

Kapitel 6

Fehlerschätzung und Adaptivität

Ziel von Adaptivität:

- *Effizienz:* Es soll mit möglichst geringem Aufwand eine Approximation u_{FE} an die gesuchte Lösung u gefunden werden, die eine vorgegebene Genauigkeitsanforderung erfüllt. Ein (grobes) Maß für den Aufwand ist die Anzahl Elemente des Gitters, welches u_{FE} festlegt. Adaptive Algorithmen konstruieren Gitter, die diesem Ideal nahekommen.
- *Robustheit:* Die Diskussion im vorangehenden Kapitel hat gezeigt hat, daß mit sehr präzisiertem Wissen über das Verhalten der gesuchten Lösung es möglich ist, gute Gitter zu konstruieren. Dies setzt jedoch einen sehr gut geschulten “Benutzer” eines FEM-Programms voraus. Zudem ist ein solches Vorgehen risikobehaftet, weil der “Benutzer” das richtige Gitter konstruiert haben muß¹. Eine automatische Konstruktion von Gittern kann helfen, Fehler zu vermeiden.

Adaptive Algorithmen basieren auf Fehlerschätzern, die die Gitterverfeinerung steuern.

Ziel von Fehlerschätzern:

- *Fehlerschätzer als Abbruchkriterium:* Schätzen des Fehlers $u - u_{FE}$, um zu entscheiden, ob eine Approximation u_{FE} hinreichend genau ist
- *Fehlerschätzer als Steuerungsmechanismen:* Schätzen des Fehlers $(u - u_{FE})|_K$ für ein Element K , um zu entscheiden, ob dieses Element verfeinert werden muß.

Wir führen die Herleitung eines *a posteriori* (“im Nachhinein”) am folgenden Modellproblem vor:

$$-\Delta u = f \quad \text{auf Polygon } \Omega \subset \mathbb{R}^2, \quad (6.1a)$$

$$u = 0 \quad \text{auf } \Gamma_D, \quad \partial_n u = g \quad \text{auf } \Gamma_N. \quad (6.1b)$$

Hier sind Γ_D, Γ_N wie in Abschnitt 4.2. Wir fordern der Einfachheit halber $|\Gamma_D| > 0; \Gamma_N = \emptyset$ ist jedoch zugelassen. Weiter fordern wir $f \in L^2(\Omega), g \in L^2(\Gamma_N)$. Die schwache Formulierung ist:

$$\text{Finde } u \in H_0^1(\Omega, \Gamma_D) \text{ s.d. } B(u, v) := \int_{\Omega} \nabla u \cdot \nabla v = l(v) := \int_{\Omega} f v + \int_{\Gamma_N} g v \quad \forall v \in H_0^1(\Omega, \Gamma_D), \quad (6.2)$$

wobei wir wiederum $H_0^1(\Omega, \Gamma_D) := \{u \in H^1(\Omega) \mid (\gamma_0 u)|_{\Gamma_D} = 0\}$ gesetzt haben.

6.1 Notation und der Clémentinterpolant

Wir betrachten nur reguläre, affine, γ -formreguläre Triangulierungen und verwenden folgende Notation:

- $\mathcal{E}(K) :=$ Menge der 3 Kanten, von $K \in \mathcal{T}$
- $\mathcal{N}(K) :=$ Menge der 3 Knoten von $K \in \mathcal{T}$

¹das Sinken der Bohrrinsel Sleipner A im Jahr 1991 mit einem Schaden von fast 1 Milliarde US-Dollar ging letztendlich auf ein falsch ausgelegtes FEM-Gitter zurück

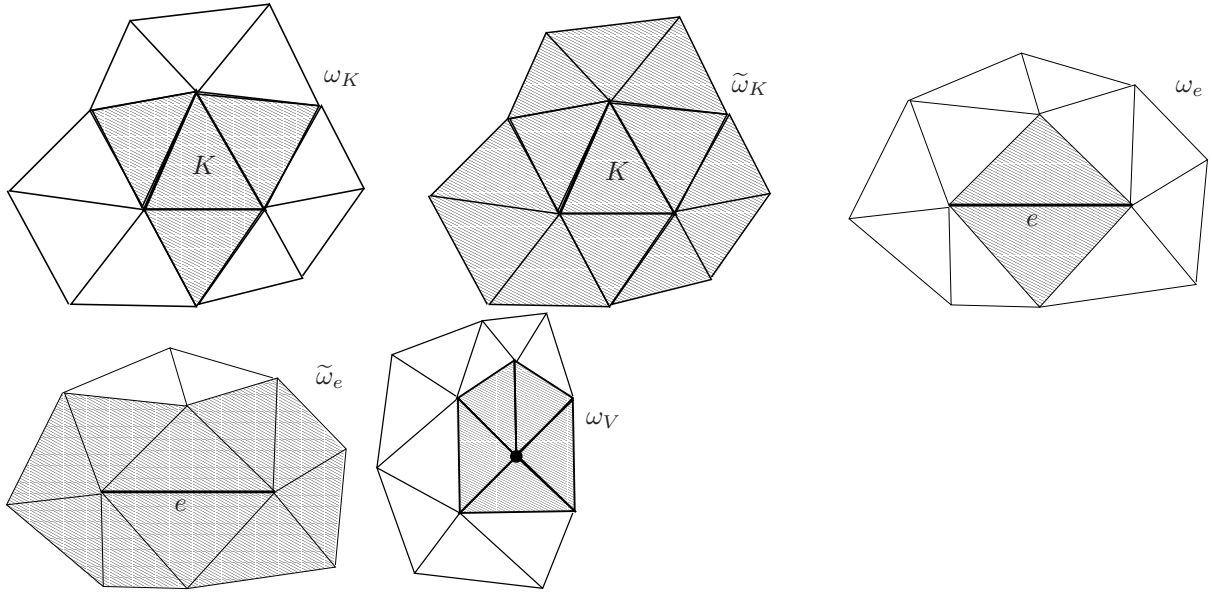


Abbildung 6.1: Gebiete $\omega_K, \tilde{\omega}_K, \omega_e, \tilde{\omega}_e, \omega_V$

- $\mathcal{E}(\mathcal{T}) = \cup_{K \in \mathcal{T}} \{e \in \mathcal{E}(K) \mid e \subset \Omega\}$ = Menge der *inneren* Kanten
- $\mathcal{E}_N(\mathcal{T}) = \cup \{e \in \mathcal{E}(K) \mid e \subset \Gamma_N\}$ = Menge der Kanten, die auf Γ_N liegen
- $\mathcal{N}(\mathcal{T})$ = Menge der Knoten von \mathcal{T}
- $\mathcal{N}(e)$ = Menge der 2 Endpunkte der Kante e

Weiter definieren wir

$$\omega_K := \left(\bigcup_{\substack{K' \in \mathcal{T} \\ \mathcal{E}(K) \cap \mathcal{E}(K')}} \overline{K'} \right)^\circ, \quad \tilde{\omega}_K := \left(\bigcup_{\substack{K' \in \mathcal{T} \\ \mathcal{N}(K) \cap \mathcal{N}(K')}} \overline{K'} \right)^\circ$$

$$\omega_e := \left(\bigcup_{\substack{K \in \mathcal{T} \\ e \in \mathcal{E}(K)}} \overline{K} \right)^\circ, \quad \tilde{\omega}_e := \left(\bigcup_{\substack{K \in \mathcal{T} \\ \mathcal{N}(e) \cap \mathcal{N}(K) \neq \emptyset}} \overline{K} \right)^\circ, \quad \omega_V := \left(\bigcup_{\substack{K \in \mathcal{T} \\ V \in \mathcal{N}(K)}} \overline{K} \right)^\circ$$

Im folgenden betrachten wir nun Triangulierungen, die zu den Randbedingungen “passen” in dem Sinn, daß die folgende Annahme erfüllt ist:

Voraussetzung 6.1 Die Triangulation \mathcal{T} ist regulär, affin, γ -formregulär und $\overline{\Gamma_D} \cap \overline{\Gamma_N} \subset \mathcal{N}(\mathcal{T})$. M.a.W.: eine Kante e mit $e \subset \partial\Omega$ erfüllt entweder $e \subset \Gamma_D$ oder $e \subset \Gamma_N$.

Bei γ -formregulären Triangulierungen sind die Elementgrößen benachbarter Elemente vergleichbar:

Lemma 6.2 Sei \mathcal{T} reguläre, affine, γ -formreguläre Triangulierung eines Polygons. Dann existiert $C > 0$, welches nur von γ abhängt, so daß

$$\begin{aligned} \text{diam } \omega_K &\leq \text{diam } \tilde{\omega}_K \leq Ch_K & \forall K \in \mathcal{T} \\ \text{diam } \omega_e &\leq \text{diam } \tilde{\omega}_e \leq Ch_e & \forall e \in \mathcal{E}(\mathcal{T}), \quad h_e := \text{diam } e \\ \text{diam } \omega_V &\leq Ch_K & \forall (V, K) \in \mathcal{N}(\mathcal{T}) \times \mathcal{T} \quad \text{mit } V \in \mathcal{N}(K) \\ \#\{K' \mid K' \subset \tilde{\omega}_K\} &\leq C & \forall K \in \mathcal{T} \\ \#\{K \mid K \subset \tilde{\omega}_e\} &\leq C & \forall e \in \mathcal{E}(\mathcal{T}) \end{aligned}$$

Beweis: Elementar. Wesentlich ist, daß die Formregularitätskonstante γ die Innenwinkel der Dreiecke kontrolliert, so daß

$$\min_{e \in \mathcal{E}(K)} h_e \sim h_K \sim \max_{e \in \mathcal{E}(K)} h_e$$

gilt. □

Wir kommen nun zum Clémentinterpolanten:

Satz 6.3 *Gelte Voraussetzung 6.1. Dann existiert ein $C > 0$, welches nur von γ und Γ_D abhängt und ein linearer Operator $I^C : H_0^1(\Omega, \Gamma_D) \rightarrow S^{1,1}(\mathcal{T}) \cap H_0^1(\Omega, \Gamma_D)$ derart, daß*

$$\begin{aligned} \|u - I^C u\|_{L^2(K)} &\leq Ch_K \|\nabla u\|_{L^2(\tilde{\omega}_K)} & \forall K \in \mathcal{T} \\ \|\nabla(u - I^C u)\|_{L^2(K)} &\leq Ch_K \|\nabla u\|_{L^2(\tilde{\omega}_K)} & \forall K \in \mathcal{T} \\ \|u - I^C u\|_{L^2(e)} &\leq C\sqrt{h_e} \|\nabla u\|_{L^2(\tilde{\omega}_e)} & \forall e \in \mathcal{E}(\mathcal{T}) \cup \mathcal{E}_N(\mathcal{T}) \end{aligned}$$

Beweis: Ein vollständiger Beweis befindet sich im Appendix. Wir illustrieren die Kernidee für den einfachen Fall, daß $\Gamma_D = \emptyset$ und ein regelmäßiges Gitter vorliegt. Wir definieren den Operator $I : H^1(\Omega) \rightarrow S^{1,1}(\mathcal{T})$ dadurch, daß für jeden Knoten $V \in \mathcal{N}(\mathcal{T})$ der Knotenwert $(Iu)(V) := u_V$ gesetzt wird mit

$$u_V := \frac{1}{|\omega_V|} \int_{\omega_V} u.$$

Eine andere Darstellung von Iu ist damit $Iu = \sum_{V \in \mathcal{N}(\mathcal{T})} u_V \varphi_V$, wobei $\varphi_V \in S^{1,1}(\mathcal{T})$ die Hutfunktion ist, die zum Knoten $V \in \mathcal{N}(\mathcal{T})$ gehört. Wir nutzen nun aus, daß für jedes Dreieck $K \in \mathcal{T}$ gilt: $\sum_{V \in \mathcal{N}(K)} \varphi_V \equiv 1$ auf K . Damit erhalten wir unter Berücksichtigung von $\|\varphi_V\|_{L^\infty(K)} \leq 1$:

$$\|u - Iu\|_{L^2(K)} = \left\| \sum_{V \in \mathcal{N}(K)} \varphi_V (u - u_V) \right\|_{L^2(K)} \leq \sum_{V \in \mathcal{N}(K)} \|u - u_V\|_{L^2(K)}$$

Nun ist für $V \in \mathcal{N}(K)$:

$$\|u - u_V\|_{L^2(K)} \leq \|u - u_V\|_{L^2(\omega_V)} \leq Ch_V \|\nabla u\|_{L^2(\omega_V)},$$

wobei $h_V = \text{diam } \omega_V$ und $C > 0$ hängt nur von der *Form* (nicht aber der Größe!) von ω_V ab. Diese letzte Abschätzung ergibt aus der 2. Poincaréschen Ungleichung (Satz 3.29) und einem Skalierungsargument: Man definiert eine affine Abbildung F_V als Verkettung einer Translation, einer Rotation und einer Streckung um $1/h_V$ derart, daß $\omega_V = F_V(\tilde{\omega}_V)$, wobei $\text{diam } \tilde{\omega}_V = 1$. Anschließend wendet man auf $\tilde{\omega}_V$ die 2. Poincarésche Ungleichung auf $\hat{u} := u \circ F_V$ an. Rücktransformation auf ω_V liefert dann die Behauptung. Die Konstante hängt nur von $\tilde{\omega}_V$ ab. Der Beweis wird nun abgeschlossen durch die Beobachtung, daß man bei regelmäßigen Gittern mit endlich vielen ‘Referenzpatches’ auskommt.

Für die Abschätzung für ∇Iu gehen wir analog vor. Auf Element K haben wir

$$\|\nabla(u - Iu)\|_{L^2(K)} = \left\| \sum_{V \in \mathcal{N}(K)} \nabla \varphi_V (u - u_V) \right\|_{L^2(K)} = \left\| \sum_{V \in \mathcal{N}(K)} (\nabla \varphi_V)(u - u_V) + \varphi_V \nabla(u - u_V) \right\|_{L^2(K)}.$$

Aus der Formregularität des Gitters folgt $\|\nabla \varphi_V\|_{L^\infty(K)} \leq Ch_V^{-1}$ für alle $V \in \mathcal{N}(K)$. Damit ergibt sich

$$\|\nabla(u - Iu)\|_{L^2(K)} \leq Ch_V^{-1} \|u - u_V\|_{L^2(K)} + C \|\nabla u\|_{L^2(K)},$$

und der erste Term wurde bereits oben abgeschätzt.

Für die Abschätzung für $\|u - Iu\|_{L^2(e)}$ für eine Kante e gehen wir wie folgt vor. Sei $K \in \mathcal{T}$ ein Dreieck der Triangulierung, welches e als Kante hat. Sei F_K die Elementabbildung für K . Dann gilt nach dem Spursatz (Satz 3.31) auf \hat{K} für den pull-back $\widehat{u - Iu}$:

$$\|\widehat{u - Iu}\|_{L^2(F_K^{-1}(e))} \leq C \|\widehat{u - Iu}\|_{H^1(\hat{K})}.$$

Transformiert man wieder auf e und K zurück, so ergibt sich

$$\|u - Iu\|_{L^2(e)} \leq Ch_e^{1/2} \|\widehat{u - Iu}\|_{L^2(F_K^{-1}(e))} \leq Ch_e^{1/2} \|\widehat{u - Iu}\|_{H^1(\hat{K})} \leq Ch_e^{1/2} (h_K^{-1} \|u - Iu\|_{L^2(K)} + |u - Iu|_{H^1(K)})$$

Nutzt man nun die Resultate für $\|u - Iu\|_{L^2(K)}$ und $|u - Iu|_{H^1(K)}$ aus, so ergibt sich die Behauptung. □

6.2 Ein residualer Fehlerschätzer

Sei u Lösung von (6.1) und sei $V_N \subset S^{1,1}(\mathcal{T}) \cap H_0^1(\Omega, \Gamma_D)$ und $u_N \in V_N$ Lösung von

$$\text{Finde } u_N \in V_N \text{ s.d. } \quad B(u_N, v) = l(v) \quad \forall v \in V_N. \quad (6.3)$$

Dann gilt:

$$B(u - u_N, v) = 0 \quad \forall v \in V_N \quad \text{Galerkinorthogonalität} \quad (6.4)$$

$$B(u - u_N, v) = R(v) := l(v) - B(u_N, v) \quad \forall v \in H_0^1(\Omega, \Gamma_D) \quad \text{Residualgleichung} \quad (6.5)$$

Insbesondere erfüllt das *Residuum* R die Orthogonalitätsbedingung

$$R(v) = 0 \quad \forall v \in V_N. \quad (6.6)$$

Weiter ist wegen der Annahme $|\Gamma_D| > 0$ die Energienorm (gegeben durch $\|v\|_E^2 = B(v, v)$) eine Norm auf $H_0^1(\Omega, \Gamma_D)$ und erfüllt

$$\|v\|_E \leq \|v\|_{H^1(\Omega)} \leq C\|v\|_E. \quad (6.7)$$

Damit können wir die Bestimmung des Fehlers $\|u - u_N\|_E$ umschreiben als ein Problem, die Norm des Residuums R zu berechnen:

$$\|u - u_N\|_E = \sup_{v \in H_0^1(\Omega, \Gamma_D)} \frac{B(u - u_N, v)}{\|v\|_E} \stackrel{(6.5)}{=} \sup_{v \in H_0^1(\Omega, \Gamma_D)} \frac{R(v)}{\|v\|_E}. \quad (6.8)$$

Wir müssen nun dieses Supremum abschätzen. Dies geschieht durch Ausnutzen der Galerkinorthogonalität und der Beobachtung (6.7). Elementweise partielle Integration liefert uns

$$\begin{aligned} R(v) &= l(v) - B(u_N, v) = \sum_K \int_K (fv - \nabla u_N \cdot \nabla v) + \sum_{e \in \mathcal{E}_N(\mathcal{T})} gv \\ &= \sum_K \int_K (f + \Delta u_N)v - \sum_{e \in \mathcal{E}(\mathcal{T})} \int_e [\partial_n u_N]v + \sum_{e \in \mathcal{E}_N(\mathcal{T})} (g - \partial_n u_N)v \end{aligned}$$

wobei der Sprung der Normalenableitung $[\partial_n u_N]$ definiert ist

$$[\partial_n u_N]|_e := n_K \cdot \nabla u_N + n_{K'} \cdot \nabla u_N, \quad (6.9)$$

wobei K, K' die beiden Elemente sind, die sich die Kante e teilen und n_K und $n_{K'}$ die äußeren Normalenvektoren von K bzw. K' sind. Wegen (6.6) gilt $R(v) = R(v - w)$ für alle $v \in H_0^1(\Omega, \Gamma_D)$ und alle $w \in V_N$. Insbesondere folgt damit für den Clémentinterpolanten $w = I^C v$ aus Satz 6.3

$$\begin{aligned} |R(v)| &= |R(v - I^C v)| = \left| \sum_K \int_K (f + \Delta u_N)(v - I^C v) - \sum_{e \in \mathcal{E}(\mathcal{T})} \int_e [\partial_n u_N](v - I^C v) + \sum_{e \in \mathcal{E}_N(\mathcal{T})} \int_e (g - \partial_n u_N)(v - I^C v) \right| \\ &\leq \sum_K \|f + \Delta u_N\|_{L^2(K)} \|v - I^C v\|_{L^2(K)} + \sum_{e \in \mathcal{E}(\mathcal{T})} \|[\partial_n u_N]\|_{L^2(e)} \|v - I^C v\|_{L^2(e)} \\ &\quad + \sum_{e \in \mathcal{E}_N(\mathcal{T})} \|g - \partial_n u_N\|_{L^2(e)} \|v - I^C v\|_{L^2(e)} \\ &\stackrel{\text{Satz 6.3}}{\leq} C \sum_K h_K \|f + \Delta u_N\|_{L^2(K)} \|\nabla v\|_{L^2(\tilde{\omega}_K)} + \sum_{e \in \mathcal{E}} \sqrt{h_e} \|[\partial_n u_N]\|_{L^2(e)} \|\nabla v\|_{L^2(\tilde{\omega}_e)} \\ &\quad + \sum_{e \in \mathcal{E}_N} \sqrt{h_e} \|g - \partial_n u_N\|_{L^2(e)} \|\nabla v\|_{L^2(\tilde{\omega}_e)} \\ &\stackrel{\text{c.s. für Summen}}{\leq} C \sqrt{\sum_K h_K^2 \|f + \Delta u_N\|_{L^2(K)}^2} \sqrt{\sum_K \|\nabla v\|_{L^2(\tilde{\omega}_K)}^2} + C \sqrt{\sum_{e \in \mathcal{E}} h_e \|[\partial_n u_N]\|_{L^2(e)}^2} \sqrt{\sum_{e \in \mathcal{E}} \|\nabla v\|_{L^2(\tilde{\omega}_e)}^2} \\ &\quad + C \sqrt{\sum_{e \in \mathcal{E}_N} h_e \|g - \partial_n u_N\|_{L^2(e)}^2} \sqrt{\sum_{e \in \mathcal{E}_N} \|\nabla v\|_{L^2(\tilde{\omega}_e)}^2} \end{aligned}$$

Wir behaupten nun, daß aufgrund der Formregularität des Gitters gilt:

$$\sum_K \|\nabla v\|_{L^2(\tilde{\omega}_K)}^2 \leq C \|\nabla v\|_{L^2(\Omega)}^2 \quad (6.10)$$

$$\sum_{e \in \mathcal{E} \cup \mathcal{E}_N} \|\nabla v\|_{L^2(\tilde{\omega}_e)}^2 \leq C \|\nabla v\|_{L^2(\Omega)}^2 \quad (6.11)$$

Wir zeigen nur (6.10):

$$\begin{aligned} \sum_K \|\nabla v\|_{L^2(\tilde{\omega}_K)}^2 &= \sum_K \sum_{K': K' \subset \tilde{\omega}_K} \|\nabla v\|_{L^2(K')}^2 = \sum_{K'} \sum_{K: K' \subset \tilde{\omega}_K} \|\nabla v\|_{L^2(K')}^2 = \sum_{K'} \|\nabla v\|_{L^2(K')}^2 \left(\sum_{K: K' \subset \tilde{\omega}_K} 1 \right) \\ &= \sum_{K'} \|\nabla v\|_{L^2(K')}^2 \sum_{\substack{K \in \mathcal{T} \\ \mathcal{N}(K) \cap \mathcal{N}(K') \neq \emptyset}} 1 = \sum_{K'} \|\nabla v\|_{L^2(K')}^2 \sum_{K: K \subset \tilde{\omega}_{K'}} 1 \stackrel{\text{Lemma 6.2}}{\leq} C \|\nabla v\|_{L^2(\Omega)}^2. \end{aligned}$$

Damit erhalten wir

$$\begin{aligned} |R(v)| &\leq C \|\nabla v\|_{L^2(\Omega)} \left(\sqrt{\sum_K h_K^2 \|f + \Delta u_N\|_{L^2(K)}^2} + \sqrt{\sum_{e \in \mathcal{E}} h_e \|[\partial_n u_N]\|_{L^2(e)}^2} + \sqrt{\sum_{e \in \mathcal{E}_N} h_e \|g - \partial_n u_N\|_{L^2(e)}^2} \right) \\ &\leq C \|\nabla v\|_{L^2(\Omega)} \sqrt{\sum_K h_K^2 \|f + \Delta u_N\|_{L^2(K)}^2 + \sum_{e \in \mathcal{E}} h_e \|[\partial_n u_N]\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}_N} h_e \|g - \partial_n u_N\|_{L^2(e)}^2} \end{aligned}$$

Wir fassen nun die Sprungterme etwas anders zusammen, indem wir die Abkürzung

$$\eta_K^2 := h_K^2 \|f + \Delta u_N\|_{L^2(K)}^2 + \frac{1}{2} \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}(\mathcal{T})} h_e \|[\partial_n u_N]\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_N(\mathcal{T})} h_e \|g - \partial_n u_N\|_{L^2(e)}^2 \quad (6.12)$$

eingeführen. Zusammenfassend erhalten wir

$$|R(v)| \leq C \|\nabla v\|_{L^2(\Omega)} \sqrt{\sum_{K \in \mathcal{T}} \eta_K^2} \quad (6.13)$$

Somit ergibt sich aus (6.8)

Satz 6.4 *Es gelte Voraussetzung 6.1. Dann existiert eine Konstante $C > 0$, welche nur von der Formregularitätskonstante γ und Γ_D abhängt, so daß die durch (6.12) definierten Elementfehlerindikatoren η_K für*

$$\text{EST} := \sqrt{\sum_{K \in \mathcal{T}} \eta_K^2} \quad (6.14)$$

die Abschätzung

$$\|\nabla(u - u_N)\|_{L^2(\Omega)} = \|u - u_N\|_E \leq C \text{EST}$$

erfüllen.

Beweis: Setze (6.13) in (6.8) ein und beobachte $\|v\|_E = \|\nabla v\|_{L^2(\Omega)}$. □

finis 24.Stunde

Bemerkung 6.5 • Der Fehlerschätzer EST ist *berechenbar*, denn es gehen nur bekannte Größen (u_N , f , g) ein.

- In der Praxis werden die Elementindikatoren η_K als Indikatoren für den Fehler auf dem Element K verwendet, um zu entscheiden, welche Elemente verfeinert werden sollen.

Allgemein nennt man einen berechenbaren Ausdruck EST einen *zuverlässigen* Fehlerschätzer, falls $\|u - u_N\|_E \leq C \text{EST}$ gilt für eine Konstante $C > 0$, die *nicht* von u_N abhängt. Ein solcher berechenbarer Ausdruck heißt *effizient*, falls $\text{EST} \leq C \|u - u_N\|_E$ für eine von u_N unabhängige Konstante gilt. Offensichtlich ist man an Fehlerschätzern interessiert, die sowohl zuverlässig als auch effizient sind, d.h. den Fehler weder zu sehr überschätzen noch unterschätzen. Der Fehlerschätzer EST aus (6.14) erfüllt diese Forderung im folgenden Sinn:

Satz 6.6 *Es gelte Voraussetzung 6.1. Dann existiert $C > 0$, welches nur von der Formregularitätskonstante γ und Γ_D abhängt, so daß EST aus (6.14) die folgende Abschätzung erfüllt:*

$$\text{EST}^2 \leq C \left\{ \|\nabla(u - u_N)\|_{L^2(\Omega)}^2 + \sum_K h_K^2 \|f - f_{\mathcal{T}}\|_{L^2(K)}^2 + \sum_{e \in \mathcal{E}_N} h_e \|g - g_{\mathcal{T}}\|_{L^2(e)}^2 \right\}. \quad (6.15)$$

Hier sind die stückweise konstanten Funktionen $f_{\mathcal{T}}$, $g_{\mathcal{T}}$ gegeben als Mittelwerte:

$$f_{\mathcal{T}}|_K = \frac{1}{|K|} \int_K f, \quad g_{\mathcal{T}}|_e = \frac{1}{|e|} \int_e g.$$

Es gilt sogar folgende Abschätzung für die Fehlerindikatoren η_K :

$$\eta_K^2 \leq C \left[\|\nabla(u - u_N)\|_{L^2(\omega_K)}^2 + h_K^2 \|f - f_{\mathcal{T}}\|_{L^2(\omega_K)}^2 + \sum_{e \in \mathcal{E}_N \cap \mathcal{E}(K)} h_e \|g - g_{\mathcal{T}}\|_{L^2(e)}^2 \right] \quad (6.16)$$

Beweis von Satz 6.6: Der Beweis von Satz 6.6 ist lang und deshalb in mehrere Teilschritte zerlegt, die als Lemmata 6.7, 6.8, 6.9 formuliert sind. \square

Bevor wir Lemmata 6.7, 6.8, 6.9 formulieren und beweisen, bemerken wir, daß die Beiträge $\sum_K h_K^2 \|f - f_{\mathcal{T}}\|_{L^2(K)}^2$ und $\sum_{e \in \mathcal{E}_N} h_e \|g - g_{\mathcal{T}}\|_{L^2(e)}^2$ “von höherer Ordnung” sind, d.h. man erwartet/hofft, daß diese Beiträge klein sind gegenüber dem Fehler $\|\nabla(u - u_N)\|_{L^2(\Omega)}^2$. Als Motivation für diese Erwartung/Hoffnung betrachten wir den Fall eines quasiuniformen Gitters mit Gitterweite h . Sind f und g stückweise glatt (also z.B. $f|_K \in H^1(K)$ für jedes $K \in \mathcal{T}$ und $\sum_{K \in \mathcal{T}} \|f\|_{H^1(K)}^2 < \infty$ sowie $g|_e \in H^1(e)$ für jedes $e \in \mathcal{E}_N$ mit $\sum_{e \in \mathcal{E}_N} \|g\|_{H^1(e)}^2$), dann sind die Beiträge von $f - f_{\mathcal{T}}$ und $g - g_{\mathcal{T}}$ in (6.15) von der Größe h^3 . Andererseits erwarten wir selbst bei exakter Lösung $u \in H^2(\Omega)$ nur $\|\nabla(u - u_N)\|_{L^2(\Omega)}^2 \sim h^2$.

Lemma 6.7 *Es existiert $C > 0$, welches nur von der Formregularitätskonstante γ abhängt, so daß für alle $K \in \mathcal{T}$*

$$h_K^2 \|f + \Delta u_N\|_{L^2(K)}^2 \leq C \left[\|\nabla(u - u_N)\|_{L^2(K)}^2 + h_K^2 \|f - f_{\mathcal{T}}\|_{L^2(K)}^2 \right].$$

Beweis: Die Kernidee des Beweises ist die Verwendung einer geeigneten Testfunktion in der Variationsformulierung. Diese Testfunktion wird als “Blasenfunktion” gewählt, so daß man sich auf die Untersuchung eines einzelnen Elementes konzentrieren kann.

Wir fixieren ein Element $K \in \mathcal{T}$. Wir definieren eine “Blasenformfunktion” $b_K \in H_0^1(\Omega)$ mit Träger $\text{supp } b_K \subset \overline{K}$ durch $b_K := \widehat{b} \circ F_K^{-1}$ und

$$\widehat{b}(\xi, \eta) := \xi \eta (1 - \xi - \eta)$$

(Hier sei daran erinnert, daß $\widehat{K} = \{(\xi, \eta) \mid 0 < \xi < 1, \quad 0 < \eta < 1 - \xi\}$). Wir behaupten, daß ein $C > 0$ gibt, welches nur von der Formregularitätskonstante γ abhängt, so daß

$$\|\sqrt{b_K}\|_{L^2(K)}^2 = \frac{1}{60} |K| \quad (6.17)$$

$$\|b_K\|_{L^2(K)}^2 = \frac{1}{2520} |K| \quad (6.18)$$

$$\|\nabla b_K\|_{L^2(K)} \leq \frac{C'}{h_K} \|b_K\|_{L^2(K)} \leq C, \quad (6.19)$$

Für die erste Gleichung bemerken wir, daß $|\det F'_K| = 2|K|$ und berechnen dann $\|\sqrt{b_K}\|_{L^2(K)}^2 = \int_K b_K = \int_{\widehat{K}} \widehat{b} |\det F'_K| = 2|K| \int_{\widehat{K}} \widehat{b} = 2/120|K|$. Die Gleichung (6.18) wird analog gezeigt. Ungleichung (6.19) ist eine *inverse Ungleichung*, dessen Beweis als Übung empfohlen ist.

Definiere (für das oben fixierte K) die Funktion

$$v := \underbrace{(f_{\mathcal{T}} + \Delta u_N)}_{\in \mathbb{R}} b_K \in H_0^1(K) \subset H_0^1(\Omega).$$

Dann folgt aus (6.18), (6.19) und der Tatsache, daß $f_{\mathcal{T}} + \Delta u_N \in \mathbb{R}$:

$$\begin{aligned} \|\nabla v\|_{L^2(K)} &= |f_{\mathcal{T}} + \Delta u_N| \|\nabla b_K\|_{L^2(K)} \leq C |f_{\mathcal{T}} + \Delta u_N| \leq C h_K^{-1} \|f_{\mathcal{T}} + \Delta u_N\|_{L^2(K)}, \\ \|v\|_{L^2(K)} &= |f_{\mathcal{T}} + \Delta u_N| \|b_K\|_{L^2(K)} = \frac{1}{\sqrt{2520}} \|f_{\mathcal{T}} + \Delta u_N\|_{L^2(K)}. \end{aligned}$$

Die Dreiecksungleichung liefert

$$h_K^2 \|f + \Delta u_N\|_{L^2(K)}^2 \leq 2h_K^2 \|f - f_{\mathcal{T}}\|_{L^2(K)}^2 + 2h_K^2 \|f_{\mathcal{T}} + \Delta u_N\|_{L^2(K)}^2.$$

Weiter schätzen wir ab wegen $(f_{\mathcal{T}} + \Delta u_N)|_K \in \mathbb{R}$

$$\begin{aligned} \|f_{\mathcal{T}} + \Delta u_N\|_{L^2(K)}^2 &= 60 \|\sqrt{b_K}(f_{\mathcal{T}} + \Delta u_N)\|_{L^2(K)}^2 = 60 \int_K (f_{\mathcal{T}} + \Delta u_N)^2 b_K = 60 \int_K (f_{\mathcal{T}} + \Delta u_N)v \\ &\stackrel{v \in H_0^1(K)}{=} 60 \left(\int_K f_{\mathcal{T}} v - \int_K \nabla u_N \cdot \nabla v \right) = 60 \left(\int_K f v - \int_K \nabla u_N \cdot \nabla v + \int_K (f_{\mathcal{T}} - f)v \right) \\ &\stackrel{v \in H_0^1(\Omega), (6.2)}{=} 60 \left(\int_{\Omega} \nabla u \cdot \nabla v - \int_K \nabla u_N \cdot \nabla v + \int_K (f_{\mathcal{T}} - f)v \right) \\ &\stackrel{\text{supp } v \subset \overline{K}}{=} 60 \left(\int_K \nabla u \cdot \nabla v - \int_K \nabla u_N \cdot \nabla v + \int_K (f_{\mathcal{T}} - f)v \right) \\ &\leq 60 \|\nabla(u - u_N)\|_{L^2(K)} \|\nabla v\|_{L^2(K)} + 60 \|f_{\mathcal{T}} - f\|_{L^2(K)} \|v\|_{L^2(K)} \\ &\leq C [h_K^{-1} \|\nabla(u - u_N)\|_{L^2(K)} + \|f_{\mathcal{T}} - f\|_{L^2(K)}] \|f_{\mathcal{T}} + \Delta u_N\|_{L^2(K)} \end{aligned}$$

Kürzen mit $\|f_{\mathcal{T}} + \Delta u_N\|_{L^2(K)}$ liefert

$$\|f_{\mathcal{T}} + \Delta u_N\|_{L^2(K)} \leq C h_K^{-1} \|\nabla(u - u_N)\|_{L^2(K)} + C \|f - f_{\mathcal{T}}\|_{L^2(K)},$$

was gerade die gewünschte Behauptung ist. \square

Lemma 6.8 *Es existiert eine Konstante $C > 0$, welche nur von der Formregularitätskonstante γ abhängt, so daß für alle inneren Kanten $e \in \mathcal{E}(\mathcal{T})$ gilt:*

$$h_e \|[\partial_n u_N]\|_{L^2(e)}^2 \leq C \left[\|\nabla(u - u_N)\|_{L^2(\omega_e)}^2 + h_e^2 \|f - f_{\mathcal{T}}\|_{L^2(\omega_e)}^2 \right]$$

Beweis: Der Beweis erfolgt ähnlich wie der von Lemma 6.7 durch geschickte Wahl einer Testfunktion. Sei $e \in \mathcal{E}(\mathcal{T})$ fixiert. Setze $\eta_{e,\mathcal{E}}^2 := h_e \|[\partial_n u_N]\|_{L^2(e)}^2$. Seien $K, K' \in \mathcal{T}$ die Elemente mit $\overline{K} \cap \overline{K'} = \overline{e}$; es ist also $K \cup K' \cup e = \omega_e$. Weiter sei $b_e \in S^{2,1}(\mathcal{T})$ die quadratische Blasenformfunktion, die zur Kante e gehört, d.h. sie erfüllt

$$\text{supp } b_e \subset \omega_e$$

$$b_e|_e \text{ ist ein quadratisches Polynom mit } b_e(m) = 1, \text{ wobei } m \text{ der Mittelpunkt von } e \text{ ist.}$$

Wir betrachten nun die Funktion

$$v := [\partial_n u_N]|_e b_e.$$

Weil $[\partial_n u_N]|_e \in \mathbb{R}$, ist $v \in H_0^1(\omega_e) \subset H_0^1(\Omega)$. Weiter zeigt eine Rechnung

$$\int_e b_e = h_e \int_0^1 4x(1-x) dx = \frac{2}{3} h_e \quad (6.20)$$

$$\|b_e\|_{L^2(K)}^2 = \frac{2}{6} |K| \quad \text{falls } e \text{ Kante von } K \in \mathcal{T} \quad (6.21)$$

$$\|v\|_{L^2(\omega_e)}^2 \leq C \eta_{e,\mathcal{E}}^2 \quad (6.22)$$

$$\|\nabla v\|_{L^2(\omega_e)} \leq C h_e^{-1} \|v\|_{L^2(\omega_e)} \leq C h_e^{-1} \eta_{e,\mathcal{E}}. \quad (6.23)$$

Wir zeigen nur (6.22)—die inverse Ungleichung (6.23) wird mit ähnlichen Techniken gezeigt. Es ist

$$\begin{aligned} \|v\|_{L^2(\omega_e)}^2 &= \int_{\omega_e} ([\partial_n u_N]|_e)^2 b_e^2 = |[\partial_n u_N]|_e|^2 \left(\int_K b_e^2 + \int_{K'} b_e^2 \right) \\ &= |[\partial_n u_N]|_e|^2 \left(2|K| \frac{1}{6} + 2|K'| \frac{1}{6} \right) \leq C h_e \underbrace{h_e |[\partial_n u_N]|_e|^2}_{\sim \|[\partial_n u_N]\|_{L^2(e)}^2} \leq C \eta_{e,\mathcal{E}}^2. \end{aligned}$$

Damit haben wir (6.22) gezeigt. Wir schätzen nun ab:

$$\eta_{e,\mathcal{E}}^2 = h_e \|[\partial_n u_N]\|_{L^2(e)}^2 = h_e |[\partial_n u_N]|^2 \int_e 1 = h_e |[\partial_n u_N]|^2 \frac{3}{2} \int_e b_e = \frac{3}{2} h_e \int_e v [\partial_n u_N]$$

Wir nutzen nun, daß $v \in H_0^1(\omega_e)$ und erhalten damit durch partielle Integration auf K und auf K'

$$\begin{aligned} \eta_{e,\mathcal{E}}^2 &= \frac{3}{2} h_e \left[\int_K \Delta u_N v + \int_K \nabla u_N \cdot \nabla v + \int_{K'} \Delta u_N v + \int_{K'} \nabla u_N \cdot \nabla v \right] \\ &= \frac{3}{2} h_e \left[\int_K (f + \Delta u_N) v + \int_K \nabla u_N \cdot \nabla v + \int_{K'} (f + \Delta u_N) v + \int_{K'} \nabla u_N \cdot \nabla v - \int_{\omega_e} f v \right]. \end{aligned}$$

Wegen $v \in H_0^1(\omega_e) \subset H_0^1(\Omega) \subset H_0^1(\Omega, \Gamma_D)$ können wir v als Testfunktion in der schwachen Formulierung (6.2) einsetzen und erhalten

$$\int_{\omega_e} \nabla u \cdot \nabla v = \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} f v + \int_{\Gamma_N} \underbrace{g v}_{=0} = \int_{\Omega} f v = \int_{\omega_e} f v$$

Damit erhalten wir

$$\begin{aligned} \eta_{e,\mathcal{E}}^2 &= \frac{3}{2} h_e \left[\int_{\omega_e} \nabla(u_N - u) \cdot \nabla v + \int_K (f + \Delta u_N) v + \int_{K'} (f + \Delta u_N) v \right] \\ &\leq \frac{3}{2} h_e \left[\|\nabla(u - u_N)\|_{L^2(\omega_e)} \|\nabla v\|_{L^2(\omega_e)} + \|f + \Delta u_N\|_{L^2(K)} \|v\|_{L^2(K)} + \|f + \Delta u_N\|_{L^2(K')} \|v\|_{L^2(K')} \right] \\ &\leq C \|\nabla(u - u_N)\|_{L^2(\omega_e)} \eta_{e,\mathcal{E}} + C h_e \|f + \Delta u_N\|_{L^2(K)} \eta_{e,\mathcal{E}} + C h_e \|f + \Delta u_N\|_{L^2(K')} \eta_{e,\mathcal{E}} \end{aligned}$$

Kürzen mit $\eta_{e,\mathcal{E}}$ und Einsetzen der Abschätzung für $\|f + \Delta u_N\|_{L^2(K)}$, $\|f + \Delta u_N\|_{L^2(K')}$ aus Lemma 6.7 liefert dann die gewünschte Aussage. \square

Lemma 6.9 *Es existiert eine Konstante $C > 0$, welche nur von der Formregularitätskonstante γ abhängt, so daß für alle Kanten $e \in \mathcal{E}_N$ gilt:*

$$h_e \|g - \partial_n u_N\|_{L^2(e)}^2 \leq C \left[\|\nabla(u - u_N)\|_{L^2(K_e)}^2 + h_e^2 \|f - f_{\mathcal{T}}\|_{L^2(K_e)}^2 + h_e \|g - g_{\mathcal{T}}\|_{L^2(e)}^2 \right],$$

wobei $K_e \in \mathcal{T}$ das Element ist mit $e \in \mathcal{E}(K_e)$.

Beweis: Der Beweis wird ähnlich geführt wie der von Lemma 6.8. Wir setzen

$$\eta_{e,\mathcal{E}_N}^2 := h_e \|g - \partial_n u_N\|_{L^2(e)}^2, \quad \tilde{\eta}_{e,\mathcal{E}_N}^2 := h_e \|g_{\mathcal{T}} - \partial_n u_N\|_{L^2(e)}^2.$$

Sei $b_e \in S^{2,1}(\mathcal{T})$ so, daß $\text{supp } b_e = \overline{K_e}$ und $b_e|_e$ ein quadratisches Polynom ist mit $b_e(m) = 1$, wobei $m \in e$ der Mittelpunkt von e ist. Wir betrachten die Funktion

$$v := (g_{\mathcal{T}}|_e - \partial_n u_N) b_e \in H_{0,\Gamma_D}^1(\Omega)$$

und bemerken $\text{supp } v \subset \overline{K_e}$. Dann ergibt sich wie im Beweis von Lemma 6.8

$$\begin{aligned} \|v\|_{L^2(K_e)} &\leq C \tilde{\eta}_{e,\varepsilon_N}, \\ \|\nabla v\|_{L^2(K_e)} &\leq C h_e^{-1} \tilde{\eta}_{e,\varepsilon_N}, \\ \|\nabla v\|_{L^2(e)} &\leq C \tilde{\eta}_{e,\varepsilon_N}. \end{aligned}$$

Damit ergibt sich weiter wie im Beweis von Lemma 6.8

$$\begin{aligned} \tilde{\eta}_{e,\varepsilon_N}^2 &= h_e \|g_{\mathcal{T}} - \partial_n u_N\|_{L^2(e)}^2 = \frac{3}{2} h_e \int_e (g_{\mathcal{T}} - \partial_n u_N) v = \frac{3}{2} h_e \int_e (g_{\mathcal{T}} - g) v + \frac{3}{2} h_e \int_e (g - \partial_n u_N) v \\ &= \frac{3}{2} h_e \int_e (g_{\mathcal{T}} - g) v + \frac{3}{2} h_e \int_e g v - \frac{3}{2} h_e \left[\int_{K_e} \nabla u_N \cdot \nabla v + \int_{K_e} \Delta u_N v \right], \end{aligned}$$

wobei der letzte Schritt durch eine partielle Integration entsteht. Weiter können wir wegen $v \in H_0^1(\Omega, \Gamma_D)$ wiederum die schwache Formulierung (6.2) verwenden und erhalten aus $\text{supp } v \subset \overline{K_e}$:

$$\int_e g v = \int_{\Gamma_N} g v = \int_{K_e} \nabla u \cdot \nabla v - \int_{K_e} f v$$

Setzt man dies ein in die obige Darstellung von $\tilde{\eta}_{e,\varepsilon_N}^2$, so ergibt sich

$$\begin{aligned} \tilde{\eta}_{e,\varepsilon_N}^2 &= \frac{3}{2} h_e \int_e (g_{\mathcal{T}} - g) v + \frac{3}{2} h_e \left[\int_{K_e} \nabla(u - u_N) \cdot \nabla v - \int_{K_e} (f + \Delta u_N) v \right] \\ &\leq C h_e \left[\|g - g_{\mathcal{T}}\|_{L^2(e)} \|v\|_{L^2(e)} + \|\nabla(u - u_N)\|_{L^2(K_e)} \|\nabla v\|_{L^2(K_e)} + \|f + \Delta u_N\|_{L^2(K_e)} \|v\|_{L^2(K_e)} \right] \\ &\leq C \tilde{\eta}_{e,\varepsilon_N} \left[h_e \|g - g_{\mathcal{T}}\|_{L^2(e)} + \|\nabla(u - u_N)\|_{L^2(K_e)} + h_e \|f + \Delta u_N\|_{L^2(K_e)} \right]. \end{aligned}$$

Berücksichtigt man nun wiederum Lemma 6.7 zum Abschätzen von $\|f + \Delta u_N\|_{L^2(K_e)}$ sowie, daß die Dreiecksungleichung $\eta_{e,\varepsilon_N} \leq \tilde{\eta}_{e,\varepsilon_N} + \|g - g_{\mathcal{T}}\|_{L^2(e)}$ liefert, so ergibt sich die gewünschte Behauptung. \square

6.3 Ein adaptiver Algorithmus

Ziel: konstruiere (möglichst einfach) ein Gitter \mathcal{T} (mit möglichst wenig Elementen), so daß $\|u - u_N\|_E \leq \text{TOL}$ ist, wobei TOL eine vorgegebene Genauigkeit ist.

In der Praxis wird man, ausgehend von einer groben Starttriangulierung \mathcal{T}_0 , eine Folge \mathcal{T}_n , $n = 0, 1, \dots$ von Gittern erzeugen und abbrechen, wenn die gewünschte Genauigkeit erreicht ist. Das Gitter \mathcal{T}_n ergibt sich dabei durch Verfeinerung des Gitters \mathcal{T}_{n-1} . Um zu entscheiden, welche Elemente K des Gitters \mathcal{T}_{n-1} verfeinert werden, macht man die Annahme, daß die Fehlerindikatoren η_K aus (6.12) den Fehler auf K gut widerspiegeln, d.h., daß eine Zerlegung der Elemente K , für die η_K groß ist, auch zu einer starken Reduktion des Gesamtfehlers $\sum_{K \in \mathcal{T}} \eta_K^2$ führt. Es ergibt sich damit

1. bestimme FE-Lösung $u_N \in S^{1,1}(\mathcal{T}_n) \cap H_0^1(\Omega, \Gamma_D)$
2. bestimme η_K für alle $K \in \mathcal{T}_n$
3. markiere alle Elemente K als zu verfeinern, für die η_K "groß" ist
4. erzeuge \mathcal{T}_{n+1}
5. gehe zu 1.

Ein typischer Algorithmus ist in Alg. 3 angegeben.

Algorithm 3 klassischer adaptiver FEM-Algorithmus (“Dörfler marking”)

 %input: Startgitter \mathcal{T}_0 , Toleranz TOL, Parameter $\theta \in (0, 1)$
 $n := 0$

- 1: **repeat**{
 - 2: Berechne FEM-Lösung $u_N \in S^{1,1}(\mathcal{T}_n) \cap H_0^1(\Omega, \Gamma_D)$
 - 3: Berechne für alle $K \in \mathcal{T}_n$ die Indikatoren η_K mittels (6.12)
 - 4: $\text{EST}^2 := \sum_{K \in \mathcal{T}} \eta_K^2$
 - 5: wähle eine minimale Menge $\mathcal{M}_n \subset \mathcal{T}_n$ mit $\sum_{K \in \mathcal{M}_n} \eta_K^2 \geq \theta \sum_{K \in \mathcal{T}_n} \eta_K^2$.
 - 6: markiere alle $K \in \mathcal{M}_n$ zum Verfeinern
 - 7: erzeuge \mathcal{T}_{n+1} , indem alle zum Verfeinern markierten Elemente $K \in \mathcal{M}_n$ verfeinert werden und hängende Knoten durch “Gitterabschluß” (siehe Alg. 4) beseitigt werden
 - 8: $n := n + 1$
 - 9: } **until** $\text{EST} < \text{TOL}$
-


 Abbildung 6.2: Die Verfeinerungsmöglichkeiten *rot*, *grün*, *blau*. Es wird auf jeden Fall eine längste Seite geteilt.

Bemerkung 6.10 Eine typische Wahl von θ ist $\theta = 0.25$. Man versucht, Qualität der erzeugten Gitter und Gesamtkosten in ein vernünftiges Verhältnis zu bringen: Einerseits erreicht man die vorgegebene Toleranz in weniger Schritten des adaptiven Algorithmus, je mehr Elemente in jedem Schritt erzeugt werden; andererseits erwartet man, daß der Algorithmus umso ineffizienter ist, je mehr Elemente pro Schritt verfeinert werden (Extremfall: $\theta = 1$ führt auf uniforme Verfeinerungen). Man beachte: Falls die Problemgrößen N_n (für Gitter \mathcal{T}_n) die Bedingung $\frac{N_{n+1}}{N_n} \geq q > 1$ erfüllen, dann sind die Gesamtkosten nur ein (festes) Vielfaches der Kosten, die auf dem akzeptierten Gitter \mathcal{N}_ν entstehen. Um dies einzusehen, sei W_n der Aufwand zum Bestimmen von u_N auf Gitter \mathcal{T}_n ; wir nehmen an, daß $W_n \sim N_n$. Dann gilt:

$$W_{ges} = \sum_{n=0}^{\nu} W_n \leq C \sum_{n=0}^{\nu} N_n \leq C \sum_{n=0}^{\nu} N_\nu q^{-(\nu-n)} \leq CN_\nu \sum_{n=0}^{\infty} q^{-n} \leq CN_\nu \sim W_\nu \quad \blacksquare$$

Wir betrachten nun das Problem des Gitterabschlusses, d.h. der Eliminierung von hängenden Knoten. Die Schwierigkeit besteht darin, sicherzustellen, daß die Gitter \mathcal{T}_n , $n = 0, 1, \dots$, formregulär bleiben. Dies wird mit Alg. 4 erreicht. Zur Vereinfachung nehmen wir an, daß anstelle von Elementen Kanten zur Verfeinerung markiert sind. Dies erreicht man z.B. so:

for $K \in \mathcal{T}$ **do** { **if** K markiert, **then** $\forall e \in \mathcal{E}(K)$ markiere e }

Um die Formregularität der Elemente sicherzustellen, darf ein Element nur nach einer der folgenden 4 Regeln verfeinert werden (siehe Fig. 6.2):

- rot** alle 3 Kanten sind markiert \rightarrow zerlege in 4 kongruente Dreiecke
- grün** genau eine Kante ist markiert, welche zudem eine längste Kante ist \rightarrow zerlege Element in 2 Dreiecke, indem der Mittelpunkt der markierten Kante mit der gegenüberliegenden Ecke verbunden wird.
- blau** genau 2 Kanten sind markiert und eine längste Kante ist unter den markierten \rightarrow zerlege Element in 3 Dreiecke, indem die Mittelpunkte der markierten Kanten verbunden werden und zudem der Mittelpunkt der längsten Kante mit der gegenüberliegenden Ecke verbunden wird.

Algorithm 4 rot-grün-blau-Abschluß

%input: Triangulierung \mathcal{T} , bei der die Kanten zur Verfeinerung markiert sind

%output: reguläres Gitter \mathcal{T}' , das durch Teilen der markierten Kanten und Beseitigung der hängenden Knoten entsteht

- 1: **while** ($\exists K \in \mathcal{T}$ s.d. mindestens eine Kante von K ist markiert, aber keine längste Kante von K)
 - 2: markiere eine längste Kante von K
 - 3: erzeuge \mathcal{T}' durch Anwenden der *Rot-, Grün-, Blauverfeinerungen* aus Abb. 6.2.
-

Bemerkung 6.11 Alg. 4 ist nicht vollständig determiniert, da ein Element mehr als eine längste Kante haben kann. ■

Ohne Beweis zeigt folgender Satz, daß Alg. 4 die Formregularität der Ausgangstriangulierung erhält. Wesentlich im Beweis ist, daß bei Verfeinerung immer eine längste Kante eines Elementes geteilt wird (d.h. insb., daß der kleinste Winkel nicht geteilt wird).

Satz 6.12 Sei \mathcal{T}_0 ein Gitter. Sei $\varepsilon > 0$ der kleinste Innenwinkel von \mathcal{T}_0 . Sei $(\mathcal{T}_n)_{n=0}^\infty$ ein Folge von Gittern, die von Alg. 4 erzeugt werden (die in jedem Schritt gewählten Kantenmarkierungen sind beliebig). Dann gilt: Für alle $n \in \mathbb{N}$ ist der kleinste Innenwinkel von \mathcal{T}_n nach unten durch $\varepsilon/2$ von Null weg beschränkt.

Beispiel 6.13

$$-u'' + u = f \quad \text{auf } \Omega = (0, 1), \quad u(0) = u(1) = 0.$$

Die Fehlerindikatoren können wie folgt definiert werden:

$$\eta_K^2 = h_K^2 \|f - (-u''_{FE} + u_{FE})\|_{L^2(K)}^2.$$

Die rechte Seite f wird so gewählt, daß die exakte Lösung

$$u(x) = x^\alpha(1 - x), \quad \alpha = 0.75$$

ist. Die FE-Räume sind gegeben durch $S_0^{1,1}(\mathcal{T})$, wobei die Gitter \mathcal{T} mit Hilfe von Alg. 3 erzeugt wird. Fig. 6.3 zeigt die Gitter in Schritt 10 und 30 sowie das Konvergenzverhalten des Algorithmus. ■

Beispiel 6.14 Wir betrachten den adaptiven Algorithmus für das Modellproblem (6.1) mit $f \equiv 1$, $\Gamma_D = \partial\Omega$ auf dem L-Gebiet. Die Folge von erzeugten Gittern ist in Fig. 6.4 dargestellt. ■

finis 25.Stunde

finis 26.Stunde

6.3.1 Weitere Verfeinerungsalgorithmen

Die “rot-grün-blau”-Verfeinerung ist nur ein mögliches Beispiel. Ein weiteres Beispiel ist die *newest vertex bisection*. Die “Datenstruktur” ist dabei so, daß für jedes Element eine der 3 Ecken als “newest vertex” gekennzeichnet ist.

Ein adaptiver Algorithmus benötigt eine Starttriangulierung \mathcal{T}_0 . Für jedes Element $K \in \mathcal{T}_0$ wählen wir eine der 3 Ecken als “newest vertex” aus. Es erweist sich als sinnvoll, diese Wahl der “newest vertices” so zu gestalten, daß folgendes gilt: Wenn zwei Elemente K, K' mit “newest vertices” $v(K)$ und $v(K')$ eine gemeinsame Kante e haben, dann ist sowohl für K als auch für K' die Ecke $v(K)$ und $v(K')$ gerade die, die der Kante e gegenüberliegt. Eine solche Startmarkierung heißt “gut” (vgl. Fig. 6.5).

Bemerkung 6.15 Eine einfache Art, eine Starttriangulierung zu konstruieren, die eine Startmarkierung zuläßt, die “gut” ist, geht von der folgenden Beobachtung aus:

- Es mögen sich die Elemente zu Paaren zusammenfassen lassen, die eine gemeinsame Kanten haben; die verbliebenen Elemente, die keinen Partner haben, mögen eine Kante auf dem Rand haben. Dann ist die Zuweisung des “newest vertex” einfach wie in Fig. 6.5 angedeutet: Für die Elemente, die in Paaren auftreten, wird die der gemeinsamen Kante gegenüberliegende Ecke markiert; für die einzelnen Elemente, die am Rand liegen, wird die Ecke markiert, die der Randkante gegenüber liegt.

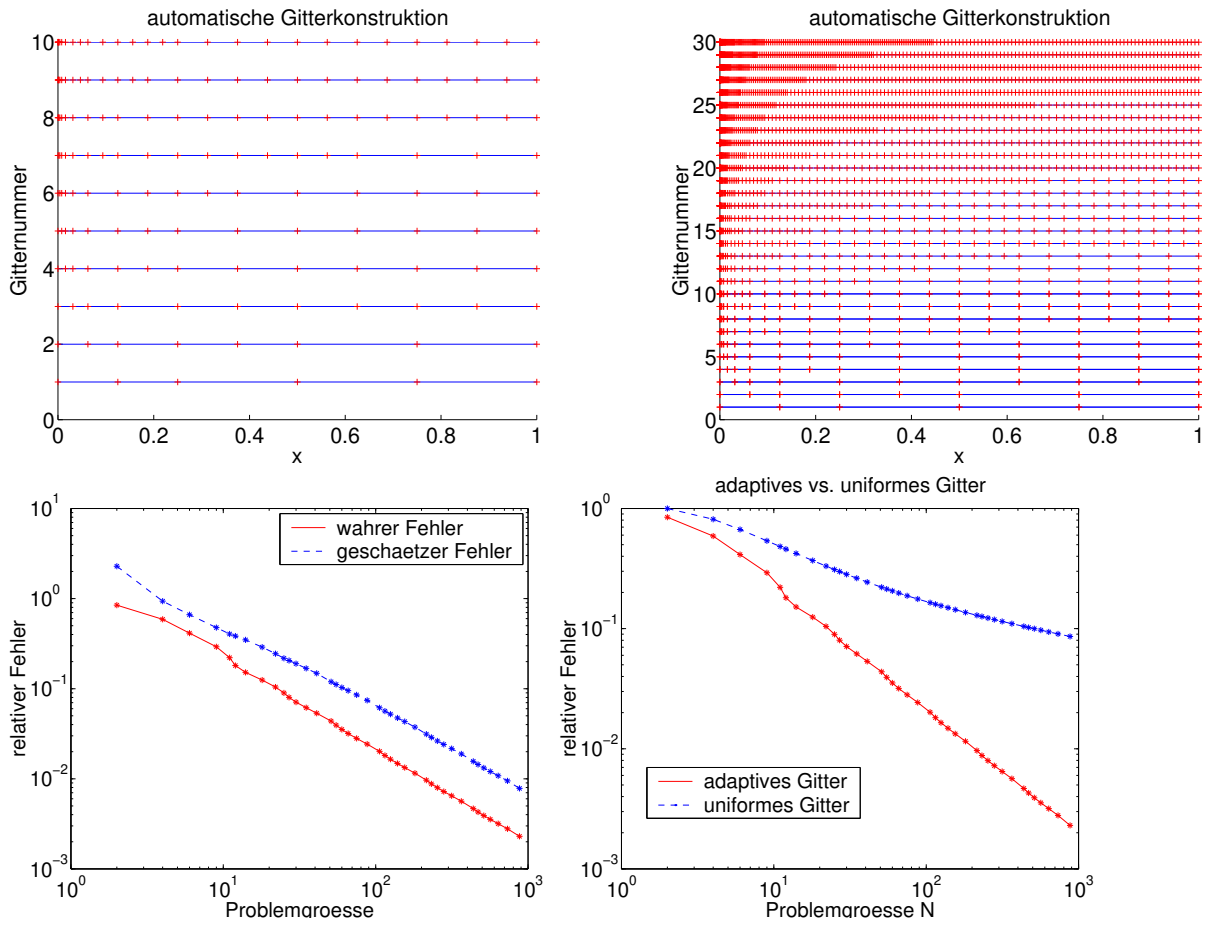


Abbildung 6.3: Verhalten des adaptiven Algorithmus für ein 1D-Beispiel (siehe Beispiel 6.13).

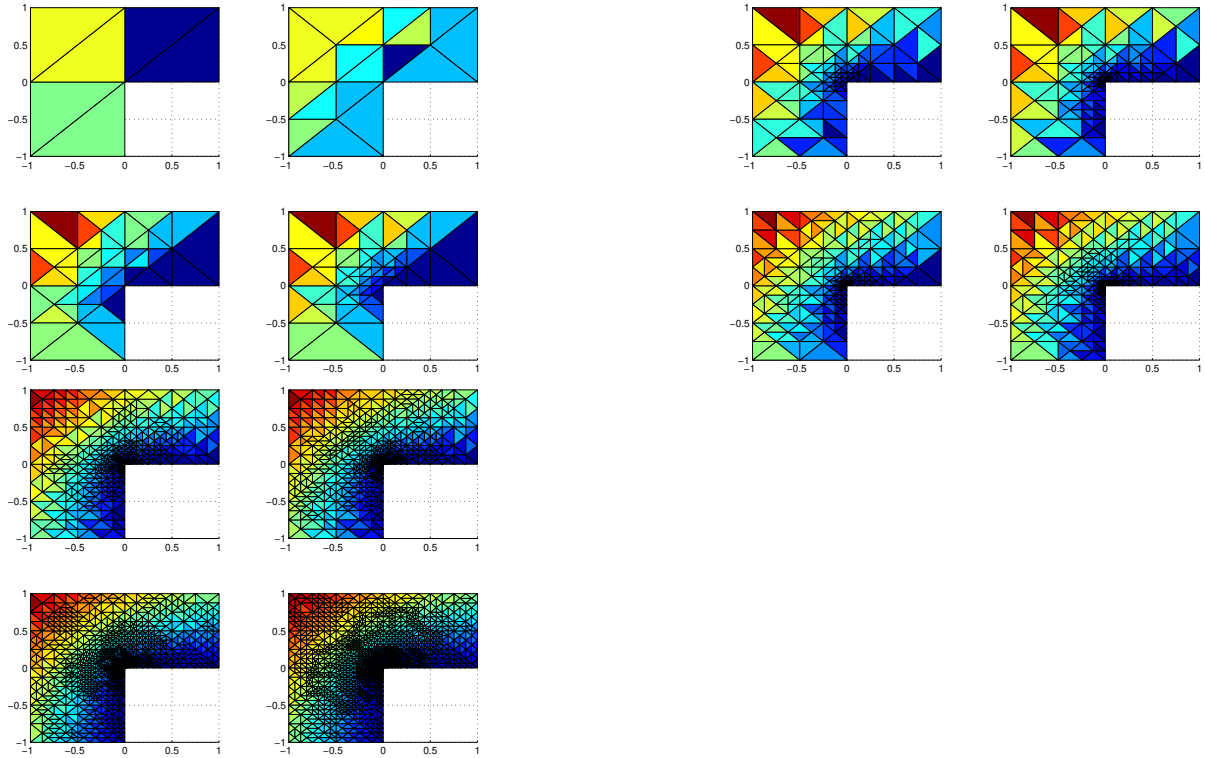


Abbildung 6.4: Verhalten des adaptiven Algorithmus für ein 2D-Beispiel (siehe Beispiel 6.14).

- Aus einer Triangulierung \mathcal{T}'_0 kann man einfach eine Triangulierung \mathcal{T}_0 erzeugen, die die obige Eigenschaft hat. Hierzu wird einfach jedes Element wie in Fig. 6.5 gezeigt zerlegt.

Die Verfeinerung geschieht dann mit Algorithmus 5:

Algorithm 5 Newest Vertex Bisection

%input: konforme Triangulierung \mathcal{T} , bei der jedes Element genau eine Ecke als “newest vertex” gekennzeichnet hat

%input: Liste \mathcal{M} mit Elementen, die zu verfeinern sind

%output: reguläre Triangulierung \mathcal{T}' , bei der alle zu verfeinernden Elemente verfeinert wurden

- 1: **while** $\mathcal{M} \neq \emptyset$
 - 2: wende “newest vertex bisection” auf alle $K \in \mathcal{M}$ an, d.h. für jedes $K \in \mathcal{M}$ erzeuge zwei Elemente K_1, K_2 durch Halbieren von K , indem die dem “newest vertex” gegenüberliegende Seite e mit Mittelpunkt m geteilt wird. Weise den Elementen K_1, K_2 als “newest vertex” den Punkt m zu
 - 3: $\mathcal{M} :=$ Menge der Dreiecke aus der so erhaltenden Triangulierung, die hängende Knoten haben
 - 4: }
-

Eine wichtige Eigenschaft von Algorithmus 5 ist, daß er terminiert, formreguläre Gitter erzeugt und zudem “nicht zuviele” zusätzliche Elemente durch “Netzabschluß” erzeugt:

Satz 6.16 Sei \mathcal{T}_0 eine Starttriangulierung, die im obigen Sinne “gut” ist. Sei $(\mathcal{T}_n)_{n \in \mathbb{N}_0}$ eine Folge von Triangulierungen, die mit Alg. 5 erzeugt werden (die Mengen \mathcal{M}_n , der markierten Elemente sind beliebig). Dann existieren $\varepsilon, C > 0$, welche nur von der Starttriangulierung \mathcal{T}_0 abhängen, so daß folgendes gilt:

(i) für alle n sind die Innenwinkel der Dreiecke $K \in \mathcal{T}_n$ größer als ε

(ii) Es gilt: $|\mathcal{T}_n| \leq |\mathcal{T}_0| + C \sum_{l=0}^{n-1} |\mathcal{M}_l|$

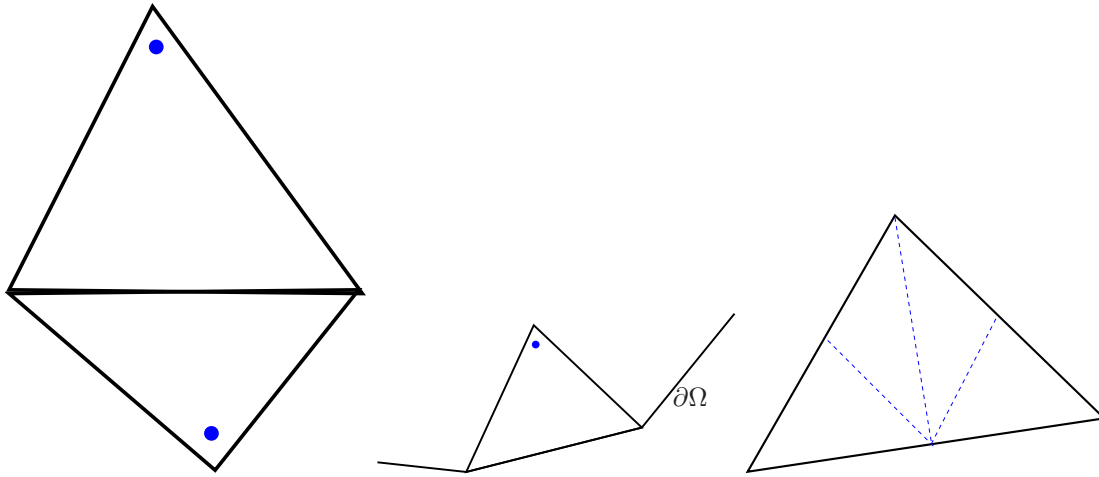


Abbildung 6.5: links: Zuweisen von “newest vertex” für Dreiecke, die ein Paar bilden. Mitte: Zuweisung des “newest vertex” an ein einzelnes Element am Rand. Rechts: Zerlegung eines Dreiecks, um eine “gute” Triangulation zu erhalten.

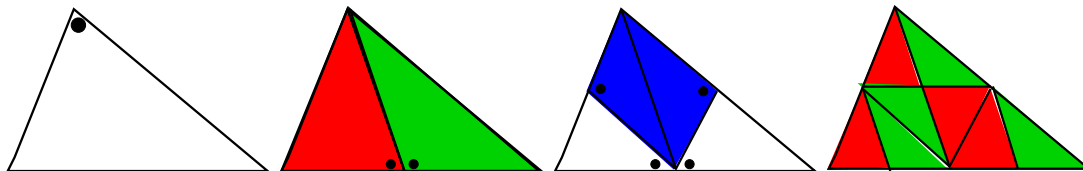


Abbildung 6.6: Durch wiederholtes Teilen eines Dreiecks können höchstens 4 verschiedene Ähnlichkeitsklassen entstehen. Der “newest vertex” jedes Dreiecks ist gekennzeichnet. Gleiche Farbe bedeutet ähnlich.

Beweis: Wir beweisen nur (i). Wir sagen, daß zwei Dreiecke ähnlich sind, wenn sie gleiche Winkel haben. Fig. 6.6 illustriert die elementare geometrische Überlegungen, daß durch wiederholtes Halbieren eines Ausgangsdreiecks mittels “newest vertex bisection” höchstens 4 Klassen von ähnlichen Dreiecken entstehen können. Damit ist die Gesamtzahl von ähnlichen Dreiecken beschränkt durch $4|\mathcal{T}_0|$. Die Aussage (ii) geht auf [4] zurück; die Erweiterung auf 3D findet sich in [16]. \square

6.4 Ein äquilibrierter Fehlerschätzer

Der residuale Fehlerschätzer (6.12), (6.14) spiegelt das Verhalten des Fehlers (in der Energienorm) korrekt wider—jedoch ist die Konstante $C > 0$ in der Abschätzung $\|u - u_N\|_E \leq C \text{EST}$ unbekannt bzw. in der Praxis zu pessimistisch. Der residuale Fehlerschätzer (6.12), (6.14) eignet sich deshalb zwar gut zum Steuern eines adaptiven Algorithmus, aber nicht so sehr als Abbruchkriterium. Hier stellen wir einen Fehlerschätzer vor, der den Fehler *ohne Konstante* beschränkt.

6.4.1 Fehlerschätzer

Wir erinnern an die Fehlerdarstellung (6.8)

$$\|u - u_N\|_E = \sup_{v \in H_0^1(\Omega, \Gamma_D)} \frac{B(u - u_N, v)}{\|v\|_E} = \sup_{v \in H_0^1(\Omega, \Gamma_D)} \frac{R(v)}{\|v\|_E},$$

$$R(v) = l(v) - B(u_N, v) = \int_{\Omega} f v + \int_{\Gamma_N} g v - B(u_N, v)$$

Motivation: Wir “leiten” nun eine Gleichung für den elementweisen Fehler $e|_K := (u - u_N)|_K$ her. Seien hierzu

$$\begin{aligned} B_K(w, v) &:= \int_K \nabla w \cdot \nabla v \\ R_K(v) &:= \int_K f v + \int_{\partial K \cap \Gamma_N} g v - B_K(u_N, v). \end{aligned}$$

Weiter bemerken wir, daß die Lösung u (falls u hinreichend glatt ist) die Gleichung

$$B_K(u, v) = \int_K f v + \int_{\partial K} \partial_n u v \quad \forall v \in H^1(K)$$

erfüllt. Damit ergibt sich für den Fehler $e|_K$:

$$B_K(e, v) = \int_K f v + \int_{\partial K} \partial_n u v - B_K(u_N, v) = R_K(v) + \int_{\partial K \setminus \Gamma_N} \partial_n u v \quad (6.24)$$

Da die Normalableitung $\partial_n u|_{\partial K}$ natürlich nicht bekannt ist, approximieren wir sie für jedes $K \in \mathcal{T}$ durch eine noch festzulegende Funktion g_K :

$$g_K \approx n_K \cdot \nabla u \quad (6.25)$$

Weil $[\partial_n u]|_e = 0$ für alle inneren Kanten $e \in \mathcal{E}(\mathcal{T})$ ist und weil $\partial_n u|_e = g|_e$ für $e \subset \Gamma_N$, fordern wir für die Flüsse g_K

$$g_K|_e = g|_e \quad \forall e \in \mathcal{E}_N \quad (6.26)$$

$$g_K + g_{K'} = 0 \quad \text{auf } \overline{K} \cap \overline{K'}. \quad (6.27)$$

Wir definieren den Raum

$$H_D^1(K) := \{v \in H^1(K) \mid v|_{\Gamma_D} = 0\};$$

in Hinblick auf die elementweise Fehlergleichung (6.24) und unsere Vorstellung (6.25) können wir damit eine “Approximation” \tilde{e}_K an den Fehler definieren als Lösung von

$$\begin{aligned} \text{Finde } \tilde{e}_K \in H_D^1(K), \text{ s.d. } B_K(\tilde{e}_K, v) &= R_K(v) + \int_{\partial K \setminus \mathcal{E}_N} g_K v \quad \forall v \in H_D^1(K) \\ &= \int_K f v - B_K(u_N, v) + \int_{\partial K} g_K v \quad \forall v \in H_D^1(K). \end{aligned} \quad (6.28)$$

Falls \tilde{e}_K existiert² können wir abschätzen:

$$\begin{aligned} R(v) &= \sum_{K \in \mathcal{T}} R_K(v) \stackrel{(6.27)}{=} \sum_{K \in \mathcal{T}} R_K(v) + \int_{\partial K \setminus \Gamma_N} g_K v = \sum_{K \in \mathcal{T}} B_K(\tilde{e}_K, v) \leq \sum_{K \in \mathcal{T}} \|\nabla \tilde{e}_K\|_{L^2(K)} \|\nabla v\|_{L^2(K)} \\ &\leq \sqrt{\sum_{K \in \mathcal{T}} \|\nabla \tilde{e}_K\|_{L^2(K)}^2} \sqrt{\sum_{K \in \mathcal{T}} \|\nabla v\|_{L^2(K)}^2} \end{aligned}$$

Damit ist $\|u - u_N\|_E = \|e\|_E = \sup_{v \in H_0^1(\Omega, \Gamma_D)} \frac{R(v)}{\|v\|_E} \leq \text{EST}^{\text{eq}}$ mit

$$\text{EST}^{\text{eq}} := \sqrt{\sum_{K \in \mathcal{T}} \|\nabla \tilde{e}_K\|_{L^2(K)}^2} \quad (6.29)$$

Wir halten das Ergebnis im folgenden Satz fest:

²Existenz von \tilde{e}_K tritt nur als Problem auf, wenn $|\partial K \cap \Gamma_D| = 0$, denn dann ist $H_D^1(K) = H^1(K)$ und B_K ist nicht koerziv auf $H^1(K)$. In diesem Fall ist das Problem (6.28) ein Neumannproblem ohne Absolutterm wie in Abschnitt 4.9 genauer diskutiert: Partielle Integration zeigt, daß \tilde{e}_K das folgende Randwertproblem löst: $-\Delta \tilde{e}_K = (f + \Delta u_N)$ auf K und $\partial_n \tilde{e}_K = g_K - \partial_n u_N$ auf ∂K . Aus Abschnitt 4.9 wissen wir, daß die rechte Seite von (6.28) die Kompatibilitätsbedingung $0 = R_K(1) + \int_{\partial K \setminus \mathcal{E}_N} g_K 1$ erfüllen muß.

Satz 6.17 Für jedes $K \in \mathcal{T}$ sei eine (auf ∂K definierte Funktion) g_K gewählt, die die Bedingungen (6.26), (6.27) erfüllt. Falls es für jedes Element $K \in \mathcal{T}$ eine Lösung \tilde{e}_K von (6.28) gibt, so erfüllt der Fehlerschätzer EST^{eq} aus (6.29) die Abschätzung $\|u - u_N\|_E \leq \text{EST}^{\text{eq}}$.

Bemerkung 6.18 Die Funktionen \tilde{e}_K müssen als Lösungen der lokalen Hilfsprobleme (6.28) bestimmt werden. In der Praxis kann man diese wiederum nur approximieren, z.B. als Lösungen von

$$\text{Finde } \tilde{e}'_K \in V_K: \quad B_K(\tilde{e}'_K, v) = R_K(v) + \int_{\partial K \setminus \Gamma_N} g_K v \quad \forall v \in V_K, \quad (6.30)$$

wobei $V_K \subset H_D^1(K)$ geeignet gewählt ist, z.B. als Raum von Polynomen höherer Ordnung (z.B. $p = 2$ oder $p = 3$). ■

6.4.2 Wahl der Flußfunktionen g_K

Struktur der g_K

Die Flußfunktionen g_K müssen die Bedingungen (6.26), (6.27) erfüllen. Weitere Bedingungen ergeben sich aus der Forderung nach Lösbarkeit von (6.28). Wir hatten zudem die Forderung (6.25) als sinnvoll erarbeitet. Die Lösbarkeit von (6.28) erfordert, daß für die Elemente $K \in \mathcal{T}$, die “nicht am Dirichletrand sind” (genauer: $|\partial K \cap \Gamma_D| = 0$), die Kompatibilitätsbedingung

$$\begin{aligned} 0 &= R_K(1) + \int_{\partial K \setminus \Gamma_N} g_K 1 \\ &= \int_K f 1 - B(u_N, 1) + \int_{\partial K} g_K 1 = \int_K (f + \Delta u_N) 1 + \int_{\partial K} (g_K - \partial_n u_N) 1 \end{aligned} \quad (6.31)$$

erfüllt ist. Wir beobachten nun:

1. Bisher forderten wir von den Funktionen g_K lediglich die Bedingungen (6.26), (6.27), (6.31)—diese 3 Bedingungen legen die Funktionen g_K nicht fest, und wir müssen (und können!) weitere Bedingungen an die Funktionen g_k stellen.
2. Die zusätzlichen Bedingungen an die Funktionen g_K dienen auch dazu, die Funktionen g_K lokal festlegen zu können—eine nichttriviale Forderung, denn (6.27) koppelt ja die Flüße benachbarter Elemente.

Um die Flußfunktionen g_K festzulegen, machen wir einen Strukturansatz. Wir betrachten hier den Ansatz, daß die Funktionen g_K auf jeder Kante (die nicht auf dem Neumannrand Γ_N liegt) ein Polynom vom Grad 1 sind (g_K muß aber nicht auf ∂K stetig sein!):

$$\begin{cases} g_K|_e = g & \forall e \in \mathcal{E}(K) \cap \mathcal{E}_N \\ g_K|_E \in \text{span}\{\varphi_V|_e \mid V \in \mathcal{N}(e)\} & \forall e \in \mathcal{E}(K) \setminus \mathcal{E}_N. \end{cases} \quad (6.32)$$

Es wird sich als zweckmäßig herausstellen, die Momente $\mu_{K,V}^e \in \mathbb{R}$ einzuführen:

$$\mu_{K,V}^e := \int_e \varphi_V g_K \quad V \in \mathcal{N}(e). \quad (6.33)$$

Wir bemerken, daß Kenntnis der Momente $\mu_{K,V}^e$, $V \in \mathcal{N}(e)$ reicht, um die g_K festzulegen:

Beispiel 6.19 Sei $e \in \mathcal{E}(K) \setminus \mathcal{E}_N$. Aus dem Ansatz (6.32) folgt $g_K|_e = \alpha_1 \varphi_{V_1} + \alpha_2 \varphi_{V_2}$, wobei $\{V_1, V_2\} = \mathcal{N}(e)$ die beiden Endpunkte von e sind. Für die Momente $\mu_{K,V_1}^e, \mu_{K,V_2}^e$ folgt dann aus der Definition (6.33)

$$\begin{pmatrix} \mu_{K,V_1}^e \\ \mu_{K,V_2}^e \end{pmatrix} = M_e \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \quad M_e = \begin{pmatrix} \int_e |\varphi_{V_1}|^2 & \int_e \varphi_{V_1} \varphi_{V_2} \\ \int_e \varphi_{V_1} \varphi_{V_2} & \int_e |\varphi_{V_2}|^2 \end{pmatrix} = \frac{h_e}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Die Massematrix M_e ist invertierbar mit Inverser $M_e^{-1} = \frac{2}{h_e} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, weshalb sich die benötigten Koeffizienten α_1, α_2 aus dem Momenten $\mu_{K,V_1}^e, \mu_{K,V_2}^e$ bestimmen lassen. ■

Festlegen der Momente

Nachdem wir die Form der Flußfunktionen g_K durch den Ansatz (6.32) festgelegt haben und gesehen haben, daß diese durch die Momente $\mu_{K,V}^e$ eindeutig bestimmt sind, stellen wir nun Gleichungen für die Momente $\mu_{K,V}^e$ auf. Unser Vorgehen wird dabei sein, für jeden festen Knoten $V \in \mathcal{N}(\mathcal{T})$ die Momente $\mu_{K,V}^e$ für die Kanten e und Elemente K zu bestimmen, für die $V \in \mathcal{N}(e)$ und $V \in \mathcal{N}(K)$. Es wird sich herausstellen, daß dies als Lösung eines kleinen LGS möglich ist.

Die Strukturannahme (6.32) zusammen mit den Bedingungen (6.26), (6.27) und (6.31) legt die Funktionen g_K immer noch nicht fest. Um weitere Bedingungen zu erhalten, betrachten wir wiederum die Kompatibilitätsbedingung (6.31). Wir verschärfen sie in zwei Arten: zum einen fordern wir eine solche Bedingung für *alle* Elemente K und zum anderen fordern wir, daß $v \mapsto R_K(v) + \int_{\partial K \setminus \Gamma_N} g_K v$ nicht nur für alle $v \in \mathcal{P}_0$, sondern auch für alle $v \in \mathcal{P}_1$ verschwindet. Mit anderen Worten: Wir fordern

$$\forall K \in \mathcal{T} \forall V \in \mathcal{N}(K) : \quad 0 = R_K(\varphi_V) + \int_{\partial K \setminus \Gamma_N} g_K \varphi_V \quad (6.34)$$

(6.34) impliziert (6.31), denn $\sum_{V \in \mathcal{N}(K)} \varphi_V \equiv 1$ auf K .³ Eine zu (6.34) äquivalente Forderung ist “knotenorientiert”:

$$\forall V \in \mathcal{T} \forall K \in \mathcal{N}(\mathcal{T}) \text{ mit } V \in \mathcal{N}(K) : \quad 0 = R_K(\varphi_V) + \int_{\partial K \setminus \Gamma_N} g_K \varphi_V = \int_K f \varphi_V - B_K(u_N, \varphi_V) + \int_{\partial K} g_K \varphi_V \quad (6.35)$$

Wir bemerken, daß für jeden Knoten $V \in \mathcal{N}(\mathcal{T})$ genau einer der folgenden vier Fälle auftritt (vgl. Fig. 6.7):

1. V ist ein innerer Knoten (“I”),
2. V ist ein Randknoten, an dem zwei Neumannkanten anstoßen (“NN”),
3. V ist ein Randknoten, an dem eine Neumannkante und eine Dirichletkante anstoßen (“ND”),
4. V ist ein Randknoten, an dem zwei Dirichletkanten anstoßen (“DD”).

Wir fassen nun die Bedingungen an die Momente $\mu_{K,V}^e$ zusammen, die wir formuliert haben: Sei $V \in \mathcal{N}(\mathcal{T})$ fest. Seien K_1, \dots, K_n die n Elemente, die an den Knoten V anstoßen (siehe Fig. 6.7). Dann haben wir bis jetzt folgende Bedingungen an die Momente $\mu_{K,V}^e$ für dieses feste V erhalten:

- Die Forderung (6.35) ergibt n Gleichungen (1 pro Element K_i , $i = 1, \dots, n$, im Patch ω_V)
- Die Forderung (6.27) ergibt 1 Gleichung pro *innere* Kante
- Für Kanten $e \subset \mathcal{E}_N$ ist g_K (und damit die Momente) bereits festgelegt.
- Für Kanten $e \subset \mathcal{E}_D$ gibt es keine Bedingungen an die Momente.

Ausgeschrieben als Gleichungssystem ergibt sich: Die Bedingung (6.35) liefert $-\left(\int_K f \varphi_V - B_K(u_N, \varphi_V)\right) = \int_{\partial K} g_K \varphi_V$ für alle $K \subset \omega_V$, d.h. für jedes Element K_i , $i = 1, \dots, n$, des Patches ω_V : finis 27.Stunde

$$-\left(\int_{K_i} f \varphi_V - B_{K_i}(u_N, \varphi_V)\right) = \mu_{K_i,V}^{e_i} + \mu_{K_i,V}^{e_{i+1}} \quad i = 1, \dots, n. \quad (6.36a)$$

Weiter liefert (6.27) die Bedingung

$$\mu_{K_i,V}^{e_i} + \mu_{K_{i+1},V}^{e_i} = 0 \quad \begin{cases} i = 1, \dots, n & \text{im Fall “I”} \\ i = 2, \dots, n & \text{in den Fällen “NN”, “ND”, “DD”} \end{cases} \quad (6.36b)$$

³Man nennt (6.31) auch die Equilibrierungsbedingung “nullter Ordnung”, weil das Funktional $v \mapsto R_K(v) + \int_{\partial K \setminus \Gamma_N} g_K v$ auf dem Raum \mathcal{P}_0 verschwinden muß; die Bedingung (6.34) heißt analog Equilibrierungsbedingung “erster Ordnung”, weil das Funktional auf \mathcal{P}_1 verschwinden muß.

Fall	# Elemente	# innere Kanten	# Momente	# Momente, die bereits festliegen	# Gleichungen	lösbar?	Lsg. eindeutig?
innerer Knoten	n	n	$2n$	0	$2n$	ja	nein
Neumann-Neumann	n	$n - 1$	$2n$	2	$2n + 1$	ja	ja
Neumann-Dirichlet	n	$n - 1$	$2n$	1	$2n$	ja	ja
Dirichlet-Dirichlet	n	$n - 1$	$2n$	0	$2n - 1$	ja	nein

Tabelle 6.1: Lösbarkeit von (6.36).

Zudem sind durch (6.26) festgelegt

$$\begin{cases} \text{Fall "I":} & \text{---} \\ \text{Fall "NN":} & \mu_{K_1, V}^{e_1}, \quad \mu_{K_n, V}^{e_{n+1}} \\ \text{Fall "ND":} & \mu_{K_n, V}^{e_{n+1}} \\ \text{Fall "DD":} & \text{---} \end{cases} \quad (6.36c)$$

Dies führt je nach Fall auf eine unterschiedliche Anzahl von Bedingungen—Tabelle 6.1 stellt die verschiedenen Fälle zusammen. Ohne die in Tabelle 6.1 aufgestellten Behauptungen über die Lösbarkeit bzw. eindeutige Lösbarkeit im einzelnen überprüfen zu wollen (siehe Übung 6.21 für den Fall "I"), wollen wir nur auf den Fall "ND" näher eingehen: Der Fall "NN" wirkt überbestimmt (mehr Gleichungen als Unbekannte)—dennoch existiert eine Lösung aufgrund der Galerkinorthogonalität der Lösung u_N , wie die folgende Bemerkung zeigt:

Bemerkung 6.20 Im Fall "NN" existiert eine Lösung von (6.36). Hierzu schreiben wir das System als $\mathbf{A}\boldsymbol{\mu} = \mathbf{f}$, wobei

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & & & & & \\ 1 & 1 & & & & & \\ 0 & 1 & 1 & & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & 1 & 1 \\ & & & & & & 1 \end{pmatrix} \in \mathbb{R}^{2n+1 \times 2n}, \quad \mathbf{f} = \begin{pmatrix} \int_{e_1} g_{K_1} \varphi_V \\ - \left(\int_{K_1} f \varphi_V - B_{K_1}(u_N, \varphi_V) \right) \\ 0 \\ - \left(\int_{K_2} f \varphi_V - B_{K_2}(u_N, \varphi_V) \right) \\ 0 \\ \vdots \\ 0 \\ \int_{e_{n+1}} g_{K_n} \varphi_V \end{pmatrix} \in \mathbb{R}^{2n+1}.$$

Dieses Gleichungssystem hat eine Lösung, falls $\mathbf{f} \perp \text{Ker} \mathbf{A}^\top$ (Dies folgt aus der linearen Algebra⁴). Eine einfache Rechnung zeigt, daß

$$\text{Ker} \mathbf{A}^\top = \text{span}\{\mathbf{k}\}, \quad \mathbf{k} = (1 \quad -1 \quad 1 \quad -1 \quad \dots \quad -1 \quad 1)^\top \in \mathbb{R}^{2n}$$

Wir berechnen $\mathbf{k} \cdot \mathbf{f}$:

$$\mathbf{k} \cdot \mathbf{f} = \sum_{k=1}^n \int_{K_i} f \varphi_V - B_{K_i}(u_N, \varphi_V) + \int_{e_1} g_{K_1} \varphi_V + \int_{e_{n+1}} g_{K_n} \varphi_V = l(\varphi_V) - B(u_N, \varphi_V) = R(\varphi_V) \stackrel{\text{Galerkin}}{=} 0$$

Damit ist die Lösbarkeit der der Momentengleichungen (6.36) in allen 4 Fällen gesichert. Da Eindeutigkeit nicht immer gegeben ist, müssen noch weitere Bedingungen gestellt werden. Hierzu erinnern wir uns an

⁴ Beh.: $Ax = f$ lösbar $\iff f \perp \text{Ker} A^\top$. Man zeigt dies in mehreren Schritten. 1. Schritt: Es gilt $R(A) \subset (\text{Ker} A^\top)^\perp$, denn $v \in \text{Ker} A^\top$ impliziert $\langle Ax, v \rangle = \langle x, A^\top v \rangle = 0$. Dies impliziert insbesondere den " \implies " Teil der Behauptung. 2. Schritt: Es gilt $R(A) = (\text{Ker} A^\top)^\perp$. Um dies zu sehen, nutzen wir, daß $R(A)$ ein abgeschlossener Teilraum von $(\text{Ker} A^\top)^\perp$ ist. Sei $f \in (\text{Ker} A^\top)^\perp$ mit $f \perp R(A)$. Dann ist $0 = \langle Ax, f \rangle = \langle x, A^\top f \rangle$ für alle x . Also ist $f \in \text{Ker} A^\top \cap (\text{Ker} A^\top)^\perp = \{0\}$. Damit ist $R(A) = (\text{Ker} A^\top)^\perp$.

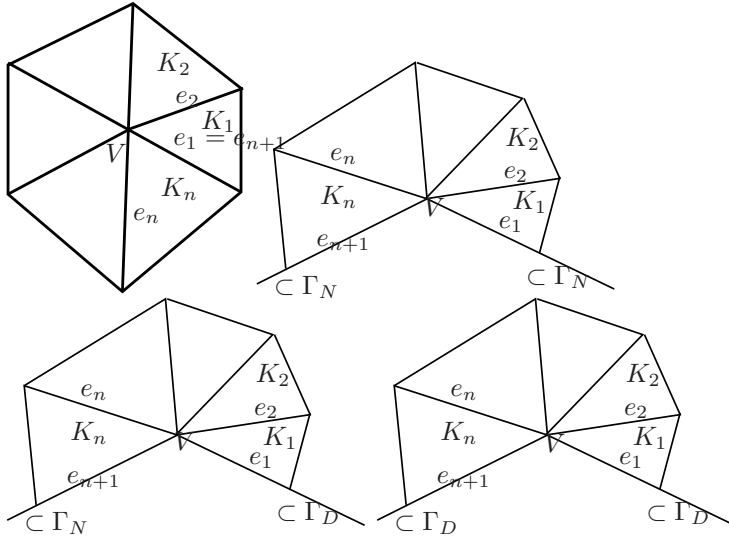


Abbildung 6.7: Für die Bestimmung der Momente $\mu_{K,V}^e$, die an einem festen Knoten $V \in \mathcal{N}(\mathcal{T})$ hängen gibt es 4 Fälle: ein innerer Knoten (“Fall I”), ein Randknoten, an dem zwei Neumannkanten anstoßen (Fall “NN”), ein Randknoten, an dem eine Neumann- und eine Dirichletkante anstoßen (“Fall DD”) und ein Randknoten, an dem zwei Dirichletkanten anstoßen (“Fall DD”).

die Forderung (6.25); in Ermangelung der genauen Kenntnis von $n \cdot \nabla u$ fordern wir, daß die Momente $\mu_{K,V}^e$ möglichst nah an

$$\tilde{\mu}_{K,V}^e := \int_e \varphi_V n_K \cdot \nabla u_N \quad (6.37)$$

sind. Damit ergeben sich für jeden Knoten $V \in \mathcal{N}(\mathcal{T})$ die Momente $\mu_{K,V}^e$, $K \subset \omega_V$ als eindeutige Lösungen des Minimierungsproblems

$$\text{Minimiere } \frac{1}{2} \sum_{K \subset \omega_V} \sum_{e \in \mathcal{E}(K): V \in \mathcal{N}(e)} |\tilde{\mu}_{K,V}^e - \mu_{K,V}^e|^2 \quad \text{unter den Nebenbedingungen (6.36)}. \quad (6.38)$$

Die Minimierungsaufgabe (6.38) kann natürlich auch als LGS formuliert werden—der Fall eines inneren Knotens ist der Inhalt der folgenden Übungsaufgabe.

Übung 6.21 (Siehe Appendix.) Betrachten Sie den Fall eines inneren Knotens $V \in \Omega$. Zeigen Sie:

- Im vorliegenden Fall “I” stellt (6.36) ein unterbestimmtes LGS dar; der Kern der zugehörigen Matrix hat Dimension 1.
- Geben Sie den Kern an.
- Formulieren Sie ein LGS zur Lösung von (6.38). Zeigen Sie, daß dieses vollen Rang hat.

■

Der oben konstruierte Fehlerschätzer stellt eine *obere Schranke* für den wahren Fehler $\|u - u_N\|_E$ dar. Zugleich überschätzt er den Fehler nicht zu sehr, wie das folgende Resultat zeigt:

Satz 6.22 *Gelte Voraussetzung 6.1. Dann existiert ein $C > 0$, welches nur von γ und Γ_D abhängt, so daß*

$$(\text{EST}^{\text{eq}})^2 \leq C \left[\|u - u_N\|_E^2 + \sum_{K \in \mathcal{T}} h_K^2 \|f - f_{\mathcal{T}}\|_{L^2(K)}^2 + \sum_{e \in \mathcal{E}_N} h_e \|g - g_{\mathcal{T}}\|_{L^2(e)}^2 \right],$$

wobei wie in Satz 6.6 die stückweise konstanten Funktionen $f_{\mathcal{T}}$ und $g_{\mathcal{T}}$ definiert sind als $f_{\mathcal{T}}|_K = \frac{1}{|K|} \int_K f$ und $g_{\mathcal{T}}|_e = \frac{1}{|e|} \int_e g$.

Beweis: Siehe [1]. □

6.5 Mittelungsbasierte Fehlerschätzer

Bei Ingenieuren sehr beliebt sich auch “mittelungsbasierte Fehlerschätzer”, oft auch einfach ZZ (“Zienkiewicz-Zhu”)-Fehlerschätzer genannt. Wesentliche Idee ist, durch (lokale) Projektion des Flusses $\nabla u_N \in (S^{0,0}(\mathcal{T}))^2$ auf $(S^{1,1}(\mathcal{T}))^2$ eine bessere Approximation $\mathbf{q}_N \in (S^{1,1}(\mathcal{T}))^2$ an ∇u zu erhalten. Der Fehlerschätzer ergibt sich dann als $\|\nabla u_N - \mathbf{q}_N\|_{L^2(\Omega)}$.

Der Einfachheit halber betrachten wir den Fall $\Gamma_N = \emptyset$, also

$$-\Delta u = f \quad \text{auf } \Omega, \quad u| = 0 \quad \text{auf } \partial\Omega. \quad (6.39)$$

Sei $u_N \in S_0^{1,1}(\mathcal{T})$ die FEM-Approximation an die Lösung von (6.39). Wir definieren

$$\eta_{M,Z}^2 := \inf_{\mathbf{q}_N \in (S^{1,1}(\mathcal{T}))^2} \|\nabla u_N - \mathbf{q}_N\|_{L^2(\Omega)}^2 \quad (6.40)$$

$$\eta_{M,Z}^2(K) := \|\nabla u_N - \mathbf{q}_N^*\|_{L^2(K)}^2 \quad \mathbf{q}_N^* \text{ ist der Minimierer aus (6.40)} \quad (6.41)$$

$$\eta_{Z,A}^2 := \|\nabla u_N - \mathcal{A}\nabla u_N\|_{L^2(\Omega)}^2 \quad (6.42)$$

$$\eta_{Z,A}^2(K) := \|\nabla u_N - \mathcal{A}\nabla u_N\|_{L^2(K)}^2 \quad (6.43)$$

wobei der *Mittelungsoperator* \mathcal{A} definiert ist als

$$\mathcal{A} : (L^2(\Omega))^2 \rightarrow (S^{1,1}(\mathcal{T}))^2, \quad \mathbf{q} \mapsto \sum_{V \in \mathcal{N}(\mathcal{T})} \varphi_V \frac{1}{|\omega_V|} \int_{\omega_V} \mathbf{q}. \quad (6.44)$$

Bemerkung 6.23 Das Auswerten von $\eta_{M,Z}$ erfordert immer noch das Lösen eines globalen Gleichungssystems und wird deshalb in der Praxis nicht gemacht. Der clémentartige Mittelungsoperator \mathcal{A} ist lokal und wird deshalb viel verwendet. ■

Satz 6.24 (Effizienz von $\eta_{M,Z}$)

$$\eta_{M,Z} \leq \|\nabla(u - u_N)\|_{L^2(\Omega)} + \min_{\mathbf{q} \in (S^{1,1}(\mathcal{T}))^2} \|\nabla u - \mathbf{q}\|_{L^2(\Omega)}.$$

Beweis: Sei $\mathbf{q} \in (S^{1,1}(\mathcal{T}))^2$ beliebig. Dann gilt:

$$\|\nabla u_N - \mathbf{q}_N^*\|_{L^2(\Omega)} \leq \inf_{\mathbf{q} \in (S^{1,1}(\mathcal{T}))^2} \|\nabla u_N - \mathbf{q}\|_{L^2(\Omega)} \leq \inf_{\mathbf{q} \in (S^{1,1}(\mathcal{T}))^2} \|\nabla(u - u_N)\|_{L^2(\Omega)} + \|\nabla u - \mathbf{q}\|_{L^2(\Omega)}. \quad \square$$

Bemerkung 6.25 In der Praxis ist der Term $\min_{\mathbf{q} \in (S^{1,1}(\mathcal{T}))^2} \|\nabla u - \mathbf{q}\|_{L^2(\Omega)}$ klein im Vergleich zu $\|\nabla(u - u_N)\|_{L^2(\Omega)}$, weil ∇u (bis auf einige Singularitäten) glatt ist. ■

Die Zuverlässigkeit von $\eta_{M,Z}$ ist schwieriger zu zeigen. Um sie zu zeigen, benötigen wir eine Variante des Clémentinterpolanten aus Satz 6.3: Ein wesentliches Hilfsmittel des Beweises ist Clémentinterpolant, der zusätzlich zu den Approximationseigenschaften aus Satz 6.3 folgende Orthogonalitätseigenschaft hat:

Lemma 6.26 (Clément-Interpolant mit zusätzlicher Orthogonalitätseigenschaft) *Es gelten die Voraussetzungen von Satz 6.3. Sei $\Gamma_D = \partial\Omega$. Dann kann der Interpolationsoperator $I : H_0^1(\Omega) \rightarrow S_0^{1,1}(\mathcal{T})$ so gewählt werden, daß die in Satz 6.3 genannten Approximationseigenschaften gelten und zudem für jedes $f \in L^2(\Omega)$ und $w \in H_0^1(\Omega)$ gilt:*

$$\left| \int_{\Omega} f(w - Iw) \right| \leq C \|\nabla w\|_{L^2(\Omega)} \sqrt{\sum_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega} h_V^2 \min_{f_V \in \mathbb{R}} \|f - f_V\|_{L^2(\omega'_V)}^2}, \quad (6.45)$$

wobei

$$\omega'_V := \omega_V \cup \bigcup_{V' \in \mathcal{N}(\mathcal{T}) \cap \partial\Omega \cap \overline{\omega_V}} \omega_{V'}.$$

Mithilfe des in Lemma 6.26 konstruierten Clémentinterpolanten kann nun die Zuverlässigkeit des Mittelungsfehlerschätzers gezeigt werden:

Satz 6.27 *Es existiert $C > 0$, so daß*

$$\|\nabla(u - u_N)\|_{L^2(\Omega)} \leq C \left[\eta_{M,Z} + \sqrt{\sum_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega} h_V^2 \inf_{f_V \in \mathbb{R}} \|f - f_V\|_{L^2(\omega'_V)}^2} \right].$$

Bemerkung 6.28 Bevor wir uns dem Beweis von Satz 6.27 zuwenden, bemerken wir, daß für “glatte” rechte Seiten f der Term $\sqrt{\sum_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega} h_V^2 \inf_{f_V \in \mathbb{R}} \|f - f_V\|_{L^2(\omega'_V)}^2}$ einen Term “höherer Ordnung” darstellt. Ist $f \in H^1(\Omega)$, so erwarten wir

$$\sum_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega} h_V^2 \inf_{f_V \in \mathbb{R}} \|f - f_V\|_{L^2(\omega'_V)}^2 \leq C \sum_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega} h_V^4 \|\nabla f\|_{L^2(\omega'_V)}^2 \leq Ch^4 \|f\|_{H^1(\Omega)}^2,$$

wobei $h = \max_{K \in \mathcal{T}} h_K$. Also liefert Satz 6.27

$$\|\nabla(u - u_N)\|_{L^2(\Omega)} \leq C [\eta_{M,Z} + h^2 \|f\|_{H^1(\Omega)}].$$

Für in der Praxis auftretenden Lösungen u und Gitter erwarten wir $\|\nabla(u - u_N)\|_{L^2(\Omega)} \geq Ch$. ■

Beweis von Satz 6.27: Für beliebiges $w \in H_0^1(\Omega)$ und $\mathbf{q} \in (S^{1,1}(\mathcal{T}))^2$ erhalten wir

$$\int_{\Omega} \nabla(u - u_N) \cdot \nabla w = \int_{\Omega} \nabla(u - u_N) \cdot \nabla(w - Iw) = \int_{\Omega} (\nabla u - \mathbf{q}) \cdot \nabla(w - Iw) + \int_{\Omega} (\mathbf{q} - \nabla u_N) \cdot \nabla(w - Iw)$$

Für den zweiten Term verwenden wir die Cauchy-Schwarzsche Ungleichung, für den ersten partielle Integration

$$\begin{aligned} \left| \int_{\Omega} (\nabla u_N - \mathbf{q}) \cdot \nabla(w - Iw) \right| &\leq \|\nabla u_N - \mathbf{q}\|_{L^2(\Omega)} \|\nabla(w - Iw)\|_{L^2(\Omega)} \leq C \|\nabla u_N - \mathbf{q}\|_{L^2(\Omega)} \|\nabla w\|_{L^2(\Omega)} \\ \int_{\Omega} (\nabla u - \mathbf{q}) \cdot \nabla(w - Iw) &= - \int_{\Omega} \nabla \cdot (\nabla u - \mathbf{q})(w - Iw) = \int_{\Omega} (f + \nabla \cdot \mathbf{q})(w - Iw) \\ &= \int_{\Omega} f(w - Iw) + \sum_{K \in \mathcal{T}} \int_K \nabla \cdot \mathbf{q}(w - Iw). \end{aligned}$$

Für den ersten Term verwenden wir nun (6.45) und erhalten

$$\left| \int_{\Omega} f(w - Iw) \right| \leq C \sqrt{\sum_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega} h_V^2 \inf_{f_V \in \mathbb{R}} \|f - f_V\|_{L^2(\omega'_V)}^2} \|\nabla w\|_{L^2(\Omega)};$$

aus dem zweiten Term wollen wir wieder einen Term der Form $\|\nabla u_N - \mathbf{q}\|_{L^2(\Omega)}$ erzeugen. Hierzu bemerken wir, daß $u_N \in S^{1,1}(\mathcal{T})$ impliziert, daß $\nabla \cdot \nabla u_N|_K = 0$ für jedes $K \in \mathcal{T}$. Also erhalten wir für jedes $K \in \mathcal{T}$

$$\left| \int_K \nabla \cdot \mathbf{q}(w - Iw) \right| = \left| \int_K \nabla \cdot (\mathbf{q} - \nabla u_N)(w - Iw) \right| \leq \|\nabla \cdot (\mathbf{q} - \nabla u_N)\|_{L^2(K)} \|w - Iw\|_{L^2(K)}$$

Die Approximationseigenschaften des Clémentinterpolanten liefern $\|w - Iw\|_{L^2(K)} \leq Ch_K \|\nabla w\|_{L^2(\tilde{\omega}_K)}$; die Tatsache, daß $\mathbf{q} \in (S^{1,1}(\mathcal{T}))^2$ erlaubt uns, die inverse Ungleichung

$$\|\nabla \cdot (\mathbf{q} - \nabla u_N)\|_{L^2(K)} \leq Ch_K^{-1} \|\mathbf{q} - \nabla u_N\|_{L^2(K)}$$

anzuwenden. Damit ergibt sich

$$\sum_{K \in \mathcal{T}} \left| \int_K \nabla \cdot \mathbf{q}(w - Iw) \right| \leq C \sum_{K \in \mathcal{T}} \|\mathbf{q} - \nabla u_N\|_{L^2(K)} \|\nabla w\|_{L^2(\tilde{\omega}_K)} \leq C \|\mathbf{q} - \nabla u_N\|_{L^2(\Omega)} \|\nabla w\|_{L^2(\Omega)},$$

wobei im letzten Schritt wiederum die Cauchy-Schwarzsche Ungleichung für Summen Anwendung fand. Fassen wir die obigen Abschätzungen zusammen, so erhalten wir für beliebiges $w \in H_0^1(\Omega)$ und beliebiges $\mathbf{q} \in (S^{1,1}(\mathcal{T}))^2$

$$\left| \int_{\Omega} \nabla(u - u_N) \cdot \nabla w \right| \leq C \|\nabla w\|_{L^2(\Omega)} \left[\|\nabla u_N - \mathbf{q}\|_{L^2(\Omega)} + \sqrt{\sum_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega} h_V^2 \inf_{f_V \in \mathbb{R}} \|f - f_V\|_{L^2(\omega'_V)}^2} \right]$$

Daraus ergibt sich die Behauptung. \square

Beweis von Lemma 6.26: Wir werden unten den Operator I explizit konstruieren. Beim Beweis des Lemmas werden wir uns jedoch auf den Beweis der Abschätzung (6.45) beschränken—die Approximationseigenschaften von I folgen mit Argumenten, die denen des Beweises von Satz 6.3 ähneln.

1. Schritt: Wesentliches Hilfsmittel ist eine Partition der Eins $(\psi_V)_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega}$, bei der die meisten Funktionen ψ_V mit den Hutfunktionen φ_V übereinstimmen. Genauer: für jeden *Randknoten* $V \in \mathcal{N}(\mathcal{T}) \cap \partial\Omega$ wähle man (willkürlich) einen inneren Knoten $I(V)$, der sich mit V ein Dreieck teilt, d.h. $I(V)$ erfüllt

$$\exists K \in \mathcal{T} \quad \text{s.d.} \quad V \in \mathcal{N}(K) \quad \text{und} \quad I(V) \in \mathcal{N}(K)$$

Weiter definieren wir für innere Knoten $V \in \mathcal{N}(\mathcal{T}) \cap \Omega$ die Funktionen ψ_V durch

$$\psi_V := \varphi_V + \sum_{z \in \mathcal{N}(\mathcal{T}) : V=I(z)} \varphi_z.$$

Wir bemerken, daß $\psi_V \neq \varphi_V$ nur für die Knoten in Randnähe gelten kann; weiter bemerken wir, daß $\text{supp } \psi_V \subset \omega'_V$.

2. Schritt: Wir definieren den Operator I durch

$$Iw := \sum_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega} \varphi_V l_V(w), \quad l_V(w) := \frac{\int_{\Omega} \psi_V w}{\int_{\Omega} \varphi_V}$$

3. Schritt: Wir behaupten nun, daß

$$\|\psi_V w - \varphi_V l_V(w)\|_{L^2(\Omega)} \leq Ch_V \|\nabla w\|_{L^2(\omega'_V)}. \quad (6.46)$$

Um dies einzusehen, betrachten wir die Fälle $\psi_V = \varphi_V$ und $\psi_V \neq \varphi_V$ getrennt. Sei zuerst $\psi_V = \varphi_V$. Dann ist

$$l_V(w) = \frac{\int_{\omega_V} \varphi_V w}{\int_{\omega_V} \varphi_V}$$

Weiter ist l_V ein lineares Funktional, für welches $l_V(1) = 1$ gilt. Damit können wir für den Mittelwert $\bar{w}_V := \frac{1}{|\omega_V|} \int_{\omega_V} w$ abschätzen:

$$\|\psi_V w - \varphi_V l_V(w)\|_{L^2(\Omega)} = \|\varphi_V(w - l_V(w))\|_{L^2(\omega_V)} = \|\varphi_V((w - \bar{w}_V) - l_V(w - \bar{w}_V))\|_{L^2(\omega_V)}.$$

Weil $|\varphi_V| \leq 1$ gilt, folgt $|\varphi_V| \leq \sqrt{|\varphi_V|}$. Somit wollen wir

$$\|\sqrt{\varphi_V}(z - l_V(z))\|_{L^2(\omega_V)}, \quad z := w - \bar{w}_V$$

abschätzen. Um die Struktur des Vorgehens besser zu sehen, führen wir die (positiv semidefinite) Bilinearform $\langle u, v \rangle_{\varphi} := \int_{\Omega} \varphi_V uv$ ein. Wir bemerken, daß l_V so definiert ist, daß die Orthogonalität

$$\langle z - l_V(z), 1 \rangle_{\varphi} = 0$$

gilt. Damit ergibt sich

$$\|\sqrt{\varphi_V}(z - l_V(z))\|_{L^2(\omega_V)}^2 = \langle z - l_V(z), z - l_V(z) \rangle_{\varphi} = \langle z, z - l_V(z) \rangle_{\varphi} \leq \|z\|_{\varphi} \|z - l_V(z)\|_{\varphi},$$

was wiederum

$$\|\sqrt{\varphi_V}(z - l_V(z))\|_{L^2(\omega_V)} \leq \|\sqrt{\varphi_V} z\|_{L^2(\omega_V)} \leq \|z\|_{L^2(\omega_V)} = \|w - \bar{w}_V\|_{L^2(\omega_V)}$$

nach sich zieht. Wenn man wie im Beweis von Satz 6.3 argumentiert⁵, dann ergibt sich $\|w - \bar{w}_V\|_{L^2(\omega_V)} \leq Ch_V \|\nabla w\|_{L^2(\omega_V)}$, wobei die Konstante $C > 0$ nur von der Formregularitätskonstante des Gitters abhängt. Damit hätten wir (6.46) für den Fall $\psi_V = \varphi_V$ gezeigt.

Wir betrachten nun den Fall $\psi_V \neq \varphi_V$. In diesem Fall besteht $\partial\omega'_V \cap \partial\Omega$ aus mindestens einer Kante. Durch Argumente wie im Beweis von Satz 6.3 kann man nun die Voraussetzung $w \in H_0^1(\Omega)$ ausnutzen, um zu schließen⁶

$$\|w\|_{L^2(\omega'_V)} \leq Ch_V \|\nabla w\|_{L^2(\omega'_V)}. \quad (6.47)$$

Damit ergibt sich mit der Formregularität des Gitters

$$\begin{aligned} \|\psi_V w - \varphi_V l_V(w)\|_{L^2(\Omega)} &= \|\psi_V w - \varphi_V l_V(w)\|_{L^2(\omega'_V)} \leq \|\psi_V w\|_{L^2(\omega'_V)} + \|\varphi_V l_V(w)\|_{L^2(\omega'_V)} \\ &\leq \|w\|_{L^2(\omega'_V)} + \sqrt{|\omega'_V|} |l_V(w)| \leq Ch_V \|\nabla w\|_{L^2(\omega'_V)} + \sqrt{|\omega'_V|} \frac{\|\psi_V\|_{L^2(\omega'_V)} \|w\|_{L^2(\omega'_V)}}{\|\varphi_V\|_{L^1(\omega'_V)}} \end{aligned}$$

Nun ist $\|\varphi_V\|_{L^1(\omega'_V)} \sim h_V^2$ und $|\omega'_V| \sim h_V^2$ und $\|\psi_V\|_{L^2(\omega'_V)} \sim h_V$. Damit ergibt sich mit Hilfe von (6.47) die gewünschte Behauptung (6.46) auch im Falle $\psi_V \neq \varphi_V$.

4. *Schritt:* Sei $f \in L^2(\Omega)$. Um die im Lemma behauptete Abschätzung zu zeigen, nutzen wir aus, daß die Funktionen ψ_V eine Partition der Eins liefern. Sei weiter für jeden inneren Knoten $V \in \mathcal{N}(\mathcal{T}) \cap \Omega$ ein $f_V \in \mathbb{R}$ gewählt. Dann gilt:

$$\begin{aligned} \left| \int_{\Omega} f(w - Iw) \right| &= \left| \int_{\Omega} f \sum_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega} \psi_V w - \varphi_V l_V(w) \right| = \left| \sum_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega} \int_{\Omega} f(\psi_V w - \varphi_V l_V(w)) \right| \\ &= \left| \sum_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega} \int_{\Omega} (f - f_V)(\psi_V w - \varphi_V l_V(w)) \right| \leq \sum_{V \in \mathcal{N}(\mathcal{T}) \cap \Omega} \|f - f_V\|_{L^2(\omega'_V)} \|\psi_V w - \varphi_V l_V(w)\|_{L^2(\omega'_V)} \end{aligned}$$

Verwendet man nun (6.46) sowie die Cauchy-Schwarzsche Ungleichung für Summen, so ergibt sich die gewünschte Behauptung des Lemma. \square

⁵im Fall von regelmäßigen Gittern entspricht dies der 2. Poincaré-Ungleichung

⁶von der Idee her ist dies die 1. Poincaré-Ungleichung

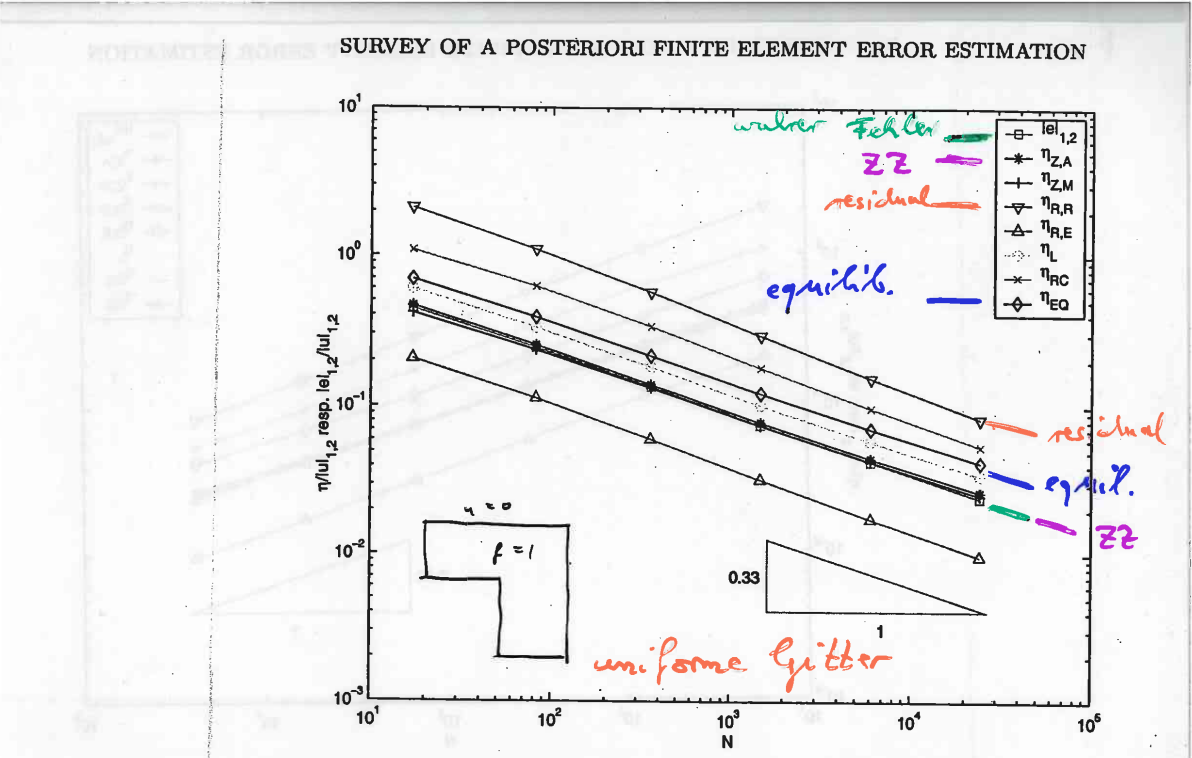


FIGURE 1. Relative error and error estimators for uniform mesh refinement.

ans: Carstensen, Bartels, Klöbe vi: survey ...

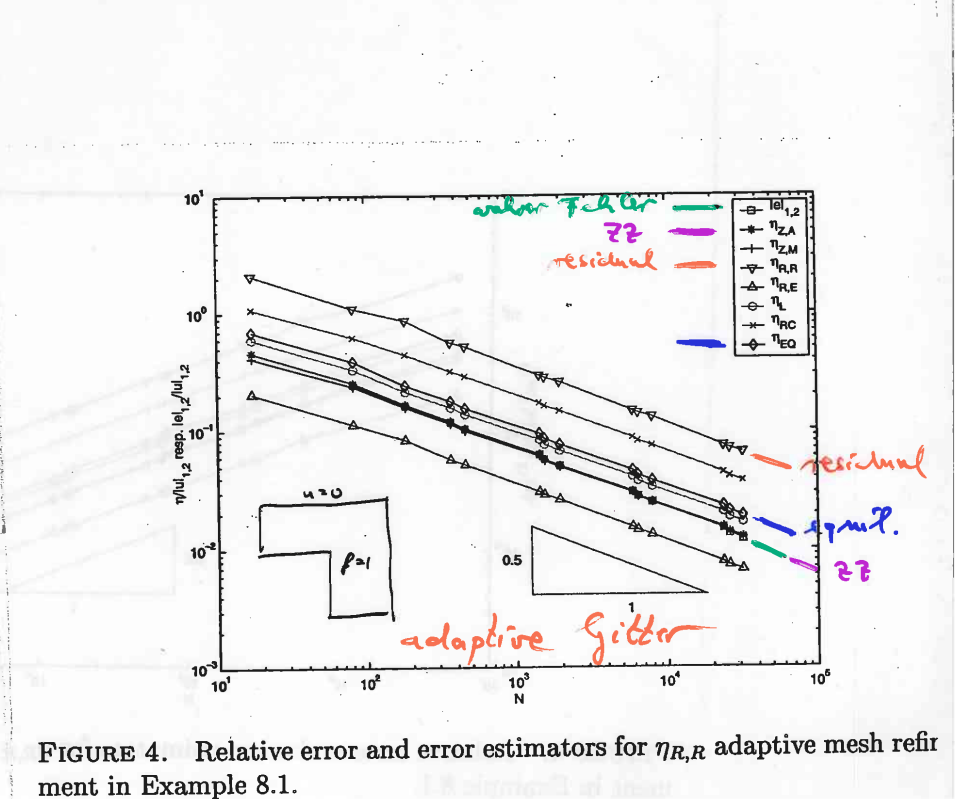


FIGURE 4. Relative error and error estimators for $\eta_{R,R}$ adaptive mesh refinement in Example 8.1.

6.6 Konvergenz von adaptiven Verfahren

Die numerischen Beispiele zeigen die Leistungsfähigkeit des adaptiven Algorithmus 3. Tatsächlich kann man für eine große Klasse von Problemen zeigen, daß diese Algorithmus eine Folge von FEM-Approximationen erzeugt, die gegen die exakte Lösung konvergiert und daß diese Konvergenz sogar mit der bestmöglichen Rate (im Sinne von “Fehler gegen Problemgröße”) erfolgt. Wir betrachten wieder das Modellproblem (6.1).

Vorab machen wir eine kleine Notationsänderung, die sich für den Beweis des folgende Satzes 6.29 als hilfreich erweisen wird:

- für jedes $K \in \mathcal{T}$ definieren wir $h_K := |K|^{1/2}$ und für jedes $e \in \mathcal{E}$ definieren wir $h_e := |e|$. Wegen der Formregulärität der Gitter ist das so definierte h_K von der Größenordnung von $\text{diam } K$.

Die wesentliche Eigenschaft dieser Definition von h_K ist, daß, falls ein Element K verfeinert, sich die Elementgröße um einen festen Faktor reduziert. Bei “newest vertex bisection” z.B. gilt bei Halbieren eines Elementes K für die “Söhne” K_1 und K_2 , daß $h_{K_1} = h_{K_2} = 2^{-1/2}h_K$.

- mit dem so definierten h_K werden die Fehlerindikatoren η_K nun so definiert:

$$\eta_K^2 := h_K^2 \|f + \Delta u_N\|_{L^2(K)}^2 + \frac{1}{2} \sum_{e \in \mathcal{E}(K) \cap \Omega} h_K \|[\partial_n u_N]\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}(K) \cap \Gamma_N} h_K \|g - \partial_n u_N\|_{L^2(e)}^2.$$

Wir bemerken an dieser Stelle bereits, daß der so definierte Fehlerindikator bis auf Konstante äquivalent zum ursprünglich definierten ist.

- Für eine Menge von Elementen \mathcal{M} definieren wir $\eta(\mathcal{M})$ durch

$$\eta(\mathcal{M})^2 := \sum_{K \in \mathcal{M}} \eta_K^2.$$

- Um die Abhängigkeit des Schätzers $\eta(\mathcal{T})$ von seinem Argument hervorzuheben, schreiben wir manchmal ausführlich

$$\begin{aligned} \eta(v, \mathcal{T})^2 &:= \sum_{K \in \mathcal{T}} h_K^2 \|f + \Delta v\|_{L^2(K)}^2 + \frac{1}{2} \sum_{e \in \mathcal{E}(K) \cap \Omega} h_K \|[\partial_n v]\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}(K) \cap \Gamma_N} h_K \|g - \partial_n v\|_{L^2(e)}^2, \\ \eta_K(v)^2 &:= h_K^2 \|f + \Delta v\|_{L^2(K)}^2 + \frac{1}{2} \sum_{e \in \mathcal{E}(K) \cap \Omega} h_K \|[\partial_n v]\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}(K) \cap \Gamma_N} h_K \|g - \partial_n v\|_{L^2(e)}^2. \end{aligned}$$

Wir erinnern ferner daran, daß die Markierungsstrategie in Alg. 3 das sog. “Dörflermarkieren” ist, d.h. zu gewähltem $\theta \in (0, 1)$ ist $\mathcal{M} \subset \mathcal{T}$ die minimale Menge, für die gilt:

$$\theta \eta^2 = \theta \eta^2(\mathcal{T}) \leq \eta^2(\mathcal{M}) \quad (6.48)$$

Satz 6.29 *Sei der Fehlerschätzer mit $h_K := |K|^{1/2}$ definiert. Sei $\theta \in (0, 1)$. Dann existieren Konstanten $\gamma > 0$ und $q \in (0, 1)$ so daß für die von Alg. 3 erzeugte Folge $(u_n)_{n \in \mathbb{N}}$ von FEM-Approximationen an die exakte Lösung u gilt:*

$$\left(\|\nabla(u - u_{n+1})\|_{L^2(\Omega)}^2 + \gamma \eta_{n+1}^2 \right)^{1/2} \leq q \left(\|\nabla(u - u_n)\|_{L^2(\Omega)}^2 + \gamma \eta_n^2 \right)^{1/2}.$$

Hier ist $\eta_n := \eta(\mathcal{T}_n)$. Insbesondere gilt deshalb:

$$\lim_{n \rightarrow \infty} \|\nabla(u - u_n)\|_{L^2(\Omega)} = 0.$$

Beweis: Es erweist sich als hilfreich, die Abkürzungen

$$e_n := \|\nabla(u - u_n)\|_{L^2(\Omega)}, \quad E_n := \|\nabla(u_{n+1} - u_n)\|_{L^2(\Omega)},$$

einzuführen. Weiters kürzen wir ab für den FEM-Raum zu Gitter \mathcal{T}_n :

$$V_n := S^{1,1}(\mathcal{T}_n) \cap H_D^1(\Omega, \Gamma_D).$$

Wegen der Einbettung der FEM-Räume $V_n \subset V_{n+1}$ haben wir die Galerkinorthogonalität

$$e_n^2 = e_{n+1}^2 + E_n^2.$$

1. *Schritt*: Wir behaupten, daß folgende “Dreiecksungleichung” gilt:

$$\eta_n(v) \leq \eta_n(w) + C_{\mathcal{T}} \|\nabla(v-w)\|_{L^2(\Omega)} \quad \forall v, w \in V_n, \quad (6.49)$$

wobei wir $\eta_n(v)$ für $\eta(v, \mathcal{T}_n)$ schreiben. Hier ist $C_{\mathcal{T}}$ eine Konstante, die nur von der Formregularität der Gitter abhängt. Um das zu sehen, berechnen wir (wir nutzen nun $\Delta v = \Delta w = 0$ auf jedem Element)

$$\begin{aligned} \eta_n(v) &= \left(\sum_{K \in \mathcal{T}_n} h_K^2 \|f\|_{L^2(K)}^2 + \sum_{e \in \mathcal{E}(K) \cap \Omega} h_K \|\partial_n v\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}(K) \cap \Gamma_N} h_K \|g - \partial_n v\|_{L^2(e)}^2 \right)^{1/2} \\ &\leq \left(\sum_{K \in \mathcal{T}_n} h_K^2 \|f\|_{L^2(K)}^2 + \sum_{e \in \mathcal{E}(K) \cap \Omega} h_K \|\partial_n w\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}(K) \cap \Gamma_N} h_K \|g - \partial_n w\|_{L^2(e)}^2 \right)^{1/2} \\ &\quad + \left(\sum_{K \in \mathcal{T}_n} \sum_{e \in \mathcal{E}(K) \cap \Omega} h_K \|\partial_n(v-w)\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}(K) \cap \Gamma_N} h_K \|\partial_n(v-w)\|_{L^2(e)}^2 \right)^{1/2}, \end{aligned}$$

wobei wir hier Varianten der klassischen Dreiecksungleichung verwendet haben (Übung). Man kann sich überlegen (weitere Übung!), daß auf dem Raum V_n folgende inverse Ungleichung gilt:

$$h_K \|\partial_n z\|_{L^2(e)}^2 \leq C' \|\nabla z\|_{L^2(\omega_e)}^2 \quad \forall z \in V_n$$

für ein C' , welches nur von der Formregularität von \mathcal{T} abhängt. Damit können wir weiters mit $z = v - w$ abschätzen

$$\eta_n(v) \leq \eta_n(w) + \sqrt{3C'} \|\nabla(v-w)\|_{L^2(\Omega)}.$$

2. *Schritt*: Beim Übergang von \mathcal{T}_n zu \mathcal{T}_{n+1} werden einige Elemente von \mathcal{T}_n verfeinert (sei es, daß sie markiert sind, sei es, daß sie verfeinert werden, um ein konformes Gitter \mathcal{T}_{n+1} zu erhalten). Entsprechend zerlegen wir

$$\mathcal{T}_n = \mathcal{R}_n \dot{\cup} \mathcal{U}_n,$$

wobei $\mathcal{U}_n = \mathcal{T}_{n+1} \cap \mathcal{T}_n$ die Menge der Elemente ist, die nicht verfeinert werden und \mathcal{R}_n die derjenigen, die verfeinert werden (“ \mathcal{R} ” steht für *refined* und “ \mathcal{U} ” für *unrefined*). Für jedes $K \in \mathcal{R}_n$ fassen wir diese “Söhne” zusammen in der Menge $\mathcal{S}_n(K)$, d.h.

$$\mathcal{S}_n(K) := \{K' \in \mathcal{T}_{n+1} : K' \subset K\}. \quad (6.50)$$

Die Menge $\mathcal{S}_n(K)$ ist ein Triangulierung von K :

$$\overline{K} = \cup_{K' \in \mathcal{S}_n(K)} \overline{K'} \quad (6.51)$$

mit der wesentlichen Eigenschaft, daß für jedes $K' \in \mathcal{S}_n(K)$ gilt: $h_{K'} \leq 2^{-1/2} h_K$. Daraus folgt nun für $K \in \mathcal{R}_n$ und $v \in V_n$:

$$\begin{aligned} \eta^2(v, \mathcal{S}_n(K)) &= \sum_{K' \in \mathcal{S}_n(K)} \left(h_{K'}^2 \|f\|_{L^2(K')}^2 + \frac{1}{2} \sum_{e \in \mathcal{E}(K') \cap \Omega} h_{K'} \|\partial_n v\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}(K') \cap \Gamma_N} h_{K'} \|g - \partial_n v\|_{L^2(e)}^2 \right) \\ &\leq 2^{-1/2} \left(h_K^2 \|f\|_{L^2(K)}^2 + \frac{1}{2} \sum_{e \in \mathcal{E}(K) \cap \Omega} h_K \|\partial_n v\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}(K) \cap \Gamma_N} h_K \|g - \partial_n v\|_{L^2(e)}^2 \right) \\ &= 2^{-1/2} \eta_K^2(v) \end{aligned}$$

3. *Schritt*: Es gilt die folgende Schätzerreduktion: Es existiert ein $\rho \in (0, 1)$ so daß für jedes $\delta > 0$ ein $C_\delta > 0$ existiert mit

$$\eta_{m+1}^2 \leq (1 + \delta)\rho\eta_n^2 + C_\delta E_n^2. \quad (6.52)$$

Wir rechnen zuerst für die FEM-Approximation u_n auf Gitter \mathcal{T}_n wegen $\mathcal{T}_{n+1} = \mathcal{U}_n \cup (\cup_{K \in \mathcal{R}_n} \mathcal{S}_n(K))$

$$\begin{aligned} \eta^2(u_n, \mathcal{T}_{n+1}) &= \eta^2(u_n, \mathcal{U}_n) + \eta^2(u_n, \cup_{K \in \mathcal{R}_n} \mathcal{S}_n(K)) \\ &= \eta^2(u_n, \mathcal{U}_n) + \sum_{K \in \mathcal{R}_n} \eta^2(u_n, \mathcal{S}_n(K)) \\ &\leq \eta^2(u_n, \mathcal{U}_n) + 2^{-1/2} \sum_{K \in \mathcal{R}_n} \eta^2(u_n, \{K\}) \\ &= \eta^2(u_n, \mathcal{T}_n) - \eta^2(u_n, \mathcal{R}_n) + 2^{-1/2} \eta^2(u_n, \mathcal{R}_n) = \eta^2(u_n, \mathcal{T}_n) - (1 - 2^{-1/2})\eta^2(u_n, \mathcal{R}_n) \\ &\leq \eta^2(u_n, \mathcal{T}_n) - (1 - 2^{-1/2})\eta^2(u_n, \mathcal{M}_n) \\ &\leq \underbrace{(1 - (1 - 2^{-1/2})\theta)}_{=: \rho} \eta_n^2, \end{aligned}$$

wobei wir im vorletzten Schritt $\mathcal{M}_n \subset \mathcal{R}_n$ und im letzten Schritt die Dörflermarkierung ausgenutzt haben. Die gewünschte Abschätzung folgt nun mit dem 1. Schritt und der Beobachtung $(a + b)^2 \leq (1 + \delta)a^2 + (1 + \delta^{-1})b^2$ für beliebige $a, b \geq 0$ und $\delta > 0$:

$$\begin{aligned} \eta_{m+1}^2 &= \eta^2(u_{n+1}, \mathcal{T}_{n+1}) \leq (\eta_{n+1}(u_n, \mathcal{T}_{n+1}) + C_{\mathcal{T}} \|\nabla(u_{n+1} - u_n)\|_{L^2(\Omega)})^2 \\ &\leq (1 + \delta)\eta_{m+1}^2(u_n, \mathcal{T}_{n+1}) + (1 + \delta^{-1})C_{\mathcal{T}}^2 \|\nabla(u_{n+1} - u_n)\|_{L^2(\Omega)}^2 \\ &\leq (1 + \delta)\rho\eta_n^2 + (1 + \delta^{-1})C_{\mathcal{T}}^2 E_n^2 \end{aligned}$$

4. *Schritt*: Sei $\gamma > 0$ beliebig (wird später gewählt). Dann folgt aus der Galerkinorthogonalität

$$e_{n+1}^2 + \gamma\eta_{m+1}^2 = e_n^2 - E_n^2 + \gamma\eta_{m+1}^2 \leq e_n^2 + (\gamma C_\delta - 1)E_n^2 + \gamma(1 + \delta)\rho\eta_n^2.$$

Unter der Annahme, daß $C_\delta\gamma < 1$, folgt daraus für beliebiges $\beta > 0$

$$\begin{aligned} e_{n+1}^2 + \gamma\eta_{m+1}^2 &\leq e_n^2 + \gamma(1 + \delta)\rho\eta_n^2 \\ &\leq e_n^2 - \beta\gamma\eta_n^2 + \gamma((1 + \delta)\rho + \beta)\eta_n^2. \end{aligned}$$

Verwendet man nun die Zuverlässigkeit des Fehlerschätzers (vgl. Satz 6.4) in der Form

$$e_n \leq C_{rel}\eta_n,$$

so ergibt sich

$$e_{n+1}^2 + \gamma\eta_{m+1}^2 \leq (1 - \beta\gamma C_{rel}^{-2})e_n^2 + \gamma((1 + \delta)\rho + \beta)\eta_n^2 \leq \underbrace{\max\{1 - \beta\gamma C_{rel}^{-2}, (1 + \delta)\rho + \beta\}}_{=: q^2} (e_n^2 + \gamma\eta_n^2)$$

Um den Beweis abzuschließen müssen wir nun die Konstanten γ, δ, β geeignet wählen. Es sind die folgenden Bedingungen zu erfüllen:

$$\begin{aligned} C_\delta\gamma &< 1 \\ (1 + \delta)\rho + \beta &< 1 \end{aligned}$$

Dies geschieht, indem zuerst δ so klein gewählt wird, daß $(1 + \delta)\rho < 1$. Anschließend wird $\gamma > 0$ so klein gewählt, daß $C_\delta\gamma < 1$. Schlußendlich wird $\beta > 0$ so gewählt, daß $(1 + \delta)\rho + \beta < 1$. \square

Wir schließen den Konvergenzbeweis ab mit der Beobachtung, daß Satz 6.29 auch die Konvergenz von Alg. 3 nach sich zieht, wenn der residuale Fehlerschätzer die Elementgröße h_K so definiert, daß $h_K \sim \text{diam } K$ gilt:

Korollar 6.30 Sei $\theta \in (0, 1)$. Sei die Elementgröße h_K so definiert, daß $h_K \sim \text{diam } K$ gilt. Dann gilt für die Folge $(u_n)_{n \in \mathbb{N}}$ der FEM-Approximationen

$$\lim_{n \rightarrow \infty} \|\nabla(u - u_n)\|_{L^2(\Omega)} = 0.$$

Beweis: Wir bezeichnen mit η_n den Fehlerschätzer, für den lediglich $h_K = \text{diam } K$ gilt und mit $\tilde{\eta}_n$ den Fehlerschätzer, der auf $h_K = |K|^{1/2}$ basiert. Die Kernbeobachtung ist, daß eine Konstante $\lambda > 0$ gibt, die nur auf der Formregularität der Gitter abhängt, so daß für alle n gilt:

$$\lambda^{-1}\eta_n \leq \tilde{\eta}_n \leq \lambda\eta_n. \quad (6.53)$$

Hier können wir $\lambda \geq 1$ annehmen. Die Mengen \mathcal{M}_n von markierten Elementen, die Alg. 3 erzeugt, erfüllen

$$\theta\eta_n^2 \leq \eta_n^2(\mathcal{M}_n).$$

Damit folgt mit (6.53)

$$(\theta\lambda^{-2})\tilde{\eta}_n^2 \leq \theta\eta_n^2 \leq \eta_n^2(\mathcal{M}_n) \leq \lambda^2\tilde{\eta}_n^2(\mathcal{M}_n),$$

so daß sich ergibt mit $\theta' := \theta\lambda^{-4} \in (0, 1)$

$$\theta'\tilde{\eta}_n^2 \leq \tilde{\eta}_n^2(\mathcal{M}_n).$$

Diese Bedingung war die entscheidende Bedingung für den Beweis von Satz 6.29. Damit liefert Satz 6.29 die Existenz von $q \in (0, 1)$ und $\gamma > 0$ so daß

$$\left(\|\nabla(u - u_{n+1})\|_{L^2(\Omega)} + \gamma\tilde{\eta}_{n+1}^2\right)^{1/2} \leq q \left(\|\nabla(u - u_n)\|_{L^2(\Omega)} + \gamma\tilde{\eta}_n^2\right)^{1/2},$$

was die gewünschte Konvergenzaussage impliziert. □