$  auf  $\chi_E$  lässt sich mit Hilfe der partiellen Integration auf ein Randintegral über  $\partial \chi_E$  reduzieren.

$$\nabla u_{\chi_E} = \frac{1}{|\chi_E|} \int_{\chi_E} \nabla u \, dx = \frac{1}{|\chi_E|} \int_{\partial \chi_E} u \mathbf{n}_{\chi_E} \, dx, \quad (2.14)$$

wobei  $\mathbf{n}_{\chi_E}$  der auf  $\partial \chi_E$  nach außen gerichtete Normalvektor ist. Um nun das rechte Integral zu approximieren, wird  $u$  auf einer Kante  $e_i$  von  $\partial \chi_E$  einfach durch den Mittelwert der Werte an den Endpunkten ersetzt. Auf der Kante  $e_1$  z.B. nähern wir mit  $(u_w + u_s)/2$  an. Für die Approximation von  $\nabla u$  auf  $\chi_E$  und somit für die Approximation  $\nabla u_E$  von  $\nabla u$  auf  $E$  erhalten wir

$$\nabla u_E = \frac{1}{|\chi_E|} \left[ \frac{u_W + u_S}{2} h_{e_1} \mathbf{n}_{e_1} + \frac{u_S + u_E}{2} h_{e_2} \mathbf{n}_{e_2} + \frac{u_E + u_N}{2} h_{e_3} \mathbf{n}_{e_3} + \frac{u_N + u_W}{2} h_{e_4} \mathbf{n}_{e_4} \right], \quad (2.15)$$

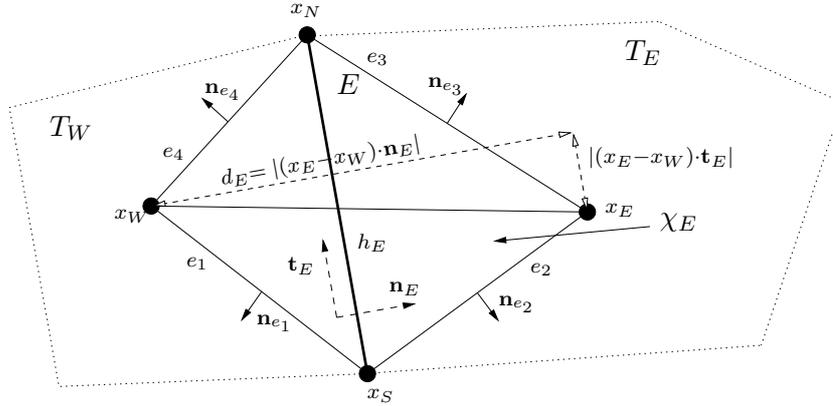


Abbildung 2.3: Diamond Path mit Fläche  $\chi_E$ , die punktierten Linien geben die Begrenzung der Kontrollvolumina  $T_W$  und  $T_E$  an.

wobei  $h_{e_i}$  die Länge der Kante  $e_i$  bezeichne und  $\mathbf{n}_{e_i} = \mathbf{n}_{\chi_E}|_{e_i}$  der dazugehörige Normalvektor bzgl.  $\partial\chi_E$  sei. Es gilt beispielsweise

$$h_{e_1} = |x_S - x_W|, \quad \mathbf{n}_{e_1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \frac{x_S - x_W}{|x_S - x_W|}.$$

**Bemerkung.** Die Abbildung

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \mathbf{x}$$

erzeugt eine Drehung des Vektors  $\mathbf{x}$  um den Winkel  $\varphi$  (im mathematisch positiven Sinn). Um die Normale auf  $\mathbf{x}$  zu erhalten, müssen wir um  $-\frac{\pi}{2}$  drehen. Somit erhalten wir eine Drehmatrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

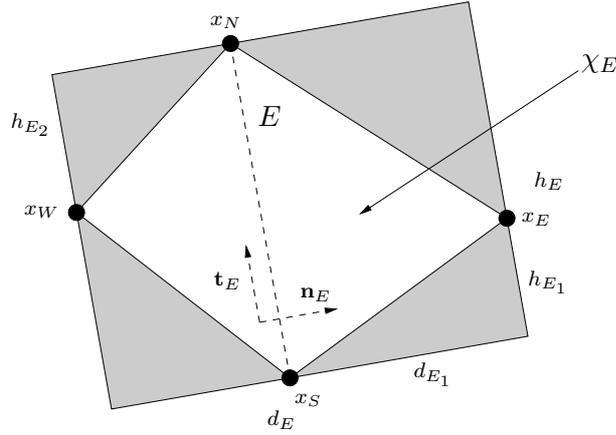
Die Werte  $h_{e_2}$ ,  $h_{e_3}$  und  $h_{e_4}$  bzw. die Vektoren  $\mathbf{n}_{e_2}$ ,  $\mathbf{n}_{e_3}$  und  $\mathbf{n}_{e_4}$  werden analog berechnet. Somit erhalten wir

$$\begin{aligned} h_{e_1} \mathbf{n}_{e_1} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (x_S - x_W), & h_{e_2} \mathbf{n}_{e_2} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (x_E - x_S), \\ h_{e_3} \mathbf{n}_{e_3} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (x_N - x_E), & h_{e_4} \mathbf{n}_{e_4} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (x_W - x_N). \end{aligned}$$

Einsetzen und Zusammenfassen der Terme bei  $u_W$ ,  $u_S$ ,  $u_E$  und  $u_N$  ergibt

$$\begin{aligned} \nabla u_E &= \frac{1}{2|\chi_E|} \left[ u_W \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (x_S - x_N) + u_S \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (x_E - x_W) \right. \\ &\quad \left. + u_E \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (x_N - x_S) + u_N \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (x_W - x_E) \right]. \end{aligned} \quad (2.16)$$

**Berechnung von  $|\chi_E|$ .** Nach Abbildung 2.4 berechnen wir die Fläche des Rechtecks und ziehen die Flächen der 4 Dreiecke (grau) ab. Dies ist in dieser einfachen Weise nur deshalb möglich, da


 Abbildung 2.4: Das Rechteck, das den Diamond Path umschließt, hat die Seiten  $d_E$  und  $h_E$ .

wir für  $\chi_E$  Konvexität voraussetzen. Die Seiten des Rechtecks haben gerade die Länge  $d_E$  und  $h_E$ .

$$\begin{aligned} |\chi_E| &= d_E h_E - \left[ \frac{h_{E_1} d_{E_1}}{2} + \frac{(h_E - h_{E_1}) d_{E_1}}{2} + \frac{h_{E_2} (d_E - d_{E_1})}{2} + \frac{(h_E - h_{E_2}) (d_E - d_{E_1})}{2} \right] \\ &= d_E h_E - \frac{1}{2} [h_{E_1} d_{E_1} + h_E d_{E_1} - h_{E_1} d_{E_1} + h_{E_2} d_E - h_{E_2} d_{E_1} \\ &\quad + h_E d_E - h_E d_{E_1} - h_{E_2} d_E + h_{E_2} d_{E_1}] = \frac{d_E h_E}{2}. \end{aligned}$$

Weiters schreiben wir

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \overbrace{(x_S - x_N)}^{=-h_E \mathbf{t}_E} = -h_E \mathbf{n}_E, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \overbrace{(x_N - x_S)}^{=h_E \mathbf{t}_E} = h_E \mathbf{n}_E.$$

Nach Abbildung 2.3 gilt

$$x_E - x_W = d_E \mathbf{n}_E + \alpha_E d_E \mathbf{t}_E \quad \text{mit } d_E = (x_E - x_W) \cdot \mathbf{n}_E, \quad \alpha_E = \frac{(x_E - x_W) \cdot \mathbf{t}_E}{(x_E - x_W) \cdot \mathbf{n}_E}.$$

Wir stellen den Vektor  $x_E - x_W$  also in der Basis  $(\mathbf{n}_E, \mathbf{t}_E)$  dar. Beachte, dass  $(x_E - x_W) \cdot \mathbf{t}_E$  positiv als auch negativ sein kann.

$$\begin{aligned} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (x_E - x_W) &= -d_E \mathbf{t}_E + \alpha_E d_E \mathbf{n}_E, \\ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (x_W - x_E) &= d_E \mathbf{t}_E - \alpha_E d_E \mathbf{n}_E. \end{aligned}$$

Aufgrund dieser Überlegungen erhalten wir für (2.16)

$$\nabla u_E = \frac{1}{d_E h_E} \left[ (u_E - u_W) h_E \mathbf{n}_E - \alpha_E (u_N - u_S) d_E \mathbf{n}_E + (u_N - u_S) d_E \mathbf{t}_E \right]$$

und somit die Darstellung des Gradienten  $\nabla u_E$  in den Basisvektoren  $(\mathbf{n}_E, \mathbf{t}_E)$

$$\nabla u_E = \left[ \frac{u_E - u_W}{d_E} - \alpha_E \frac{u_N - u_S}{h_E} \right] \mathbf{n}_E + \frac{u_N - u_S}{h_E} \mathbf{t}_E. \quad (2.17)$$

### 2.3.4 Berechnung von $\psi_T(a)$

Für die Flussapproximation brauchen wir die Werte von  $u$  auch an den Knoten. Diese interpolieren wir linear aus  $u_h$ . Im Folgenden wird gezeigt, wie sich die Gewichte  $\psi_T(a)$  berechnen lassen. In diesem Abschnitt bezeichnet  $x_T$  bzw.  $y_T$  die  $x$ - bzw.  $y$ -Koordinate des Mittelpunkts  $m_T$  von  $T$ . Weiters ist  $x_a$  bzw.  $y_a$  die  $x$ - bzw.  $y$ -Koordinate des Punktes  $a \in \mathcal{N}$ . Für jeden Knoten  $a \in (\mathcal{N}^* \cup \mathcal{N}_h)$  einer Triangulierung  $\mathcal{T}$  ist nach Abschnitt 1.1.2 der Knotenpatch  $\omega_a$  die Vereinigung aller Elemente, die  $a$  enthalten.

Zur stückweisen konstanten Funktion  $u_h$  suchen wir jene affine Funktion  $\phi$  auf  $\omega_a$ , die

$$J(\phi) = \sum_{T \in \tilde{\omega}_a} |\phi(m_T) - u_T|^2 \quad \text{minimiert.} \quad (2.18)$$

Für die affine Funktion  $\phi$  setzen wir allgemein an:

$$\phi(x) = M \cdot x + c \quad M = (m_1, m_2) \in \mathbb{R}^{1 \times 2}, \quad c \in \mathbb{R}, \quad x \in \mathbb{R}^{2 \times 1}. \quad (2.19)$$

Wir setzen  $\phi$  in (2.18) ein und bestimmen das Minimum von  $J(\phi)$ .

$$\begin{aligned} \frac{\partial J(\phi)}{\partial m_1} &= 2 \sum_{T \in \tilde{\omega}_a} (m_1 x_T + m_2 y_T + c - u_T) x_T = 0, \\ \frac{\partial J(\phi)}{\partial m_2} &= 2 \sum_{T \in \tilde{\omega}_a} (m_1 x_T + m_2 y_T + c - u_T) y_T = 0, \\ \frac{\partial J(\phi)}{\partial c} &= 2 \sum_{T \in \tilde{\omega}_a} (m_1 x_T + m_2 y_T + c - u_T) = 0. \end{aligned} \quad (2.20)$$

Wir fassen in jeder Gleichung die Koeffizienten der Unbekannten  $m_1, m_2, c$  zusammen und erhalten drei Gleichungen mit den drei Unbekannten, die sogenannten Gaußschen Normalgleichungen.  $n$  bezeichnet dabei die Anzahl der Elemente in  $\omega_a$ .

$$\begin{aligned} m_1 \sum_{T \in \tilde{\omega}_a} x_T^2 + m_2 \sum_{T \in \tilde{\omega}_a} x_T y_T + c \sum_{T \in \tilde{\omega}_a} x_T &= \sum_{T \in \tilde{\omega}_a} u_T x_T, \\ m_1 \sum_{T \in \tilde{\omega}_a} x_T y_T + m_2 \sum_{T \in \tilde{\omega}_a} y_T^2 + c \sum_{T \in \tilde{\omega}_a} y_T &= \sum_{T \in \tilde{\omega}_a} u_T y_T, \\ m_1 \sum_{T \in \tilde{\omega}_a} x_T + m_2 \sum_{T \in \tilde{\omega}_a} y_T + cn &= \sum_{T \in \tilde{\omega}_a} u_T. \end{aligned} \quad (2.21)$$

Um die Übersicht zu bewahren, schreiben wir für die Summen folgende Abkürzungen:

$$\begin{aligned} k_1 &= \sum_{T \in \tilde{\omega}_a} x_T^2, & k_2 &= \sum_{T \in \tilde{\omega}_a} x_T y_T, & k_3 &= \sum_{T \in \tilde{\omega}_a} x_T, & k_4 &= \sum_{T \in \tilde{\omega}_a} u_T x_T, \\ k_5 &= \sum_{T \in \tilde{\omega}_a} y_T^2, & k_6 &= \sum_{T \in \tilde{\omega}_a} y_T, & k_7 &= \sum_{T \in \tilde{\omega}_a} u_T y_T, & k_8 &= \sum_{T \in \tilde{\omega}_a} u_T. \end{aligned}$$

Wir erhalten folgende Lösung

$$\begin{aligned} m_1 &= \frac{1}{N} (-nk_2 k_7 - k_6^2 k_4 + k_6 k_3 k_7 + nk_5 k_4 - k_5 k_3 k_8 + k_2 k_6 k_8), \\ m_2 &= \frac{1}{N} (-nk_2 k_4 + k_2 k_3 k_8 + k_6 k_3 k_4 + nk_7 k_1 - k_7 k_3^2 - k_6 k_8 k_1), \\ c &= \frac{1}{N} (k_3 k_2 k_7 - k_3 k_5 k_4 + k_6 k_2 k_4 - k_6 k_7 k_1 + k_8 k_5 k_1 - k_8 k_2^2), \\ &\quad \text{mit } N = -k_6^2 k_1 + nk_5 k_1 - k_5 k_3^2 - nk_2^2 + 2k_2 k_6 k_3. \end{aligned} \quad (2.22)$$

Die Lösung  $g$  hängt aber auch linear von  $(u_T|T \in \omega_a) \in \mathbb{R}^n$  ( $n = |\tilde{\omega}_a|$ ) ab. Daher gilt für einen Knoten  $a$

$$u_a = \phi(a) = m_1 x_a + m_2 y_a + c = \sum_{T \in \omega_a} \psi_T(a) u_T, \tag{2.23}$$

wobei  $u_a$  der approximierte Wert von  $u$  am Knoten  $a$  ist.

Wenn wir nun die Gewichte  $\psi_T(a)$  berechnen wollen, müssen wir nur einen Koeffizientenvergleich durchführen. Die  $u_T$  kommen nur in den Zählern von  $m_1, m_2, c$  vor.

$$\begin{aligned} \psi_T(a) = \frac{1}{N} & \left( (-nk_2 y_T - k_6^2 x_T + k_6 k_3 y_T + nk_5 x_T - k_5 k_3 + k_2 k_6) a_x + \right. \\ & \left. (-nk_2 x_T + k_2 k_3 + k_6 k_3 x_T + nk_1 y_T - k_5^2 y_T - k_6 k_1) a_y + \right. \\ & \left. (k_3 k_2 y_T - k_3 k_5 x_T + k_6 k_2 x_T - k_6 k_1 y_T + k_5 k_1 - k_2^2) \right) \end{aligned} \tag{2.24}$$

Diese Formel ist sehr kompliziert und benötigt sehr viel Rechenzeit. Deshalb ist es günstig, gewisse Anforderungen an die Triangulierung zu stellen, so dass nur endlich viele Anordnungen vorkommen. Somit können die Gewichte a priori berechnet werden und müssen nicht für jeden Knoten aufs Neue berechnet werden. Außerdem werden sie nur für die inneren Knoten  $a \in \mathcal{N}^* \cup \mathcal{N}^h$  benötigt.

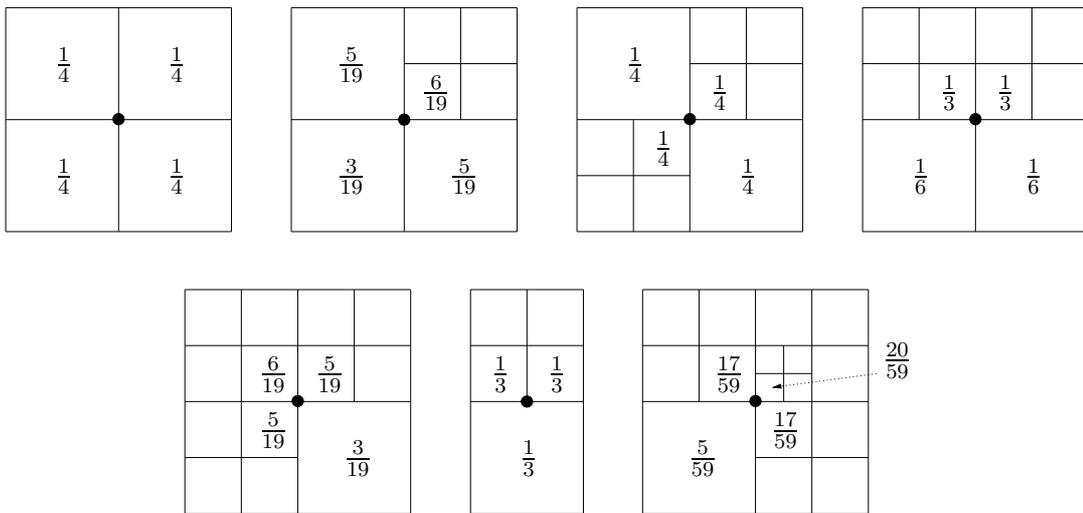


Abbildung 2.5: Bei einer Triangulierung  $\mathcal{T}$  in lauter Quadrate mit höchstens einem hängenden Knoten pro Elementkante, gibt es bis auf Drehungen 7 verschiedene Gewichtskombinationen  $\psi_T(a)$  für einen Knotenpatch  $T \in \tilde{\omega}_a$ , die a priori berechnet werden können. Für eine Dreieckstriangulierung ist es schwer, eine geeignete Einschränkung zu finden.

# Kapitel 3

## Residualschätzer

In diesem Kapitel beweisen wir die Zuverlässigkeit und Effizienz eines Residual-Fehlerschätzers, der auf **Morley-Interpolation** basiert. Wir beschränken uns dabei auf das Laplace Problem mit reinem Dirichlet-Rand. Außerdem sei eine fastreguläre Triangulierung  $\mathcal{T}$  nach (1.1.4) in Dreiecke oder Rechtecke vorausgesetzt.

### 3.1 Das Laplace Problem mit reinem Dirichlet-Rand

Gegeben sei ein  $f \in L^2(\Omega)$ . Finde ein  $u \in H^1(\Omega)$  mit

$$\begin{cases} -\Delta u = f & \text{auf } \Omega, \\ u = u_D & \text{auf } \Gamma = \Gamma_D, \end{cases} \quad (3.1)$$

wobei der Differentialoperator  $\Delta$  in schwacher Form verstanden wird und  $u_D \in L^2(\Gamma_D)$  ist. Da wir im Folgenden ein reines Dirichlet-Randproblem betrachten, gilt

$$H_D^1(\Omega) := \{v \in H^1(\Omega) \mid v|_{\Gamma_D} = 0\} = H_0^1(\Omega). \quad (3.2)$$

**Die schwache Form.** Die schwache Form des Laplace Problems lautet

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx, \quad \forall v \in H_D^1(\Omega) \quad (= H_0^1(\Omega)). \quad (3.3)$$

Die Diskretisierung von (3.1) erfolgt gemäß Kapitel 2. Beachte, dass für das Laplace-Problem kein konvektiver Fluss auftritt. Zusammenfassend erhalten wir folgende Finite Volumen Diskretisierung des Laplace-Problems mit reinem Dirichlet-Rand nach (3.1):

Finde die Lösung  $u_h \in \mathcal{P}_0(\mathcal{T})$  für  $T \in \mathcal{T}$  von

$$-\sum_{E \in \mathcal{E}_T} \sigma_{T,E} F_E^D(u_h) = \int_T f \, dx \quad \forall T \in \mathcal{T}, \quad (3.4)$$

wobei  $F_E^D(u_h)$ , der diffusive numerische Fluss, nach Abschnitt 2.3.2 berechnet wird.

### 3.2 Das Morley-Element

Die Idee der Morley-Interpolation beruht auf dem sogenannten **Morley-Element**. Dazu benötigen wir zunächst eine Definition eines Finiten Elements. Wir folgen der Charakterisierung nach P. Ciarlet.

**Definition 3.2.1 (Finites Element nach P. Ciarlet).** Das Tripel  $(T, \mathcal{P}, \Sigma)$  heißt *Finites Element*, wenn

1.  $T$  ein beschränktes Lipschitz-Gebiet im  $\mathbb{R}^d$  ist,
2.  $\mathcal{P}$  ein endlich-dimensionaler Vektorraum mit Dimension  $m$  von Funktionen  $p : T \rightarrow \mathbb{R}$  ist,
3.  $\Sigma$  eine Basis für den dualen Vektorraum  $\mathcal{P}^* = \{F : \mathcal{P} \rightarrow \mathbb{R} \mid F \text{ linear}\}$  von  $\mathcal{P}$  ist.

**Bemerkung.**

- Die Definition ist in dem Sinne unvollständig, als dass wir unten annehmen, dass  $F \in \mathcal{P}^*$  sich von  $\mathcal{P}$  auf gewisse (eben hier *nicht* bezeichnete) Sobolev-Räume fortsetzen lässt. Tatsächlich sind  $\mathcal{P}$  meist analytische Funktionen, so dass diese technische Anmerkung marginal wird.
- Es gilt  $\dim \mathcal{P}^* = \dim \mathcal{P} = m$ .
- Da  $\mathcal{P}$  endlich-dimensional ist, sind alle Funktionen  $F \in \mathcal{P}^*$  stetig.

**Definition 3.2.2 (Nodale Basis).** Wenn  $(T, \mathcal{P}, \Sigma)$  ein Finites Element ist mit  $\Sigma = \{S_1, \dots, S_m\}$ , dann heißt  $(\varphi_1, \dots, \varphi_m)$  **nodale Basis** von  $(T, \mathcal{P}, \Sigma)$ , falls  $\varphi_j \in \mathcal{P}$  für alle  $1 \leq j \leq m$  und

$$S_i(\varphi_j) = \delta_{ij} \quad \text{für } i, j = 1 \dots m \quad (3.5)$$

gelten, wobei  $\delta_{ij}$  das Kronecker-Symbol bezeichne.

Wenn ein Tripel  $(T, \mathcal{P}, \Sigma)$  gegeben ist, dann ist meist unmittelbar klar, dass die Eigenschaften 1. und 2. von Definition 3.2.1 erfüllt sind. Aus der Theorie der dualen Vektorräume wissen wir, dass eine Basis des Dualraums  $\mathcal{P}^*$  gerade durch (3.5) definiert wird. Somit ist es also hinreichend, eine nodale Basis anzugeben, um die Eigenschaft 3. von Definition 3.2.1 zu zeigen.

**Lemma 3.2.3.** Für ein Rechteck  $T = \text{conv}\{a_1, a_2, a_3, a_4\} \subseteq \mathbb{R}^2$  und  $\mathcal{P} = \mathcal{P}_2 \oplus \text{Span}\{x^3 - 3xy^2, y^3 - 3yx^2\}$  sowie  $\Sigma = (S_1, \dots, S_8)$  mit

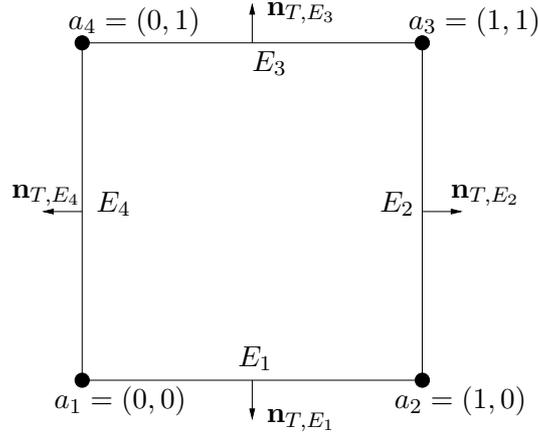
$$S_i(p) = \begin{cases} p(a_i), & i = 1 \dots 4, \\ \int_{E_{i-4}} \partial p / \partial \mathbf{n}_{T, E_{i-4}} ds, & i = 5 \dots 8, \end{cases} \quad (3.6)$$

ist  $(T, \mathcal{P}, \Sigma)$  ein Finites Element, das **Morley-Element** für Rechtecke.

**Beweis.** Es genügt, den Beweis für das Referenzelement nach Abbildung 3.1 zu zeigen.

Es ist unmittelbar klar, dass die Eigenschaften 1. und 2. von Definition (3.2.1) erfüllt sind. Für Eigenschaft 3. suchen wir eine nodale Basis, sodass (3.5) gilt. Es sei nun  $\varphi \in \mathcal{P}$  und  $\varphi(x, y) := k_1 + k_2x + k_3y + k_4xy + k_5x^2 + k_6y^2 + k_7(x^3 - 3xy^2) + k_8(y^3 - 3x^2y)$  mit  $\mathbf{k} = (k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8) \in \mathbb{R}^8$ . Dann gilt

$$\begin{aligned} \varphi(a_1) &= k_1, \\ \varphi(a_2) &= k_1 + k_2 + k_5 + k_7, \\ \varphi(a_3) &= k_1 + k_2 + k_3 + k_4 + k_5 + k_6 - 2k_7 - 2k_8, \\ \varphi(a_4) &= k_1 + k_3 + k_6 + k_8. \end{aligned}$$

Abbildung 3.1: Das Referenzelement  $[0, 1] \times [0, 1]$ 

Mit

$$\begin{aligned} \nabla \varphi(x, y) &= (k_2 + k_4 y + 2k_5 x + 3k_7(x^2 - y^2) - 6k_8 xy, \\ &\quad k_3 + k_4 x + 2k_6 y - 6k_7 xy + 3k_8(y^2 - x^2)), \\ \mathbf{n}_{T,E_1} &= (0, -1), \quad \mathbf{n}_{T,E_2} = (1, 0), \quad \mathbf{n}_{T,E_3} = (0, 1), \quad \mathbf{n}_{T,E_4} = (-1, 0) \end{aligned}$$

und dem Transformationssatz berechnet sich

$$\begin{aligned} \int_{E_1} \nabla \varphi \cdot \mathbf{n}_{T,E_1} ds &= \int_{E_1} (-k_3 - k_4 x - 2k_6 y + 6k_7 xy - 3k_8(y^2 - x^2)) ds_{(x,y)} \\ &= \int_0^1 (-k_3 - k_4 t + 3k_8 t^2) dt = -k_3 - \frac{1}{2}k_4 + k_8, \end{aligned}$$

$$\begin{aligned} \int_{E_2} \nabla \varphi \cdot \mathbf{n}_{T,E_2} ds &= \int_{E_2} (k_2 + k_4 y + 2k_5 x + 3k_7(x^2 - y^2) - 6k_8 xy) ds_{(x,y)} \\ &= \int_0^1 (k_2 + k_4 t + 2k_5 + 3k_7(1 - t^2) - 6k_8 t) dt \\ &= k_2 + \frac{1}{2}k_4 + 2k_5 + 2k_7 - 3k_8, \end{aligned}$$

$$\begin{aligned} \int_{E_3} \nabla \varphi \cdot \mathbf{n}_{T,E_3} ds &= \int_{E_3} (k_3 + k_4 x + 2k_6 y - 6k_7 xy + 3k_8(y^2 - x^2)) ds_{(x,y)} \\ &= \int_0^1 (k_3 + k_4(1 - t) + 2k_6 - 6k_7(1 - t) + 3k_8(1 - (1 - t)^2)) dt \\ &= k_3 + \frac{1}{2}k_4 + 2k_6 - 3k_7 + 2k_8, \end{aligned}$$

$$\begin{aligned} \int_{E_4} \nabla \varphi \cdot \mathbf{n}_{T,E_4} ds &= \int_{E_4} (-k_2 - k_4 y - 2k_5 x - 3k_7(x^2 - y^2) + 6k_8 xy) ds_{(x,y)} \\ &= \int_0^1 (-k_2 - k_4(1 - t) + 3k_7(1 - t)^2) dt = -k_2 - \frac{1}{2}k_4 + k_7. \end{aligned}$$

Somit erhalten wir

$$\begin{pmatrix} S_1(\varphi) \\ S_2(\varphi) \\ S_3(\varphi) \\ S_4(\varphi) \\ S_5(\varphi) \\ S_6(\varphi) \\ S_7(\varphi) \\ S_8(\varphi) \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & -2 & -2 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & -0.5 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0.5 & 2 & 0 & 2 & -3 \\ 0 & 0 & 1 & 0.5 & 0 & 2 & -3 & 2 \\ 0 & -1 & 0 & -0.5 & 0 & 0 & 1 & 0 \end{pmatrix}}_A \underbrace{\begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \\ k_5 \\ k_6 \\ k_7 \\ k_8 \end{pmatrix}}_{\mathbf{k}}.$$

Elementare Rechnung zeigt,  $\det(A) = 4$ . Die Matrix  $A \in \mathbb{R}^{8 \times 8}$  ist daher regulär und das Gleichungssystem  $A\mathbf{k} = \mathbf{b}$  für jedes  $\mathbf{b} \in \mathbb{R}^8$  eindeutig lösbar. Somit finden wir eine nodale Basis  $\varphi_1, \dots, \varphi_8 \in \mathcal{P}$ , so dass

$$S_i(\varphi_j) = \delta_{ij} \quad (i, j = 1, \dots, 8)$$

gilt. ■

Das Morley-Element für Dreiecke lässt sich ähnlich bestimmen. In der Literatur wird oft nur dieses Element über Dreiecke Morley-Element genannt.

**Lemma 3.2.4.** *Für ein Dreieck  $T = \text{conv}\{a_1, a_2, a_3\} \subseteq \mathbb{R}^2$  und  $\mathcal{P} = \mathcal{P}_2$  sowie  $\Sigma = (S_1, \dots, S_6)$  mit*

$$S_i(p) = \begin{cases} p(a_i), & i = 1 \dots 3, \\ \partial p / \partial \mathbf{n}_{T, E_{i-3}}(m_{E_{i-3}}), & i = 4 \dots 6, \end{cases} \quad (3.7)$$

wobei  $m_{E_{i-3}}$  der Mittelpunkt der Kante  $E_{i-3}$  ( $i = 4 \dots 6$ ) sei, ist  $(T, \mathcal{P}, \Sigma)$  ein *Finite Element*, das **Morley-Element** für Dreiecke. Beachte, dass

$$\int_{E_{i-3}} \frac{\partial p}{\partial \mathbf{n}_{T, E_{i-3}}} ds = |E_{i-3}| \frac{\partial p}{\partial \mathbf{n}_{T, E_{i-3}}}(m_{E_{i-3}}) \quad \forall p \in \mathcal{P}_2$$

gilt. ■

In der Literatur zur Finite Elemente Methode werden die Finite Elemente oft durch kleine Prinzipskizzen veranschaulicht. Punktfunktionale werden dabei als Punkte gezeichnet, die Richtungsableitung in Normalrichtung im Punkt  $m$  durch einen Pfeil dargestellt. Abbildung 3.2 zeigt die Prinzipskizze für das Morley-Dreieck mit den Momenten aus (3.7). Beim Morley-Rechteck haben wir jedoch Schwierigkeiten, eine Prinzipskizze zu fertigen, die die Momente in (3.6) illustriert.

### 3.3 Der Morley-Interpolant

Im Folgenden definieren wir einen Interpolanten, dessen Eigenschaften wir später zur Herleitung eines Fehlerschätzers nützen können.

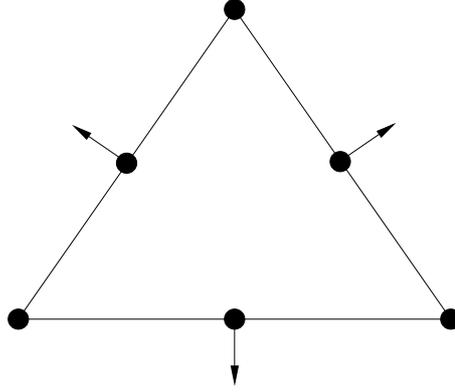


Abbildung 3.2: Das (Standard-) Morley-Element

**Lemma 3.3.1.** Gegeben sei die Lösung  $u_h$  auf  $\mathcal{T}$  von (3.4). Dann existiert eine eindeutige elementweise definierte Funktion  $I_M u_h|_T \in \mathcal{P}$  für  $T \in \mathcal{T}$  ( $\mathcal{P} = \mathcal{P}_2$  für Dreiecke,  $\mathcal{P} = \mathcal{P}_2 \oplus \text{Span}\{x^3 - 3xy^2, y^3 - 3yx^2\}$  für Rechtecke) mit den Eigenschaften:

$$I_M u_h|_T(a) = \sum_{T_a \in \tilde{\omega}_a} \psi_{T_a}(a) u_{T_a} \quad \forall a \in \mathcal{N}_T \setminus (\mathcal{N}^h \cup \mathcal{N}_\Gamma), \quad (3.8)$$

$$I_M u_h|_T(a) = u_D(a) \quad \forall a \in \mathcal{N}_T \cap \mathcal{N}_\Gamma, \quad (3.9)$$

$$I_M u_h|_T(a) = I_M u_h|_{T_a}(a) \quad \forall a \in \mathcal{N}_T \cap \mathcal{N}^h, \quad a \subseteq (E \in \mathcal{E}^h \cap \mathcal{E}_{T_a}), \quad T_a \in \mathcal{T}, \quad (3.10)$$

$$\int_E \frac{\partial I_M u_h|_T}{\partial \mathbf{n}_E} ds = F_E^D(u_h) \quad \forall E \in \mathcal{E}_T. \quad (3.11)$$

$I_M u_h$  bezeichnen wir als **Morley-Interpolant**.

**Beweis.**  $I_M u_h$  wird gerade auf dem Morley-Element für Dreiecke oder Rechtecke definiert. Da das Morley-Element ein Finites Element ist, sind die Funktionale linear unabhängig. Somit ist das  $6 \times 6$  Gleichungssystem für Dreiecke bzw.  $8 \times 8$  Gleichungssystem für Rechtecke eindeutig lösbar und  $I_M u_h|_T$  für alle  $T \in \mathcal{T}$  eindeutig bestimmt. ■

**Bemerkung.**

- $\psi_{T_a}(a)$  sind die Gewichte nach Abschnitt 2.3.4.  $F_E^D(u_h)$  ist der numerische diffusive Fluss nach Abschnitt 2.3.2. Für eine hängende Kante  $E \in \mathcal{E}^h$  ist  $F_E^D(u_h)$  nach (2.5) die Summe der Teilflüsse der dazugehörigen Elementarkanten.
- Die Eigenschaften (3.8)–(3.10) stellen sicher, dass wir in allen Knoten  $a \in \mathcal{N}$  Stetigkeit des Morley-Interpolanten haben. Damit dies auch auf hängenden Knoten gilt, müssen wir nach (3.10) also zuerst den Morley-Interpolanten für  $T_a$  berechnen, damit wir für  $T$  den Funktionswert nehmen können. Die Eigenschaft (3.11) stellt sicher, dass die Konservativität des Flusses auf  $E \in \mathcal{E}^o$  erhalten bleibt. Auf einer Elementarkante, die Teil einer hängenden Kante ist, haben wir jedoch keine Konservativitätseigenschaft mehr!

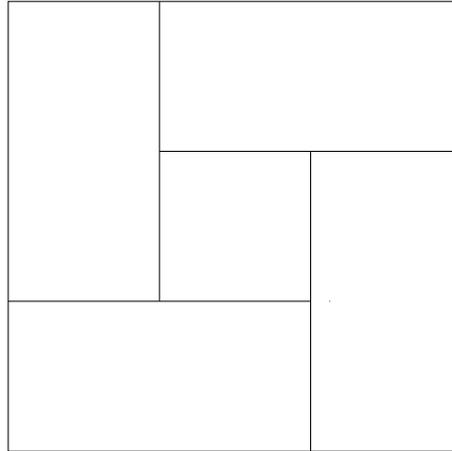


Abbildung 3.3: Eine fastreguläre Triangulierung, für die die Rekursion zur Bestimmung des Morley-Interpolanten nicht abbricht!

- Bei der Eigenschaft (3.10) wird also zuerst  $T_a$  berechnet. Dies führt zu einer Rekursion, deren Abbruchbedingung damit erfüllt wird, indem ein Element  $T'_a \in \mathcal{T}$  erreicht wird, dessen Ecken nicht hängend sind. Abbildung 3.3 zeigt eine fastreguläre Triangulierung, wo die Rekursion nicht abbricht! Diese müssen somit vermieden werden.

Für die Finite Elemente Methode hat die diskrete Lösung  $\tilde{u}_h \in \mathcal{P}_1 \cap \mathcal{C}(\Omega)$  die sogenannte Galerkin-Orthogonalität

$$\int_T \nabla(u - \tilde{u}_h) \cdot \nabla v_h \, dx = 0 \quad \forall v_h \in \mathcal{P}_1 \cap \mathcal{C}(\Omega), \quad v_h|_\Gamma = 0.$$

Beim Finite Volumen Verfahren haben wir in Verbindung mit der Morley-Interpolation eine Orthogonalitätseigenschaft in  $L^2$  für das Residuum  $-\Delta(u - I_M u_h) = f + \Delta I_M u_h$ .

**Lemma 3.3.2.** Sei  $u_h$  eine Lösung von (3.4), dann gilt für alle  $T \in \mathcal{T}$

$$\int_T (f + \Delta(I_M u_h)) \, dx = 0, \tag{3.12}$$

d.h.  $f + \Delta I_M u_h$  ist  $L^2$ -orthogonal zu  $\mathcal{P}_0(\mathcal{T})$ .

**Beweis.** Der Gaußsche Integralsatz zeigt

$$\begin{aligned} \int_T \Delta(I_M u_h) \, dx &= \int_T \operatorname{div} \nabla(I_M u_h) \, dx = \int_{\partial T} \underbrace{\nabla(I_M u_T) \cdot \mathbf{n}_E}_{= \frac{\partial I_M u_T}{\partial \mathbf{n}_E}} \, ds \\ &= \sum_{E \in \mathcal{E}_T} \sigma_{T,E} \underbrace{\int_E \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \, ds}_{\text{nach (3.11)} = F_E^D(u_h)} \stackrel{(3.4)}{=} - \int_T f(x) \, dx. \quad \blacksquare \end{aligned}$$

### 3.4 A posteriori Fehlerschätzer

Wir wollen den Fehler  $e := u - I_M u_h$  abschätzen. Im Folgenden wird ein Fehlerschätzer  $\eta$  definiert, der auf dem Residuum  $R := f + \Delta I_M u_h$  basiert. Wir beweisen seine Zuverlässigkeit und Effizienz, d.h. es existieren Konstanten  $C_1, C_2 > 0$ , die nur von  $\kappa_{\mathcal{T}}$  (also von der Form der Elemente  $T \in \mathcal{T}$ ) abhängen, mit

$$C_1(\eta - h\|f - f_{\mathcal{T}}\|_{L^2(\Omega)}) \leq \|\nabla_{\mathcal{T}} e\|_{L^2(\Omega)} \leq C_2 \eta.$$

Die Analysis adaptiert Techniken aus dem Kontext der Finite Elemente Methode.

**Definition 3.4.1 (Kantensprung).** Sei  $g \in H^1(T)$ ,  $T \in \mathcal{T}$  und  $E \in \mathcal{E}_T$ . Dann existiert die Spur  $g|_{T,E} := g|_E$ .

Es sei nun  $E \in \mathcal{E}^*$ ,  $T_W, T_E \in \mathcal{T}$  mit  $\overline{T_W} \cap \overline{T_E} = E$  ( $\mathbf{n}_E$  zeigt nach Definition von  $T_W$  nach  $T_E$ ) und  $g \in H^1(T)$ . Mit  $[g]_E$  bezeichnen wir den Sprung von  $g$  über die Kante  $E$ , der punktweise durch

$$[g]_E(x) := g|_{T_E, E}(x) - g|_{T_W, E}(x) \quad \text{für } x \in E \quad \text{definiert ist.} \quad (3.13)$$

Weiters definieren wir für  $E \in \mathcal{E}^h$

$$[g]_E(x) := [g]_{E_i}(x) \quad \text{für } x \in E_i, \quad (3.14)$$

wobei  $E = \bigcup_{i=1}^n E_i$ ,  $E_1, \dots, E_n \in \mathcal{E}^*$  ist.

**Definition 3.4.2.** Sei  $\chi \subseteq \mathbb{R}^d$  eine messbare Teilmenge. Für  $g \in L^1(\chi)$  definieren wir das **Integralmittel** von  $g$  über  $\chi$  mittels

$$g_{\chi} := \frac{1}{|\chi|} \int_{\chi} g \, dx, \quad (3.15)$$

wobei  $|\chi|$  das Lebesgue-Maß von  $\chi$  bezeichne.

**Bemerkung.**  $g_{\chi}$  von (3.15) ist die Orthogonalprojektion von  $g \in L^1(\chi)$  auf  $\mathcal{P}_0(\chi)$ . Für Kanten verwenden wir dieselbe Notation. Außerdem sei  $g_{\mathcal{T}} \in \mathcal{P}^0(\mathcal{T})$  das  $\mathcal{T}$ -stückweise Integralmittel, d.h.  $g_{\mathcal{T}|T} := g_T$  mit  $T \in \mathcal{T}$ .

Für das Residuum gilt  $R = f + \Delta I_M u_h = -\Delta u + \Delta I_M u_h$ , und nach obiger Bemerkung ist

$$\|R\|_{L^2(T)}^2 = \|R - R_{\mathcal{T}}\|_{L^2(T)}^2 + \|R_{\mathcal{T}}\|_{L^2(T)}^2 = \|f - f_{\mathcal{T}}\|_{L^2(T)}^2 + \|R_{\mathcal{T}}\|_{L^2(T)}^2. \quad (3.16)$$

Nach Lemma 3.3.2 ist  $\|R_{\mathcal{T}}\|_{L^2(T)} = 0$ . Den Term  $\|f - f_{\mathcal{T}}\|_{L^2(T)}$  betrachten wir als einen von höherer Ordnung.

**Definition 3.4.3.** Für  $T \in \mathcal{T}$  definieren wir die Verfeinerungsindikatoren

$$\begin{aligned} \eta_T^2 := & h_T \left( h_T \|f - f_{\mathcal{T}}\|_{L^2(T)}^2 + \sum_{E \in (\mathcal{E}_T \setminus \mathcal{E}_{\Gamma})} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)}^2 \right. \\ & \left. + \sum_{E \in (\mathcal{E}_T \setminus \mathcal{E}_{\Gamma})} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right\|_{L^2(E)}^2 + \sum_{E \in (\mathcal{E}_T \cap \mathcal{E}_{\Gamma})} \left\| \frac{\partial (I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}^2 \right). \end{aligned} \quad (3.17)$$

Der Residual-Fehlerschätzer  $\eta$  sei dann

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

Nachdem wir nun alle notwendigen Bezeichnungen eingeführt haben, beweisen wir die Zuverlässigkeit dieses Residual-Fehlerschätzers, der auf Morley-Interpolation basiert.

### 3.4.1 Zuverlässigkeit

Für den Beweis benötigen wir einige wichtige Eigenschaften.

**Lemma 3.4.4.** *Für die Sprünge der Richtungsableitung in Normalrichtung bzw. Tangentialrichtung gelten*

$$\int_E \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right] \right]_E ds = 0 \quad \forall E \in \mathcal{E}^o, \quad (3.18)$$

$$\int_E \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right] \right]_E ds = 0 \quad \forall E \in \mathcal{E}^*, \quad (3.19)$$

$$\int_E \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} ds = 0 \quad \forall E \in \mathcal{E}_\Gamma. \quad (3.20)$$

**Beweis.** Zuerst beweisen wir (3.18). Wegen  $E \in \mathcal{E}^o$  gilt  $E = \bigcup_{i=1}^n E_i$  mit  $E_1, \dots, E_n \in \mathcal{E}^*$ ,  $E_i = \overline{T_{W_i}} \cap \overline{T_E}$  und  $\mathbf{n}_E$  zeigt vom Element  $T_{W_i}$  nach  $T_E$ . Somit erhalten wir

$$\begin{aligned} \int_E \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right] \right]_E ds &= \sum_{i=1}^n \int_{E_i} \underbrace{\left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right] \right]_{E_i}}_{= \frac{\partial(I_M u_h)|_{T_E, E}}{\partial \mathbf{n}_E} - \frac{\partial(I_M u_h)|_{T_{W_i}, E}}{\partial \mathbf{n}_E}} ds \\ &= \int_E \frac{\partial(I_M u_h)|_{T_E, E}}{\partial \mathbf{n}_E} ds - \sum_{i=1}^n \int_{E_i} \frac{\partial(I_M u_h)|_{T_{W_i}, E}}{\partial \mathbf{n}_E} ds \\ &= F_E^D(u_h) - \sum_{i=1}^n F_{E_i}^D(u_h) = 0. \end{aligned}$$

Für (3.19) gilt nach dem Hauptsatz der Differential- und Integralrechnung für  $E \in \mathcal{E}^*$  mit  $a_S$  Anfangs- und  $a_N$  Endknoten von  $E$

$$\begin{aligned} \int_E \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right] \right]_E ds &= \int_{a_S}^{a_N} \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right] \right]_E ds = \int_{a_S}^{a_N} \left( \frac{\partial(I_M u_h)|_{T_E, E}}{\partial \mathbf{t}_E} - \frac{\partial(I_M u_h)|_{T_W, E}}{\partial \mathbf{t}_E} \right) \\ &= [I_M u_h]_E(a_N) - [I_M u_h]_E(a_S). \end{aligned}$$

Aufgrund der Stetigkeit des Morley-Interpolanten in den Knoten, vergleiche (3.8)–(3.10), gilt  $[I_M u_h]_E(a_N) = 0 = [I_M u_h]_E(a_S)$ .

Für (3.20) gilt

$$\int_E \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} ds = \int_{a_S}^{a_N} \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} ds = (I_M u_h - u_D)(a_N) - (I_M u_h - u_D)(a_S)$$

mit  $a_S$  Anfangs- und  $a_N$  Endknoten von  $E \in \mathcal{E}_\Gamma$ . Wegen (3.9) für Randknoten gilt  $(I_M u_h - u_D)(a_N) = 0 = (I_M u_h - u_D)(a_S)$ .  $\blacksquare$

**Satz 3.4.5 (Spurgleichung für Rechtecke und Quader).** *Es seien  $T \subseteq \mathbb{R}^d$  ein Quader,  $E$  eine Seite von  $T$  und  $w \in W^{1,1}(T)$ . Dann gilt*

$$\frac{1}{|T|} \int_T w(x) dx = \frac{1}{|E|} \int_E w(x) ds_x - \frac{1}{|T|} \int_T \mathbf{n}_E \cdot (x - P) \mathbf{n}_E \cdot \nabla w(x) dx, \quad (3.21)$$

wobei  $\mathbf{n}_E$  den (äußeren) Normalvektor von  $\partial T$  auf  $E$  bezeichne und  $P \in \mathbb{R}^d$  ein Eckpunkt von  $T$  sei, der nicht auf  $E$  liegt.

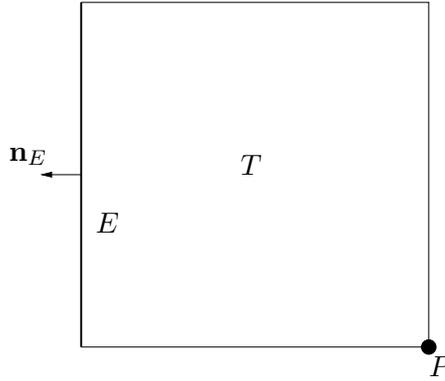


Abbildung 3.4: Veranschaulichung der Voraussetzungen zum Beweis von Satz 3.4.5.

**Beweis.** Zunächst wird der Spezialfall  $T = \hat{T} := [0, 1]^d$  und  $E = \hat{E} = 0 \times [0, 1]^{d-1}$  betrachtet. Für  $t \in [0, 1]$  und  $y \in Q := [0, 1]^{d-1}$  definieren wir  $g(t, y) := (t - 1)w(t, y)$ . Dann gilt nach dem Hauptsatz der Differential- und Integralrechnung

$$w(0, y) = (g(1, y) - g(0, y)) = \int_0^1 \frac{\partial}{\partial t} g(t, y) dt = \int_0^1 w(t, y) dt + \int_0^1 (t - 1) \frac{\partial w}{\partial t}(t, y) dt.$$

Wir schreiben daher für

$$\begin{aligned} \int_{\hat{E}} w(x) ds_x &= \int_Q w(0, y) dy = \int_Q \int_0^1 w(t, y) dt dy + \int_Q \int_0^1 (t - 1) \frac{\partial w}{\partial t}(t, y) dt dy \\ &= \int_{\hat{T}} w(x) dx + \int_{\hat{T}} (x_1 - 1) \frac{\partial w}{\partial x_1}(x) dx. \end{aligned} \quad (3.22)$$

Da  $T$  ein Quader ist, existieren zur Ecke  $P \in T \setminus E$  weitere Ecken  $a^{(1)}, \dots, a^{(d)} \in \mathbb{R}^d$  von  $T$ , so dass

$$\tau : [0, 1]^d \rightarrow T, x \mapsto a^{(1)} + x_1(P - a^{(1)}) + x_2(a^{(2)} - a^{(1)}) + \dots + x_d(a^{(d)} - a^{(1)})$$

eine affine bijektive Transformation von  $\hat{T}$  auf  $T$  mit  $E = \tau(\hat{E})$  ist. Nach dem Transformationsatz gelten für jede integrierbare Funktion  $g : T \rightarrow \mathbb{R}$  bzw.  $g : E \rightarrow \mathbb{R}$

$$\int_T g(y) dy = |T| \int_{\hat{T}} g(\tau(x)) dx \quad \text{und} \quad \int_E g(y) ds_y = |E| \int_{\hat{E}} g(\tau(x)) ds_x.$$

Nun folgt mit (3.22) angewandt auf  $(w \circ \tau) \in W^{1,1}(\hat{T})$

$$\begin{aligned} \frac{1}{|T|} \int_T w(y) dy &= \int_{\hat{T}} (w \circ \tau)(x) dx = \int_{\hat{E}} (w \circ \tau)(x) ds_x - \int_{\hat{T}} (x_1 - 1) \frac{\partial (w \circ \tau)}{\partial x_1}(x) dx \\ &= \frac{1}{|E|} \int_E w(y) ds_y - \int_{\hat{T}} (x_1 - 1) \frac{\partial (w \circ \tau)}{\partial x_1}(x) dx. \end{aligned}$$

Wir müssen nun nur noch das letzte Integral umformulieren, um auf Gleichung (3.21) zu kommen. Nach der Kettenregel für partielle Ableitungen gilt

$$\frac{\partial (w \circ \tau)}{\partial x_1}(x) = \sum_{j=1}^d \partial_j w(\tau(x)) \underbrace{\frac{\partial \tau_j}{\partial x_1}(x)}_{P_j - a_j^{(1)}} = (P - a^{(1)}) \cdot \nabla w(\tau(x)).$$

Da  $T$  ein Quader ist, sind die Vektoren  $(P - a^{(1)}), (a^{(2)} - a^{(1)}), \dots, (a^{(d)} - a^{(1)})$  paarweise orthogonal. Die Vektormultiplikation von  $\tau(x)$  mit  $(P - a^{(1)})$  für ein fixes  $x \in \hat{T}$  liefert  $\tau(x) \cdot (P - a^{(1)}) = a^{(1)} \cdot (P - a^{(1)}) + x_1 |P - a^{(1)}|^2$  und somit

$$x_1 = \frac{(\tau(x) - a^{(1)}) \cdot (P - a^{(1)})}{|P - a^{(1)}|^2}.$$

Für den Integranden des letzten Integrals folgt wegen  $\mathbf{n}_E = -\frac{(P - a^{(1)})}{|P - a^{(1)}|^2}$

$$\begin{aligned} (x_1 - 1) \frac{\partial(w \circ \tau)}{\partial x_1}(x) &= \frac{(\tau(x) - a^{(1)}) \cdot (P - a^{(1)}) - (P - a^{(1)}) \cdot (P - a^{(1)})}{|P - a^{(1)}|^2} (P - a^{(1)}) \cdot \nabla w(\tau(x)) \\ &= \frac{(\tau(x) - P) \cdot (P - a^{(1)})}{|P - a^{(1)}|^2} (P - a^{(1)}) \cdot \nabla w(\tau(x)) \\ &= \frac{(P - a^{(1)})}{|P - a^{(1)}|^2} \cdot (\tau(x) - P) \frac{(P - a^{(1)})}{|P - a^{(1)}|^2} \cdot \nabla w(\tau(x)) \\ &= \mathbf{n}_E \cdot (\tau(x) - P) \mathbf{n}_E \cdot \nabla w(\tau(x)). \end{aligned}$$

Daraus ergibt sich wegen des Transformationsatzes

$$\begin{aligned} \int_{\hat{T}} (x_1 - 1) \frac{\partial(w \circ \tau)}{\partial x_1}(x) dx &= \int_{\hat{T}} \mathbf{n}_E \cdot (\tau(x) - P) \mathbf{n}_E \cdot \nabla w(\tau(x)) dx \\ &= \frac{1}{|T|} \int_T \mathbf{n}_E \cdot (y - P) \mathbf{n}_E \cdot \nabla w(y) dy \end{aligned}$$

und insgesamt die Behauptung (3.21). ■

Als direkte Folgerung der Spurgleichung erhalten wir die sogenannte Spurgleichung.

**Lemma 3.4.6 (Spurgleichung für Rechtecke).** *Es sei  $T \subseteq \mathbb{R}^2$  ein Rechteck,  $E$  eine Seite von  $T$  und  $v \in H^1(T)$ . Dann gilt*

$$\|v - v_E\|_{L^2(E)} \leq Ch_T^{1/2} \|\nabla v\|_{L^2(T)} \quad (3.23)$$

mit der Konstanten  $C > 0$ , wobei  $C = 2^{1/4} \kappa_T^{1/2} (1 + 2\pi)^{1/2} \pi^{-1}$ .  $C$  hängt nur von der Rechtecksform ab.

**Bemerkung.** Anders als in der Literatur betrachten wir hier auf der rechten Seite lediglich die Norm über  $T$  und nicht über den Kantenpatch  $\omega_E$  (vergleiche dazu Ainsworth-Oden [1, Satz 1.7]). Dort sind auch immer nur reguläre Triangulierungen zugelassen.

**Beweis.** Es sei nun  $g \in H^1(T)$ . Dann ist  $|g|^2 = g \cdot g \in W^{1,1}(T)$  und  $\nabla |g|^2 = \nabla(g \cdot g) = 2g \nabla g$ . Mit der Spurgleichung (3.21) gilt deshalb

$$\frac{1}{|E|} \int_E |g|^2 ds = \frac{1}{|T|} \int_T |g|^2 dx + \frac{1}{|T|} \int_T \mathbf{n}_E \cdot (x - P) \mathbf{n}_E \cdot 2g \nabla g dx$$

und daher

$$\int_E |g|^2 ds \leq \frac{|E|}{|T|} \left\{ \int_T |g|^2 dx + \int_T |\mathbf{n}_E \cdot (x - P) \mathbf{n}_E \cdot 2g \nabla g| dx \right\}.$$

Nach Abschnitt 1.1.4 gilt  $|E| \leq h_T$  und  $|T| \geq (1/\sqrt{2})h_T \varrho_T$  und somit  $|E|/|T| \leq \sqrt{2}\varrho_T^{-1} = \sqrt{2}\kappa_T h_T^{-1}$ .  $\kappa_T$  hängt nur von der Form von  $T$  ab! Außerdem ist

$$|\mathbf{n}_E \cdot (x - P) \mathbf{n}_E \cdot 2g \nabla g| = \underbrace{|\mathbf{n}_E \cdot (x - P)|}_{\leq h_T} \underbrace{|\mathbf{n}_E \cdot 2g \nabla g|}_{\leq |2g \nabla g|} = h_T |2g \nabla g|,$$

und somit erhalten wir

$$\|g\|_{L^2(E)}^2 \leq \sqrt{2}\kappa_T h_T^{-1} \left\{ \|g\|_{L^2(T)}^2 + 2h_T \int_T |g \nabla g| dx \right\}.$$

Das letzte Integral schätzen wir mit der Cauchy-Schwarz Ungleichung ab und erhalten

$$\|g\|_{L^2(E)}^2 \leq \sqrt{2}\kappa_T h_T^{-1} \left\{ \|g\|_{L^2(T)}^2 + 2h_T \|g\|_{L^2(T)} \|\nabla g\|_{L^2(T)} \right\}.$$

Da  $v - v_T \in H^1(T)$  schreiben wir  $g := v - v_T$ , wenden die Poincaré Ungleichung (nach Bebendorf [2] ist die Konstante für konvexe Gebiete  $h_T/\pi$ ) an und erhalten

$$\begin{aligned} \|v - v_T\|_{L^2(E)}^2 &\leq \sqrt{2}\kappa_T h_T^{-1} \left\{ \underbrace{\|v - v_T\|_{L^2(T)}^2}_{\leq (h_T^2/\pi^2) \|\nabla v\|_{L^2(T)}^2} + 2h_T \underbrace{\|v - v_T\|_{L^2(T)} \|\nabla v\|_{L^2(T)}}_{\leq (h_T/\pi) \|\nabla v\|_{L^2(T)}} \right\} \\ &\leq \sqrt{2}\kappa_T h_T^{-1} \left\{ (1 + 2\pi) \frac{h_T^2}{\pi^2} \|\nabla v\|_{L^2(T)}^2 \right\}. \end{aligned}$$

Da  $v_E$  Bestapproximation von  $v$  in  $\mathcal{P}_0(E)$  ist, gilt  $\|v - v_E\|_{L^2(E)} \leq \|v - v_T\|_{L^2(E)}$  und somit folgt

$$\|v - v_E\|_{L^2(E)} \leq 2^{1/4} \kappa_T^{1/2} (1 + 2\pi)^{1/2} \frac{h_T^{1/2}}{\pi} \|\nabla v\|_{L^2(T)}. \quad \blacksquare$$

**Bemerkung.** Für ein Quadrat  $T$  können wir die Konstante  $C$  in (3.23) konkret angeben. Dabei ist die Abschätzung  $|E|/|T|$  schärfer, als im obigen Beweis. Es sei  $a$  die Seite des Quadrats, dann gilt  $h_T = \sqrt{2}a$ ,  $|T| = a^2 = h_T^2/2$  und  $|E| = a = h_T/\sqrt{2}$ . Die Abschätzung  $|E|/|T|$  im obigen Beweis wird deshalb durch

$$\frac{|E|}{|T|} = \frac{h_T/\sqrt{2}}{h_T^2/2} = \sqrt{2} h_T^{-1}$$

ersetzt. Beachte, dass  $k_T$  nicht mehr auftaucht. Wir erhalten deshalb eine schärfere Ungleichung, als wenn wir einfach  $\kappa_T = \sqrt{2}$  (bei Quadraten) in (3.23) einsetzen. Die Spurgleichung für Quadrate  $T$  lautet deshalb

$$\|v - v_E\|_{L^2(E)} \leq \underbrace{2^{1/4} (1 + 2\pi)^{1/2} \pi^{-1}}_{\approx 1.02157} h_T^{1/2} \|\nabla v\|_{L^2(T)}.$$

Das Gleiche gilt auch für Tetraeder  $T \subseteq \mathbb{R}^d$ , wobei lediglich die Spurgleichung modifiziert wird.

**Satz 3.4.7 (Spurgleichung für Dreiecke und Tetraeder).** *Es seien  $T \subseteq \mathbb{R}^d$  ein beschränktes Lipschitz-Gebiet,  $H$  eine Hyperebene in  $\mathbb{R}^d$  und  $E := \partial T \cap H$ . Ferner existiere ein Punkt  $P \in \mathbb{R}^d \setminus H$  mit  $T = \text{conv}(E \cup \{P\})$ . Dann gilt für alle  $w \in W^{1,1}(T)$  die Gleichheit*

$$\frac{1}{|T|} \int_T w(x) dx = \frac{1}{|E|} \int_E w(x) ds_x - \frac{1}{d} \frac{1}{|T|} \int_T (x - P) \cdot \nabla w(x) dx.$$

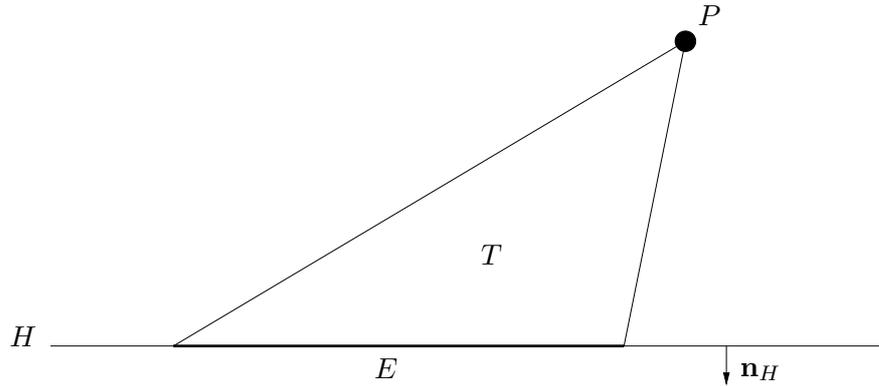


Abbildung 3.5: Veranschaulichung der Voraussetzungen zum Satz 3.4.7.

**Beweis.** Siehe Carstensen-Vorlesungsskriptum [5]. ■

**Lemma 3.4.8 (Spurungleichung für Dreiecke).** Es sei  $T \subseteq \mathbb{R}^2$  ein Dreieck,  $E$  eine Seite von  $T$  und  $v \in H^1(T)$ . Dann gilt

$$\|v - v_E\|_{L^2(E)} \leq Ch_T^{1/2} \|\nabla v\|_{L^2(T)} \quad (3.24)$$

mit  $C > 0$ , wobei  $C = 2^{1/2} \kappa_T^{1/2} (1 + \pi)^{1/2} \pi^{-1}$ .  $C$  hängt also nur von der Dreiecksform ab. ■

Mit Hilfe dieser Eigenschaften beweisen wir nun die Zuverlässigkeit unseres Residualfehlerschätzers.

**Satz 3.4.9 (Zuverlässigkeit).** Der Fehler  $e := u - I_M u_h \in H^1(T)$  erfüllt

$$\|\nabla_{\mathcal{T}} e\|_{L^2(\Omega)} \leq C \left( \sum_{T \in \mathcal{T}} \eta_T^2 \right)^{1/2}, \quad (3.25)$$

wobei die Konstante  $C > 0$  nur von der Form der Elemente in  $\mathcal{T}$  ( $\kappa_{\mathcal{T}}$ ), nicht aber der Größe (Durchmesser  $h_{\mathcal{T}}$ ) oder der Anzahl der Elemente in  $\mathcal{T}$  abhängt. Der Gradient  $\nabla_{\mathcal{T}}$  wird dabei  $\mathcal{T}$ -stückweise betrachtet:

$$(\nabla_{\mathcal{T}} e)|_T = \nabla e \quad \text{auf } T \in \mathcal{T}.$$

**Beweis.** Wir wenden die Helmholtz-Zerlegung nach Braess [3, Satz 6.1] auf  $\nabla_{\mathcal{T}} e \in L^2(\Omega)^2$  an,

$$\nabla_{\mathcal{T}} e = \nabla \rho + \text{curl } \xi \quad \rho \in H_0^1(\Omega), \quad \xi \in H^1(\Omega) \text{ mit } \int_{\Omega} \xi \, dx = 0. \quad (3.26)$$

Wir zeigen zuerst  $\nabla \rho \perp \text{curl } \xi$ :

Da  $\mathcal{C}^\infty(\overline{\Omega})$  in  $H^1(\Omega)$  dicht ist, betrachten wir zunächst die Funktion  $\xi \in \mathcal{C}^\infty(\overline{\Omega})$ . Mit partieller Integration erhalten wir

$$\langle \nabla \rho, \text{curl } \xi \rangle_{\Omega} = -\langle \rho, \text{div}(\text{curl } \xi) \rangle_{\Omega} + \int_{\partial \Omega} \rho \, \text{curl } \xi \cdot \mathbf{n} \, ds.$$

Aufgrund von  $\rho|_{\partial\Omega} = 0$  und

$$\operatorname{div}(\operatorname{curl} \xi) = \left( \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2} \right) \cdot \left( \frac{\partial \xi}{\partial x_2}, -\frac{\partial \xi}{\partial x_1} \right) = \frac{\partial^2 \xi}{\partial x_1 \partial x_2} - \frac{\partial^2 \xi}{\partial x_2 \partial x_1} = 0$$

folgt  $\langle \nabla \rho, \operatorname{curl} \xi \rangle_\Omega = 0$ . Nun betrachte ein beliebiges  $\xi \in H^1(\Omega)$ . Der Operator

$$\Theta : H^1(\Omega) \rightarrow \mathbb{R}, \xi \mapsto \langle \nabla \rho, \operatorname{curl} \xi \rangle$$

ist linear und stetig. Die Linearität ist unmittelbar klar. Die Stetigkeit folgt aus der Cauchy-Schwarz-Ungleichung,

$$|\Theta(\xi)| \leq \|\nabla \rho\|_{L^2(\Omega)} \|\operatorname{curl} \xi\|_{L^2(\Omega)} = \|\nabla \rho\|_{L^2(\Omega)} \|\nabla \xi\|_{L^2(\Omega)} \leq \|\nabla \rho\|_{L^2(\Omega)} \|\xi\|_{H^1}.$$

Da  $\mathcal{C}^\infty(\bar{\Omega})$  in  $H^1(\Omega)$  dicht ist, existiert eine Folge  $\xi_j \in \mathcal{C}^\infty(\bar{\Omega})$  mit  $\xi_j \rightarrow \xi$ . Aufgrund der Stetigkeit von  $\Theta$  und da  $\Theta|_{\mathcal{C}^\infty(\bar{\Omega})} = 0$  folgt

$$\Theta(\xi) = \lim_j \Theta(\xi_j) = 0.$$

Somit gilt  $\langle \nabla \rho, \operatorname{curl} \xi \rangle_\Omega = 0$ . Daher erhalten wir

$$\|\nabla_{\mathcal{T}e}\|_{L^2(\Omega)}^2 = \|\nabla \rho + \operatorname{curl} \xi\|_{L^2(\Omega)}^2 = \|\nabla \rho\|_{L^2(\Omega)}^2 + \|\operatorname{curl} \xi\|_{L^2(\Omega)}^2$$

und insbesondere

$$\|\nabla \rho\|_{L^2(\Omega)} \leq \|\nabla_{\mathcal{T}e}\|_{L^2(\Omega)} \text{ sowie } \|\operatorname{curl} \xi\|_{L^2(\Omega)} = \|\nabla \xi\|_{L^2(\Omega)} \leq \|\nabla_{\mathcal{T}e}\|_{L^2(\Omega)}. \quad (3.27)$$

Diese Abschätzungen werden wir am Ende des Beweises brauchen.

Zunächst schreiben wir mit (3.26)

$$\|\nabla_{\mathcal{T}e}\|_{L^2(\Omega)}^2 = \int_{\Omega} \nabla_{\mathcal{T}e} \cdot \overbrace{(\nabla \rho + \operatorname{curl} \xi)}^{=\nabla_{\mathcal{T}e}} dx = \int_{\Omega} \nabla_{\mathcal{T}e} \cdot \nabla \rho dx + \int_{\Omega} \nabla_{\mathcal{T}e} \cdot \operatorname{curl} \xi dx \quad (3.28)$$

Wir betrachten nun das erste und das zweite Integral auf der rechten Seite getrennt und schreiben für *das erste Integral*

$$\begin{aligned} \int_{\Omega} \nabla_{\mathcal{T}e} \cdot \nabla \rho dx &= \sum_{T \in \mathcal{T}} \int_T \nabla(u - I_M u_h) \cdot \nabla \rho dx \\ &= \sum_{T \in \mathcal{T}} \int_T \nabla u \cdot \nabla \rho dx - \sum_{T \in \mathcal{T}} \int_T \nabla(I_M u_h) \cdot \nabla \rho dx. \end{aligned}$$

Wir setzen (3.3) beim ersten Integral ein und wenden die partielle Integration beim zweiten Integral an und erhalten

$$= \sum_{T \in \mathcal{T}} \int_T f \rho dx + \sum_{T \in \mathcal{T}} \int_T \Delta(I_M u_h) \rho dx - \sum_{T \in \mathcal{T}} \int_{\partial T} \frac{\partial(I_M u_h)}{\partial \mathbf{n}_T} \rho ds.$$

Die ersten zwei Summen der rechten Seite fassen wir mit  $R = f + \Delta I_M u_h$  zusammen. Beim letzten Summanden verschwindet das Integral für Randkanten  $\mathcal{E}_\Gamma$  wegen  $\rho \in H_0^1(\Omega)$ . Jede Elementarkante, die nicht Teil einer hängenden Kante ist (also  $E \in \mathcal{E}^*$  mit  $E \not\subseteq \mathcal{E}^h$ ), tritt genau

zweimal auf ( $E$  grenzt an zwei Elemente  $T_W$  und  $T_E$ ), jedoch jeweils mit unterschiedlichem Vorzeichen des Normalvektors  $\sigma_{T_W,E} \mathbf{n}_E$  und  $\sigma_{T_E,E} \mathbf{n}_E$ . Vergleiche dazu Abbildung 1.8. Aufgrund der Unstetigkeit entlang der Kante  $E \in \mathcal{E}^*$  mit  $E \not\subseteq \mathcal{E}^h$  der Richtungsableitung in Normalrichtung von  $I_M u_h$  tritt der Kantensprung nach (3.13) auf. Für hängende Kanten  $E \in \mathcal{E}^h$  treten auch immer die dazugehörigen Elementarkanten  $E_j \in \mathcal{E}^*$  mit  $E = \bigcup_{i=1}^n E_i$  auf ( $E_i$  grenzt an die Elemente  $T_{W_i}$  und  $T_E$ ). Hier tritt wegen des unterschiedlichen Vorzeichens des Normalvektors und der Unstetigkeit von  $I_M u_h$  auf der Kante  $E \in \mathcal{E}^h$  der Kantensprung nach (3.14) auf. Daher gilt

$$\int_{\Omega} \nabla_{\mathcal{T}e} \cdot \nabla \rho \, dx = \sum_{T \in \mathcal{T}} \int_T R \rho \, dx - \sum_{E \in \mathcal{E}^o} \int_E \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right] \right]_E \rho \, ds.$$

Da  $\rho_T \in \mathbb{R}$  und  $\rho_E \in \mathbb{R}$  und mit Hilfe der Eigenschaft (3.12) bzw. (3.18) schreiben wir

$$\begin{aligned} &= \sum_{T \in \mathcal{T}} \int_T R(\rho - \rho_T) \, dx - \sum_{E \in \mathcal{E}^o} \int_E \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right] \right]_E (\rho - \rho_E) \, ds \\ &\leq \sum_{T \in \mathcal{T}} \int_T |R(\rho - \rho_T)| \, dx + \sum_{E \in \mathcal{E}^o} \int_E \left| \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right] \right]_E (\rho - \rho_E) \right| \, ds. \end{aligned}$$

Wir wenden die Cauchy-Schwarz-Ungleichung und anschließend die Poincaré-Ungleichung (nach Beberdorf [2] ist die Konstante für konvexe Gebiete  $h_T/\pi$ ) auf den ersten Summanden und die Spurgleichung (3.23) auf den zweiten Summanden an, wobei die Konstante  $C_1$  nur von der Form der Elemente, nicht aber von  $h_T$  oder der Anzahl der Elemente in  $\mathcal{T}$  abhängt.

$$\begin{aligned} &\leq \sum_{T \in \mathcal{T}} \|R\|_{L^2(T)} \|(\rho - \rho_T)\|_{L^2(T)} + \sum_{E \in \mathcal{E}^o} \left\| \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right] \right]_E \right\|_{L^2(E)} \|(\rho - \rho_E)\|_{L^2(E)} \\ &\leq \sum_{T \in \mathcal{T}} \|R\|_{L^2(T)} \frac{h_T}{\pi} \|\nabla \rho\|_{L^2(T)} + C_1 \sum_{E \in \mathcal{E}^o} \left\| \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right] \right]_E \right\|_{L^2(E)} h_{T_E}^{1/2} \|\nabla \rho\|_{L^2(T_E)}. \end{aligned}$$

Bei der Spurgleichung betrachten wir dabei immer  $T_E$  zu  $E$ , da nach Definition eine Kante  $E \in \mathcal{E}^o$  immer Kante von  $T_E$  ist, auch wenn die Kante hängend ist (siehe z.B. Abbildung 1.7)! Nun wenden wir wieder die Cauchy-Schwarz-Ungleichung an

$$\begin{aligned} &\leq \frac{1}{\pi} \left( \sum_{T \in \mathcal{T}} \|R\|_{L^2(T)}^2 h_T^2 \right)^{1/2} \underbrace{\left( \sum_{T \in \mathcal{T}} \|\nabla \rho\|_{L^2(T)}^2 \right)^{1/2}}_{=\|\nabla \rho\|_{L^2(\Omega)}} \\ &\quad + C_1 \left( \sum_{E \in \mathcal{E}^o} \left\| \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right] \right]_E \right\|_{L^2(E)}^2 h_{T_E} \right)^{1/2} \underbrace{\left( \sum_{E \in \mathcal{E}^o} \|\nabla \rho\|_{L^2(T_E)}^2 \right)^{1/2}}_{\leq 2\|\nabla \rho\|_{L^2(\Omega)}} \\ &\leq \left\{ \frac{1}{\pi} \left( \sum_{T \in \mathcal{T}} \|R\|_{L^2(T)}^2 h_T^2 \right)^{1/2} + 2C_1 \left( \sum_{E \in \mathcal{E}^o} \left\| \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right] \right]_E \right\|_{L^2(E)}^2 h_{T_E} \right)^{1/2} \right\} \|\nabla \rho\|_{L^2(\Omega)}, \end{aligned}$$

und mit  $(a + b) \leq \sqrt{2}(a^2 + b^2)^{1/2}$  erhalten wir schlussendlich für *das erste Integral* von (3.28)

$$\begin{aligned} \int_{\Omega} \nabla_{\mathcal{T}} e \cdot \nabla \rho \, dx &\leq \sqrt{2} \left\{ \frac{1}{\pi^2} \sum_{T \in \mathcal{T}} \|R\|_{L^2(T)}^2 h_T^2 \right. \\ &\quad \left. + 4C_1^2 \sum_{E \in \mathcal{E}^o} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right] \right\|_{L^2(E)}^2 h_{T_E} \right\}^{1/2} \|\nabla \rho\|_{L^2(\Omega)}. \end{aligned} \quad (3.29)$$

Für *das zweite Integral* von (3.28) schreiben wir

$$\int_{\Omega} \nabla_{\mathcal{T}} e \cdot \operatorname{curl} \xi \, dx = \sum_{T \in \mathcal{T}} \int_T \nabla e \cdot \operatorname{curl} \xi \, dx.$$

Da  $\nabla e \in H(\operatorname{div}, T)$ , existiert nach Girault-Raviart [11] die Spur  $\nabla e|_{\partial T}$ . Mit der partiellen Integration, mit  $\nabla e \cdot \mathbf{t}_T = -\operatorname{curl} e \cdot \mathbf{n}_T$  und  $\xi \in H^1(\Omega)$  folgt

$$\int_T \nabla e \cdot \operatorname{curl} \xi \, dx = \int_{\partial T} \nabla e \cdot \mathbf{t}_T \xi \, ds = - \int_{\partial T} \operatorname{curl} e \cdot \mathbf{n}_T \xi \, dx$$

und somit

$$\int_{\Omega} \nabla_{\mathcal{T}} e \cdot \operatorname{curl} \xi \, dx = \sum_{T \in \mathcal{T}} \int_T \nabla e \cdot \operatorname{curl} \xi \, dx = \sum_{T \in \mathcal{T}} - \int_{\partial T} \operatorname{curl} e \cdot \mathbf{n}_T \xi \, ds.$$

Da  $\operatorname{curl} e \cdot \mathbf{n}_T = \nabla e \cdot \mathbf{t}_T = \partial e / \partial \mathbf{t}_T$  gilt

$$\begin{aligned} \int_{\Omega} \nabla_{\mathcal{T}} e \cdot \operatorname{curl} \xi \, dx &= - \sum_{T \in \mathcal{T}} \int_{\partial T} \frac{\partial e}{\partial \mathbf{t}_T} \xi \, ds = - \sum_{T \in \mathcal{T}} \int_{\partial T} \frac{\partial}{\partial \mathbf{t}_T} (u - I_M u_h) \xi \, ds \\ &= - \sum_{T \in \mathcal{T}} \int_{\partial T} \frac{\partial u}{\partial \mathbf{t}_T} \xi \, ds + \sum_{T \in \mathcal{T}} \int_{\partial T} \frac{\partial I_M u_h}{\partial \mathbf{t}_T} \xi \, ds. \end{aligned}$$

Vom ersten Summanden bleibt nur der Anteil der Randkanten übrig: Für eine hängende Kante  $E \in \mathcal{E}^h$  können wir das Integral über  $E$  auch als Summe der Integrale über seine dazugehörigen Elementarkanten  $E_i \in \mathcal{E}^*$ ,  $E_i \subseteq E$  ( $i = 1, \dots, n$ ) schreiben. Darum verschwindet die neue Summe für jede Kante  $E \in \mathcal{E}^*$ , da  $u$  in Tangentialrichtung nicht springt und jede Kante genau zweimal auftritt, jedoch jeweils mit unterschiedlichem Vorzeichen des Tangentialvektors. Für den zweiten Summanden können wir hängende Kanten wie oben aufteilen und betrachten wieder Kanten aus  $E \in \mathcal{E}^*$ , die genau zweimal vorkommen. Aufgrund der Unstetigkeit entlang der Kante aus  $\mathcal{E}^*$  der Richtungsableitung in Tangentialrichtung von  $I_M u_h$  tritt der Kantensprung nach (3.13) auf. Den Randanteil des zweiten Summanden verknüpfen wir mit dem Anteil des ersten Summanden und erhalten somit

$$\int_{\Omega} \nabla_{\mathcal{T}} e \cdot \operatorname{curl} \xi \, dx = \sum_{E \in \mathcal{E}^*} \int_E \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \xi \, ds + \sum_{E \in \mathcal{E}_\Gamma} \int_E \left( \frac{\partial (I_M u_h - u_D)}{\partial \mathbf{t}_E} \right) \xi \, ds.$$

Wieder können wir die Integrale wegen (3.19) bzw. (3.20) mit  $\xi_E \in \mathbb{R}$  ergänzen und erhalten

$$\begin{aligned} &= \sum_{E \in \mathcal{E}^*} \int_E \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E (\xi - \xi_E) \, ds + \sum_{E \in \mathcal{E}_\Gamma} \int_E \left( \frac{\partial (I_M u_h - u_D)}{\partial \mathbf{t}_E} \right) (\xi - \xi_E) \, ds \\ &\leq \sum_{E \in \mathcal{E}^*} \int_E \left| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E (\xi - \xi_E) \right| \, ds + \sum_{E \in \mathcal{E}_\Gamma} \int_E \left| \left( \frac{\partial (I_M u_h - u_D)}{\partial \mathbf{t}_E} \right) (\xi - \xi_E) \right| \, ds. \end{aligned}$$

Wir wenden die Cauchy-Schwarz-Ungleichung und anschließend die Spurgleichung (3.23) an, wobei die Konstanten  $C_2$  nur von der Form der Elemente in  $\mathcal{T}$  abhängt, nicht aber von  $h_T$  oder der Anzahl der Elemente in  $\mathcal{T}$ .

$$\begin{aligned} \int_{\Omega} \nabla_{\mathcal{T}e} \cdot \operatorname{curl} \xi \, dx &\leq \sum_{E \in \mathcal{E}^*} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right\|_{L^2(E)} \|\xi - \xi_E\|_{L^2(E)} \\ &\quad + \sum_{E \in \mathcal{E}_{\Gamma}} \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)} \|\xi - \xi_E\|_{L^2(E)} \\ &\leq C_2 \left( \sum_{E \in \mathcal{E}^*} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right\|_{L^2(E)} h_{T_W}^{1/2} \|\nabla \xi\|_{L^2(T_W)} \right. \\ &\quad \left. + \sum_{E \in \mathcal{E}_{\Gamma}} \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)} h_W^{1/2} \|\nabla \xi\|_{L^2(T_W)} \right). \end{aligned}$$

Bei der ersten Abschätzung durch die Spurgleichung betrachten wir dabei zu  $E \in \mathcal{E}^*$  immer  $T_W$ . Dies ist auch dann möglich, wenn  $E$  Teil einer hängenden Kante ist, da per Definition  $E \in \mathcal{E}_{T_W}$  gilt. Bei der zweiten Abschätzung für die Randkanten existiert  $T_W$  immer, da nach Definition der Normalvektor nach außen zeigt. Nach Anwendung der Cauchy-Schwarz-Ungleichung schreiben wir

$$\begin{aligned} &\leq C_2 \left\{ \left( \sum_{E \in \mathcal{E}^*} h_{T_W} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right\|_{L^2(E)}^2 \right)^{1/2} \underbrace{\left( \sum_{E \in \mathcal{E}^*} \|\nabla \xi\|_{L^2(T_W)}^2 \right)^{1/2}}_{\leq 2\|\nabla \xi\|_{L^2(\Omega)}} \right. \\ &\quad \left. + \left( \sum_{E \in \mathcal{E}_{\Gamma}} h_W \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}^2 \right)^{1/2} \underbrace{\left( \sum_{E \in \mathcal{E}_{\Gamma}} \|\nabla \xi\|_{L^2(T_W)}^2 \right)^{1/2}}_{\leq 2\|\nabla \xi\|_{L^2(\Omega)}} \right\}. \end{aligned}$$

Die 2 in der Abschätzung der zwei Summen kommt daher, da jedes Element höchstens viermal betrachtet wird (Dreieck oder Rechteck!). Somit erhalten wir für *das zweite Integral* von (3.28) insgesamt

$$\begin{aligned} \int_{\Omega} \nabla_{\mathcal{T}e} \cdot \operatorname{curl} \xi \, dx &\leq 2C_2 \left\{ \left( \sum_{E \in \mathcal{E}^*} h_{T_W} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right\|_{L^2(E)}^2 \right)^{1/2} \right. \\ &\quad \left. + \left( \sum_{E \in \mathcal{E}_{\Gamma}} h_T \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}^2 \right)^{1/2} \right\} \|\nabla \xi\|_{L^2(\Omega)}. \end{aligned} \quad (3.30)$$

Damit bekommen wir für (3.28) mit (3.29) und (3.30) die Abschätzung

$$\begin{aligned} \|\nabla_{\mathcal{T}e}\|_{L^2(\Omega)}^2 &\leq \sqrt{2} \left\{ \frac{1}{\pi^2} \sum_{T \in \mathcal{T}} \|R\|_{L^2(T)}^2 h_T^2 + 4C_1^2 \sum_{E \in \mathcal{E}^o} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)}^2 h_{T_E} \right\}^{1/2} \|\nabla \rho\|_{L^2(\Omega)} \\ &\quad + 2C_2 \left\{ \left( \sum_{E \in \mathcal{E}^*} h_{T_W} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right\|_{L^2(E)}^2 \right)^{1/2} \right. \\ &\quad \left. + \left( \sum_{E \in \mathcal{E}_{\Gamma}} h_T \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}^2 \right)^{1/2} \right\} \|\nabla \xi\|_{L^2(\Omega)}. \end{aligned}$$

Nach Anwendung von (3.27) erhalten wir

$$\begin{aligned} \|\nabla_{\mathcal{T}} e\|_{L^2(\Omega)} &\leq \sqrt{2} \left\{ \frac{1}{\pi^2} \sum_{T \in \mathcal{T}} \|R\|_{L^2(T)}^2 h_T^2 + 4C_2^2 \sum_{E \in \mathcal{E}^\circ} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)}^2 h_{T_E} \right\}^{1/2} \\ &\quad + 2C_2 \left\{ \left( \sum_{E \in \mathcal{E}^*} h_{T_W} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right\|_{L^2(E)}^2 \right)^{1/2} \right. \\ &\quad \left. + \left( \sum_{E \in \mathcal{E}_\Gamma} h_T \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}^2 \right)^{1/2} \right\}. \end{aligned}$$

Wir wenden  $a + b + c \leq \sqrt{3}(a^2 + b^2 + c^2)^{1/2}$  an, und daher gilt

$$\begin{aligned} \|\nabla_{\mathcal{T}} e\|_{L^2(\Omega)} &\leq \sqrt{3} \left\{ \frac{2}{\pi^2} \sum_{T \in \mathcal{T}} \|R\|_{L^2(T)}^2 h_T^2 + 8C_1^2 \sum_{E \in \mathcal{E}^\circ} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)}^2 h_{T_E} \right. \\ &\quad + 4C_2^2 \sum_{E \in \mathcal{E}^*} h_{T_W} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right\|_{L^2(E)}^2 \\ &\quad \left. + 4C_2^2 \sum_{E \in \mathcal{E}_\Gamma} h_T \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}^2 \right\}^{1/2}. \end{aligned}$$

Schlussendlich schreiben wir für

$$\begin{aligned} \sum_{E \in \mathcal{E}^\circ} h_{T_E} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)}^2 &\leq \sum_{T \in \mathcal{T}} h_T \sum_{E \in \mathcal{E}_T \setminus \mathcal{E}_\Gamma} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)}^2, \\ \sum_{E \in \mathcal{E}^*} h_{T_W} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right\|_{L^2(E)}^2 &\leq \sum_{T \in \mathcal{T}} h_T \sum_{E \in \mathcal{E}_T \setminus \mathcal{E}_\Gamma} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right\|_{L^2(E)}^2, \\ \sum_{E \in \mathcal{E}_\Gamma} h_T \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}^2 &\leq \sum_{T \in \mathcal{T}} h_T \sum_{E \in \mathcal{E}_T \cap \mathcal{E}_\Gamma} \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}^2, \end{aligned}$$

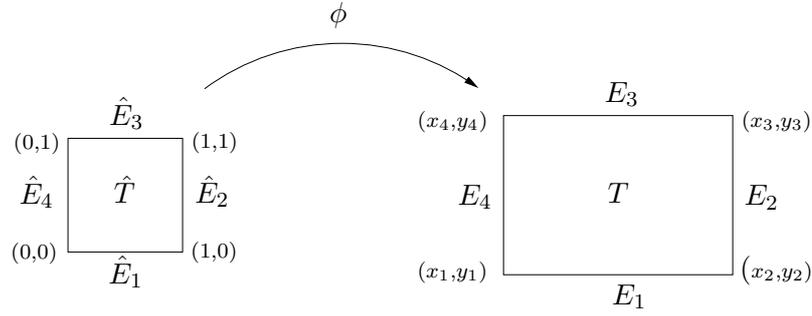
und da

$$\|R\|_{L^2(T)}^2 = \|f - f_{\mathcal{T}}\|_{L^2(T)}^2 + \|R_{\mathcal{T}}\|_{L^2(T)}^2 = \|f - f_{\mathcal{T}}\|_{L^2(T)}^2$$

(nach Lemma 3.3.2 ist  $\|R_{\mathcal{T}}\|_{L^2(T)} = 0$ ) gilt, erhalten wir

$$\begin{aligned} \|\nabla_{\mathcal{T}} e\|_{L^2(\Omega)} &\leq \sqrt{3} \left\{ \sum_{T \in \mathcal{T}} \left[ h_T \left( \frac{2}{\pi^2} h_T \|f - f_{\mathcal{T}}\|_{L^2(T)}^2 \right. \right. \right. \\ &\quad + 8C_1^2 \sum_{E \in \mathcal{E}_T \setminus \mathcal{E}_\Gamma} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)}^2 + 4C_2^2 \sum_{E \in \mathcal{E}_T \setminus \mathcal{E}_\Gamma} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right\|_{L^2(E)}^2 \\ &\quad \left. \left. + 4C_2^2 \sum_{E \in \mathcal{E}_T \cap \mathcal{E}_\Gamma} \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}^2 \right) \right\}^{1/2} \\ &\leq C \left( \sum_{T \in \mathcal{T}} \eta_T^2 \right)^{1/2} = C\eta. \end{aligned}$$

■


 Abbildung 3.6: Die affine bijektive Abbildung  $\phi : \hat{T} \rightarrow T$ 

### 3.4.2 Effizienz

Bevor wir die Effizienz unseres Fehlerschätzers beweisen können, führen wir die sogenannten Bubble-Funktionen sowie eine Extension-Funktion ein und zeigen einige wichtige Abschätzungen.

#### Bubble-Funktion auf Rechtecken

Im Folgenden sei  $\hat{T} \subseteq \mathbb{R}^2$  das Referenzelement mit den Eckpunkten  $\hat{P}_1 = (0,0)$ ,  $\hat{P}_2 = (1,0)$ ,  $\hat{P}_3 = (1,1)$ ,  $\hat{P}_4 = (0,1)$ , die gegen den Uhrzeigersinn angeordnet sind. Die Kante  $\hat{E}_1$  gehe von  $\hat{P}_1$  nach  $\hat{P}_2$  usw. Die Funktionen

$$\hat{\varphi}_1(\hat{x}, \hat{y}) = (1 - \hat{x})(1 - \hat{y}), \quad \hat{\varphi}_2(\hat{x}, \hat{y}) = \hat{x}(1 - \hat{y}), \quad \hat{\varphi}_3(\hat{x}, \hat{y}) = \hat{x}\hat{y}, \quad \hat{\varphi}_4(\hat{x}, \hat{y}) = (1 - \hat{x})\hat{y}$$

mit  $0 \leq \hat{x}, \hat{y} \leq 1$  bilden eine Basis von  $\mathcal{Q}_1(\hat{T})$  und auch den Ansatzraum auf  $\hat{T}$  bei der  $\mathcal{Q}_1$ -FEM (nodale Basisfunktionen). Wir betrachten nun eine affine bijektive Abbildung  $\phi : \hat{T} \rightarrow T$ , die das Referenzelement auf ein beliebiges Rechteck  $T \in \mathcal{T}$  mit den Eckpunkten  $P_j = (x_j, y_j)$  ( $1 \leq j \leq 4$ ) abbildet (vergleiche dazu Abbildung 3.6). Wir setzen die Funktion  $\phi$  allgemein an:

$$\phi(\hat{x}, \hat{y}) : \hat{T} \rightarrow T : (\hat{x}, \hat{y}) \mapsto M \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} + \mathbf{c} \quad \text{mit} \quad M = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix} \in \mathbb{R}^{2 \times 2},$$

$$\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \in \mathbb{R}^2.$$

$\phi$  bildet eine Ecke von  $\hat{T}$  auf  $T$  ab, somit können wir die Matrix  $M$  und den Vektor  $\mathbf{c}$  bestimmen:

$$\phi(0,0) = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \mathbf{c},$$

$$\phi(1,0) = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \Rightarrow \begin{pmatrix} m_1 \\ m_3 \end{pmatrix} = \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix},$$

$$\phi(0,1) = \begin{pmatrix} x_4 \\ y_4 \end{pmatrix} = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \Rightarrow \begin{pmatrix} m_2 \\ m_4 \end{pmatrix} = \begin{pmatrix} x_4 - x_1 \\ y_4 - y_1 \end{pmatrix}.$$

Da  $\phi$  affiner Diffeomorphismus, bildet  $\phi$  den Rand  $\partial\hat{T}$  auf  $\partial T$ ,  $\hat{E}_j$  auf  $E_j$  und  $\hat{P}_j$  auf  $P_j$  ab. Die Funktionen

$$\varphi_i(x, y) := \hat{\varphi}_i(\phi^{-1}(x, y)) \quad i = 1 \dots 4, \quad (x, y) \in T \tag{3.31}$$

bilden eine Basis von  $\mathcal{Q}_1(T)$ . Wie leicht zu erkennen ist, erfüllt  $\varphi_i$  für  $P_j = (x_j, y_j)$  und  $(x, y) \in T$  folgende Eigenschaften:

$$\varphi_i(x_j, y_j) = \delta_{ij} \quad i, j = 1 \dots 4$$

und

$$\begin{aligned} \varphi_1(x, y) = 0 \quad \text{auf} \quad E_2, E_3, \quad \varphi_2(x, y) = 0 \quad \text{auf} \quad E_3, E_4, \\ \varphi_3(x, y) = 0 \quad \text{auf} \quad E_1, E_4, \quad \varphi_4(x, y) = 0 \quad \text{auf} \quad E_1, E_2. \end{aligned}$$

**Bemerkung.** Die in (3.31) definierten Funktionen werden bei der  $\mathcal{Q}_1$ -FEM auch zur Konstruktion der sogenannten **Hutfunktionen** verwendet: Es gibt auf einer regulären Triangulierung mit  $z \in \mathcal{N}$  genau eine global stetige Funktion  $\eta_z \in \mathcal{C}(\Omega) \cap \mathcal{Q}_1(\mathcal{T})$ , die in  $z$  den Wert 1 annimmt, den Träger  $\omega_z$  (Knotenpatch) besitzt und auf jedem  $T \in \mathcal{T}$  affin ist, d.h.

$$\begin{aligned} \eta_z(z) &= 1, \\ \eta_z(a) &= 0, \quad a \in \mathcal{N} \setminus \{z\}. \end{aligned}$$

$\eta_z$  heißt Hutfunktion und wird elementweise konstruiert. Für jedes  $T \in \mathcal{T}$  mit  $T \subseteq \omega_z$  ist  $z$  Eckpunkt von  $T$ . Sei  $P_i$  ( $1 \leq i \leq 4$ ) derjenige Knoten von  $T$ , der mit  $z$  ident ist, also  $P_i = z$ . Dann folgt  $\eta_z|_T = \varphi_i$ . Die Hutfunktion  $\eta_z$  ist stetig, denn auf Elementkanten ist  $\eta_z$  affin und an den Ecken stetig. Folglich ist  $\eta_z \in H^1(\Omega)$ .

**Definition 3.4.10 (Bubble-Funktion  $b_T$ ).** *Es sei  $T$  ein Rechteck. Die Bubble-Funktion  $b_T$  definieren wir mit*

$$b_T := 16 \varphi_1 \varphi_3 \quad \text{auf} \quad T, \tag{3.32}$$

wobei  $\varphi_1$  und  $\varphi_3$  die nach (3.31) definierten Funktionen sind.

Die Bubble-Funktion  $b_T$  hat folgende Eigenschaften auf der fastregulären Triangulierung  $\mathcal{T}$ , wie elementar nachgerechnet werden kann:

$$b_T = 0 \quad \text{auf} \quad \partial T, \quad 0 \leq b_T \leq 1, \quad \max_{x \in T} b_T(x) = 1, \quad T \in \mathcal{T}.$$

Insbesondere ist  $b_T$  ein Polynom vom totalen Grad 4, und  $b_T$  ist auf  $T \in \mathcal{T}$  wohldefiniert.

**Definition 3.4.11 (Bubble-Funktion  $b_E$  für  $E \notin \mathcal{E}_\Gamma$ ).** *Es seien nun  $T_1$  und  $T_2$  zwei Rechtecke, die eine gemeinsame Kante  $E = T_1 \cap T_2$  haben. Die Knoten von  $T_1$  bzw.  $T_2$  seien so angeordnet, dass die Knoten (Anfangs- und Endknoten) von  $E$  zuerst nummeriert sind (Abbildung 3.7).  $\varphi_{1,T_1}$  und  $\varphi_{2,T_1}$  seien die nach (3.31) definierten Funktionen auf  $T_1$ ,  $\varphi_{1,T_2}$  und  $\varphi_{2,T_2}$  die auf  $T_2$ . Die Bubble-Funktion  $b_E$  auf  $T_1 \cup T_2$  definieren wir mit*

$$b_E := \begin{cases} 4\varphi_{1,T_1}\varphi_{2,T_1} & \text{auf} \quad T_1, \\ 4\varphi_{1,T_2}\varphi_{2,T_2} & \text{auf} \quad T_2. \end{cases} \tag{3.33}$$

$b_E|_{T_1}$  und  $b_E|_{T_2}$  sind Polynome vom totalen Grad 4.  $b_E$  ist nach Konstruktion auf  $E$  stetig und insgesamt in  $H_0^1(T_1 \cup T_2)$ .

**Bemerkung.** Beachte, dass die Bubble-Funktion  $b_E$  nur auf regulären Triangulierungen für alle inneren Kanten definiert ist, da für zwei Elemente  $T_1$  und  $T_2$ ,  $E = T_1 \cap T_2$  eine *gemeinsame* Kante ist. Bei fastregulären Triangulierungen muss das nicht sein! Bei einer fastregulären Triangulierung ist  $b_E$  auf  $\omega_E^*$  für alle  $E \in \mathcal{E}^*$  wohldefiniert. Für hängende Kanten  $E$  ist  $b_E$  nicht definiert!

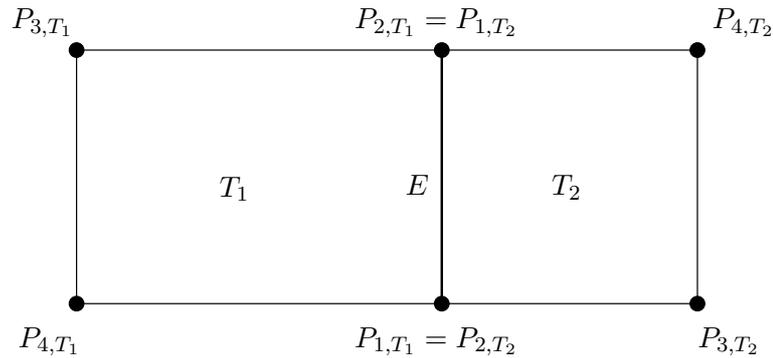


Abbildung 3.7: Bezeichnungen für die Definition der Bubble-Funktion  $b_E$  auf zwei Rechtecke.

**Definition 3.4.12 (Bubble-Funktion  $b_E$  für  $E \in \mathcal{E}_\Gamma$ ).** Sei  $E \in \mathcal{E}_\Gamma$  eine Randkante von  $T \in \mathcal{T}$ , dann gilt

$$b_E := 4\varphi_1\varphi_2 \quad \text{auf } T, \tag{3.34}$$

wobei  $P_1$  und  $P_2$  die Knoten (Anfangs- bzw. Endknoten) von  $E$  sind.  $\varphi_1$  und  $\varphi_2$  seien die nach (3.31) definierten Funktionen auf  $T$ .

Die Bubble-Funktion  $b_E$  hat somit folgende Eigenschaften auf der fastregulären Triangulierung  $\mathcal{T}$ :

$$b_E = 0 \quad \text{auf } \partial T \setminus E \quad \text{mit } T \in \tilde{\omega}_E^*, \quad 0 \leq b_E \leq 1, \quad \max_{x \in E} b_E(x) = 1, \quad E \in (\mathcal{E}^* \cup \mathcal{E}_\Gamma).$$

Daher folgt direkt,

$$b_E = 0 \quad \text{auf } \partial \omega_E^* \quad \text{für } E \in \mathcal{E}^*.$$

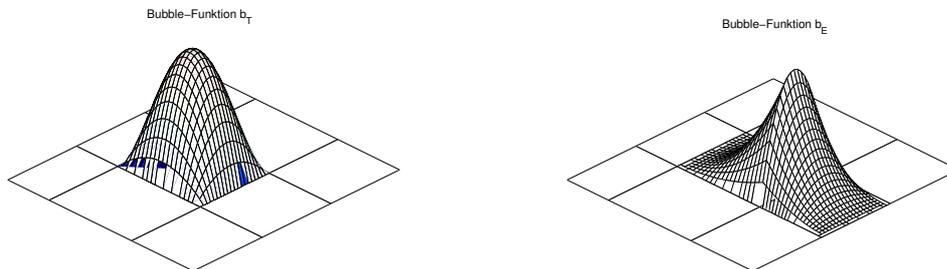


Abbildung 3.8: Die Bubble-Funktionen  $b_T$  (links) und  $b_E$  (rechts) auf einer Triangulierung mit Rechtecken.

**Bubble-Funktion auf Dreiecken**

Das Referenzelement Element  $\hat{T} \subseteq \mathbb{R}^2$  für Dreiecke hat die Eckpunkte  $\hat{P}_1 = (0, 0)$ ,  $\hat{P}_2 = (1, 0)$ ,  $\hat{P}_3 = (0, 1)$ , die gegen den Uhrzeigersinn angeordnet sind. Die (Ansatz-) Funktionen

$$\hat{\varphi}_1(\hat{x}, \hat{y}) = (1 - \hat{x} - \hat{y}), \quad \hat{\varphi}_2(\hat{x}, \hat{y}) = \hat{x}, \quad \hat{\varphi}_3(\hat{x}, \hat{y}) = \hat{y} \quad (3.35)$$

mit  $0 \leq \hat{x}, \hat{y} \leq 1$  bilden eine Basis von  $\mathcal{P}_1(\hat{T})$  (nodale Basis). Damit werden die Bubblefunktionen  $b_T$  auf  $T \in \mathcal{T}$  ( $\mathcal{T}$  fastreguläre Triangulierung) und  $b_E$  für  $E \in (\mathcal{E}^* \cup \mathcal{E}_\Gamma)$  auf Dreiecken konstruiert mit den Eigenschaften:

$$\begin{aligned} b_T &= 0 \quad \text{auf} \quad \partial T, \quad 0 \leq b_T \leq 1, \quad \max_{x \in T} b_T(x) = 1, \quad T \in \mathcal{T}, \\ b_E &= 0 \quad \text{auf} \quad \partial T \setminus E \quad \text{mit} \quad T \in \tilde{\omega}_E^*, \quad 0 \leq b_E \leq 1, \quad \max_{x \in E} b_E(x) = 1, \quad E \in (\mathcal{E}^* \cup \mathcal{E}_\Gamma). \end{aligned}$$

**Bemerkung.** In der Literatur (z.B. Verfürth [20]) werden die Bubblefunktionen auf Dreiecken oft auch über die *Baryzentrischen Koordinaten* definiert. Diese sind auf dem Dreieck eindeutig (nicht auf einem Rechteck<sup>1</sup>). Sei  $T \in \mathcal{T}$  ein Dreieck mit den Ecken  $P_1, P_2$  und  $P_3$ , dann gilt

$$T = \text{conv}\{P_1, P_2, P_3\} = \left\{ \sum_{j=1}^3 \lambda_j P_j : 0 \leq \lambda_j \leq 1, \sum_{j=1}^3 \lambda_j = 1 \right\}. \quad (3.36)$$

Für  $(x, y) \in T$  sind die  $\lambda_j \doteq \lambda_j(x, y)$  eindeutig definiert und heißen Baryzentrische Koordinaten. Für  $T \in \mathcal{T}$  mit den Baryzentrischen Koordinaten  $\lambda_{T,1}, \lambda_{T,2}$  und  $\lambda_{T,3}$  wird die Dreiecks-Bubblefunktion definiert

$$b_T := 27 \lambda_{T,1} \lambda_{T,2} \lambda_{T,3} \quad \text{auf} \quad T.$$

Für  $E \in \mathcal{E}^*$  mit  $\omega_E^* = T_1 \cup T_2$  werden die Ecken von  $T_1$  und  $T_2$  so angeordnet, dass die Knoten von  $E$  zuerst nummeriert sind. Dann definiert

$$b_E := 4 \lambda_{1,T_i} \lambda_{2,T_i} \quad \text{auf} \quad T_i, \quad i = 1, 2$$

die Bubblefunktion auf  $E \in \mathcal{E}^*$ . Für  $E \in \mathcal{E}_\Gamma$  definieren wir  $b_E := 4 \lambda_{1,T} \lambda_{2,T}$  mit  $E \subseteq \partial T$ . Die Baryzentrischen Koordinaten  $\hat{\lambda}_j$  ( $j = 1 \dots 3$ ) auf dem Referenzdreieck  $\hat{T}$  sind gerade die nodalen Basisfunktionen

$$\hat{\lambda}_1(\hat{x}, \hat{y}) = (1 - \hat{x} - \hat{y}) = \hat{\varphi}_1(\hat{x}, \hat{y}), \quad \hat{\lambda}_2(\hat{x}, \hat{y}) = \hat{x} = \hat{\varphi}_2(\hat{x}, \hat{y}), \quad \hat{\lambda}_3(\hat{x}, \hat{y}) = \hat{y} = \hat{\varphi}_3(\hat{x}, \hat{y}),$$

mit  $0 \leq \hat{x}, \hat{y} \leq 1$ . Es sei  $\phi(\mathbf{x}) : \hat{T} \rightarrow T : \mathbf{x} \mapsto M \cdot \mathbf{x} + \mathbf{c}$  eine affine Transformation, wobei  $\hat{P}_1 = (0, 0)$ ,  $\hat{P}_2 = (1, 0)$  und  $\hat{P}_3 = (0, 1)$  die Ecken des Referenzdreiecks sind, und  $\hat{P} \in \hat{T}$  habe die Darstellung

$$\hat{P} = \sum_{j=1}^3 \hat{\lambda}_j \hat{P}_j.$$

---

<sup>1</sup>Werden die Baryzentrischen Koordinaten auf dem Rechteck geeignet gewählt, so ist die Konstruktion der Bubblefunktionen auch über diese möglich.

Ein Bildpunkt unter einer affinen Abbildung hat dieselben Baryzentrischen Koordinaten. Es gilt also

$$P = \phi(\hat{P}) = \sum_{j=1}^3 M\hat{\lambda}_j \hat{P}_j + \mathbf{c} = \sum_{j=1}^3 \hat{\lambda}_j \underbrace{(M\hat{P}_j + \mathbf{c})}_{=P_j} - \mathbf{c} \underbrace{\sum_{j=1}^3 \hat{\lambda}_j}_{=1} + \mathbf{c} = \sum_{j=1}^3 \hat{\lambda}_j P_j,$$

wobei  $P_j$  die Ecken von  $T$  sind und  $P \in T$ . Somit ist klar, dass diese Definition der Bubblefunktionen mit Baryzentrischen Koordinaten auf dasselbe hinausführt wie mittels Ansatzfunktionen.

### Extension-Operator $F_{ext}$

Der Extension-Operator  $F_{ext} : C(E) \rightarrow C(T)$  setzt jede auf einer Kante  $E$  definierte Funktion auf  $T$  fort. Wir betrachten wieder zuerst das Referenzrechteck  $\hat{T} := \text{conv}\{(0, 0), (1, 0), (1, 1), (0, 1)\}$  bzw. das Referenzdreieck  $\hat{T} := \text{conv}\{(0, 0), (1, 0), (0, 1)\}$ , und  $\hat{E}$  sei die Kante  $[0, 1] \times \{0\}$  von  $\hat{T}$ .  $\phi$  sei wieder die affine Abbildung, die  $\hat{T}$  auf  $T$  und  $\hat{E}$  auf  $E$  abbildet. Wir definieren den Extension-Operator  $\hat{F}_{ext} : C(\hat{E}) \rightarrow C(\hat{T})$  mit

$$\hat{F}_{ext}(\hat{v})(x, y) := \hat{v}(x) \quad \text{für alle } \hat{v} \in C(\hat{E}), \quad (x, y) \in \hat{T}.$$

Dann ist  $F_{ext}$  definiert mit

$$F_{ext}(v) := [\hat{F}_{ext}(v \circ \phi)] \circ \phi^{-1} \quad \text{für alle } v \in C(E).$$

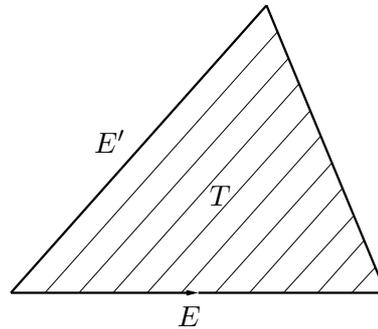


Abbildung 3.9: Niveaulinien bei der Erweiterung einer Funktion durch  $F_{ext}$  von  $E$  auf  $T$ .

**Bemerkung.** Beachte, dass  $F_{ext}(v)$  konstant entlang Linien ist, die parallel zu  $E'$  verlaufen, wobei  $E'$  diejenige Kante von  $T$  ist, welche als rechten Endpunkt den linken Endpunkt von  $E$  hat (siehe Abbildung 3.9).

### Abschätzungen

Wir betrachten nun die affine Abbildung  $\phi : \hat{T} \rightarrow T$ , die das Referenzelement auf das Element  $T$  abbildet. Mit der Bezeichnung von Abbildung 3.10 gilt für Dreiecke

$$\begin{aligned} \phi_{\Delta} : \hat{T} \rightarrow T : \mathbf{x} \mapsto M_{\Delta} \cdot \mathbf{x} + \mathbf{c} \quad \text{mit} \quad M_{\Delta} &= \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}, \\ \mathbf{c} &= \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \in \mathbb{R}^{2 \times 1} \end{aligned} \tag{3.37}$$

und für Rechtecke

$$\begin{aligned} \phi_{\square} : \hat{T} \rightarrow T : \mathbf{x} \mapsto M_{\square} \cdot \mathbf{x} + \mathbf{c} \quad \text{mit} \quad M_{\square} = \begin{pmatrix} x_2 - x_1 & x_4 - x_1 \\ y_2 - y_1 & y_4 - y_1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}, \\ \mathbf{c} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \in \mathbb{R}^{2 \times 1}. \end{aligned} \quad (3.38)$$

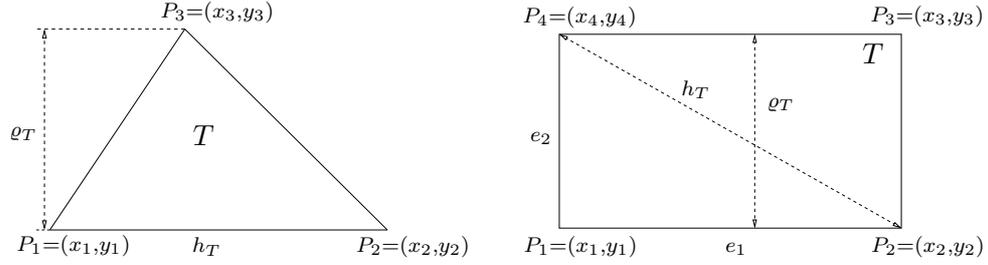


Abbildung 3.10: Der Einfluss der Form der Elemente

Beim Dreieck erhalten wir

$$|\det(D\phi_{\Delta})| = |\det(M_{\Delta})| = 2|T| = 2\frac{1}{2}h_T\rho_T = h_T\rho_T.$$

Beim Rechteck gilt wegen  $e_2 = \rho_T$  ( $e_1$  sei die längere Seite,  $e_1 \leq h_T$ ,  $e_1 \geq (1/\sqrt{2})h_T$ )

$$|\det(D\phi_{\square})| = |\det(M_{\square})| = |T| = e_1e_2 \leq h_T\rho_T,$$

$$|\det(D\phi_{\square})| = |\det(M_{\square})| = |T| = e_1e_2 \geq (1/\sqrt{2})h_T\rho_T.$$

Wir schreiben daher allgemein für eine affine bijektive Abbildung

$$\phi : \hat{T} \rightarrow T : \mathbf{x} \mapsto M \cdot \mathbf{x} + \mathbf{c}, \quad (3.39)$$

die das Referenzdreieck bzw. Referenzrechteck auf ein Dreieck bzw. Rechteck abbildet,

$$\frac{1}{\sqrt{2}}h_T\rho_T \leq |\det(D\phi)| = |\det(M)| \leq h_T\rho_T. \quad (3.40)$$

Das folgende Lemma beschreibt einige nützliche Abschätzungen, die wir für den Beweis der Effizienz unseres Fehlerschätzers benötigen. Sofern die Beweise nicht direkt folgen, beweisen wir die Aussagen zuerst auf dem Referenzelement  $\hat{T}$ , danach zeigen wir mittels Transformation, wie sich die Abschätzung auf  $T$  überträgt. Diese Vorgehensweise hat den Vorteil, dass die auftretenden Konstanten in den Abschätzungen nur von der Form der Elemente  $T$ , nicht aber von  $h_T$  bzw.  $h_E$  abhängen. Für diese Technik ist die Bezeichnung **Skalierungsargument** üblich.

**Lemma 3.4.13.** *Es sei  $p \in \mathbb{N}_0$ ,  $E \in (\mathcal{E}^* \cup \mathcal{E}_{\Gamma})$ ,  $T \in \omega_E^*$ ,  $v \in \mathcal{P}_p(T)$ ,  $w \in \mathcal{P}_p(E)$ ,  $b_T$  und  $b_E$  die Bubble-Funktionen auf  $T$  bzw.  $E$ , dann gilt:*

$$C_1\|v\|_{L^2(T)} \leq \|vb_T\|_{L^2(T)} \leq \|v\|_{L^2(T)} \quad (3.41)$$

$$C_2\|w\|_{L^2(E)} \leq \|wb_E\|_{L^2(E)} \leq \|w\|_{L^2(E)} \quad (3.42)$$

$$C_3h_E^{1/2}\|w\|_{L^2(E)} \leq \|F_{ext}(w)b_E\|_{L^2(T)} \leq C_4h_E^{1/2}\|w\|_{L^2(E)} \quad (3.43)$$

Die Konstanten  $C_1 > 0$  und  $C_2 > 0$  hängen nur von  $p$  und die Konstanten  $C_3 > 0$  und  $C_4 > 0$  von  $p$  und der Form von  $T$  ab, nicht aber von  $h_T$ .

**Beweis.** (3.41)a beweisen wir zuerst auf dem Referenzelement  $\hat{T}$ . Es sei  $\phi : \hat{T} \rightarrow T$  die affine Abbildung aus (3.39). Dann gilt  $\hat{v} := v \circ \phi \in \mathcal{P}_p(\hat{T})$ ,  $b_{\hat{T}} := b_T \circ \phi$  und  $\|\hat{v}\| := \|\hat{v}b_{\hat{T}}\|_{L^2(\hat{T})}$  definiert eine Norm auf  $\mathcal{P}_p(\hat{T})$ :

1. Die Definitheit gilt, da  $b_{\hat{T}}$  im Inneren von  $\hat{T}$  nicht verschwindet. Das heißt, für alle  $\hat{v} \in \mathcal{P}_p(\hat{T})$  gilt  $\|\hat{v}\| = \|\hat{v}b_{\hat{T}}\|_{L^2(\hat{T})} = 0$  genau dann, wenn  $\hat{v} = 0$ .
2. Die Homogenität folgt aus  $\|\lambda \hat{v}\| = \|\lambda \hat{v}b_{\hat{T}}\|_{L^2(\hat{T})} = |\lambda| \|\hat{v}b_{\hat{T}}\|_{L^2(\hat{T})} = |\lambda| \|\hat{v}\|$  für alle  $\lambda \in \mathbb{R}$  und für alle  $\hat{v} \in \mathcal{P}_p(\hat{T})$ .
3. Wegen

$$\begin{aligned} \|\hat{v}_1 + \hat{v}_2\| &= \|(\hat{v}_1 + \hat{v}_2)b_{\hat{T}}\|_{L^2(\hat{T})} = \|\hat{v}_1b_{\hat{T}} + \hat{v}_2b_{\hat{T}}\|_{L^2(\hat{T})} \\ &\leq \|\hat{v}_1b_{\hat{T}}\|_{L^2(\hat{T})} + \|\hat{v}_2b_{\hat{T}}\|_{L^2(\hat{T})} = \|\hat{v}_1\| + \|\hat{v}_2\| \end{aligned}$$

für alle  $\hat{v}_1, \hat{v}_2 \in \mathcal{P}_p(\hat{T})$  gilt auch die Dreiecksungleichung.

Aufgrund der Äquivalenz der Normen auf endlich-dimensionalen Vektorräumen gilt

$$\|\hat{v}\| = \|\hat{v}b_{\hat{T}}\|_{L^2(\hat{T})} \geq \hat{C} \|\hat{v}\|_{L^2(\hat{T})}, \quad (3.44)$$

wobei  $\hat{C} > 0$  nur von  $p$  abhängt.

Wegen des Transformationssatzes und da  $\det(D\phi)$  bei affinen Abbildungen  $\phi$  konstant ist, gilt

$$\begin{aligned} \|vb_T\|_{L^2(T)} &= \left( \int_T (vb_T)^2 dx \right)^{1/2} = \left( \int_{\hat{T}} ((vb_T) \circ \phi)^2 |\det(D\phi)| dy \right)^{1/2} \\ &= |\det(D\phi)|^{1/2} \left( \int_{\hat{T}} (\hat{v}b_{\hat{T}})^2 dy \right)^{1/2} = |\det(D\phi)|^{1/2} \|\hat{v}b_{\hat{T}}\|_{L^2(\hat{T})} \end{aligned}$$

und

$$\|v\|_{L^2(T)} = |\det(D\phi)|^{1/2} \|\hat{v}\|_{L^2(\hat{T})}.$$

Mit (3.44) folgt die Behauptung (3.41)a.

Die Abschätzung (3.41)b folgt direkt wegen  $0 \leq b_T \leq 1$ . Somit ist (3.41) bewiesen.

(3.42)a beweisen wir zuerst auf der Referenzkante  $\hat{E} = [0, 1]$ . Es seien  $a_S$  und  $a_N$  der Anfangs- und Endknoten der Kante  $E$ . Dann ist  $[a_S, a_N] \subseteq \mathbb{R}^2$  und

$$\Upsilon : [0, 1] \rightarrow [a_S, a_N] : t \mapsto a_S + (a_N - a_S)t$$

bildet  $\hat{E}$  auf  $E$  ab. Somit gilt  $\hat{w} := w \circ \Upsilon \in \mathcal{P}_p(\hat{E})$ ,  $b_{\hat{E}} := b_E \circ \Upsilon$  und  $\|\hat{w}\| := \|\hat{w}b_{\hat{E}}\|_{L^2(\hat{E})}$  definiert eine Norm auf  $\mathcal{P}_p(\hat{E})$ , da  $b_{\hat{E}}$  im Inneren der Kante nicht verschwindet. Somit gilt aufgrund der Äquivalenz der Normen auf endlich-dimensionalen Vektorräumen

$$\|\hat{w}\| = \|\hat{w}b_{\hat{E}}\|_{L^2(\hat{E})} \geq \hat{C} \|\hat{w}\|_{L^2(\hat{E})}, \quad (3.45)$$

wobei  $\hat{C} > 0$  nur von  $p$  abhängt. Nach dem Transformationssatz in 1D gilt

$$\begin{aligned} \|wb_E\|_{L^2(E)} &= \left( \int_E (wb_E)^2 ds \right)^{1/2} = \left( \int_{a_S}^{a_N} (wb_E)^2 ds \right)^{1/2} \\ &= \left( \int_0^1 ((wb_E) \circ \Upsilon(t))^2 \|\Upsilon'(t)\|_2 dt \right)^{1/2} \\ &= \|a_N - a_S\|_2^{1/2} \left( \int_0^1 (\hat{w}b_{\hat{E}})^2 dt \right)^{1/2} = h_E^{1/2} \|\hat{w}b_{\hat{E}}\|_{L^2(\hat{E})} \end{aligned}$$

und ebenso

$$\|w\|_{L^2(E)} = h_E^{1/2} \|\hat{w}\|_{L^2(\hat{E})}. \quad (3.46)$$

Mit (3.45) folgt die Behauptung (3.42)a.

(3.42)b folgt wegen  $0 \leq b_E \leq 1$  ebenfalls direkt. Somit ist (3.42) bewiesen.

Für (3.43) definiert  $\|\hat{w}\| := \|b_{\hat{E}} \hat{F}_{ext}(\hat{w})\|_{L^2(\hat{T})}$  eine Norm auf  $\mathcal{P}_p(\hat{E})$ :

1. Da  $b_{\hat{E}}$  im Inneren von  $\hat{T}$  nicht verschwindet, gilt für alle  $\hat{w} \in \mathcal{P}_p(\hat{E})$   $\hat{F}_{ext}(\hat{w}) = 0$  auf  $\hat{T}$  nur dann, wenn  $\hat{w} = 0$  auf  $\hat{E}$  ist. Somit folgt die Definitheit.
2. Homogenität ist klar, da  $\hat{F}_{ext}(\lambda \hat{w}) = \lambda \hat{F}_{ext}(\hat{w})$  für  $\lambda \in \mathbb{R}$  und  $\hat{w} \in \mathcal{P}_p(\hat{E})$  gilt.
3. Wegen  $\hat{F}_{ext}(\hat{w}_1 + \hat{w}_2) = \hat{F}_{ext}(\hat{w}_1) + \hat{F}_{ext}(\hat{w}_2)$  für  $\hat{w}_1, \hat{w}_2 \in \mathcal{P}_p(\hat{E})$  folgt auch die Dreiecksungleichung.

Es gilt somit wegen Äquivalenz der Normen

$$\hat{C}_1 \|\hat{w}\|_{L^2(\hat{E})} \leq \|b_{\hat{E}} \hat{F}_{ext}(\hat{w})\|_{L^2(\hat{T})} \leq \hat{C}_2 \|\hat{w}\|_{L^2(\hat{E})}, \quad (3.47)$$

wobei  $\hat{C}_1 > 0$  und  $\hat{C}_2 > 0$  nur von  $p$  abhängen. Wie oben erhalten wir mit  $\phi : \hat{T} \rightarrow T$ , dem Transformationssatz und  $|\det(D\phi)| \geq (1/\sqrt{2})h_T \varrho_T = (1/\sqrt{2})(1/\kappa_T)h_T^2 \geq C'_1 h_E^2$  bzw.  $|\det(D\phi)| \leq h_T \varrho_T = (1/\kappa_T)h_T^2 \leq C'_2 h_E^2$  ( $C'_1$  und  $C'_2$  hängen nur von der Form von  $T$  ab,  $\kappa_T = h_T/\varrho_T$ )

$$\|b_E F_{ext}(w)\|_{L^2(T)} = |\det(D\phi)|^{1/2} \|b_{\hat{E}} \hat{F}_{ext}(\hat{w})\|_{L^2(\hat{T})} \geq C_1^* h_E \|b_{\hat{E}} \hat{F}_{ext}(\hat{w})\|_{L^2(\hat{T})}$$

bzw.

$$\|b_E F_{ext}(w)\|_{L^2(T)} = |\det(D\phi)|^{1/2} \|b_{\hat{E}} \hat{F}_{ext}(\hat{w})\|_{L^2(\hat{T})} \leq C_2^* h_E \|b_{\hat{E}} \hat{F}_{ext}(\hat{w})\|_{L^2(\hat{T})}$$

und nach (3.46)

$$\|w\|_{L^2(E)} = h_E^{1/2} \|\hat{w}\|_{L^2(\hat{E})}.$$

Daraus folgt mit (3.47)  $C_3 h_E^{1/2} \|w\|_{L^2(E)} \leq \|b_E F_{ext}(w)\|_{L^2(T)} \leq C_4 h_E^{1/2} \|w\|_{L^2(E)}$ , wobei  $C_3$  und  $C_4$  nur von  $p$  und der Form von  $T$ , nicht aber von  $h_T$  abhängen!  $\blacksquare$

Den folgenden Satz werden wir als Hilfsmittel für den Beweis weiterer wichtiger Abschätzungen benötigen.

**Satz 3.4.14 (Transformationsformel).** *Eine affine Abbildung  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$  bilde das beschränkte Lipschitz-Gebiet  $\Omega_1$  bijektiv auf das beschränkte Lipschitz-Gebiet  $\Omega_2$  ab. Es seien  $M \in \mathbb{R}^{d \times d}$  und  $y \in \mathbb{R}^d$  mit  $\phi(x) = Mx + y$  (insbesondere sei  $M$  regulär). Dann gilt für alle  $m \in \mathbb{N}_0$  und  $v \in H^m(\Omega_2)$*

$$\|D^m(v \circ \phi)\|_{L^2(\Omega_1)} \leq |\det(M)|^{-1/2} \|M\|_F^m \|D^m v\|_{L^2(\Omega_2)}. \quad (3.48)$$

Dabei bezeichnet  $\|\cdot\|_F$  die Frobenius-Norm, d.h. die  $\ell_2$ -Norm auf  $\mathbb{R}^{d \times d}$ ,

$$\|A\|_F := \left( \sum_{j,k=1}^d a_{jk}^2 \right)^{1/2} \quad A \in \mathbb{R}^{d \times d}.$$

**Beweis.** Siehe Carstensen-Vorlesungsskriptum [6, Seite 206] oder Braess [3, Satz 6.6]. ■

Bei Dreiecken gilt mit den Bezeichnungen von (3.37)

$$\begin{aligned} \|M_\Delta\|_F &= \left( (x_2 - x_1)^2 + (y_2 - y_1)^2 + (x_3 - x_1)^2 + (y_3 - y_1)^2 \right)^{1/2} \\ &= \left( \underbrace{|P_2 - P_1|^2}_{\leq h_T^2} + \underbrace{|P_3 - P_1|^2}_{\leq h_T^2} \right)^{1/2} \leq \sqrt{2}h_T. \end{aligned}$$

Für Rechtecke erhalten wir mit den Bezeichnungen aus (3.38)

$$\begin{aligned} \|M_\square\|_F &= \left( (x_2 - x_1)^2 + (y_2 - y_1)^2 + (x_4 - x_1)^2 + (y_4 - y_1)^2 \right)^{1/2} \\ &= \left( \underbrace{|P_2 - P_1|^2}_{\leq e_1^2} + \underbrace{|P_4 - P_1|^2}_{\leq e_2^2} \right)^{1/2} = h_T. \end{aligned}$$

Wir schreiben wieder allgemein mit der Bezeichnung von (3.39)

$$\|M\|_F \leq \sqrt{2}h_T. \quad (3.49)$$

Für die inverse Matrix  $M^{-1}$  von  $M$  gilt

$$M^{-1} = \frac{1}{\det(M)} \begin{pmatrix} m_4 & -m_2 \\ -m_3 & m_1 \end{pmatrix} \quad \text{mit} \quad M = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix}.$$

Daher ist wegen  $\det(M) \geq (1/\sqrt{2})h_T\varrho_T$  und (3.49)

$$\|M^{-1}\|_F = \frac{1}{\det(M)} \|M\|_F \leq \frac{1}{(1/\sqrt{2})h_T\varrho_T} \sqrt{2}h_T = \frac{2}{\varrho_T}. \quad (3.50)$$

**Lemma 3.4.15 (Inverse Abschätzungen).** *Es sei  $p \in \mathbb{N}_0$ ,  $E \in (\mathcal{E}^* \cup \mathcal{E}_\Gamma)$ ,  $T \in \omega_E^*$ ,  $v \in \mathcal{P}_p(T)$ ,  $w \in \mathcal{P}_p(E)$ ,  $b_T$  und  $b_E$  die Bubble-Funktionen auf  $T$  bzw.  $E$ , dann gilt:*

$$\|\nabla(vb_T)\|_{L^2(T)} \leq C_1 h_T^{-1} \|v\|_{L^2(T)} \quad (3.51)$$

$$\|\nabla(F_{ext}(w)b_E)\|_{L^2(T)} \leq C_2 h_E^{-1/2} \|w\|_{L^2(E)} \quad (3.52)$$

Die Konstanten  $C_1 > 0$  und  $C_2 > 0$  hängen nur von  $p$  und der Form von  $T$  ab, nicht aber von  $h_T$ .

**Beweis.** Für (3.51) betrachten wir wieder die affine Abbildung  $\phi : \hat{T} \rightarrow T$  von (3.39),  $\hat{v} := v \circ \phi \in \mathcal{P}_p(\hat{T})$  und  $b_{\hat{T}} := b_T \circ \phi$ . Wir beweisen die Abschätzung zunächst auf dem Referenzelement  $\hat{T}$ .  $\mathcal{P}_p(\hat{T})$  ist ein endlich dimensionaler Raum,  $|\cdot|$  sei eine Halbnorm auf  $\mathcal{P}_p(\hat{T})$ ,  $\|\cdot\|$  eine Norm auf  $\mathcal{P}_p(\hat{T})$ ,  $Y := \{y \in \mathcal{P}_p(\hat{T}) \mid |y| = 0\}$  und  $X := \mathcal{P}_p(\hat{T})$ . Dann gilt:

$$|x + Y| := |x| \quad \text{definiert eine Norm auf dem Faktorraum } X/Y,$$

$$\|x + Y\| := \inf_{y \in Y} \|x + y\| \quad \text{definiert eine Norm auf dem Faktorraum } X/Y.$$

Da  $\dim(X/Y) < \infty$  ist, und alle Normen auf  $X/Y$  äquivalent sind, existiert eine Konstante  $C > 0$ , die nur von  $p$  abhängt mit  $|\cdot + Y| \leq C\|\cdot + Y\|$ . Es gilt daher

$$|x| = |x + Y| \leq C\|x + Y\| \leq C\|x\| \quad \forall x \in \mathcal{P}_p(\hat{T}).$$

$|\hat{v}| := \|\nabla(\hat{v}b_{\hat{T}})\|_{L^2(\hat{T})}$  definiert offensichtlich eine Halbnorm auf  $\mathcal{P}_p(\hat{T})$ . Wir erhalten die Abschätzung auf dem Referenzelement  $\hat{T}$

$$|\hat{v}| = \|\nabla(\hat{v}b_{\hat{T}})\|_{L^2(\hat{T})} \leq \hat{C}\|\hat{v}\|_{L^2(\hat{T})}, \quad (3.53)$$

wobei  $\hat{C} > 0$  nur von  $p$  und der Form von  $\hat{T}$  abhängt. Mit der Transformationsformel (3.48) gilt

$$\|\nabla(vb_T)\|_{L^2(T)} = \|\nabla((\hat{v}b_{\hat{T}}) \circ \phi^{-1})\|_{L^2(T)} \leq \underbrace{|\det(M^{-1})|^{-1/2}}_{|\det(M)|^{1/2}} \|M^{-1}\|_F \|\nabla(\hat{v}b_{\hat{T}})\|_{L^2(\hat{T})}$$

und dem Transformationssatz

$$\|v\|_{L^2(T)} = \underbrace{|\det(D\phi)|^{1/2}}_{|\det(M)|^{1/2}} \|\hat{v}\|_{L^2(\hat{T})}.$$

Wegen (3.53) und (3.50) gilt

$$\|\nabla(vb_T)\|_{L^2(T)} \leq \hat{C}\|M^{-1}\|_F \|v\|_{L^2(T)} \leq \hat{C} \underbrace{\frac{2}{\varrho_T}}_{=(2\kappa_T)/h_T} \|v\|_{L^2(T)} = C_1 h_T^{-1} \|v\|_{L^2(T)},$$

wobei  $C_1$  nur von  $p$  und der Form von  $T$  abhängt.

Für (3.52) definiert  $|\hat{w}| := \|\nabla(\hat{F}_{ext}(\hat{w})b_{\hat{E}})\|_{L^2(\hat{T})}$  offensichtlich eine Halbnorm auf  $\mathcal{P}_p(\hat{E})$ . Mit dem abstrakten Argument von oben folgt

$$\|\hat{w}\| = \|\nabla(\hat{F}_{ext}(\hat{w})b_{\hat{E}})\|_{L^2(\hat{T})} \leq \hat{C}\|\hat{w}\|_{L^2(\hat{E})}, \quad (3.54)$$

wobei  $\hat{C} > 0$  nur von  $p$  abhängt. Mit der Transformationsformel (3.48) erhalten wir

$$\begin{aligned} \|\nabla(F_{ext}(w)b_E)\|_{L^2(T)} &= \|\nabla((\hat{F}_{ext}(\hat{w})b_{\hat{E}}) \circ \phi^{-1})\|_{L^2(T)} \\ &\leq \underbrace{|\det(M^{-1})|^{-1/2}}_{|\det(M)|^{1/2}} \|M^{-1}\|_F \|\nabla(\hat{F}_{ext}(\hat{w})b_{\hat{E}})\|_{L^2(\hat{T})} \end{aligned}$$

und dem Transformationssatz (vergleiche (3.46))

$$\|w\|_{L^2(E)} = h_E^{1/2} \|\hat{w}\|_{L^2(\hat{E})}.$$

Wegen (3.40) und (3.50) erhalten wir für

$$|\det(M)|^{1/2} \|M^{-1}\|_F \leq h_T^{1/2} \varrho_T^{1/2} \frac{2}{\varrho_T} = 2h_T^{1/2} \frac{1}{\varrho_T^{1/2}} = 2h_T^{1/2} \frac{\kappa_T^{1/2}}{h_T^{1/2}} = C',$$

wobei  $C'$  nur von der Form von  $T$  abhängt. Schlussendlich erhalten wir mit (3.54)

$$\|\nabla(F_{ext}(w)b_E)\|_{L^2(T)} \leq C' \|\nabla(\hat{F}_{ext}(\hat{w})b_{\hat{E}})\|_{L^2(\hat{T})} \leq C_2 \|\hat{w}\|_{L^2(\hat{E})} = C_2 h_E^{-1/2} \|w\|_{L^2(E)},$$

wobei  $C_2$  nur von  $p$  und der Form von  $T$  abhängt. ■

**Effizienzbeweis**

Wir können nun die Effizienz unseres Residualschätzers beweisen. In (3.25) gilt auch eine entsprechende Abschätzung nach oben, die sogar lokal ist.

**Satz 3.4.16 (Effizienz).** *Der Fehler  $e := u - I_M u_h \in H^1(\mathcal{T})$  erfüllt für alle  $T \in \mathcal{T}$*

$$\eta_T^2 \leq C(\|\nabla_T e\|_{L^2(\omega_T)}^2 + h_T^2 \|f - f_T\|_{L^2(\omega_T)}^2), \quad (3.55)$$

wobei die Konstante  $C > 0$  nur von der Form der Elemente abhängt und  $\omega_T$  nach Abschnitt 1.1.2 der Elementpatch ist.

**Beweis.** Der Beweis gliedert sich in 4 Schritte. Zuerst schätzen wir das Residuum, dann den Jump in Normalrichtung, als dritten Schritt den Jump in Tangentialrichtung und zum Schluss den Jump in Tangentialrichtung für Randkanten ab.

**Schritt 1.** Es sei  $v_T = R_T b_T \in H_0^1(T)$ , wobei  $b_T$  die oben definierte Bubble-Funktion auf der Fläche  $T$  ist. Wir schreiben deshalb

$$\begin{aligned} \int_T R_T v_T dx &= \int_T (R_T - R) v_T dx + \int_T R v_T dx \\ &= \int_T (R_T - R) v_T dx - \int_T \underbrace{\Delta(u - I_M u_h)}_{=e} v_T dx. \end{aligned}$$

Wenn wir nun beim Integral auf der linken Seite  $v_T$  einsetzen und das zweite Integral auf der rechten Seite partiell integrieren (ohne Randterme, da  $v_T = 0$  auf  $\partial T$ ), erhalten wir

$$\int_T R_T^2 b_T dx = \int_T (R_T - R) v_T dx + \int_T \nabla e \cdot \nabla v_T dx. \quad (3.56)$$

Wir betrachten die Integrale getrennt. Für das Erste ergibt sich, nach Anwendung von (3.41)a,

$$\int_T R_T^2 b_T dx = \|R_T b_T^{1/2}\|_{L^2(T)}^2 \geq C_1 \|R_T\|_{L^2(T)}^2$$

mit  $C_1 > 0$ . Wegen der Cauchy-Schwarz-Ungleichung und (3.41)b bekommen wir

$$\int_T (R_T - R) v_T dx \leq \|R - R_T\|_{L^2(T)} \|R_T b_T\|_{L^2(T)} \leq \|R - R_T\|_{L^2(T)} \|R_T\|_{L^2(T)},$$

und schlussendlich mit der Cauchy-Schwarz-Ungleichung und (3.51)

$$\int_T \nabla e \cdot \nabla v_T dx \leq \|\nabla e\|_{L^2(T)} \|\nabla(R_T b_T)\|_{L^2(T)} \leq C_2 \|\nabla e\|_{L^2(T)} h_T^{-1} \|R_T\|_{L^2(T)}.$$

$C_2$  hängt nur von der Form von  $T$  ab. Mit diesen Abschätzungen erhalten wir für (3.56)

$$h_T \|R_T\|_{L^2(T)} \leq C_3 (\|\nabla e\|_{L^2(T)} + h_T \|R - R_T\|_{L^2(T)}), \quad (3.57)$$

wobei  $C_3$  nur von der Form von  $T$  abhängt. Wegen (3.16) und  $a + b \leq \sqrt{2}(a^2 + b^2)^{1/2}$  gilt

$$h_T^2 \|R_T\|_{L^2(T)}^2 = h_T^2 (\|R\|_{L^2(T)}^2 - \|R - R_T\|_{L^2(T)}^2) \leq C_4 (\|\nabla e\|_{L^2(T)}^2 + h_T^2 \|R - R_T\|_{L^2(T)}^2)$$

und daher

$$h_T^2 \|R\|_{L^2(T)}^2 \leq C_5 (\|\nabla e\|_{L^2(T)}^2 + h_T^2 \|R - R_T\|_{L^2(T)}^2) \quad \text{für alle } T \in \mathcal{T}, \quad (3.58)$$

wobei  $C_5$  nur von der Form der Elemente  $T$  abhängt.

**Schritt 2.** Es sei nun  $E \in \mathcal{E}^*$  fixiert und

$$v_E := F_{ext} \left( \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right) b_E \in H_0^1(\omega_E^*).$$

Da  $u \in H^1(\Omega)$  ist, ergibt sich

$$\left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E = \left[ \frac{\partial}{\partial \mathbf{n}_E} (I_M u_h - u) \right]_E = \left[ -\frac{\partial e}{\partial \mathbf{n}_E} \right]_E \quad \text{wegen} \quad \left[ \frac{\partial u}{\partial \mathbf{n}_E} \right]_E = 0.$$

Da  $v_E = 0$  auf  $\partial\omega_E^*$  schreiben wir für den Jump

$$\int_E \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E v_E dx = - \int_E \left[ \frac{\partial e}{\partial \mathbf{n}_E} \right]_E v_E dx = - \sum_{T \in \tilde{\omega}_E^*} \int_{\partial T} \frac{\partial e}{\partial \mathbf{n}_T} v_E dx.$$

Partielle Integration mit Randtermen ergibt

$$= - \sum_{T \in \tilde{\omega}_E^*} \left( \int_T \nabla e \cdot \nabla v_E dx + \int_T \underbrace{(-f - \Delta I_M u_h)}_{=-R} v_E dx \right)$$

mit  $\Delta e = -f - \Delta I_M u_h$ . Durch Anwendung der Cauchy-Schwarz-Ungleichung bekommen wir

$$\leq \sum_{T \in \tilde{\omega}_E^*} (\|\nabla e\|_{L^2(T)} \|\nabla v_E\|_{L^2(T)} + \|R\|_{L^2(T)} \|v_E\|_{L^2(T)}).$$

Mit (3.52) und (3.43)b erhalten wir

$$\leq C_6 \sum_{T \in \tilde{\omega}_E^*} \left( \|\nabla e\|_{L^2(T)} h_E^{-1/2} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)} + \|R\|_{L^2(T)} h_E^{1/2} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)} \right)$$

und somit

$$\int_E \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E v_E dx \leq C_6 h_E^{-1/2} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)} \sum_{T \in \tilde{\omega}_E^*} (\|\nabla e\|_{L^2(T)} + h_E \|R\|_{L^2(T)}). \quad (3.59)$$

$C_6$  hängt nur von der Form der Elemente  $T \in \tilde{\omega}_E^*$  ab. Für das linke Integral von (3.59) (wir befinden uns auf  $E$ !) schreiben wir mit Hilfe von (3.42)a

$$\begin{aligned} \int_E \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E v_E dx &= \int_E \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E^2 b_E dx = \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E b_E^{1/2} \right\|_{L^2(E)}^2 \\ &\geq C_7 \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)}^2. \end{aligned}$$

Außerdem ist

$$\sum_{T \in \tilde{\omega}_E^*} \|\nabla e\|_{L^2(T)} \leq \sqrt{2} \|\nabla e\|_{L^2(\omega_E^*)} \leq \sqrt{2} \|\nabla e\|_{L^2(\omega_E)},$$

und wegen  $\|R\|_{L^2(T)} \leq \|R_T\|_{L^2(T)} + \|R - R_T\|_{L^2(T)}$  und Anwendung von (3.57) gilt

$$\begin{aligned} \sum_{T \in \tilde{\omega}_E^*} h_E \|R\|_{L^2(T)} &\leq \sum_{T \in \tilde{\omega}_E} h_E \|R\|_{L^2(T)} \leq \sum_{T \in \tilde{\omega}_E} h_E \|R_T\|_{L^2(T)} + \sum_{T \in \tilde{\omega}_E} h_E \|R - R_T\|_{L^2(T)} \\ &\leq C_8 \left( \sum_{T \in \tilde{\omega}_E} (\|\nabla e\|_{L^2(T)} + h_E \|R - R_T\|_{L^2(T)}) \right) \\ &\quad + \sum_{T \in \tilde{\omega}_E} h_E \|R - R_T\|_{L^2(T)} \\ &\leq C_9 \left( \|\nabla e\|_{L^2(\omega_E)} + \sum_{T \in \tilde{\omega}_E} h_E \|R - R_T\|_{L^2(T)} \right). \end{aligned}$$

Daher ergibt sich durch Einsetzen in (3.59)

$$h_E^{1/2} \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)} \leq C_{10} \left( \|\nabla e\|_{L^2(\omega_E)} + h_E \sum_{T \in \tilde{\omega}_E} \|R - R_T\|_{L^2(T)} \right).$$

Somit erhalten wir für alle  $E \in \mathcal{E}^*$

$$\begin{aligned} h_E \left\| \left[ \frac{\partial I_M u_h}{\partial \mathbf{n}_E} \right]_E \right\|_{L^2(E)}^2 &\leq C_{11} \left( \|\nabla e\|_{L^2(\omega_E)}^2 + \left( h_E \sum_{T \in \tilde{\omega}_E} \|R - R_T\|_{L^2(T)} \right)^2 \right) \\ &\leq C_{12} \left( \|\nabla e\|_{L^2(\omega_E)}^2 + \underbrace{h_E^2 \sum_{T \in \tilde{\omega}_E} \|R - R_T\|_{L^2(T)}^2}_{=\|R - R_T\|_{L^2(\omega_E)}^2} \right), \end{aligned} \quad (3.60)$$

wobei  $C_{12}$  nur von der Form der Elemente abhängt.

**Schritt 3.** Für die Abschätzung des Tangentialsprungs im Inneren fixieren wir  $E \in \mathcal{E}^*$  und definieren

$$v_E := F_{ext} \left( \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E \right) b_E \in H_0^1(\omega_E^*).$$

Für  $u \in H^1(\Omega)$  und  $v_E \in H_0^1(\omega_E^*)$  gilt

$$\int_{\omega_E^*} \nabla u \cdot \operatorname{curl} v_E \, dx = 0. \quad (3.61)$$

(3.61) kann analog wie auf Seite 31 mit einem Dichtigkeitsargument gezeigt werden. Anstatt  $\Omega$  betrachten wir  $\omega_E^*$ , und die Randterme verlieren wir diesmal wegen  $v_E \in H_0^1(\omega_E^*)$ . Für den Tangentialsprung gilt für alle  $E \in \mathcal{E}^*$  mit partieller Integration mit Randtermen (wie oben)

$$\int_E \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right]_E v_E \, dx = \sum_{T \in \tilde{\omega}_E^*} \int_{\partial T} \frac{\partial I_M u_h}{\partial \mathbf{t}_T} v_E \, dx = \sum_{T \in \tilde{\omega}_E^*} \int_T \nabla I_M u_h \cdot \operatorname{curl} v_E \, dx.$$

Mit (3.61), der Cauchy-Schwarz-Ungleichung und  $\|\operatorname{curl} v_E\|_{L^2(T)} = \|\nabla v_E\|_{L^2(T)}$  erhalten wir

$$= - \sum_{T \in \tilde{\omega}_E^*} \int_T \underbrace{\nabla (u - I_M u_h)}_{=e} \cdot \operatorname{curl} v_E \, dx \leq \sum_{T \in \tilde{\omega}_E^*} \|\nabla e\|_{L^2(T)} \|\nabla v_E\|_{L^2(T)}.$$

Auf  $E$  gilt mit (3.42)a

$$\begin{aligned} \int_E \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right] \right]_E v_E dx &= \int_E \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right] \right]_E^2 b_E dx = \left\| \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right] \right]_E b_E^{1/2} \right\|_{L^2(T)}^2 \\ &\geq C_{13} \left\| \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right] \right]_E \right\|_{L^2(E)}^2. \end{aligned}$$

Mit (3.52) erhalten wir deshalb

$$\begin{aligned} C_{13} \left\| \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right] \right]_E \right\|_{L^2(E)}^2 &\leq \sum_{T \in \tilde{\omega}_E^*} \|\nabla e\|_{L^2(T)} \|\nabla v_E\|_{L^2(T)} \\ &\leq \sum_{T \in \tilde{\omega}_E^*} \|\nabla e\|_{L^2(T)} C_{14} h_E^{-1/2} \left\| \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right] \right]_E \right\|_{L^2(E)} \end{aligned}$$

Die Konstanten  $C_{13}$  und  $C_{14}$  hängen nur von der Form der Elemente ab. Somit ist

$$h_E^{1/2} \left\| \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right] \right]_E \right\|_{L^2(E)} \leq C_{15} \sum_{T \in \tilde{\omega}_E^*} \|\nabla e\|_{L^2(T)} \leq C_{16} \|\nabla e\|_{L^2(\omega_E^*)} \leq C_{16} \|\nabla e\|_{L^2(\omega_E)}.$$

Schlussendlich gilt

$$h_E \left\| \left[ \left[ \frac{\partial I_M u_h}{\partial \mathbf{t}_E} \right] \right]_E \right\|_{L^2(E)}^2 \leq C_{17} \|\nabla e\|_{L^2(\omega_E)}^2 \quad \text{für alle } E \in \mathcal{E}^*, \quad (3.62)$$

wobei  $C_{17} = C_{16}^2$  nur von der Form der Elemente abhängt.

**Schritt 4.** Für die Abschätzung des Tangentialsprunges am Rand fixieren wir  $E \in \mathcal{E}_T$  und definieren

$$v_E := F_{ext} \left( \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right) b_E \in H_0^1(\omega_E),$$

wobei  $b_E$  die Bubble-Funktion für die Randkante  $E$  ist,  $\omega_E = T$  mit  $E \subseteq \partial T$  und  $b_E = 0$  auf  $\partial T \setminus E$ . Daher ist  $v_E = 0$  auf  $\partial T \setminus E$ , und es gilt

$$\int_E \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} v_E dx = \int_{\partial T} \overbrace{-\frac{\partial(u - I_M u_h)}{\partial \mathbf{t}_T}}^{=e} v_E dx.$$

Wie auf Seite 34 gezeigt wurde, ist

$$\int_{\partial T} \frac{\partial e}{\partial \mathbf{t}_T} v_E dx = \int_T \nabla e \cdot \text{curl } v_E dx.$$

Cauchy-Schwarz-Ungleichung und (3.52) ergibt

$$\leq \|\nabla e\|_{L^2(T)} \|\nabla v_E\|_{L^2(T)} \leq C_{18} \|\nabla e\|_{L^2(T)} h_E^{-1/2} \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}.$$

Auf  $E$  gilt mit der Abschätzung (3.42)a

$$\begin{aligned} \int_E \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} v_E dx &= \int_E \left( \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right)^2 b_E dx \\ &= \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} b_E^{1/2} \right\|_{L^2(E)}^2 \geq C_{19} \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}^2. \end{aligned}$$

Weshalb die folgende Abschätzung für den Tangentialsprung auf Randkanten gilt

$$h_E \left\| \frac{\partial(I_M u_h - u_D)}{\partial \mathbf{t}_E} \right\|_{L^2(E)}^2 \leq C_{20} \|\nabla e\|_{L^2(T)}^2 \quad \text{für alle } E \in \mathcal{E}_\Gamma, \quad (3.63)$$

wobei  $C_{20}$  nur von der Form der Elemente abhängt.

Die Behauptung folgt nun direkt durch Hochsummieren der Abschätzungen (3.58), (3.60), (3.62) und (3.63). Zu beachten gilt, dass eine hängende Kante die Vereinigung von Elementarkanten ist und  $h_E \geq C_{21} h_T$  ist, wobei  $C_{21}$  nur von der Form von  $T$  abhängt. Insgesamt bekommen wir durch Zusammenfassen aller Konstanten eine Konstante  $C$ , die nur von der Form der Elemente  $T$  abhängt. ■

## Kapitel 4

# Numerische Experimente

In diesem Kapitel werden die Zuverlässigkeit und Effizienz des Fehlerschätzers anhand von Beispielen mit bekannter exakter Lösung numerisch verifiziert. Wir werden dabei eine adaptive Netzverfeinerungsstrategie verfolgen und die Vor- und Nachteile gegenüber einer uniformen Verfeinerung sehen.

**Festlegung der Triangulierung.** Für die folgenden Beispiele wird eine fastreguläre Triangulierung in lauter Quadrate verwendet, die höchstens einen hängenden Knoten pro Elementkante hat. Die Triangulierung hat somit Ordnung 2.

Als *Vaterelement*  $T$  des hängenden Knotens  $a$  bezeichnen wir jenes Element, wenn zwar  $a \subseteq T$  gilt,  $a$  aber nicht Ecke von  $T$  ist. Nach Abbildung 2.5 gibt es für einen Knoten  $a \in \mathcal{N}^* \cup \mathcal{N}^h$  bis auf Drehungen genau 7 verschiedene Kombinationen der Interpolationsgewichte  $\psi_T(a)$  für einen Knotenpatch  $T \in \tilde{\omega}_a$ . Für  $x_T$  wird der Schwerpunkt verwendet, womit bei dieser Triangulierung klar ist, dass der Diamond Path *immer* konvex sein muss. Die Rekursion, die zur Bestimmung des Morley-Interpolanten nötig ist (Lemma 3.3.1), bricht nach endlich vielen Schritten ab: Die Abbruchbedingung der Rekursion ist, dass ein Element erreicht wird, dessen Ecken nicht hängend sind. Nach Konstruktion unserer Triangulierung muss es dieses Element immer geben, da die Rekursion immer das Vaterelement  $T_a$  des hängenden Knotens  $a$  ( $a$  sei Ecke von  $T$ ) aufruft.  $T_a$  ist immer größer als  $T$ , deshalb kann ein Element in der Rekursion nie doppelt aufgerufen werden!

Alle numerischen Experimente wurden in MATLAB 6.5 berechnet.

### 4.1 Algorithmen und Datenstrukturen

In diesem Abschnitt werden die Algorithmen zur Lösung des Laplace-Problems mit adaptiver Netzverfeinerung näher erläutert. Gleich im ersten Unterabschnitt werden die verwendeten Datenstrukturen kurz erklärt. Im Zweiten erklären wir die Idee eines adaptiven Algorithmus. Im dritten Unterabschnitt erfolgt die Diskussion einer adaptiven Netzverfeinerung, wobei die oben festgelegte Triangulierung erhalten bleibt. Der Vierte zeigt die Berechnung der Finite Volumen Lösung und der fünfte Abschnitt die des Morley-Interpolanten. Die Algorithmen werden jeweils mittels Pseudocodes erklärt, vereinzelt auch mit MATLAB-Code.

#### 4.1.1 Datenstrukturen

Für die Abspeicherung und Verfeinerung des Netzes und für die Berechnung des Morley-Interpolanten werden Datenstrukturen benötigt. In diesem Abschnitt wird erläutert, wie die

Datenstrukturen *MESH* für das Netz und *MORLEY* für den Morley-Interpolanten gebildet werden können. Es sei darauf hingewiesen, dass diese Möglichkeit eine von vielen ist. Alle Knoten einer Triangulierung werden durchnummeriert und in *MESH.coordinates* abgespeichert. Wir verwenden für die *Knoten* das Format

Knotennummer   x-Koordinate   y-Koordinate   Status.

In einem zweidimensionalen Feld kann der Eintrag Knotennummer entfallen, da die erste Zeile dem ersten Knoten usw. entspricht. Für den Statuseintrag gilt

$$\text{Status} := \begin{cases} -1 & \text{falls der Knoten nicht am Rand liegt und nicht hängend ist,} \\ 0 & \text{wenn der Knoten am Rand liegt,} \\ > 0 & \text{hängender Knoten, Status = Elementnummer des Vaterlements.} \end{cases}$$

Die *Elemente* werden in *MESH.elements* mit Hilfe des Formats

Elementnummer   Knoten 1   Knoten 2   Knoten 3   Knoten 4

gespeichert, wobei bei einem zweidimensionalen Feld der erste Eintrag wieder entfallen kann (erste Zeile erstes Element usw.). Die Knoten eines Elements seien gegen den Uhrzeigersinn angeordnet. Für die Verfeinerung der Triangulierung und für die Berechnung des Morley-Interpolanten wird immer wieder der Knotenpatch eines Knotens  $a$  gebraucht. Dieser wird in einem zweidimensionalen Feld *MESH.nodes2patch* mit vier Einträgen pro Zeile abgespeichert. Die Einträge sind dabei mathematisch positiv angeordnet, beginnend links unten. Mit Hilfe des Elementenfeldes *MESH.elements* lässt sich diese Information leicht gewinnen, wie folgender MATLAB-Code zeigt.

**Algorithmus 4.1.1 (Knotenpatch).** *Der Algorithmus erzeugt ein zweidimensionales Feld *MESH.nodes2patch* mit vier Einträgen, das den Knotenpatch eines Knotens enthält:*

```
MESH.nodes2patch = sparse(size(MESH.coordinates, 1), 4);
for  $i = 1 : \text{size}(\textit{MESH.elements}, 1)$ 
    MESH.nodes2patch(MESH.elements( $i, :$ ), [3 4 1 2]) =
        MESH.nodes2patch(MESH.elements( $i, :$ ), [3 4 1 2]) +  $i * \textit{eye}(4)$ ;
end
```

**Bemerkung.** Wenn  $a$  ein hängender Knoten ist, so steht das Vaterlement nach obigem Algorithmus nicht in der entsprechenden Zeile von *MESH.nodes2patch* (wir haben dann nur zwei Einträge  $\neq 0$ ). Das Vaterlement ist jedoch in *MESH.coordinates(:,3)* abgespeichert. Die Erzeugung dieses Feldes erfolgt in  $\mathcal{O}(N)$ , wobei  $N := \#\mathcal{T}$  die Anzahl der Elemente in  $\mathcal{T}$  ist. Wir benötigen außerdem nur die Informationen aus *MESH.elements* und *MESH.coordinates*.

Aus *MESH.nodes2patch* und mit *MESH.coordinates(:,3)* können dann für eine Kante  $E$  ( $a_S$  Anfangs- und  $a_N$  Endknoten) leicht die zwei benachbarten Elemente  $T_W$  und  $T_E$  mit  $T_W \cap T_E = E$  bestimmt werden. Weiters benötigen wir im Hinblick auf Netzverfeinerung die Information, ob eine Kante eines Elements hängend ist oder nicht. Wir erzeugen deshalb ein zweidimensionales Feld *MESH.elements2hanging* mit vier Einträgen. Der erste Eintrag enthält die Nummer des hängenden Knotens, falls die erste Kante hängend ist, oder 0, falls sie elementar ist usw.

**Algorithmus 4.1.2 (Hängende Kante).** *Der Algorithmus erzeugt ein zweidimensionales Feld *MESH.elements2hanging* mit vier Einträgen, wobei ein Eintrag die Information enthält, ob die Kante hängend ist (Wert=Nummer des hängenden Knotens) oder nicht (Wert=0). Die Funktion *edgeside* berechnet mit Hilfe von *MESH.nodes2patch* auf welcher Kante des Vaterlements der*

hängende Knoten liegt:

```

MESH.elements2hanging = sparse(size(MESH.elements, 1), 4);
for i = 1 : size(MESH.coordinates, 1)
    father = MESH.coordinates(i, 3);
    if father > 0
        edge = edgeside(MESH.nodes2patch(i, :));
        MESH.elements2hanging(father, edge) = i;
    end
end
end

```

**Bemerkung.** Die Erzeugung dieses Feldes erfolgt in  $\mathcal{O}(\#\mathcal{N})$ , wobei  $\#\mathcal{N}$  die Anzahl der Knoten in  $\mathcal{T}$  ist. Zur Berechnung des Feldes reichen die Informationen aus  $MESH.elements$ ,  $MESH.coordinates$  und  $MESH.nodes2patch$  aus. Die Funktion `edgeside` wird einfach durch Vergleichen der Einträge in  $MESH.nodes2patch(i, :)$  realisiert. Bei einem hängenden Knoten haben wir, wie oben erwähnt, nur zwei Einträge  $\neq 0$ . Die Position dieser zwei Einträge hängt von der Lage (also erste Kante usw.) des hängenden Knotens ab.

Ebenso können die a priori bestimmbaren Gewichte für einen Knoten  $a \in \mathcal{N}$  in einem Feld ( $MESH.weights$ ) gespeichert werden. Zur Erzeugung dieses Feldes reichen die Informationen aus  $MESH.elements$ ,  $MESH.coordinates$  und  $MESH.nodes2patch$ . Die Zuordnung der Gewichte (vergleiche Abbildung 2.5) erfolgt dann durch Vergleichen der Elementdurchmesser  $h_{T'}$  für Elemente  $T' \in \tilde{\omega}_T$  des Elementpatches.

Als Format für den Morley-Interpolanten  $I_M u_h|_T$  wählen wir

Knotennummer    Koeffizient 1    ...    Koeffizient  $n$ .

Die Datenstruktur *MORLEY* besteht eigentlich nur aus einem zweidimensionalen Feld. Die Zeilennummer entspricht der Knotennummer und die einzelnen Einträge der Zeile entsprechen den Polynomkoeffizienten.

#### 4.1.2 Adaptiver Algorithmus

Unsere a posteriori Fehlerabschätzung (3.17) lässt sich elementweise auswerten. Da sich unsere Fehlerabschätzung (Zuverlässigkeit und Effizienz) eben genau aus diesen *lokalen*  $\eta_T$  zusammensetzt, geben die Abschätzungen (3.25) und (3.55) einen Hinweis, dass  $\eta_T$  vielleicht auch die Größe des Fehlers in der Nähe von  $T$  zu *schätzen* vermag. Zwar haben wir keine strenge Rechtfertigung, aber doch betrachten wir ein relativ großes  $\eta_T$  als Indiz dafür, dass der lokale Anteil  $\|\nabla_{\mathcal{T}} e\|_{L^2(T)}$  am Fehler  $\|\nabla_{\mathcal{T}} e\|_{L^2(\Omega)}$  relativ groß ist. Es sei darauf hingewiesen, dass die Zuverlässigkeit des Fehlerschätzers nur global bewiesen wurde und nur die Effizienz auch lokal gilt.  $\|\nabla_{\mathcal{T}} e\|_{L^2(T)}$  ist nicht notwendig durch  $C\eta_T$  nach oben beschränkt und streng genommen auch nicht nach unten. Mit unserer üblichen Bezeichnungsweise ist die *lokale* Größe  $\eta_T$  ein **Verfeinerungsindikator** (oder **Fehlerindikator**), die *globale* Größe

$$\eta := \left( \sum_{T \in \mathcal{T}} \eta_T^2 \right)^{1/2} \quad (4.1)$$

ein zuverlässiger und effizienter **Fehlerschätzer**, wie in Satz 3.4.9 und Satz 3.4.16 bewiesen wurde. Das Verfeinerungskriterium benutzt die Heuristik, dass ein Element  $T$  verfeinert werden soll, wenn sein Beitrag  $\eta_T$  zu  $\eta$  groß ist.

**Bemerkung.** Die Fehlerabschätzung erfolgt immer noch durch  $\eta$ , somit bleibt die Abschätzung global. Zusammenfassend stellen wir fest, dass mit  $\eta_T$  das Gitter verfeinert wird und mit  $\eta$  die zuverlässige und effiziente Fehlerabschätzung erfolgt.

**Definition 4.1.3 (Verfeinerungsregel).** Für einen Parameter  $\theta, 0 \leq \theta \leq 1$ , verfeinern wir  $T \in \mathcal{T}$ , falls gilt

$$\eta_T \geq \theta \max_{T' \in \mathcal{T}} \eta_{T'}. \quad (4.2)$$

Mit der Anwendung dieser Verfeinerungsregel konstruieren wir adaptive Algorithmen zur automatischen Netzgenerierung. Folgendes Schema wird dabei verwendet, wobei die einzelnen Teilschritte in den weiteren Abschnitten beschrieben werden. Das Feld *markiert* gibt an, ob ein Element  $T \in \mathcal{T}$  markiert ist ( $\text{markiert}[T] = 1$ ) oder nicht ( $\text{markiert}[T] = 0$ ).

**Algorithmus 4.1.4 (Adaptive Netzgenerierung).** Gegeben sei eine fastreguläre Triangulierung  $\mathcal{T}_0$  mit Ordnung 2 in lauter Quadrate,  $k = 0$  und  $\text{markiert}[T] = 0$  für alle  $T \in \mathcal{T}$ :

1. Berechne die diskrete Lösung  $u_h \in \mathcal{P}_0(\mathcal{T})$ .
2. Berechne den Morley-Interpolanten  $I_M u_h$  für  $\mathcal{T}_k$ .
3. Berechne  $\eta_T$  ( $T \in \mathcal{T}_k$ ) nach (3.17).
4. Entscheide, ob Abbruch oder Verfeinerung.
5. Sofern Verfeinerung, markiere Elemente  $T$  zur Verfeinerung ( $\text{markiert}[T] = 1$ ) falls (4.2) für  $T \in \mathcal{T}$  gilt.
6. Verfeinere alle markierten ( $\text{markiert}[T] = 1$ ) Elemente  $T \in \mathcal{T}$ , verfeinere gegebenenfalls auch zusätzliche Elemente, um die Ordnung der Fastregularität zu erhalten, erhöhe  $k$  ( $\mathcal{T}_k \rightarrow \mathcal{T}_{k+1}$ ), setze  $\text{markiert}[T] = 0$  für alle  $T \in \mathcal{T}$  und gehe zu 1.

Als Output liefert der Algorithmus eine (endliche) Folge von fastregulären Triangulierungen  $(\mathcal{T}_0, \mathcal{T}_1, \dots)$  der Ordnung 2 mit dem dazugehörigen Morley-Interpolant  $(I_M u_h)_k$ .

**Bemerkung.** Der Parameter  $\theta = 0$  führt zu einer Verfeinerung jedes Elements, also zu einer *uniformen* Netzverfeinerung. Ist  $\theta = 1$  werden im Allgemeinen nur wenige Elemente verfeinert (oft nur eines). Was die beste Wahl ist, lässt sich schwer sagen (und sollte problemabhängig sein). In der Regel wird mit  $\theta = 1/2$  gestartet.  $\theta$  kann natürlich im Laufe der Rechnung nach unten oder oben korrigiert werden.

### 4.1.3 Netzverfeinerung

Um die Bedingung der fastregulären Triangulierung mit Ordnung 2 in lauter Quadrate zu erfüllen, müssen bei einer Verfeinerung eines Elements (ein Quadrat wird also in 4 gleichgroße Quadrate verfeinert) eventuell auch Nachbarelemente verfeinert werden. Dies führt auf eine Rekursion. Abbildung 4.1 zeigt eine mögliche Verfeinerung.

Das Rekursionsprinzip, und somit die Abbruchbedingung, ist dasselbe wie bei der Berechnung des Morley-Interpolanten (vergleiche dazu die Bemerkungen am Anfang des Kapitels). Mit der Datenstruktur *MESH* aus Abschnitt 4.1.1 erhalten wir folgenden Algorithmus.

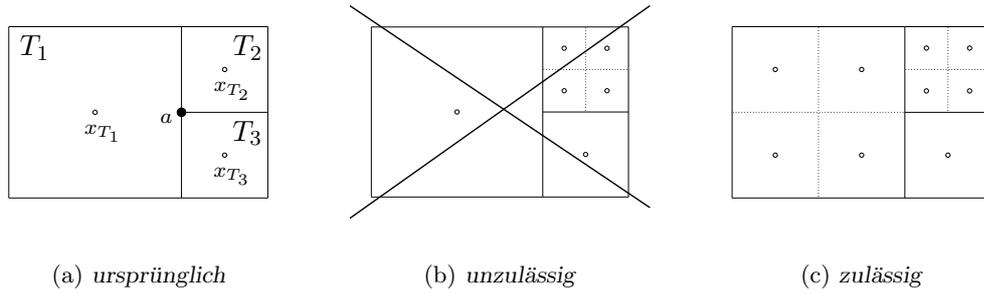


Abbildung 4.1: Bei der Verfeinerung von  $T_2$  muss auch  $T_1$  verfeinert werden, da sonst eine Kante von  $T_1$  zwei hängende Knoten hat, was der Festlegung unserer Triangulierung widerspricht!

**Algorithmus 4.1.5 (Verfeinerung).** Gegeben sei eine fastreguläre Triangulierung  $\mathcal{T}$  mit Ordnung 2 in lauter Quadrate in der Datenstruktur  $MESH$  und  $\text{markiert}[T] = 1$  für alle  $T \in \mathcal{T}$ , die verfeinert werden sollen. Der Algorithmus ruft die Funktion **refinement** auf, die weiter unten beschrieben wird.

```

for  $T \in \mathcal{T}$ 
  if  $\text{markiert}[T] == 1$ 
    call  $[\text{markiert}, MESH] = \text{refinement}\{MESH, T, \text{markiert}\}$ 
  end
end

```

Der Algorithmus liefert eine geänderte Datenstruktur  $MESH$ , die die neue Triangulierung darstellt.

Die rekursive Funktion **refinement** durchläuft bei einem Element  $T$  zuerst alle Eckknoten. Ist ein Eckknoten  $a$  hängend, so wird zuerst das Vaterelement von  $a$  verfeinert.

**Algorithmus 4.1.6 (function refinement).** Die Funktion benötigt die Datenstruktur  $MESH$ , das zu verfeinernde Element  $T$  und das Feld  $\text{markiert}$ . Sie liefert die geänderte Datenstruktur  $MESH$  und das Feld  $\text{markiert}$ . Für alle in **refinement** verfeinerten Elemente  $T$  gilt  $\text{markiert}[T] = 0$ .

```

function  $[\text{markiert}, MESH] = \text{refinement}\{MESH, T, \text{markiert}\}$ 
  for  $a$  ist Eckknoten von  $T$ 
    if  $a$  ist hängender Knoten
       $T_a = \text{Vaterelement von } a$ 
      call  $[\text{markiert}, MESH] = \text{refinement}\{MESH, T_a, \text{markiert}\}$ 
    end
  end
  Verfeinere das Element  $T$  in vier gleichgroße Quadrate, aktualisiere die Datenstruktur  $MESH$  und setze  $\text{markiert}[T] = 0$ .
  return  $\text{markiert}, MESH$ 

```

Wir erhalten die neue bzw. ergänzte Datenstruktur  $MESH$  und das geänderte Feld  $\text{markiert}$ , wobei für alle  $T \in \mathcal{T}$ , die rekursiv verfeinert wurden, jetzt  $\text{markiert}[T] = 0$  ist.

**Bemerkung.** Beachte, dass wir aufgrund der Aktualisierung von *MESH* in der Routine *refinement* uns nicht merken müssen, ob ein Element schon verfeinert wurde oder nicht. Es sei z.B.  $T_1$  und  $T_2$  aus Abbildung 4.1 markiert ( $\text{markiert}[T_1] = 1$ ,  $\text{markiert}[T_2] = 1$  und  $\text{markiert}[T_3] = 0$ ). Wird zuerst  $T_1$  verfeinert, so ist  $a$  kein hängender Knoten mehr und bei der anschließenden Verfeinerung von  $T_2$  erfolgt keine Rekursion in *refinement*. Der Algorithmus 4.1.5 wird zweimal durchlaufen. Verfeinern wir allerdings zuerst  $T_2$ , so wird  $T_1$  rekursiv in *refinement* verfeinert, *refinement* setzt  $\text{markiert}[T_1] = 0$  und  $\text{markiert}[T_2] = 0$ . Der Algorithmus 4.1.5 wird also nur einmal durchlaufen.

#### 4.1.4 Berechnung der Finite Volumen Lösung

Wir stellen ein Gleichungssystem  $A\mathbf{u}_h = \mathbf{b}$  nach Formel (3.4) auf. Wir erhalten die diskrete Finite Volumen Lösung  $\mathbf{u}_h$ . Der Vektor  $\mathbf{u}_h$  enthält sämtliche  $u_T$  und es gilt  $u_h|_T := u_T$  (Achtung,  $u_h$  ist die Finite Volumen Methode Lösung mit  $u_h \in \mathcal{P}_0(\mathcal{T})$ , also kein Vektor).

**Algorithmus 4.1.7 (Diskrete Lösung).** Gegeben sei eine fastreguläre Triangulierung  $\mathcal{T}$  mit Ordnung 2 in lauter Quadrate in der Datenstruktur *MESH* und ein Problem nach (3.1).

1. Stelle die Matrix  $A$  von (3.4) mit der Flussapproximation nach (2.13) auf. Die Unbekannten an den Knoten werden nach Abschnitt 2.3.2 durch lineare Kombination mit den Gewichten  $\psi_T(a)$  der Unbekannten im Schwerpunkt ersetzt.
2. Berechne die rechte Seite  $\mathbf{b}$  von (3.4). Achtung, für inhomogenen Rand bekommen wir für  $\mathbf{b}$  auch von (2.13) einen Anteil.
3. Löse das Gleichungssystem  $A\mathbf{u}_h = \mathbf{b}$ , wir erhalten die diskrete Lösung  $\mathbf{u}_h$ .

Als Output erhalten wir die diskrete Finite Volumen Lösung  $\mathbf{u}_h$ .

#### 4.1.5 Berechnung des Morley-Interpolanten

Wir berechnen den Morley-Interpolanten und speichern diesen in der Datenstruktur *MORLEY* (vergleiche Abschnitt 4.1.1) ab. Das Feld *markiert* gibt an, für welches Element  $T \in \mathcal{T}$  der Morley-Interpolant berechnet werden muss ( $\text{markiert}[T] = 1$ ) und für welches nicht mehr ( $\text{markiert}[T] = 0$ ).

**Algorithmus 4.1.8 (Morley-Interpolant).** Gegeben sei eine fastreguläre Triangulierung  $\mathcal{T}$  mit Ordnung 2 in lauter Quadrate in der Datenstruktur *MESH*,  $\mathbf{u}_h$  sei die diskrete Finite Volumen Methode Lösung und  $\text{markiert}[T] = 1$  für alle  $T \in \mathcal{T}$ . Der Algorithmus ruft die Funktion *morleyinter* auf.

```

for  $T \in \mathcal{T}$ 
  if  $\text{markiert}[T] == 1$ 
    call [ $\text{markiert}$ , MORLEY] = morleyinter{MORLEY, MESH,  $T$ ,  $\text{markiert}$ ,  $\mathbf{u}_h$ }
  end
end

```

Als Output erhalten wir die Datenstruktur *MORLEY*, die den Morley-Interpolanten enthält.

Die Berechnung von  $I_M u_h|_T \in \mathcal{P}$  erfolgt rekursiv nach Lemma 3.3.1 durch Aufstellen eines Gleichungssystems in den Unbekannten der Polynomkoeffizienten von  $I_M u_h|_T$ . Das Rekursionsprinzip ist wie bei der Netzgenerierung, allerdings müssen wir uns merken, für welches Element  $T$   $I_M u_h|_T$  schon berechnet wurde. Dies geschieht durch das Feld *markiert*.

**Algorithmus 4.1.9 (function morleyinter).** Die Funktion benötigt die Datenstrukturen *MORLEY* und *MESH*, das Element  $T$ , das Feld *markiert* und die Finite Volumen Lösung  $\mathbf{u}_h$ . Sie liefert die geänderte Datenstruktur *MORLEY* und das geänderte Feld *markiert* zurück.

```
function [markiert, MORLEY] = morleyinter{MORLEY, MESH, T, markiert,  $\mathbf{u}_h$ }
for a ist Eckknoten von T
  if a ist hängender Knoten
     $T_a$  = Vaterelement von a
    if markiert[ $T_a$ ] == 1
      call [markiert, MORLEY] = morleyinter{MORLEY, MESH,  $T_a$ , markiert,  $\mathbf{u}_h$ }
    end
  end
end
end
Berechne die Koeffizienten von  $I_M u_h|_T$  durch Aufstellen eines Gleichungssystems nach Lemma 3.3.1, speichere sie in der Datenstruktur MORLEY ab und setze markiert[T] = 0.
return markiert, MORLEY
```

Wir erhalten die neue bzw. ergänzte Datenstruktur *MORLEY* und das Feld *markiert*, wobei für alle in *morleyinter* berechneten  $I_M u_h|_T$  *markiert*[ $T$ ] = 0 gilt.

## 4.2 Uniforme und adaptive Resultate

Wir diskutieren in diesem Abschnitt zwei Beispiele mit glatter Lösung in  $\Omega$ , beim ersten Beispiel haben wir Nullranddaten und beim zweiten inhomogenen Dirichlet-Rand. Das dritte Beispiel hat eine Singularität (und inhomogenen Dirichlet-Rand) in  $\Omega$ . Verglichen werden die Lösungen von uniformer ( $\theta = 0$ ) und adaptiver ( $\theta = 0.5$ ) Verfeinerung. Der Fehler  $e$  läßt sich durch

$$e(N) = \beta N^{-\alpha} \quad (4.3)$$

darstellen, wobei  $N$  die Anzahl der Elemente ist. Dabei entspricht  $2\alpha$  der Konvergenzordnung  $\mathcal{O}(h)$ . Die Konvergenzkonstanten  $\beta$  und  $\alpha$  lassen sich somit aus zwei Fehlerwerten für unterschiedliche Anzahl von Elementen berechnen. Es gilt

$$\alpha = -\frac{\ln(e(N_1)/e(N_2))}{\ln(N_1/N_2)}$$

und  $\beta$  erhalten wir durch Einsetzen in (4.3). Beachte, dass für die Konvergenzordnung  $\mathcal{O}(h) = 2\alpha$  gilt. Wenn  $u \in H^1(\Omega)$  ist, dann betrachten wir den Fehler  $e = u - I_M u_h$ . In den Fehlerkurven vergleichen wir den Fehlerschätzer  $\eta$  mit der  $H^1$ -Halbnorm von  $e$  über die Anzahl der Elemente  $N$ . Wir plotten für uniforme und adaptive Verfeinerung jeweils zwei Fehlerkurven aus, die den exakten Fehler und den Fehlerschätzer über  $N$  zeigen. Bei der zweiten Grafik skalieren wir den Fehlerschätzer mit dem letzten berechneten Wert, d.h. mit  $\|\nabla e\|_{L^2(\Omega)}/\eta$ . Die Kurven zeigen die Zuverlässigkeit und Effizienz unseres Fehlerschätzers. Eine Grafik zeigt die Konvergenz (wir ändern die Form der Elemente nicht) von  $\|\nabla e\|_{L^2(\Omega)}/\eta$  gegen die Zuverlässigkeitskonstante  $C$ .

Weiters zeigen wir noch die Entwicklung der Triangulierungen bei adaptiver Verfeinerung und dann noch den direkten Vergleich der Entwicklung des exakten Fehlers  $e$  bei uniformer und adaptiver Verfeinerung in der  $H^1$ -Halbnorm sowie des diskreten Fehlers  $e_h$  in der diskreten  $H^1$ -Norm. Mit der diskreten  $H^1$ -Norm

$$\|e_h\|_{1,h} = \left( \sum_{E \in \mathcal{E}^* \cup \mathcal{E}_T} \frac{|e_{hE} - e_{hW}|^2}{d_E} h_E \right)^{1/2} \quad (4.4)$$

messen wir den Fehler der diskreten Finite Volumen Lösung. Dabei ist  $e_h = u_{\mathcal{T}} - u_h$ , wobei  $u_{\mathcal{T}} \in \mathcal{P}_0(\mathcal{T})$  das  $\mathcal{T}$ -stückweise Integralmittel der exakten Lösung  $u$  ist,  $d_E = (x_E - x_W) \cdot \mathbf{n}_E$  und  $h_E$  die Länge der Kante bezeichnet. Für uniforme Verfeinerungen ist  $d_E = |x_E - x_W|$ , und  $\|e_h\|_{1,h}$  ist dann die diskrete  $H^1$ -Norm für zulässige Triangulierungen (Definition 2.3.1) aus dem Handbook of Numerical Analysis [10, Definition 9.3]. Für eine hinreichend glatte exakte Lösung  $u \in H^2(\Omega)$  konvergiert das Verfahren, und es gilt  $\|u_{\mathcal{T}} - u_h\|_{1,h} = \mathcal{O}(h)$  (Coudière und Villedieu [9]).

#### 4.2.1 Beispiel mit glatter Lösung und Nullranddaten

Wir betrachten das Einheitsquadrat  $\Omega = ]0, 1[ \times ]0, 1[$ . Zunächst wählen wir eine Lösung mit Nullranddaten  $u(x, y) = xy(1-x)(1-y)$  und somit  $-\Delta u(x, y) = f(x, y) = 2(x+y) - 2(x^2 + y^2)$ . Die Aufgabe für das *Standard Laplace Problem mit Nullranddaten* lautet: Finde ein  $u \in H^1(\Omega)$ , so dass

$$\begin{aligned} -\Delta u(x, y) &= 2(x+y) - 2(x^2 + y^2) \quad \text{in } \Omega, \\ u &= 0 \quad \text{auf } \Gamma = \partial\Omega \end{aligned} \quad (4.5)$$

gilt. Abbildung 4.2 zeigt die exakte Lösung auf einem uniformen Gitter. Wir erwarten bei diesem

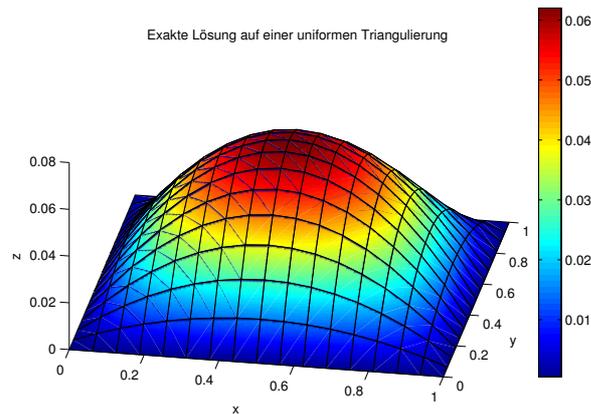


Abbildung 4.2: Die exakte Lösung von Beispiel 4.2.1 berechnet auf einem uniformen Gitter.

Beispiel eine Konvergenzordnung  $\mathcal{O}(h)$  ( $\alpha = 1/2$ ) der diskreten Lösung. Wie in der Arbeit von Nicaise [13] bekommen wir sogar Konvergenzordnung  $\mathcal{O}(h^2)$  ( $\alpha = 1$ ) für die Morley-Lösung und  $\mathcal{O}(h^{3/2})$  ( $\alpha = 3/4$ ) für die diskrete Lösung bei uniformer Verfeinerung. Bei adaptiver Verfeinerung ist die Konvergenzordnung jedoch schlechter als bei uniformer Verfeinerung.

$k$	$N$	Fehler			Fehlerabschätzung		
		$\ \nabla e\ _{L^2(\Omega)}$	$\beta$	$\alpha$	$\eta$	$\beta$	$\alpha$
0	4	$2.946e-02$	$8.97e-02$	0.80	$6.437e-02$	$1.66e-01$	0.68
1	16	$9.676e-03$	$1.31e-01$	0.94	$2.491e-02$	$2.64e-01$	0.85
2	64	$2.634e-03$	$1.57e-01$	0.98	$7.652e-03$	$3.89e-01$	0.94
3	256	$6.739e-04$	$1.68e-01$	1.00	$2.066e-03$	$4.66e-01$	0.98
4	1024	$1.695e-04$	$1.72e-01$	1.00	$5.331e-04$	$5.09e-01$	0.99
5	4096	$4.243e-05$	$1.73e-01$	1.00	$1.351e-04$	$5.32e-01$	1.00
6	16384	$1.061e-05$			$3.400e-05$		

Tabelle 4.1: Der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm und der Fehlerschätzer  $\eta$  in Beispiel 4.2.1 unter Vernachlässigung der Konstanten in der Fehlerabschätzung sowie der Terme höherer Ordnung bei uniformer Verfeinerung mit den Konstanten  $\beta$  und  $\alpha$ . Wir erhalten also eine Konvergenzordnung  $\mathcal{O}(h^2)$  ( $\alpha = 1$ ).

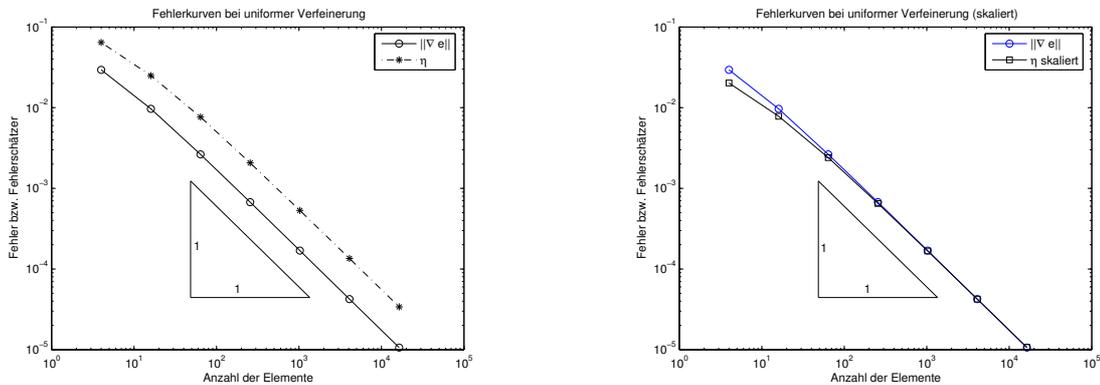


Abbildung 4.3: Der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm und Fehlerschätzer  $\eta$  bei uniformer Verfeinerung in Beispiel 4.2.1: Links stehen absolute Werte in Abhängigkeit der Anzahl der Elemente in  $\mathcal{T}$   $N = 4, 16, \dots, 16384$  unter Vernachlässigung der Konstanten in der Fehlerabschätzung sowie der Terme höherer Ordnung. Im rechten Bild wird der Fehlerschätzer so skaliert, dass für  $N = 16384$  die Werte von  $e$  und  $\eta$  übereinstimmen.

$k$	$N$	Fehler			Fehlerabschätzung		
		$\ \nabla e\ _{L^2(\Omega)}$	$\beta$	$\alpha$	$\eta$	$\beta$	$\alpha$
0	4	$2.946e-02$	$1.31e-01$	0.94	$6.437e-02$	$2.64e-01$	0.85
2	64	$2.634e-03$	$9.17e-01$	1.21	$7.652e-03$	$5.52e+00$	1.29
4	280	$1.005e-03$	$1.36e-06$	-1.11	$3.819e-03$	$1.03e-06$	-1.37
6	364	$9.719e-04$	$1.86e-07$	-1.44	$3.329e-03$	$2.82e-07$	-1.58
8	496	$1.411e-03$	$1.90e-03$	0.04	$4.998e-03$	$1.96e-03$	-0.16
10	724	$1.437e-03$	$1.75e-03$	0.09	$5.609e-03$	$9.11e-03$	0.14
12	1228	$8.957e-04$	$3.53e+01$	1.45	$3.342e-03$	$1.72e+01$	1.17
14	1900	$6.370e-04$	$1.03e+00$	1.00	$2.443e-03$	$2.83e+02$	1.55
16	2464	$4.335e-04$	$3.98e-02$	0.58	$1.592e-03$	$2.75e-02$	0.37
18	2932	$3.913e-04$	$3.43e-03$	0.28	$1.487e-03$	$6.24e-03$	0.18
20	3412	$3.652e-04$	$4.59e+01$	1.42	$1.403e-03$	$4.07e+02$	1.52
22	4612	$2.933e-04$	$6.89e-02$	0.65	$1.126e-03$	$4.72e-01$	0.71
24	5680	$2.576e-04$	$6.96e-02$	0.65	$9.875e-04$	$5.60e-01$	0.73
26	6616	$2.320e-04$	$1.74e-02$	0.49	$8.815e-04$	$2.00e-01$	0.62
28	7468	$2.199e-04$	$1.98e+01$	1.28	$8.272e-04$	$1.57e+03$	1.62
30	8236	$1.983e-04$	$1.76e-01$	0.75	$7.351e-04$	$3.18e-01$	0.67
32	8800	$1.889e-04$	$2.65e+00$	1.05	$7.038e-04$	$1.74e+00$	0.86
34	9472	$1.768e-04$	$4.87e-02$	0.62	$6.647e-04$	$1.40e-01$	0.59
36	10132	$1.665e-04$	$1.39e-02$	0.48	$6.310e-04$	$1.69e-03$	0.11
38	10888	$1.589e-04$	$2.58e-07$	-0.66	$6.107e-04$	$1.33e-06$	-0.63
40	11752	$1.224e-04$	$2.05e+00$	1.06	$4.878e-04$	$3.43e-01$	0.72
42	13120	$9.216e-05$	$6.17e+03$	1.89	$3.887e-04$	$4.64e+04$	1.95
44	15460	$7.193e-05$			$3.107e-04$		

Tabelle 4.2: Der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm und der Fehlerschätzer  $\eta$  in Beispiel 4.2.1 unter Vernachlässigung der Konstanten in der Fehlerabschätzung sowie der Terme höherer Ordnung bei adaptiver Verfeinerung mit den Konstanten  $\beta$  und  $\alpha$ . Wir erhalten also eine Konvergenzordnung  $\mathcal{O}(h)$  ( $\alpha = 1/2$ ).

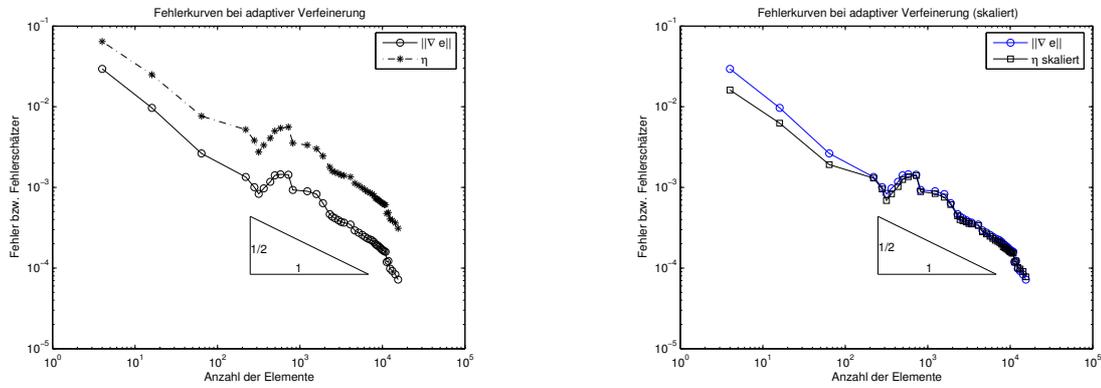


Abbildung 4.4: Der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm und Fehlerschätzer  $\eta$  bei adaptiver Verfeinerung in Beispiel 4.2.1: Links stehen absolute Werte in Abhängigkeit der Anzahl der Elemente in  $\mathcal{T}$   $N = 4, 16, \dots, 15460$  unter Vernachlässigung der Konstanten in der Fehlerabschätzung sowie der Terme höherer Ordnung. Im rechten Bild wird der Fehlerschätzer so skaliert, dass für  $N = 15460$  die Werte von  $e$  und  $\eta$  übereinstimmen.

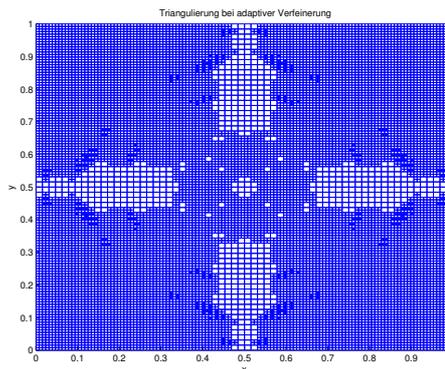
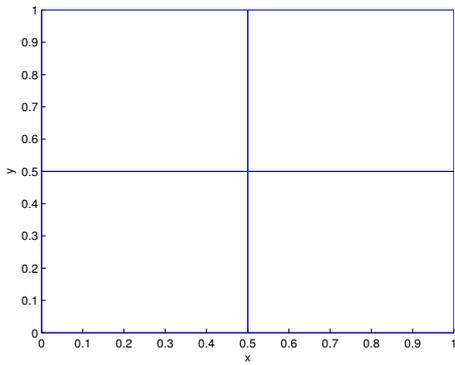
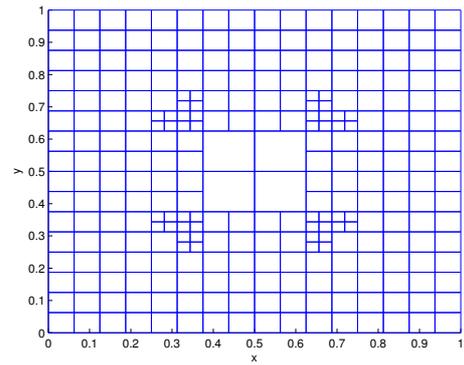


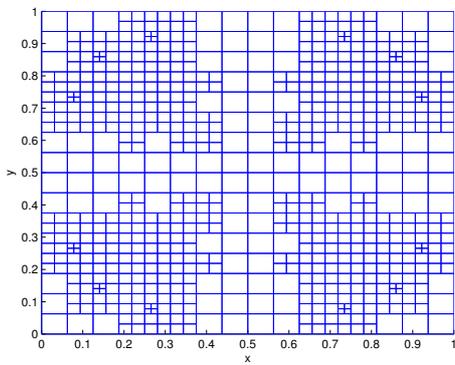
Abbildung 4.5: Die Triangulierung nach  $k = 44$  adaptiven Verfeinerungen in Beispiel 4.2.1.



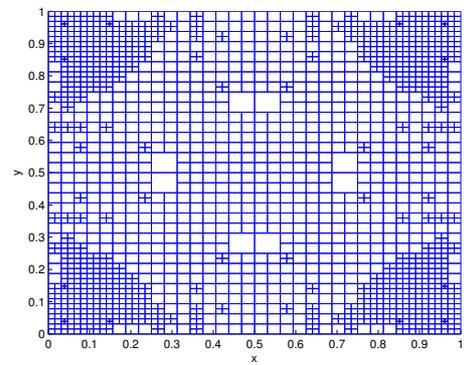
(a)  $k = 0$



(b)  $k = 4$



(c)  $k = 10$



(d)  $k = 14$

Abbildung 4.6: Verschiedene Triangulierungen nach  $k$  adaptiven Verfeinerungen in Beispiel 4.2.1.

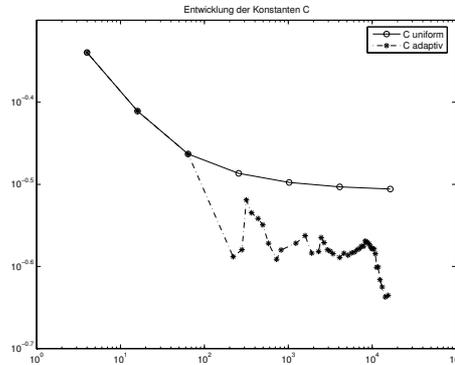


Abbildung 4.7: Das Verhältnis  $\|\nabla e\|_{L^2(\Omega)}/\eta$  bei uniformer und adaptiver Verfeinerung in Beispiel 4.2.1 konvergiert gegen die Zuverlässigkeitskonstante  $C$ .

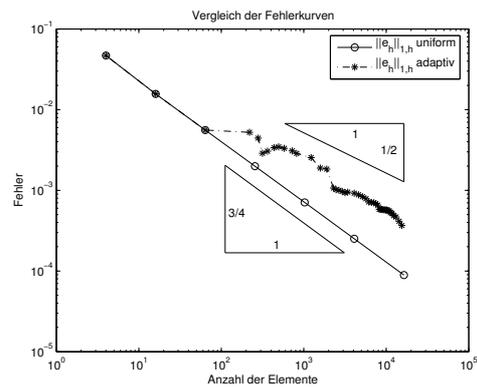
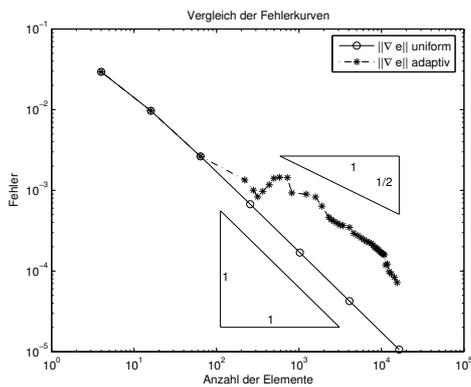


Abbildung 4.8: Fehlervergleich bei uniformer und adaptiver Verfeinerung in Beispiel 4.2.1: Links der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm, rechts der diskrete Fehler  $e_h = u_{\mathcal{T}} - u_h$  in der diskreten  $H^1$ -Norm.

### 4.2.2 Beispiel mit glatter Lösung und inhomogenen Dirichlet-Daten

Es sei nun  $\Omega = ]0, 1/2[ \times ]0, 1/2[$ . Wir wollen das Laplace Problem mit der exakten Lösung  $u(x, y) = \sinh(2\pi x) \cos(2\pi y)$  numerisch lösen. Offensichtlich haben wir inhomogenen Dirichlet-Rand. Es ist

$$\nabla u = (2\pi \cosh(2\pi x) \cos(2\pi y), -2\pi \sinh(2\pi x) \sin(2\pi y)),$$

und  $\Delta u = 0$ . Abbildung 4.9 zeigt die exakte Lösung auf einem uniformen Gitter. Die Richtungs-

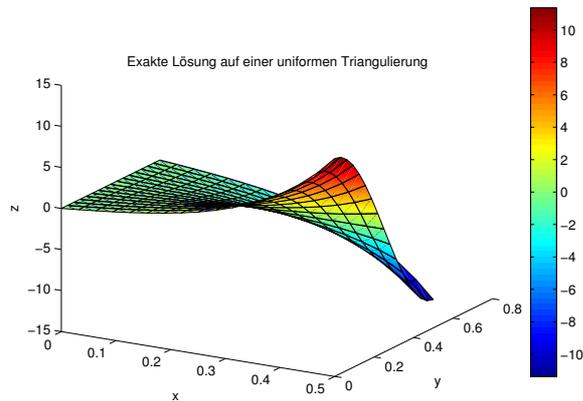


Abbildung 4.9: Die exakte Lösung von Beispiel 4.2.2 berechnet auf einem uniformen Gitter.

ableitung von  $u = u_D$  in Tangentialrichtung am Rand wird für die Berechnung des Tangentialsprunges beim Fehlerschätzer benötigt. Wir approximieren diese mit

$$\frac{u_D(a_N) - u_D(a_S)}{h_E} \quad \text{für } E \in \mathcal{E}_\Gamma,$$

wobei  $a_S$  der Anfangs- und  $a_N$  der Endknoten von  $E$  sind. Hier erhalten wir bei uniformer und adaptiver Netzverfeinerung die gleiche Konvergenzordnung  $\mathcal{O}(h^{3/2})$  ( $\alpha = 3/4$ ).

$k$	$N$	Fehler			Fehlerabschätzung		
		$\ \nabla e\ _{L^2(\Omega)}$	$\beta$	$\alpha$	$\eta$	$\beta$	$\alpha$
0	4	$2.910e + 00$	$6.60e + 00$	0.59	$7.453e + 00$	$1.48e + 01$	0.50
1	16	$1.283e + 00$	$9.40e + 00$	0.72	$3.747e + 00$	$1.83e + 01$	0.57
2	64	$4.741e - 01$	$1.14e + 01$	0.77	$1.696e + 00$	$2.76e + 01$	0.67
3	256	$1.641e - 01$	$1.15e + 01$	0.77	$6.692e - 01$	$3.48e + 01$	0.71
4	1024	$5.668e - 02$	$1.10e + 01$	0.76	$2.492e - 01$	$3.97e + 01$	0.73
5	4096	$1.975e - 02$	$1.06e + 01$	0.76	$9.042e - 02$	$4.28e + 01$	0.74
6	16384	$6.926e - 03$			$3.239e - 02$		

Tabelle 4.3: Der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm und der Fehlerschätzer  $\eta$  in Beispiel 4.2.2 unter Vernachlässigung der Konstanten in der Fehlerabschätzung sowie der Terme höherer Ordnung bei uniformer Verfeinerung mit den Konstanten  $\beta$  und  $\alpha$ . Wir erhalten also eine Konvergenzordnung von  $\mathcal{O}(h^{3/2})$  ( $\alpha = 3/4$ ).

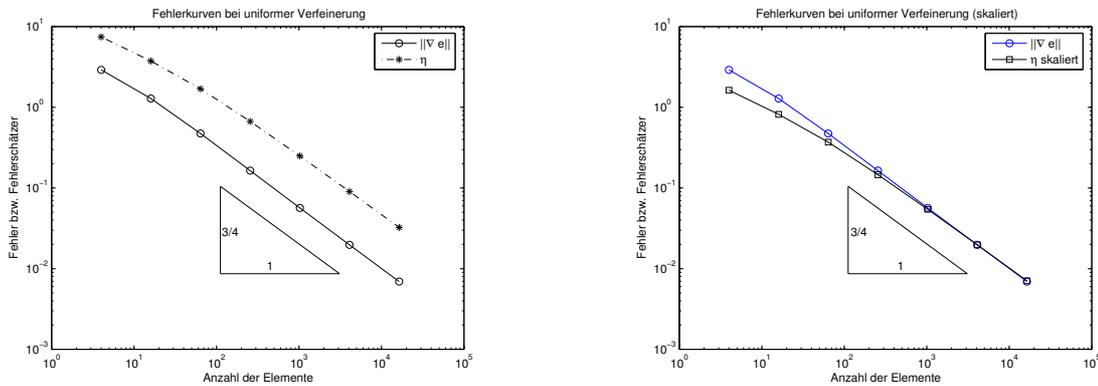


Abbildung 4.10: Der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm und Fehlerschätzer  $\eta$  bei uniformer Verfeinerung in Beispiel 4.2.2: Links stehen absolute Werte in Abhängigkeit der Anzahl der Elemente in  $\mathcal{T}$   $N = 4, 16, \dots, 16384$  unter Vernachlässigung der Konstanten in der Fehlerabschätzung sowie der Terme höherer Ordnung. Im rechten Bild wird der Fehlerschätzer so skaliert, dass für  $N = 16384$  die Werte von  $e$  und  $\eta$  übereinstimmen.

$k$	$N$	Fehler			Fehlerabschätzung		
		$\ \nabla e\ _{L^2(\Omega)}$	$\beta$	$\alpha$	$\eta$	$\beta$	$\alpha$
0	4	2.910e + 00	5.46e + 00	0.61	7.453e + 00	1.03e + 01	0.40
2	34	6.400e - 01	3.54e + 00	0.54	2.480e + 00	9.51e + 00	0.44
4	100	2.965e - 01	4.95e + 01	1.02	1.272e + 00	1.22e + 02	0.92
6	238	1.874e - 01	1.79e + 01	0.84	7.801e - 01	1.19e + 02	0.92
8	364	1.284e - 01	1.64e + 00	0.44	5.170e - 01	1.10e + 01	0.52
10	532	1.014e - 01	1.71e + 01	0.82	4.131e - 01	8.23e + 01	0.85
12	748	7.508e - 02	1.93e - 01	0.15	3.052e - 01	5.70e - 01	0.10
14	952	7.126e - 02	5.30e + 00	0.62	2.916e - 01	1.03e + 01	0.52
16	1270	6.233e - 02	3.28e + 01	0.87	2.579e - 01	1.68e + 02	0.89
18	1648	5.392e - 02	7.73e + 01	0.98	2.221e - 01	4.43e + 02	1.03
20	2014	4.375e - 02	1.00e + 01	0.71	1.799e - 01	8.73e + 01	0.81
22	2356	3.912e - 02	8.21e + 01	0.97	1.588e - 01	1.62e + 02	0.88
24	2878	3.570e - 02	3.13e + 04	1.71	1.437e - 01	1.98e + 04	1.48
26	3346	2.937e - 02	6.00e + 00	0.66	1.199e - 01	4.68e + 01	0.74
28	3784	2.620e - 02	7.13e + 03	1.52	1.064e - 01	7.03e + 03	1.35
30	4150	2.331e - 02	7.88e - 03	-0.13	9.565e - 02	1.48e - 01	0.05
32	4576	2.356e - 02	1.61e + 02	1.05	9.493e - 02	3.01e + 02	0.96
34	5044	2.142e - 02	7.19e + 01	0.95	8.719e - 02	6.22e + 01	0.77
36	5458	2.003e - 02	9.81e - 01	0.46	8.236e - 02	3.35e + 00	0.44
38	5932	1.823e - 02	5.26e + 02	1.18	7.517e - 02	1.19e + 03	1.11
40	6508	1.665e - 02	2.00e + 01	0.81	6.908e - 02	1.20e + 02	0.85
42	7078	1.545e - 02	5.78e + 05	1.96	6.386e - 02	2.73e + 05	1.72
44	7654	1.368e - 02	1.07e - 01	0.23	5.719e - 02	1.31e + 00	0.35
46	8200	1.320e - 02	1.86e + 01	0.81	5.451e - 02	5.49e + 01	0.77
48	8806	1.189e - 02	2.38e + 05	1.85	4.974e - 02	2.30e + 06	1.94
50	9400	1.086e - 02	4.48e - 02	0.16	4.522e - 02	7.12e - 01	0.30
52	9952	1.055e - 02	2.26e + 01	0.83	4.347e - 02	1.84e + 02	0.91
54	10582	1.022e - 02			4.176e - 02		

Tabelle 4.4: Der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm und der Fehlerschätzer  $\eta$  in Beispiel 4.2.2 unter Vernachlässigung der Konstanten in der Fehlerabschätzung sowie der Terme höherer Ordnung bei adaptiver Verfeinerung mit den Konstanten  $\beta$  und  $\alpha$ . Wir erhalten also eine Konvergenzordnung von  $\mathcal{O}(h^{3/2})$  ( $\alpha = 3/4$ ).

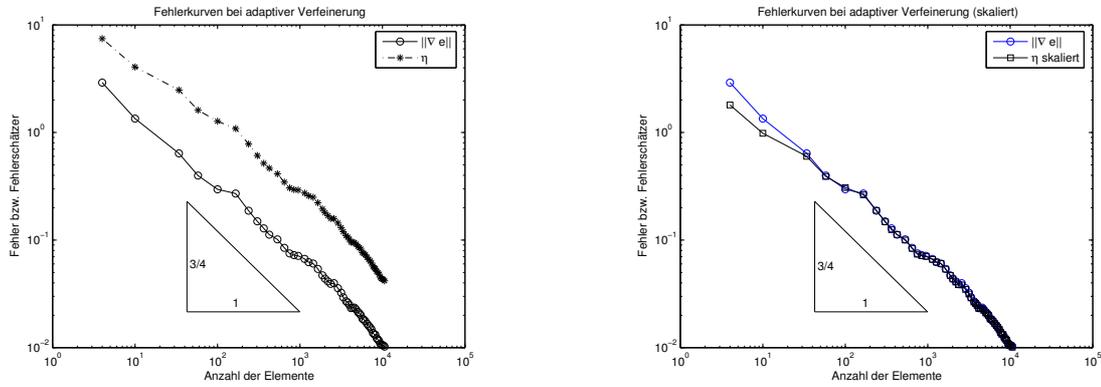


Abbildung 4.11: Der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm und Fehlerschätzer  $\eta$  bei adaptiver Verfeinerung in Beispiel 4.2.2: Links stehen absolute Werte in Abhängigkeit der Anzahl der Elemente in  $\mathcal{T}$   $N = 4, 10, \dots, 10582$  unter Vernachlässigung der Konstanten in der Fehlerabschätzung sowie der Terme höherer Ordnung. Im rechten Bild wird der Fehlerschätzer so skaliert, dass für  $N = 10582$  die Werte von  $e$  und  $\eta$  übereinstimmen.

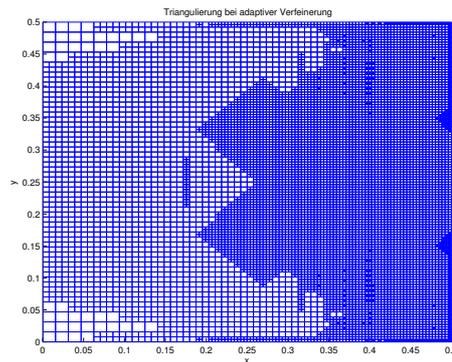
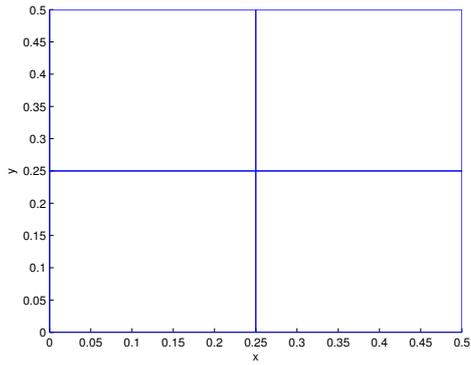
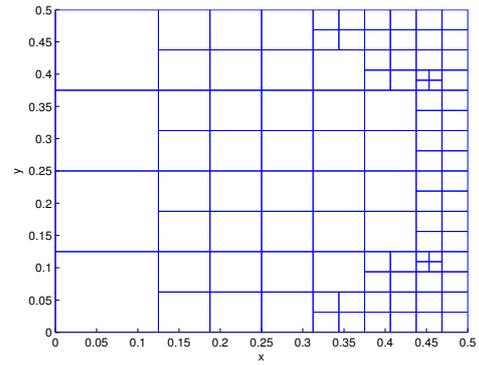


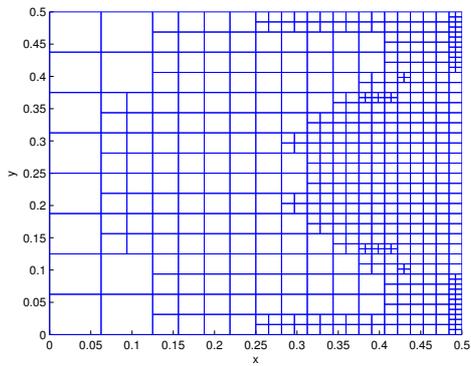
Abbildung 4.12: Die Triangulierung nach  $k = 54$  adaptiven Verfeinerungen in Beispiel 4.2.2.



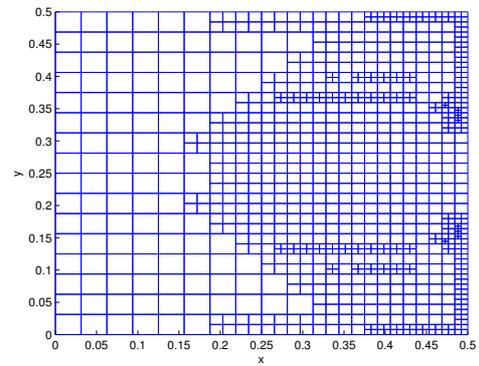
(a)  $k = 0$



(b)  $k = 4$



(c)  $k = 10$



(d)  $k = 14$

Abbildung 4.13: Verschiedene Triangulierungen nach  $k$  adaptiven Verfeinerungen in Beispiel 4.2.2.

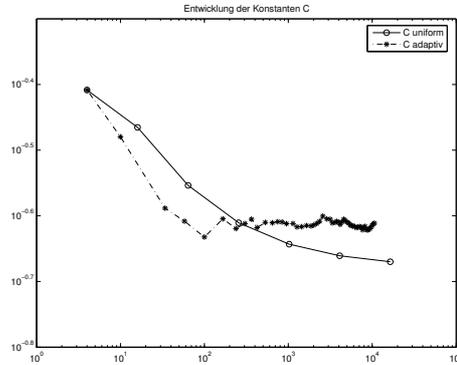


Abbildung 4.14: Das Verhältnis  $\|\nabla e\|_{L^2(\Omega)}/\eta$  konvergiert bei uniformer und adaptiver Verfeinerung in Beispiel 4.2.2 gegen die Zuverlässigkeitskonstante  $C$ .

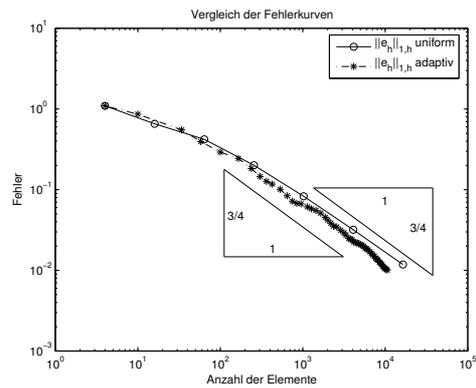
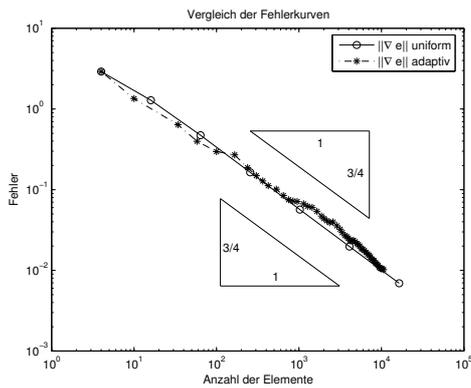


Abbildung 4.15: Fehlervergleich bei uniformer und adaptiver Verfeinerung in Beispiel 4.2.2: Links der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm, rechts der diskrete Fehler  $e_h = u_T - u_h$  in der diskreten  $H^1$ -Norm.

### 4.2.3 Beispiel mit Eckensingularität und inhomogenen Dirichlet-Daten

Diesmal sei  $\Omega = ]-1, 1[ \times ]-1, 1[ \setminus ]0, 1[ \times ]-1, 0[$  (also ein L-shaped domain). Wir betrachten wieder das Laplace Problem  $-\Delta u = f(x, y)$  auf  $\Omega$ , diesmal mit inhomogenen Dirichlet-Rand. Als Lösung sei diesmal  $u(x, y) = r^{2/3} \sin(2\varphi/3)$  in Polarkoordinaten angenommen. Für die Polarkoordinaten  $r > 0$ ,  $\varphi \in [0, 2\pi[$  setzen wir  $(x, y) := r(\cos \varphi, \sin \varphi)$ . Diese Funktion hat eine Singularität der ersten Ableitungen am Ursprung. Abbildung 4.16 zeigt die exakte Lösung auf einem uniformen Gitter.

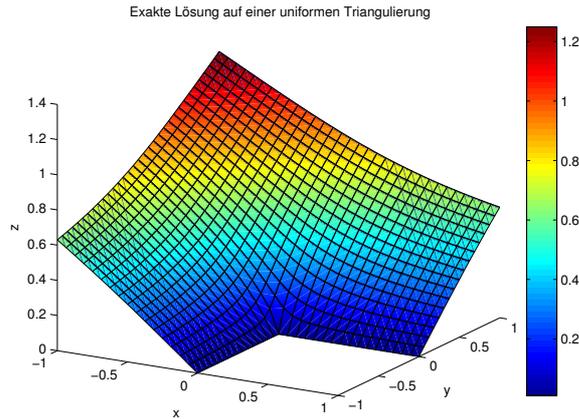


Abbildung 4.16: Die exakte Lösung von Beispiel 4.2.3 berechnet auf einem uniformen Gitter. Im Ursprung befindet sich die Singularität der ersten Ableitungen.

**Satz 4.2.1.** Die Funktion  $u(x, y) = r^{2/3} \sin(2\varphi/3)$  ist harmonisch, d.h.  $\Delta u = 0$ .

**Beweis.** Mit  $z = x + iy = r \cos \varphi + ir \sin \varphi$  ist

$$z^{2/3} = (re^{i\varphi})^{2/3} = r^{2/3} \cos(2\varphi/3) + ir^{2/3} \sin(2\varphi/3)$$

holomorph. Aus der Funktionentheorie wissen wir (siehe z.B. Remmert [16]), dass für eine holomorphe Funktion sowohl sie selbst, ihr Realteil, als auch ihr Imaginärteil harmonisch sind. Da  $u(x, y) = \text{Im}(z^{2/3}) = r^{2/3} \sin(2\varphi/3)$ , gilt  $\Delta u = 0$ . ■

Somit ist  $f(x, y) = 0$ . Weiter unten werden wir auch den Gradienten von  $u(x, y)$  brauchen. In kartesischen Koordinaten berechnen wir

$$\nabla u = \left( \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right).$$

Wir rechnen den Gradienten nun in Polarkoordinaten um. Mit der Kettenregel erhalten wir

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial u}{\partial \varphi} \frac{\partial \varphi}{\partial x}, \quad \frac{\partial u}{\partial y} = \frac{\partial u}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial u}{\partial \varphi} \frac{\partial \varphi}{\partial y}.$$

Da  $r^2 = x^2 + y^2$  und  $\varphi = \text{Arctan}(y/x) + c$  mit  $c = 0$  für  $(x, y)$  im I.Quadranten,  $c = \pi$  für  $(x, y)$

im II. und III.Quadranten und  $c = 2\pi$  für  $(x, y)$  im IV.Quadranten gilt, erhalten wir

$$\begin{aligned}\frac{\partial r}{\partial x} &= \frac{x}{r} = \cos \varphi, & \frac{\partial r}{\partial y} &= \frac{y}{r} = \sin \varphi. \\ \frac{\partial \varphi}{\partial x} &= \frac{1}{1 + (y^2/x^2)} \left( -\frac{y}{x^2} \right) = -\frac{y}{r^2} = -\frac{\sin \varphi}{r}, \\ \frac{\partial \varphi}{\partial y} &= \frac{1}{1 + (y^2/x^2)} \frac{1}{x} = \frac{x}{r^2} = \frac{\cos \varphi}{r}.\end{aligned}$$

Mit

$$\frac{\partial u}{\partial r} = \frac{2}{3} r^{-1/3} \sin(2\varphi/3) \quad \text{und} \quad \frac{\partial u}{\partial \varphi} = \frac{2}{3} r^{2/3} \cos(2\varphi/3)$$

erhalten wir für den Gradienten

$$\begin{aligned}\nabla u &= \frac{2}{3} r^{-1/3} (\sin(2\varphi/3) \cos \varphi - \cos(2\varphi/3) \sin \varphi, \sin(2\varphi/3) \sin \varphi + \cos(2\varphi/3) \cos \varphi) \\ &= \frac{2}{3} r^{-1/3} (-\sin(\varphi/3), \cos(\varphi/3)).\end{aligned}\tag{4.6}$$

Die Richtungsableitung von  $u = u_D$  in Tangentialrichtung am Rand wird für die Berechnung des Tangentialsprunges beim Fehlerschätzer benötigt. Wir approximieren diese mit

$$\frac{u_D(a_N) - u_D(a_S)}{h_E} \quad \text{für } E \in \mathcal{E}_\Gamma,$$

wobei  $a_S$  der Anfangs- und  $a_N$  der Endknoten von  $E$  sind.

Für die Berechnung des exakten Fehlers  $\|\nabla e\|_{L^2(\Omega)}$  wird in MATLAB die Quadratur *dblquad* verwendet. Bei adaptiver Verfeinerung mit der Singularität im Ursprung muss die Genauigkeit von  $10^{-6}$  (Standard) für die Berechnung des Integrals erhöht werden. Es reicht dabei aus, die Berechnungsgenauigkeit nur für die drei Elemente  $T \in \mathcal{T}$  zu erhöhen, die als Eckpunkt den Ursprung enthalten. Für die untenstehenden Berechnungen wird deshalb die Genauigkeit auf  $10^{-9}$  erhöht. Die exakte Lösung ist in  $u \in H^{1+2/3-\epsilon}(\Omega)$  für alle  $\epsilon > 0$ . Bei der Finite Elemente Methode fällt die Konvergenzordnung bei uniformer Netzverfeinerung auf  $\mathcal{O}(h^{2/3})$  ( $\alpha = 1/3$ ). Diese experimentielle Konvergenzordnung ist auch bei unserer Finite Volumen Lösung bei uniformer Netzverfeinerung sichtbar. Durch adaptive Verfeinerung erhalten wir sogar eine Konvergenzordnung  $\mathcal{O}(h^{3/2})$  ( $\alpha = 3/4$ ). Die Abbildungen der adaptiven Netzverfeinerung demonstrieren eindrucksvoll, wie zur Singularität hin verfeinert wird.

$k$	$N$	Fehler			Fehlerabschätzung		
		$\ \nabla e\ _{L^2(\Omega)}$	$\beta$	$\alpha$	$\eta$	$\beta$	$\alpha$
0	12	$1.431e-01$	$3.08e-01$	0.31	$7.891e-01$	$1.75e+00$	0.32
1	48	$9.328e-02$	$3.24e-01$	0.32	$5.067e-01$	$1.81e+00$	0.33
2	192	$5.972e-02$	$3.36e-01$	0.33	$3.211e-01$	$1.83e+00$	0.33
3	768	$3.788e-02$	$3.42e-01$	0.33	$2.028e-01$	$1.85e+00$	0.33
4	3072	$2.393e-02$	$3.46e-01$	0.33	$1.279e-01$	$1.85e+00$	0.33
5	12288	$1.509e-02$			$8.058e-02$		

Tabelle 4.5: Der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm und der Fehlerschätzer  $\eta$  in Beispiel 4.2.3 unter Vernachlässigung der Konstanten in der Fehlerabschätzung sowie der Terme höherer Ordnung bei uniformer Verfeinerung mit den Konstanten  $\beta$  und  $\alpha$ . Wir erhalten also eine Konvergenzordnung  $\mathcal{O}(h^{2/3})$  ( $\alpha = 1/3$ ).

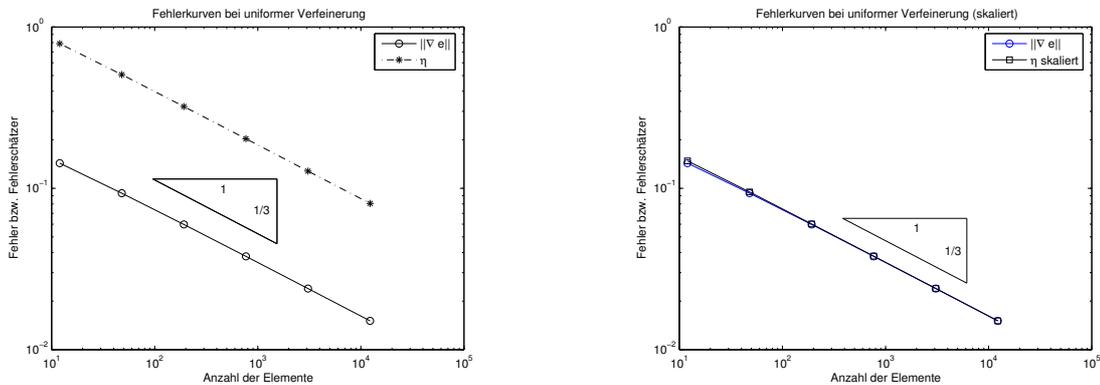


Abbildung 4.17: Der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm und Fehlerschätzer  $\eta$  bei uniformer Verfeinerung in Beispiel 4.2.3: Links stehen absolute Werte in Abhängigkeit der Anzahl der Elemente in  $\mathcal{T} N = 12, 48, \dots, 12288$  unter Vernachlässigung der Konstanten in der Fehlerabschätzung sowie der Terme höherer Ordnung. Im rechten Bild wird der Fehlerschätzer so skaliert, dass für  $N = 15460$  die Werte von  $e$  und  $\eta$  übereinstimmen.

$k$	$N$	Fehler			Fehlerabschätzung		
		$\ \nabla e\ _{L^2(\Omega)}$	$\beta$	$\alpha$	$\eta$	$\beta$	$\alpha$
0	12	1.431e-01			7.891e-01		
2	30	6.742e-02	2.35e+00	1.04	3.409e-01	1.83e+01	1.17
4	48	3.958e-02	2.35e+00	1.06	1.708e-01	5.44e+01	1.49
6	75	2.955e-02	4.12e-01	0.61	1.115e-01	3.41e+00	0.79
8	144	1.457e-02	1.05e+00	0.86	5.932e-02	6.34e+00	0.94
10	285	1.068e-02	2.28e-01	0.54	4.217e-02	9.13e-01	0.54
12	519	5.931e-03	1.40e+01	1.24	2.356e-02	3.19e+01	1.15
14	825	4.514e-03	4.45e-01	0.68	1.780e-02	1.60e+00	0.67
16	1137	3.580e-03	1.51e+00	0.86	1.427e-02	3.73e+00	0.79
18	1452	2.573e-03	1.47e+04	2.14	1.054e-02	1.03e+04	1.89
20	1893	2.179e-03	4.21e-01	0.70	8.690e-03	3.61e+00	0.80
22	2208	1.944e-03	2.17e-02	0.31	7.744e-03	2.06e-01	0.43
24	2655	1.859e-03	3.84e-03	0.09	7.328e-03	4.62e-02	0.23
26	3138	1.649e-03	2.04e+00	0.88	6.597e-03	3.72e+00	0.79
28	3813	1.522e-03	5.05e-03	0.15	6.044e-03	2.83e-02	0.19
30	4608	1.312e-03	2.90e-01	0.64	5.184e-03	2.70e+00	0.74
32	4971	1.016e-03	1.83e+05	2.23	4.223e-03	1.27e+05	2.02
34	5685	9.951e-04	1.37e-03	0.04	4.024e-03	3.69e-02	0.26
36	6519	9.161e-04	2.59e-01	0.64	3.670e-03	1.38e+00	0.68
38	7323	8.244e-04	1.78e+00	0.86	3.293e-03	7.87e+00	0.87
40	8169	7.775e-04	5.77e-02	0.48	3.095e-03	2.98e-01	0.51
42	9153	7.172e-04	7.40e-02	0.51	2.875e-03	1.02e-01	0.39
44	10452	6.922e-04	2.16e-03	0.12	2.732e-03	2.60e-02	0.24
46	11661	6.348e-04	4.05e+01	1.18	2.520e-03	7.55e+01	1.10
48	13062	6.237e-04	3.27e-03	0.17	2.453e-03	2.94e-02	0.26
49	13791	6.055e-04	1.08e-01	0.54	2.381e-03	4.48e-01	0.55

Tabelle 4.6: Der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm und der Fehlerschätzer  $\eta$  in Beispiel 4.2.3 unter Vernachlässigung der Konstanten in der Fehlerabschätzung sowie der Terme höherer Ordnung bei adaptiver Verfeinerung mit den Konstanten  $\beta$  und  $\alpha$ . Wir erhalten also eine Konvergenzordnung  $\mathcal{O}(h^{3/2})$  ( $\alpha = 3/4$ ).

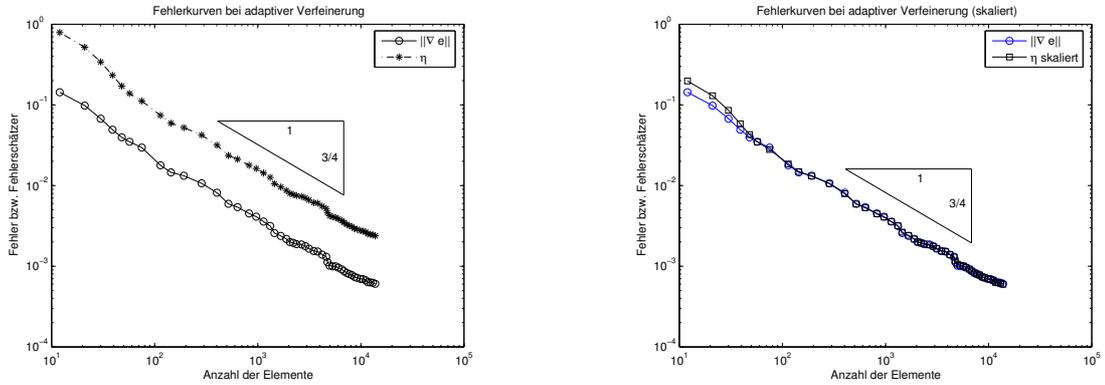


Abbildung 4.18: Der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm und Fehlerschätzer  $\eta$  bei adaptiver Verfeinerung in Beispiel 4.2.3: Links stehen absolute Werte in Abhängigkeit der Anzahl der Elemente in  $\mathcal{T}$   $N = 12, 21, \dots, 13791$  unter Vernachlässigung der Konstanten in der Fehlerabschätzung sowie der Terme höherer Ordnung. Im rechten Bild wird der Fehlerschätzer so skaliert, dass für  $N = 13791$  die Werte von  $e$  und  $\eta$  übereinstimmen.

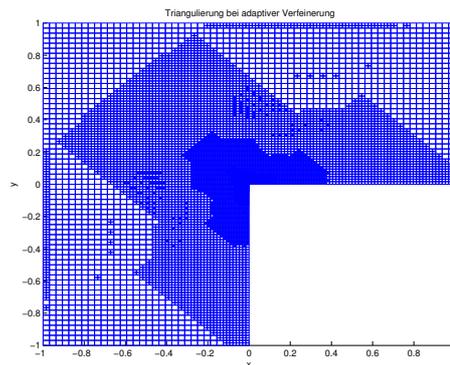
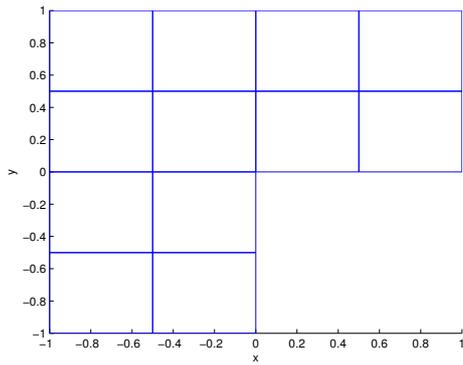
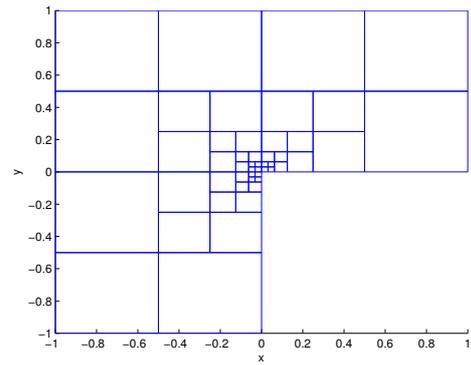


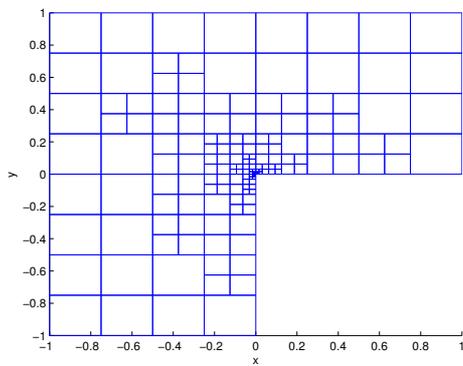
Abbildung 4.19: Die Triangulierung nach  $k = 49$  adaptiven Verfeinerungen in Beispiel 4.2.3.



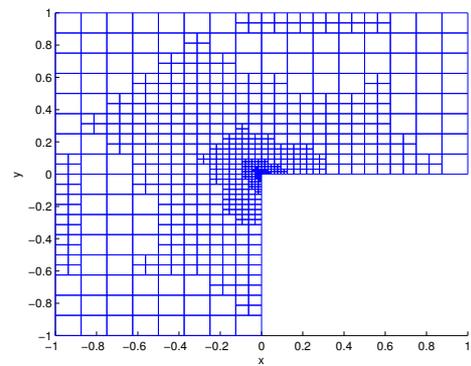
(a)  $k = 0$



(b)  $k = 4$



(c)  $k = 10$



(d)  $k = 14$

Abbildung 4.20: Verschiedene Triangulierungen nach  $k$  adaptiven Verfeinerungen in Beispiel 4.2.3. Es wird dabei zur Singularität hin verfeinert.

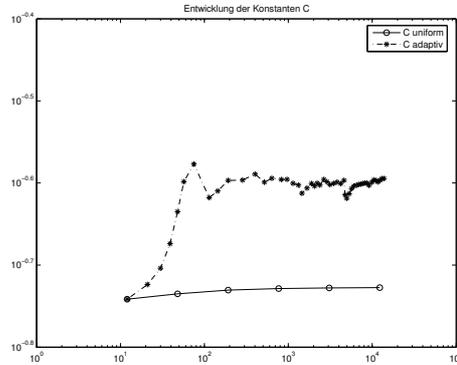


Abbildung 4.21: Das Verhältnis  $\|\nabla e\|_{L^2(\Omega)}/\eta$  konvergiert bei uniformer und adaptiver Verfeinerung in Beispiel 4.2.3 gegen die Zuverlässigkeitskonstante  $C$ .

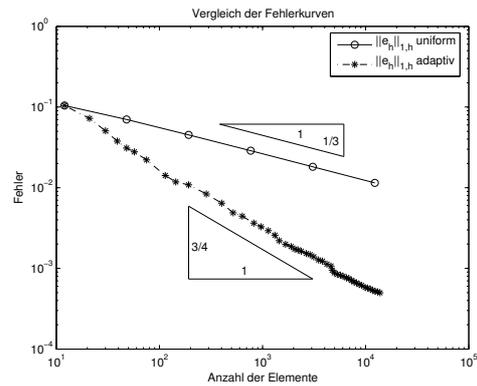
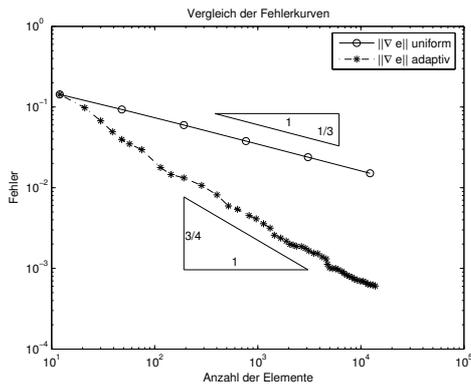


Abbildung 4.22: Fehlervergleich bei uniformer und adaptiver Verfeinerung in Beispiel 4.2.3: Links der Fehler  $e = u - I_M u_h$  in der  $H^1$ -Halbnorm, rechts der diskrete Fehler  $e_h = u_T - u_h$  in der diskreten  $H^1$ -Norm.

# Symbolverzeichnis

## Allgemeine Notationen

---

$\mathbb{N}$	natürliche Zahlen ohne 0
$\mathbb{N}_0$	natürliche Zahlen vereinigt mit 0
$\mathbb{R}$	reelle Zahlen
$\mathbb{R}^d$	Euklidischer Raum von (Spalten-) Vektoren in $d$ Komponenten
$\mathbb{R}^{m \times d}$	Raum der $m \times d$ -Matrizen mit reellen Einträgen
$ \cdot $	Euklidische Länge eines Vektors bzw. Betrag einer reellen oder komplexen Zahl
$\mathbf{x} \cdot \mathbf{y}$	Euklidisches Skalarprodukt zweier Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$
$\langle \cdot, \cdot \rangle$	Skalarprodukt in $L^2$
$\delta_{ij}$	Kronecker-Symbol
$\phi$	affine Abbildung, $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$
$\partial A$	Rand der Menge $A$
$\text{int } A$	Inneres der Menge $A$
$\overline{A}$	Abschluss der Menge $A$
$\text{conv}(\cdot)$	konvexe Hülle
$X^*$	Dualraum des normierten Raums $X$
$C^m$	Raum aller stetigen Funktionen, die stetige (partielle) Ableitungen bis zur Ordnung $m$ besitzen
$\ell_p$	Raum aller (reellen) Folgen, die in $p$ -ter Potenz summierbar sind
$L^p$	Lebesguescher Raum aller Lebesgue-messbaren Funktionen, die in $p$ -ter Potenz integrierbar sind
$W^{1,p}$	Sobolev Raum aller Funktionen in $L^p$ , deren schwache Ableitung in $L^p$ liegt
$H^1$	$H^1 = W^{1,2}$
$W^{m,p}$	Sobolev-Raum $m$ -ter Ordnung, d.h. alle partielle Ableitungen bis zur Ordnung $m$ existieren im schwachen Sinn und liegen in $L^p$
$H^m$	$H^m = W^{m,2}$
$H^{1/2}$	Raum der Spuren von $H^1$ -Funktionen
$\mathcal{P}_m$	Raum der (algebraischen) Polynome vom totalen Grad $\leq m$
$\mathcal{Q}_m$	Raum der (algebraischen) Polynome vom partiellen Grad $\leq m$

---

**Notationen für Modellproblem und Differentialoperatoren**


---

$\Omega$	beschränktes Lipschitz-Gebiet im $\mathbb{R}^2$
$\Gamma$	Rand des Gebietes $\Omega$ , $\Gamma = \partial\Omega$
$\Gamma_D$	Dirichlet-Rand des Gebietes $\Omega$
$\Gamma_N$	Neumann-Rand des Gebietes $\Omega$
$\nabla g$	Gradient als Zeilenvektor, $\nabla g = (\partial g/\partial x_1, \partial g/\partial x_2)$ in $\mathbb{R}^2$
$u$	exakte Lösung, $u \in H^1(\Omega)$
$u_D$	Dirichlet-Daten auf $\Gamma_D$
$f$	rechte Seite der Modellgleichung, $f \in L^2(\Omega)$
$\text{curl } g$	Curl-Operator als Zeilenvektor, $\text{curl } g = (\partial g/\partial x_2, -\partial g/\partial x_1)$ in $\mathbb{R}^2$
$\Delta g$	Laplace-Operator, $\Delta g = \partial^2 g/\partial x_1^2 + \partial^2 g/\partial x_2^2$ in $\mathbb{R}^2$
$\text{div } g$	Divergenz, $\text{div } g = \partial g/\partial x_1 + \partial g/\partial x_2$ in $\mathbb{R}^2$
$H_D^1$	Raum aller Funktionen in $H^1$ , die auf dem Dirichlet-Rand $\Gamma_D$ verschwinden
$H_0^1$	$H_0^1 = H_D^1$ für $\Gamma_D = \Gamma$
$H(\text{div}, \cdot)$	$H(\text{div}, \cdot) := \{v \in L^2 : \text{div } v \in L^2\}$

---

**Notationen für Netzverwaltung**


---

$\mathcal{T}$	fastreguläre Triangulierung des Gebietes $\Omega$
$N$	Anzahl der Elemente, $N := \#\mathcal{T}$
$\mathcal{N}$	alle Knoten von $\mathcal{T}$
$\mathcal{N}_T$	Ecken von $T \in \mathcal{T}$
$\mathcal{N}_\Gamma$	$\mathcal{N}_\Gamma := \{a \in \mathcal{N}   a \in \Gamma\}$
$\mathcal{N}^h$	Menge aller hängenden Knoten
$\mathcal{N}^*$	$\mathcal{N}^* := \mathcal{N} \setminus \{\mathcal{N}^h \cup \mathcal{N}_\Gamma\}$
$\mathcal{E}$	alle Kanten von $\mathcal{T}$
$\mathcal{E}_T$	$\mathcal{E}_T := \{E \in \mathcal{E}   E \subseteq \partial T \text{ und } \forall E' \in \mathcal{E}, E' \subseteq \partial T : E \subseteq E' \Rightarrow E = E'\}$
$\mathcal{E}_\Gamma$	$\mathcal{E}_\Gamma := \{E \in \mathcal{E}   E \subseteq \Gamma\}$
$\mathcal{E}^h$	Menge aller Kanten, auf denen ein hängender Knoten liegt
$\mathcal{E}^*$	$\mathcal{E}^* := \mathcal{E} \setminus \{\mathcal{E}^h \cup \mathcal{E}_\Gamma\}$
$\mathcal{E}^\circ$	$\mathcal{E}^\circ := \{E \in \mathcal{E} \setminus \mathcal{E}_\Gamma   \nexists E' \in \mathcal{E} \setminus \mathcal{E}_\Gamma : E \subset E'\}$
$T$	Element von $\mathcal{T}$
$E$	Kante
$a$	Knoten der Triangulierung $\mathcal{T}$ , $a \in \mathcal{N}$
$x$	Punkt in $\Omega$
$\omega_a$	Knotenpatch für $a \in \mathcal{N}$ , $\omega_a := \bigcup \{T \in \mathcal{T}   a \subseteq \partial T\}$
$\tilde{\omega}_a$	$\tilde{\omega}_a := \{T \in \mathcal{T}   T \subseteq \omega_a\}$
$\omega_E$	Kantenpatch, $\omega_E := \bigcup \{T \in \mathcal{T}   E \subseteq \partial T\}$ für $E \in \mathcal{E} \setminus \mathcal{E}^h$ und $\omega_E := \bigcup_{i=1}^n \omega_{E_i}$ mit $E = \bigcup_{i=1}^n E_i$ , $E_i \in \mathcal{E}^*$ für $E \in \mathcal{E}^h$
$\tilde{\omega}_E$	$\tilde{\omega}_E := \{T \in \mathcal{T}   T \subseteq \omega_E\}$
$\omega_E^*$	elementarer Kantenpatch
$\tilde{\omega}_E^*$	$\tilde{\omega}_E^* := \{T \in \mathcal{T}   T \subseteq \omega_E^*\}$
$\omega_T$	Elementpatch von $T \in \mathcal{T}$ , $\omega_T := \bigcup \{T' \in \mathcal{T}   T \cap T' \in \mathcal{E}\}$
$\tilde{\omega}_T$	$\tilde{\omega}_T := \{T' \in \mathcal{T}   T \subseteq \omega_{T'}\}$
$h_T$	$h_T := \text{diam}(T) := \sup_{x,y \in T}  x - y $
$h$	$h := \max_{T \in \mathcal{T}} h_T$

---

**Notationen für Netzverwaltung (Fortsetzung)**


---

$h_E$	Länge der Kante $E$
$\varrho_T$	die Höhe von $T$
$\kappa_T$	$\kappa_T := h_T/\varrho_T$
$\kappa_{\mathcal{T}}$	formreguläre Konstante, $\kappa_{\mathcal{T}} := \max_{T \in \mathcal{T}_h} \kappa_T$
$\mathbf{n}_E$	a priori festgelegter Normalvektor auf der Kante $E \in \mathcal{E}^* \cup \mathcal{E}_{\Gamma}$
$d_E$	$(x_E - x_W) \cdot \mathbf{n}_E$ , wobei $E = T_E \cap T_W$
$\mathbf{t}_E$	Tangentialvektor einer Kante $E$ orthogonal zu $\mathbf{n}_E$ (mathematisch positiv)
$\sigma_{T,E}$	$\sigma_{T,E} \pm 1$ , abhängig von der Richtung des a priori festgelegten Normalvektors auf $E$
$\mathbf{n}_{T,E}$	nach außen gerichteter Normalvektor des Elements $T$ auf der Kante $E$
$\mathbf{t}_{T,E}$	Tangentialvektor der Kante $E$ orthogonal zu $\mathbf{n}_{T,E}$ (mathematisch positiv)
$g_{\mathcal{T}}$	das $\mathcal{T}$ -stückweise Integralmittel, d.h. $g_{\mathcal{T} T} := g_T$ mit $T \in \mathcal{T}$
$\mathcal{P}_0(\mathcal{T})$	$\mathcal{T}$ -stückweise konstante Funktionen

---

**Notationen für Finite Volumen Methode und Fehlerschätzung**


---

$\Phi_E^C(u)$	konvektiver Fluss über Kante $E$
$\Phi_E^D(u)$	diffusiver Fluss über Kante $E$
$u_h$	diskrete Finite Volumen Lösung
$u_T$	$u_T = u_h _T$
$F_E^C(u_h)$	konvektiver numerischer Fluss über Kante $E$
$F_E^D(u_h)$	diffusiver numerischer Fluss über Kante $E$
$\psi_T(a)$	Interpolationsgewicht von $T$ bzgl. des Knotens $a$
$I_M u_h$	Morley-Interpolant
$[\![ \cdot ]\!]_E$	Kantensprung über die Kante $E$
$b_T$	Bubble-Funktion auf dem Element $T$
$b_E$	Bubble-Funktion auf der Kante $E$
$F_{ext}$	Extension-Operator
$e$	Fehler, $e := u - I_M u_h$
$R$	Residuum, $R := f + \Delta I_M u_h$
$\eta_T$	Verfeinerungsindikatoren
$\eta$	Fehlerschätzer, $\eta := (\sum_{T \in \mathcal{T}} \eta_T^2)^{1/2}$
$e_h$	diskreter Fehler, $e_h := u_{\mathcal{T}} - u_h$
$\  \cdot \ _{1,h}$	diskrete $H^1$ -Norm
$\mathcal{O}(\cdot)$	Konvergenzordnung

# Literaturverzeichnis

- [1] Mark Ainsworth und J. Tinsley Oden. *A Posteriori Error Estimation in Finite Element Analysis*. John Wiley & Sons, erste Edition, 2000.
- [2] Mario Bebendorf. A Note on the Poincaré Inequality for Convex Domains. *Zeitschrift für Analysis und ihre Anwendungen*, 22(04):751–756, 2003.
- [3] Dietrich Braess. *Finite Elemente*. Springer, Berlin, zweite Edition, 1997.
- [4] Susanne C. Brenner und L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods*. Springer, New-York, 1994.
- [5] Carsten Carstensen. Adaptive Finite Elemente Methode. Skript zur Vorlesung an der TU-Wien im Wintersemester 01/02, 2001.
- [6] Carsten Carstensen. Finite Elemente. Skript zur Vorlesung an der CAU Kiel, 2003.
- [7] William John Coirier. *An Adaptively-Refined, Cartesian, Cell-Based Scheme for the Euler and Navier-Stokes Equations*. PhD thesis, The University of Michigan, 1994.
- [8] Yves Coudière, Jean-Paul Vila und Philippe Villedieu. Convergence rate of a finite volume scheme for a two dimensional convection-diffusion problem. *ESAIM: M2AN*, 33(3):493–516, 2000.
- [9] Yves Coudière und Philippe Villedieu. Convergence rate of a finite volume scheme for the linear convection-diffusion equation on locally refined meshes. *ESAIM: M2AN*, 34(6):1123–1149, 2000.
- [10] Robert Eymard, Thierry Gallouët und Raphaële Herbin. Finite Volume Methods. In *Handbook of Numerical Analysis*, Volume 7. Elsevier Science B.V., erste Edition, 1999.
- [11] Vivette Girault und Pierre-Arnaud Raviart. *Finite Element Methods for Navier-Stokes Equations*. Springer, Berlin, 1986.
- [12] Peter Knabner und Lutz Angermann. *Numerik partieller Differentialgleichungen*. Springer, Berlin und Heidelberg, erste Edition, 2000.
- [13] Serge Nicaise. A posteriori error estimations of some cell-centered finite volume methods. Preprint 2003.
- [14] Robert Plato. *Numerische Mathematik kompakt*. Vieweg, Braunschweig/Wiesbaden, erste Edition, 2000.
- [15] Dirk Praetorius. *Analysis, Numerik und Simulation eines relaxierten Modellproblems zum Mikromagnetismus*. PhD thesis, Technische Universität Wien, 2003.

- [16] Reinhold Remmert. *Funktionentheorie 1*. Springer, Berlin, dritte Edition, 1992.
- [17] Robert Schaback und Helmut Werner. *Numerische Mathematik*. Springer, Berlin, 1993.
- [18] Michael Schäfer. *Numerik im Maschinenbau*. Springer, Berlin, erste Edition, 1999.
- [19] Rüdiger Verfürth. A posteriori error estimation and adaptive mesh-refinement techniques. *Journal of Computational and Applied Mathematics*, 50:67–83, 1994.
- [20] Rüdiger Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, 1996.