Chapter 4

A Posteriori Analysis

4.1 Introduction

We consider the model problem

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma_D, \\ \partial u/\partial n &= \phi \quad \text{on } \Gamma_N. \end{aligned} \tag{4.1}$$

Let $u \in H^1_D(\Omega)$ be the weak solution of (4.1) and $u_h \in \mathcal{S}^1_D(\mathcal{T})$ be the P1-FEM approximation of u. In the previous chapter, we aimed to control the error $\|u-u_h\|_{H^1(\Omega)}$ by a priori knowledge, e.g., regularity of the given data and the exact solution (but essentially independent of the discrete solution u_h). Since u is unknown, in general, the **a priori analysis** provides a qualitative understanding of the FEM, e.g., convergence with certain rates, but the derived bounds are non-computable in practice. In this chapter, we aim to derive *numerically computable* bounds $\eta = \eta(u_h, f, \phi, \mathcal{T})$ for the error $\|u-u_h\|_{H^1(\Omega)}$, which may depend on u_h, the triangulation $\mathcal{T}$, and the given data f and ϕ (but *not* on the exact solution u). The quantity η is referred to as **(a posteriori) error estimator**, and emphasis is laid on the fact that η can be computed algorithmically as soon as the discrete solution $u_h \in \mathcal{S}^1_D(\mathcal{T})$ has been computed. An error estimator η is called **reliable** provided that

$$\|u - u_h\|_{H^1(\Omega)} \le C_{\rm rel}\, \eta. \tag{4.2}$$

Usually, the information η provides, is used to steer a mesh-refinement that leads to a sequence $\mathcal{T}_\ell$ of regular meshes with nested spaces $\mathcal{S}^1_D(\mathcal{T}_\ell) \subseteq \mathcal{S}^1_D(\mathcal{T}_{\ell+1})$, i.e., $\mathcal{T}_{\ell+1}$ is a certain refinement of $\mathcal{T}_\ell$. If η is reliable the (numerically or algorithmically observed) decrease of η to zero implies the convergence of u_h towards u. However, it might (formally) occur that u_h tends to u, while η does not tend to zero. Therefore, an error estimator η is called **efficient** provided that

$$C_{\rm eff}\, \eta \le \|u - u_h\|_{H^1(\Omega)}. \tag{4.3}$$

For an efficient error estimator η, the convergence of u_h to u necessarily implies the convergence of η to zero. Finally, if η is reliable and efficient, we observe for η the same order of convergence as for $\|u - u_h\|_{H^1(\Omega)}$.

The aim of a posteriori error estimates is twofold:

- We want to control the accuracy $\|u - u_h\|_{H^1(\Omega)}$ of a discrete solution u_h and stop the computation if u_h is sufficiently accurate.
- The mesh-refinement should be steered automatically by the algorithm so that we are led to the highest possible accuracy with the lowest number of degrees of freedom.

Remark. Throughout, we allow the cases $\Gamma_D = \Gamma$ as well as $\Gamma_D = \emptyset$. In the latter case, (4.1) becomes the Neumann problem, for which we have to assume the compatability condition $\int_\Omega f\,dx + \int_\Gamma \phi\,ds = 0$. Then, $u \in H^1_*(\Omega)$ and, even more important, the test space $H^1_*(\Omega)$ in the weak formulation can equivalently be replaced by the entire space $H^1(\Omega)$. The same holds for the P1-FEM, where $u_h \in \mathcal{S}^1_*(\mathcal{T})$ and where the discrete test is $\mathcal{S}^1_*(\mathcal{T})$ or equivalently $\mathcal{S}^1(\mathcal{T})$. □

4.2 Scott-Zhang Projection

Since H^1-functions are in general not continuous, nodal interpolation requires additional regularity assumptions. In this section, we aim to provide some quasi-interpolation operator which is well-defined for all $u \in H^1(\Omega)$ and also has the projection property. We start with the following elementary lemma

Lemma 4.1. *For $z \in \mathcal{K}$, choose an edge $E_z \in \mathcal{E}$ with $z \in E_z$. Then, there is a unique dual function $\psi_z \in \mathcal{P}^1(E_z)$ such that*

$$\int_{E_z} \psi_z \zeta_{z'}\,ds = \delta_{zz'} \quad \text{for all } z' \in \mathcal{K}. \tag{4.4}$$

Moreover, it holds that $\|\psi_z\|_{L^\infty(E_z)} \le C\,|E_z|^{-1}$ for some generic constant $C > 0$, which is in particular independent of z and $\mathcal{T}$.

Proof. According to the Riesz theorem, there is a unique function $\widehat{\psi} \in \mathcal{P}^1[0,1]$ such that

$$\int_0^1 \widehat{\psi}\,\widehat{\phi}\,dt = \widehat{\phi}(0) \quad \text{for all } \widehat{\phi} \in \mathcal{P}^1[0,1].$$

Let $\Phi_z : [0,1] \to E_z$ be an affine parametrization of the edge E_z with $\Phi_z(0) = z$. We define

$$\psi_z := \frac{1}{|E_z|}\,\widehat{\psi} \circ \Phi_z^{-1} \in \mathcal{P}^1(E_z).$$

Clearly, $\|\psi_z\|_{L^\infty(E_z)} \le \|\widehat{\psi}\|_{L^\infty(0,1)}\,|E_z|^{-1}$. Note that $|\Phi_z'| = |E_z|$ and hence

$$\int_{E_z} \psi_z \zeta_{z'}\,ds = \int_0^1 (\psi_z \circ \Phi_z)\,(\zeta_{z'} \circ \Phi_z)\,|\Phi_z'|\,dt = \int_0^1 \widehat{\psi}(t)\,(\zeta_{z'} \circ \Phi_z)(t)\,dt = \zeta_{z'}\big(\Phi_z(0)\big) = \zeta_{z'}(z).$$

This concludes the proof. ■

Definition. For each node $z \in \mathcal{K}$ of $\mathcal{T}$, we choose an edge $E_z \in \mathcal{E}$ such that

- $E_z \subseteq \overline{\Gamma}_D$ for $z \in \overline{\Gamma}_D$,

- $E_z \subseteq \Gamma$ for $z \in \Gamma$,
- E_z arbitrary for $z \in \Omega$.

Note that the precise choice is immaterial for the following analysis. For $z \in \mathcal{K}$, let $\psi_z \in \mathcal{P}^1(E_z)$ be the corresponding dual function. Then, the **Scott-Zhang projection** is defined by

$$J_h v := \sum_{z\in\mathcal{K}} \Big(\int_{E_z} \psi_z v \, ds \Big) \zeta_z. \tag{4.5}$$

Clearly, $J_h : H^1(\Omega) \to \mathcal{S}^1(\mathcal{T})$ is well-defined and linear. Our first proposition states that J_h is in fact a projection which preserves discrete boundary data.

Proposition 4.2. *For $v \in H^1(\Omega)$ and $v_h \in \mathcal{S}^1(\mathcal{T})$, the following properties* (i)–(iii) *are true:*

(i) $J_h v_h = v_h$.

(ii) $(J_h v)|_\omega$ *depends only on the trace* $v|_\omega$ *for* $\omega \in \{\Gamma, \Gamma_D\}$.

(iii) $v|_\omega = v_h|_\omega$ *implies that* $(J_h v)|_\omega = v|_\omega$ *for* $\omega \in \{\Gamma, \Gamma_D\}$.

Proof. (i) Note that $v_h = \sum_{z'\in\mathcal{K}} v_h(z')\zeta_{z'}$. By choice of ψ_z, this shows

$$\int_{E_z} \psi_z v_h \, ds = \sum_{z'\in\mathcal{K}} v_h(z') \int_{E_z} \psi_z \zeta_{z'} \, ds = v_h(z).$$

With this, we deduce

$$J_h v_h = \sum_{z\in\mathcal{K}} \Big(\int_{E_z} \psi_z v_h \, ds \Big) \zeta_z = \sum_{z\in\mathcal{K}} v_h(z) \, \zeta_z = v_h.$$

(ii) follows from the choice of the edges E_z. (iii) We consider only $\omega = \Gamma_D$. We first note that

$$(J_h u)|_{\Gamma_D} = \sum_{z\in\mathcal{K}} \Big(\int_{E_z} \psi_z u \, ds \Big) \zeta_z|_{\Gamma_D} = \sum_{z\in\mathcal{K}\cap\overline{\Gamma}_D} \Big(\int_{E_z} \psi_z u \, ds \Big) \zeta_z|_{\Gamma_D} \quad \text{for all } u \in H^1(\Omega).$$

For $z \in \mathcal{K} \cap \overline{\Gamma}_D$, it holds that $E_z \subseteq \overline{\Gamma}_D$ and hence $\int_{E_z} \psi_z v_h \, ds = \int_{E_z} \psi_z v \, ds$. Together with the last equation and the projection property, we obtain that

$$(J_h v)|_{\Gamma_D} = (J_h v_h)|_{\Gamma_D} = v_h|_{\Gamma_D}.$$

This concludes the proof. ∎

Exercise 18. Show that Lemma 4.1 holds for any dimension $d \geq 2$. □

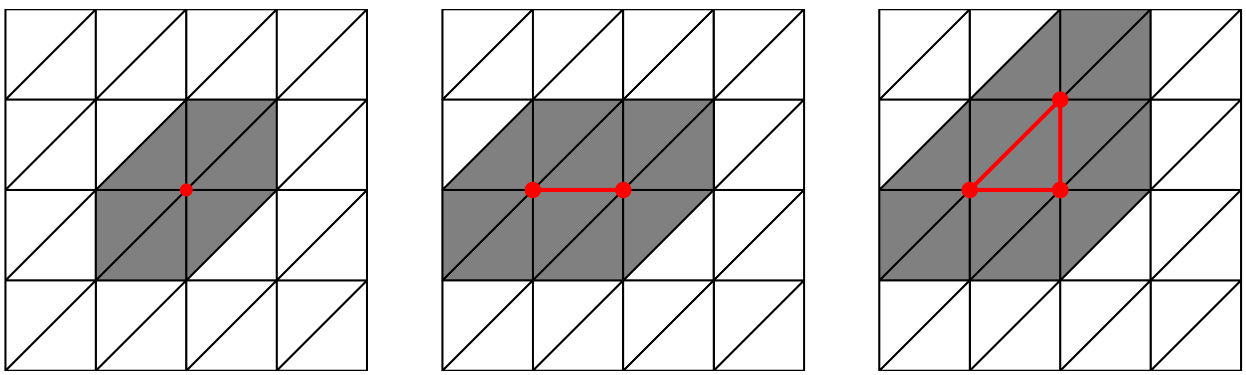
Note that the Scott-Zhang projection $J_h v$ is not defined for general L^2-functions, since $L^2(T)$ does not provide traces. However, one can define an appropriate variant as follows:

Exercise 19. Construct a linear projection $P_h : L^2(\Omega) \to \mathcal{S}^1(\mathcal{T})$ which satisfies

- $\|P_h v\|_{L^2(\Omega)} \leq C \, \|v\|_{L^2(\Omega)}$ for all $v \in L^2(\Omega)$.
- $P_h v_h = v_h$ for all $v_h \in \mathcal{S}^1(\mathcal{T})$.

The constant $C > 0$ may only depend on $\sigma(\mathcal{T})$. **Hint.** Proceed as for the standard Scott-Zhang projection. Instead of an edge E_z, associate with each node $z \in \mathcal{K}$ an arbitrary element $T_z \in \mathcal{T}$ with $z \in T_z$. □

Next, we aim to show that the Scott-Zhang projection has local stability and approximation properties. Unlike nodal interpolation, this will require appropriate patches.



Patch Ω_z of a node $z \in \mathcal{K}$ *Patch Ω_E of an edge $E \in \mathcal{E}$* *Patch Ω_T of an element $T \in \mathcal{T}$*

FIGURE 4.1. *For the a posteriori analysis, we need three types of patches $\omega \subseteq \Omega$, namely patches of nodes, edges, and elements, respectively. Note that the patch of an edge (or of an element) just is the union of the patches of its nodes.*

Definition. For the a posteriori analysis, we need certain unions of elements, called **patches**, cf. Figure 4.1: For a node $z \in \mathcal{K}$, we define

$$\widetilde{\Omega}_z := \big\{T \in \mathcal{T} \,\big|\, z \in \mathcal{K}_T\big\} \quad \text{as well as} \quad \Omega_z := \bigcup \widetilde{\Omega}_z := \big\{x \in \mathbb{R}^2 \,\big|\, \exists T \in \widetilde{\Omega}_z \quad x \in T\big\}. \tag{4.6}$$

For an edge $E \in \mathcal{E}$, we define

$$\widetilde{\Omega}_E := \big\{T \in \mathcal{T} \,\big|\, \mathcal{K}_E \cap T \neq \emptyset\big\} = \big\{T \in \widetilde{\Omega}_z \,\big|\, z \in \mathcal{K}_E\big\} \quad \text{as well as} \quad \Omega_E := \bigcup \widetilde{\Omega}_E. \tag{4.7}$$

Finally, for an element $T \in \mathcal{T}$, we define

$$\widetilde{\Omega}_T := \big\{T' \in \mathcal{T} \,\big|\, \mathcal{K}_T \cap T' \neq \emptyset\big\} = \big\{T' \in \widetilde{\Omega}_z \,\big|\, z \in \mathcal{K}_T\big\} \quad \text{as well as} \quad \Omega_T := \bigcup \widetilde{\Omega}_T. \tag{4.8}$$

The patches Ω_z, Ω_E, and Ω_T are visualized in Figure 4.1.

Lemma 4.3. *There is a constant $C > 0$ which depends only on $\sigma(\mathcal{T})$, such that*

- $\#\widetilde{\Omega}_z \leq C$ *for all $z \in \mathcal{K}$,*
- $\#\widetilde{\Omega}_E \leq C$ *for all $E \in \mathcal{E}$,*

- *$\#\widetilde{\Omega}_T \le C$ for all $T \in \mathcal{T}$,*

i.e., the number of elements per patch is uniformly bounded. Moreover,

- *$\#\{T' \in \mathcal{T} \,|\, T \in \widetilde{\Omega}_{T'}\} \le C$ for all $T \in \mathcal{T}$,*

i.e., an element $T \in \mathcal{T}$ belongs only to finitely many patches.

Proof. Note that $\sigma(\mathcal{T})$ provides a bound for the minimal interior angle of all elements $T \in \mathcal{T}$; see Exercise 13. Consequently, there is a maximal number $C > 0$ of elements in $\widetilde{\Omega}_z$, for all nodes $u \in \mathcal{K}$. By definition, there follows $\#\widetilde{\Omega}_E \le 2C$ as well as $\#\widetilde{\Omega}_T \le 3C$. ■

An essential consequence of Lemma 4.3 is that

$$\|v\|_{L^2(\Omega)} \le \Big(\sum_{T\in\mathcal{T}} \|v\|^2_{L^2(\Omega_T)} \Big)^{1/2} \le C_{\text{patch}} \, \|v\|_{L^2(\Omega)} \quad \text{for all } v \in L^2(\Omega),$$

where $C_{\text{patch}} > 0$ depends only on $\sigma(\mathcal{T})$. Another consequence of Lemma 4.3 is that the diameter $\text{diam}(\Omega_T)$ of a patch is proportional to $h_T = \text{diam}(T)$. This is stated in the following lemma.

Lemma 4.4. *For a regular triangulation, it holds that*

- $\text{diam}(\Omega_z) \le C\, h_T$ *for all $z \in \mathcal{K}$ and $T \in \widetilde{\Omega}_z$,*
- $\text{diam}(\Omega_E) \le C\, h_E \le C\, h_T$ *for all $E \in \mathcal{E}$ and $T \in \widetilde{\Omega}_E$,*
- $\text{diam}(\Omega_{T'}) \le C\, h_T$ *for all $T' \in \mathcal{T}$ and $T \in \widetilde{\Omega}_{T'}$.*

The constant $C > 0$ depends only on $\sigma(\mathcal{T})$.

Proof. **1. step.** Note that $h_T \le \sigma(\mathcal{T})\varrho_T \le \sigma(\mathcal{T})h_E$ for all $T \in \mathcal{T}$ and all edges $E \in \mathcal{E}_T$.

2. step. Patch of a node $z \in \mathcal{K}$: For $\widetilde{\Omega}_z = \{T_1, \dots, T_n\}$, we may choose a numbering such that T_{j-1}, T_j are neighbours, i.e., $T_{j-1} \cap T_j \in \mathcal{E}$. From step 1, we derive $h_{T_{j-1}} \le \sigma(\mathcal{T})h_{T_j}$, whence $h_{T'} \le \sigma(\mathcal{T})^{n-1}h_T$ for all $T, T' \in \widetilde{\Omega}_z$. This yields that

$$\text{diam}(\Omega_z) \le 2 \max_{T' \in \widetilde{\Omega}_z} h_{T'} \le 2\sigma(\mathcal{T})^{n-1}h_T \quad \text{for all } T \in \widetilde{\Omega}_z.$$

3. step. Patch of an edge $E \in \mathcal{E}$: With $E = \text{conv}\{z_1, z_2\}$ for some $z_1, z_2 \in \mathcal{K}$, it holds that $\widetilde{\Omega}_E = \widetilde{\Omega}_{z_1} \cup \widetilde{\Omega}_{z_2}$ as well as $\widetilde{\Omega}_{z_1} \cap \widetilde{\Omega}_{z_2} \neq \emptyset$. Let $T \in \widetilde{\Omega}_E$ and $n := \max\{\#\widetilde{\Omega}_{z_1}, \#\widetilde{\Omega}_{z_2}\}$. Without loss of generality, we may assume $T \in \widetilde{\Omega}_{z_1}$. Choose $T' \in \widetilde{\Omega}_{z_1} \cap \widetilde{\Omega}_{z_2}$. Then,

$$\text{diam}(\Omega_E) \le \text{diam}(\Omega_{z_1}) + \text{diam}(\Omega_{z_2}) \le 2\sigma(\mathcal{T})^{n-1}(h_T + h_{T'}) \le 2\sigma(\mathcal{T})^{n-1}(1 + \sigma(\mathcal{T})^{n-1})h_T.$$

4. step. Patch of an element $T \in \mathcal{T}$: Simply use the same arguments as in step 3. ■

The Scott-Zhang projection is locally H^1-stable and has a local first-order approximation property.

Proposition 4.5. *For all $T \in \mathcal{T}$, it holds that*

$$\|v - J_h v\|_{L^2(T)} + h_T \|\nabla J_h v\|_{L^2(T)} \le C\, h_T\, \|\nabla v\|_{L^2(\Omega_T)} \quad \text{for all } v \in H^1(\Omega). \tag{4.9}$$

The constant $C > 0$ depends only on γ-shape regularity of $\mathcal{T}$.

The proof requires the following technical lemmata which are also valid in any dimension $d \ge 2$ as the proofs reveal.

Theorem 4.6 (Trace Inequality). *Let $T = \operatorname{conv}\{z_0, \dots, z_d\} \subset \mathbb{R}^d$ be a simplex in $\mathbb{R}^d$ with $|T| > 0$ and diameter $h_T := \operatorname{diam}(T)$. Let $E = \operatorname{conv}\{z_1, \dots, z_d\}$ denote one particular side of the simplex. Then, for $v \in H^1(T)$, it holds*

$$\|v\|_{L^2(E)}^2 \le \frac{|E|}{|T|} \big(\|v\|_{L^2(T)}^2 + \frac{2}{d}\, h_T \|v \nabla v\|_{L^1(T)} \big) \le \frac{|E|}{|T|} \,(1 + 2\, h_T/d)\, \|v\|_{H^1(T)}^2. \tag{4.10}$$

With the integral means $v_T := |T|^{-1} \int_T v\, dx$ and $v_E := h_E^{-1} \int_E v\, ds$, it holds that

$$\|v - v_E\|_{L^2(E)}^2 \le \|v - v_T\|_{L^2(E)}^2 \le C\, \frac{|E|\, h_T^2}{|T|}\, \|\nabla v\|_{L^2(T)}^2, \tag{4.11}$$

where $C > 0$ depends only on the reference element T_{ref} and the dimension d.

The proof of the trace inequalities (4.10)–(4.11) is done with the help of the following lemma. In particular, we shall see that both estimates are sharp. Note that, for $d = 2$, it holds that $|E|/|T| \le 2\varrho_T^{-1}$ and $|E|\, h_T^2/|T| \le 2\sigma(T) h_T$.

Lemma 4.7 (Trace Identity). *Let $T = \operatorname{conv}\{z_0, \dots, z_d\} \subset \mathbb{R}^d$ be a simplex in $\mathbb{R}^d$ with $|T| > 0$. Let $E = \operatorname{conv}\{z_1, \dots, z_d\}$ denote one particular side of the simplex. Then,*

$$\frac{1}{|E|} \int_E w\, ds = \frac{1}{|T|} \int_T w\, dx + \frac{1}{d|T|} \int_T (x - z_0) \cdot \nabla w\, dx \tag{4.12}$$

for all $w \in C^1(\overline{T})$.

Proof. We apply the Gauss Divergence Theorem to the function $f(x) := w(x)(x - z_0)$. With $\operatorname{div} f(x) = \nabla w(x) \cdot (x - z_0) + d w(x)$, we obtain that

$$d \int_T w\, dx + \int_T (x - z_0) \cdot \nabla w(x)\, dx = \int_T \operatorname{div} f\, dx = \int_{\partial T} f \cdot n\, ds.$$

Note that $(x - z_0) \cdot n(x) = 0$ for $x \in \partial T \backslash E$, whereas $(x - z_0) \cdot n(x) = \operatorname{dist}(z_0, H)$, where $H \subset \mathbb{R}^d$ denotes the hyperplane with $E \subseteq H$. Therefore, the boundary integral simplifies to $\int_{\partial T} f \cdot n\, ds = \operatorname{dist}(z_0, H) \int_E w\, ds$ and the latter equality reads

$$\frac{1}{|T|} \int_T w\, dx + \frac{1}{d|T|} \int_T (x - z_0) \cdot \nabla w\, dx = \frac{\operatorname{dist}(z_0, H)|E|}{d|T|}\, \frac{1}{|E|} \int_E w\, ds,$$

which holds for any $w \in C^1(\overline{T})$. The special choice $w = 1$ can be used to determine the w-independent constant $\frac{\mathrm{dist}(z_0,H)|E|}{d|T|} = 1$. This concludes the proof. ■

Remark. Note that Lemma 4.7 holds for any $w \in W^{1,1}(T) := \{w \in L^1(T) \text{ weakly differentiable} \,\big|\, \nabla w \in L^1(T)^d\}$, even with the same proof. □

Proof of Theorem 4.6. According to standard density arguments, it suffices to consider $v \in C^1(\overline{T})$. Plugging $w := v^2 \in C^1(\overline{T})$ into the trace identity (4.12), we see that

$$\frac{1}{|E|}\int_E v^2\,ds = \frac{1}{|T|}\int_T v^2\,dx + \frac{1}{d|T|}\int_T (x - z_0)\cdot(2v\nabla v)\,dx.$$

This is rewritten in the form

$$\begin{aligned}\frac{|T|}{|E|}\,\|v\|^2_{L^2(E)} = \|v\|^2_{L^2(T)} + \frac{2}{d}\int_T (x - z_0)\cdot(v\nabla v)\,dx &\le \|v\|^2_{L^2(T)} + \frac{2}{d}\,h_T\|v\nabla v\|_{L^1(T)}\\ &\le (1 + 2\,h_T/d)\,\|v\|^2_{H^1(T)}\end{aligned}$$

which proves (4.10). For the proof of (4.11), we simply replace v by $v - v_T$ and apply the Poincaré inequality. This leads to

$$\begin{aligned}\|v - v_T\|^2_{L^2(E)} &\le \frac{|E|}{|T|}\big(\|v - v_T\|^2_{L^2(T)} + \frac{2}{d}\,h_T\|v - v_T\|_{L^2(T)}\|\nabla v\|_{L^2(T)}\big)\\ &\le \frac{|E|}{|T|}\big(C_P^2 h_T^2\|\nabla v\|^2_{L^2(T)} + \frac{2}{d}\,C_P h_T^2\|\nabla v\|^2_{L^2(T)}\big)\\ &= (C_P^2 + 2C_P/d)\,\frac{|E|\,h_T^2}{|T|}\,\|\nabla v\|^2_{L^2(T)}.\end{aligned}$$

The remaining estimate $\|v - v_E\|_{L^2(E)} \le \|v - v_T\|_{L^2(E)}$ follows from the L^2-best approximation property of the integral mean. ■

Lemma 4.8 (Generalized Poincaré-Friedrichs inequality). *Let $v \in H^1(\Omega)$, $T \in \mathcal{T}$, $T' \in \widetilde{\Omega}_T$, and $E' \in \mathcal{E}_{T'}$. Define the integral means $v_T := (1/|T|)\int_T v\,dx$, $v_{T'} := (1/|T'|)\int_{T'} v\,dx$, and $v_{E'} := (1/|E'|)\int_{E'} v\,ds$. Then,*

$$\|v_T - v_{T'}\|_{L^2(T)} + \|v_T - v_{E'}\|_{L^2(T)} \le C\,h_T\|\nabla v\|_{L^2(\Omega_T)}. \tag{4.13}$$

In particular, this implies that

$$\|v - v_{T'}\|_{L^2(\Omega_T)} + \|v - v_{E'}\|_{L^2(\Omega_T)} \le C\,h_T\|\nabla v\|_{L^2(\Omega_T)}. \tag{4.14}$$

In either estimate, the constant $C > 0$ depends only on γ-shape regularity of $\mathcal{T}$, but is independent of Ω and the shape of Ω_T.

Proof. To ease the notation, let $v_E := (1/|E|)\int_E v\,ds$ also denote the integral mean over edges.

1. step. For any $w \in H^1(T)$, it holds that $|w_T - w_E| \le C\,\|\nabla w\|_{L^2(T)}$, where $C > 0$ depends only on shape regularity: With the trace inequality (4.11) on T, we see that

$$\begin{aligned}|w_T - w_E| \le |E|^{-1}\,\|w - w_T\|_{L^1(E)} &\le |E|^{-1/2}\,\|w - w_T\|_{L^2(E)} \\ &\le C\,\frac{h_T}{|T|^{1/2}}\,\|\nabla w\|_{L^2(T)} =: C\,\|\nabla w\|_{L^2(T)}\end{aligned}$$

2. step. For all $T, T' \in \mathcal{T}$ with $E := T \cap T' \in \mathcal{E}$, shape regularity and the triangle inequality yield that

$$\begin{aligned}\|v_T - v_{T'}\|_{L^2(T)} = |T|^{1/2}|v_T - v_{T'}| &\lesssim |T|^{1/2}|v_T - v_E| + |T'|^{1/2}|v_E - v_{T'}| \\ &= \|v_T - v_E\|_{L^2(T)} + \|v_{T'} - v_E\|_{L^2(T')}.\end{aligned}$$

With Step 1, we thus see that

$$\|v_T - v_{T'}\|_{L^2(T)} \lesssim h_T\|\nabla v\|_{L^2(T)} + h_{T'}\|\nabla v\|_{L^2(T')} \lesssim h_T\,\|\nabla v\|_{L^2(T\cup T')},$$

where the hidden constant now depends on $C > 0$ from step 1 and from shape regularity of $\mathcal{T}$.

3. step. If $T \cap T' \neq \emptyset$, there is a minimal $n \in \mathbb{N}$ and elements $T_0, \dots, T_n \in \mathcal{T}$ with $T_0 = T$, $T_j \cap T_{j-1} \in \mathcal{E}$ and $T_j \subseteq \Omega_T$ for all $j = 1, \dots, n$, and $T_n = T'$. Note that n is uniformly bounded in terms of the γ-shape regularity of $\mathcal{T}$. Iterating the argument from Step 2, we conclude (4.13) with $\bigcup_{j=0}^n T_j \subseteq \Omega_T$. The overall constant then depends on $C > 0$ and γ.

4. step. For each element $T'' \in \widetilde{\Omega}_T$, the Poincaré inequality and (4.13) show

$$\begin{aligned}&\|v - v_{T'}\|_{L^2(T'')} + \|v - v_{E'}\|_{L^2(T'')} \\ &\lesssim \|v - v_{T''}\|_{L^2(T'')} + \|v_T - v_{T'}\|_{L^2(T'')} + \|v_T - v_{T''}\|_{L^2(T'')} + \|v_T - v_{E'}\|_{L^2(T'')} \\ &\simeq \|v - v_{T''}\|_{L^2(T'')} + \|v_T - v_{T'}\|_{L^2(T)} + \|v_T - v_{T''}\|_{L^2(T)} + \|v_T - v_{E'}\|_{L^2(T)} \\ &\lesssim h_{T''}\,\|\nabla v\|_{L^2(T'')} + h_T\,\|\nabla v\|_{L^2(\Omega_T)} \\ &\lesssim h_T\,\|\nabla v\|_{L^2(\Omega_T)}.\end{aligned}$$

Summing this estimate over all $T'' \in \widetilde{\Omega}_T$, we obtain that

$$\|v - v_{T'}\|_{L^2(\Omega_T)} + \|v - v_{E'}\|_{L^2(\Omega_T)} \lesssim h_T\,\|\nabla v\|_{L^2(\Omega_T)},$$

where the hidden constants depends only on γ-shape regularity of $\mathcal{T}$. ■

Proof of Proposition 4.5 (H^1-stability). For $z \in \mathcal{K}$, let $E_z \subset T_z \in \mathcal{T}$ and $h_z := \operatorname{diam}(T_z)$. Note that $T_z \subseteq \Omega_T$ for $z \in T$. The trace inequality (4.10) yields that

$$\|v\|_{L^2(E_z)}^2 \lesssim h_z^{-1}\big(\|v\|_{L^2(T_z)}^2 + h_z\,\|v\|_{L^2(T_z)}\|\nabla v\|_{L^2(T_z)}\big) \lesssim h_z^{-1}\,\big(\|v\|_{L^2(T_z)}^2 + h_z^2\,\|\nabla v\|_{L^2(T_z)}^2\big)$$

With this and Lemma 4.1, we see that

$$\begin{aligned}\Big|\int_{E_z} \psi_z v\,ds\Big| \le \|\psi_z\|_{L^\infty(E_z)}\|v\|_{L^1(E_z)} &\lesssim |E_z|^{-1/2}\,\|v\|_{L^2(E_z)} \\ &\lesssim |E_z|^{-1/2}\,h_z^{-1/2}\,\big(\|v\|_{L^2(T_z)} + h_z\|\nabla v\|_{L^2(T_z)}\big).\end{aligned}$$

For any hat function, an inverse estimate shows

$$\|\nabla\zeta_z\|_{L^2(T)} \lesssim h_T^{-1}\,\|\zeta_z\|_{L^2(T)} \leq |T|^{1/2}h_T^{-1}.$$

Together with $|E_z|\,h_z \simeq |T_z| \simeq |T|$ and $h_z \simeq h_T$, we therefore obtain that, for all $v \in H^1(\Omega)$,

$$\|\nabla J_h v\|_{L^2(T)} \leq \sum_{z\in\mathcal{K}\cap T}\Big|\int_{E_z}\psi_z v\,ds\Big|\|\nabla\zeta_z\|_{L^2(T)} \lesssim \sum_{z\in\mathcal{K}\cap T}\big(h_z^{-1}\,\|v\|_{L^2(T_z)} + \|\nabla v\|_{L^2(T_z)}\big). \tag{4.15}$$

With the integral mean $v_T := (1/|T|)\int_T v\,dx$ and the projection property $J_h v_T = v_T$, we apply the last estimate for $w := v - v_T$ and see that

$$\|\nabla J_h v\|_{L^2(T)} = \|\nabla J_h(v - v_T)\|_{L^2(T)} \lesssim \sum_{z\in\mathcal{K}\cap T}\big(h_z^{-1}\,\|v - v_T\|_{L^2(T_z)} + \|\nabla v\|_{L^2(T_z)}\big).$$

According to the Poincaré inequality and Lemma 4.8, it holds that for all $z \in \mathcal{K}\cap T$,

$$\|v - v_T\|_{L^2(T_z)} \leq \|v - v_{T_z}\|_{L^2(T_z)} + \|v_{T_z} - v_T\|_{L^2(T_z)} \lesssim h_z\,\|\nabla v\|_{L^2(\Omega_T)}. \tag{4.16}$$

Combining the last two estimates, we thus conclude $\|\nabla J_h v\|_{L^2(T)} \lesssim \|\nabla v\|_{L^2(\Omega_T)}$. ∎

Proof of Proposition 4.5 (approximation property). We adopt the notation from the proof of local H^1-stability. Arguing as for (4.15), we see that

$$\|J_h v\|_{L^2(T)} \leq \sum_{z\in\mathcal{K}\cap T}\Big|\int_{E_z}\psi_z v\,ds\Big|\|\zeta_z\|_{L^2(T)} \lesssim \sum_{z\in\mathcal{K}\cap T}\big(\|v\|_{L^2(T_z)} + h_z\,\|\nabla v\|_{L^2(T_z)}\big). \tag{4.17}$$

With the integral mean $v_T := (1/|T|)\int_T v\,dx$ and the projection property $J_h v_T = v_T$, we apply the last estimate for $w := v - v_T$ and see that

$$\begin{aligned}\|v - J_h v\|_{L^2(T)} &= \|(v - v_T) - J_h(v - v_T)\|_{L^2(T)}\\ &\leq \|v - v_T\|_{L^2(T)} + \|J_h(v - v_T)\|_{L^2(T)}\\ &\lesssim h_T\,\|\nabla v\|_{L^2(T)} + \sum_{z\in\mathcal{K}\cap T}\big(\|v - v_T\|_{L^2(T_z)} + h_z\,\|\nabla v\|_{L^2(T_z)}\big)\end{aligned}$$

Finally, we employ (4.16) and $h_z \simeq h_T$ to conclude $\|v - J_h v\|_{L^2(T)} \lesssim h_T\,\|\nabla v\|_{L^2(\Omega_T)}$. ∎

The following theorem concludes the main properties of the Scott-Zhang projection:

Theorem 4.9. *The Scott-Zhang projection $J_h : H^1(\Omega) \to \mathcal{S}^1(\mathcal{T})$ has the following properties* (i)–(vii)*:*

(i) *J_h is linear and continuous with respect to the H^1-norm, i.e.,*

$$\|J_h v\|_{H^1(\Omega)} \leq C\,(1 + \mathrm{diam}(\Omega))\,\|v\|_{H^1(\Omega)} \quad \textit{for all } v \in H^1(\Omega). \tag{4.18}$$

(ii) *J_h is a projection onto $\mathcal{S}^1(\mathcal{T})$, i.e.,*

$$J_h v_h = v_h \quad \textit{for all } v_h \in \mathcal{S}^1(\mathcal{T}). \tag{4.19}$$

(iii) *J_h preserves discrete boundary data, i.e., for $\omega \in \{\Gamma_D, \Gamma\}$ it holds that*

$$(J_h v)|_\omega = v|_\omega \quad \text{for all } v \in H^1(\Omega) \text{ with } v|_\omega \in \mathcal{S}^1(\mathcal{T}|_\omega). \tag{4.20}$$

(iv) *J_h is locally H^1-stable, i.e.,*

$$\|\nabla J_h v\|_{L^2(T)} \le C \|\nabla v\|_{L^2(\Omega_T)} \quad \text{for all } v \in H^1(\Omega) \text{ and } T \in \mathcal{T}. \tag{4.21}$$

(v) *J_h has a local first-order approximation property, i.e.,*

$$\|(1 - J_h)v\|_{L^2(T)} \le C h_T \|\nabla v\|_{L^2(\Omega_T)} \quad \text{for all } v \in H^1(\Omega) \text{ and } T \in \mathcal{T}. \tag{4.22}$$

(vi) *J_h is quasi-optimal in the sense of the Céa lemma, i.e.,*

$$\|(1 - J_h)v\|_{H^1(\Omega)} \le C(1 + \operatorname{diam}(\Omega)) \min_{v_h \in \mathcal{S}^1(\mathcal{T})} \|v - v_h\|_{H^1(\Omega)} \quad \text{for all } v \in H^1(\Omega). \tag{4.23}$$

(vii) *For all $\alpha \in \mathbb{R}$, J_h is quasi-optimal in the sense of*

$$\|h^\alpha \nabla(1 - J_h)v\|_{L^2(\Omega)} \le C \min_{v_h \in \mathcal{S}^1(\mathcal{T})} \|h^\alpha \nabla(v - v_h)\|_{L^2(\Omega)}. \tag{4.24}$$

The constant $C > 0$ in (i)–(vii) *depends only on γ-shape regularity of $\mathcal{T}$.*

Proof. (ii)–(v) have already been shown, and (i) is a direct consequence of (vi) and the triangle inequality. (vii) Let $v_h \in \mathcal{S}^1(\mathcal{T})$. With the projection property of J_h and (iv), we see that, for all $T \in \mathcal{T}$,

$$\|\nabla(1 - J_h)v\|_{L^2(T)} = \|\nabla(1 - J_h)(v - v_h)\|_{L^2(T)} \lesssim \|\nabla(v - v_h)\|_{L^2(\Omega_T)}.$$

With γ-shape regularity and hence $h_T \simeq h_{T'}$ for all $T' \subseteq \Omega_T$, we infer

$$\|h^\alpha \nabla(1 - J_h)v\|_{L^2(T)} \lesssim \|h^\alpha \nabla(v - v_h)\|_{L^2(\Omega_T)}.$$

Using the γ-shape regularity again, this results in

$$\begin{aligned}\|h^\alpha \nabla(1 - J_h)v\|^2_{L^2(\Omega)} = \sum_{T \in \mathcal{T}} \|h^\alpha \nabla(1 - J_h)v\|^2_{L^2(T)} &\lesssim \sum_{T \in \mathcal{T}} \|h^\alpha \nabla(v - v_h)\|^2_{L^2(\Omega_T)} \\ &\lesssim \|h^\alpha \nabla(v - v_h)\|^2_{L^2(\Omega)}.\end{aligned}$$

This proves (vii) with an infimum on the right-hand side. Due to finite dimension, this infimum is, in fact, attained. To prove (vi), it remains to estimate the L^2-part and use $\alpha = 0$ in (vii). With the projection property of J_h and (v), shape regularity yields that

$$\begin{aligned}\|(1 - J_h)v\|^2_{L^2(\Omega)} = \sum_{T \in \mathcal{T}} \|(1 - J_h)(v - v_h)\|^2_{L^2(T)} &\lesssim \sum_{T \in \mathcal{T}} h_T^2 \|\nabla(v - v_h)\|^2_{L^2(\Omega_T)} \\ &\lesssim \operatorname{diam}(\Omega)^2 \|\nabla(v - v_h)\|^2_{L^2(\Omega)}.\end{aligned}$$

Altogether, we thus see that

$$\|(1 - J_h)v\|^2_{H^1(\Omega)} \lesssim (1 + \mathrm{diam}(\Omega)^2)\, \|\nabla(v - v_h)\|^2_{L^2(\Omega)}. \lesssim (1 + \mathrm{diam}(\Omega))^2\, \|v - v_h\|^2_{H^1(\Omega)}.$$

This concludes the proof of (vi). ■

Remark. Theorem 4.9 holds for any dimension $d \geq 2$ and for any fixed polynomial degree $p \geq 1$. □

One drawback of the Scott-Zhang projection is that it is not positivity conserving, i.e., $v \geq 0$ does not necessarily imply that $J_h v \geq 0$.

Exercise 20. Suppose that $\mathcal{T}$ is a regular triangulation of $\Omega := [0, 1]^2$ into 2 triangles. Find an example of a function $v \in H^1(\Omega)$ with $v \geq 0$ such that there exists some $x \in \Omega$ with $J_h v < 0$.
Hint. Compute the function $\widehat{\psi} \in \mathcal{P}^1(0, 1)$ from Lemma 4.1 explicitly. □

Exercise 21. Extend the approach of Exercise 19 and construct an operator $P_h : L^2(\Omega) \to \mathcal{S}^1_D(\mathcal{T})$ with the following properties:

(i) $P_h : L^2(\Omega) \to \mathcal{S}^1_D(\mathcal{T})$ is a well-defined linear projection,

$$P_h v_h = v_h \quad \text{for all } v_h \in \mathcal{S}^1_D(\mathcal{T}).$$

(ii) P_h is locally L^2-stable, i.e., for all $T \in \mathcal{T}$, it holds that

$$\|(1 - P_h)v\|_{L^2(T)} \leq C\, \|v\|_{L^2(\Omega_T)} \quad \text{for all } v \in L^2(\Omega).$$

(iii) P_h is locally H^1_D-stable, i.e., for all $T \in \mathcal{T}$, it holds that

$$\|\nabla(1 - P_h)v\|_{L^2(T)} \leq C\, \|\nabla v\|_{L^2(\Omega_T)} \quad \text{for all } v \in H^1_D(\Omega).$$

(iv) P_h has a local first-order approximation property

$$\|(1 - P_h)v\|_{L^2(T)} \leq C h_T\, \|\nabla v\|_{L^2(\Omega_T)} \quad \text{for all } v \in H^1_D(\Omega).$$

(v) $P_h : L^2(\Omega) \to L^2(\Omega)$ as well as $P_h : H^1_D(\Omega) \to H^1_D(\Omega)$ are bounded linear operators.

(vi) P_h is quasi-optimal in the sense of the Céa lemma, i.e.,

$$\|(1 - P_h)v\|_{H^1(\Omega)} \leq C \min_{v_h \in \mathcal{S}^1_D(\mathcal{T})} \|v - v_h\|_{H^1(\Omega)} \quad \text{for all } v \in H^1_D(\Omega).$$

(vii) For all $\alpha \in \mathbb{R}$, P_h is quasi-optimal in the sense of

$$\|h^\alpha(1 - P_h)v\|_{L^2(\Omega)} \leq C \min_{v_h \in \mathcal{S}^1_D(\mathcal{T})} \|h^\alpha(v - v_h)\|_{L^2(\Omega)} \quad \text{for all } v \in L^2(\Omega).$$

(viii) For all $\alpha \in \mathbb{R}$, P_h is quasi-optimal in the sense of

$$\|h^\alpha \nabla(1 - P_h)v\|_{L^2(\Omega)} \le C \min_{v_h \in \mathcal{S}^1_D(\mathcal{T})} \|h^\alpha \nabla(v - v_h)\|_{L^2(\Omega)} \quad \text{for all } v \in H^1_D(\Omega).$$

The constant $C > 0$ in (i)–(viii) depends only on γ-shape regularity of $\mathcal{T}$. **Hint.** Let $\mathcal{K}_F := \mathcal{K}\backslash\Gamma_D$ denote the free nodes, where possibly $\mathcal{K}_F = \mathcal{K}$ for $\Gamma_D = \emptyset$. Then, P_h can be chosen as

$$P_h v = \sum_{z \in \mathcal{K}_F} \Big(\int_{T_z} v\psi_z \, dx \Big) \zeta_z$$

with appropriate $T_z \in \mathcal{T}$ and $\psi_z \in \mathcal{P}^1(T_z)$. □

Definition. The Scott-Zhang projection is just a special example of a Clément-type quasi-interpolation operator: We say that an operator $J_h : H^1_D(\Omega) \to \mathcal{S}^1_D(\mathcal{T})$ is a **Clément-type quasi-interpolation operator** if, for all $v \in H^1_D(\Omega)$ and all $T \in \mathcal{T}$, it holds that

- it is locally H^1-stable

$$\|\nabla(1 - J_h)v\|_{L^2(T)} \le C \, \|\nabla v\|_{L^2(\Omega_T)}, \tag{4.25}$$

- and has a local first-order approximation property

$$\|(1 - J_h)v\|_{L^2(T)} \le C h_T \, \|\nabla v\|_{L^2(\Omega_T)}. \tag{4.26}$$

The constant $C > 0$ may only depend on γ-shape regularity of $\mathcal{T}$ (and possibly the shapes of possible patches in $\mathcal{T}$).

For the a posteriori error analysis, we shall need the following simple consequence.

Lemma 4.10. *Suppose that $J_h : H^1_D(\Omega) \to \mathcal{S}^1_D(\mathcal{T})$ is a Clément-type operator, i.e.,* (4.25)–(4.26) *hold. Let $T \in \mathcal{T}$ and $E \in \mathcal{E}_T$. Then, it holds that*

$$\|(1 - J_h)v\|_{L^2(E)} \le C h_E^{1/2} \, \|\nabla v\|_{L^2(\Omega_T)} \quad \textit{for all } v \in H^1_D(\Omega). \tag{4.27}$$

The constant $C > 0$ depends only on γ-shape regularity of $\mathcal{T}$.

Proof. We apply the trace inequality

$$\|w\|^2_{L^2(E)} \lesssim h_T^{-1} \big(\|w\|^2_{L^2(T)} + h_T \, \|w\|_{L^2(T)} \|\nabla w\|_{L^2(T)} \big)$$

for $w := (1 - J_h)v \in H^1_D(\Omega)$. With the Clément properties (4.25)–(4.26), this yields that

$$\|(1 - J_h)v\|^2_{L^2(E)} \lesssim h_T \, \|\nabla v\|^2_{L^2(\Omega_T)}.$$

Shape regularity and hence $h_T \simeq h_E$ concludes the proof. ■

The following example is one further *classical* example of a Clément-type operator. The analysis will be left to the reader, but requires the following simple observation:

Exercise 22. Use a scaling argument to show that

$$C^{-1} h_T \, \|\nabla v_h\|_{L^\infty(T)} \le \|\nabla v_h\|_{L^2(T)} \le \frac{h_T}{\sqrt{2}} \|\nabla v_h\|_{L^\infty(T)} \quad \text{for all } v_h \in \mathcal{P}^m(T),$$

where the constant $C > 0$ only depends on $\sigma(\mathcal{T})$ and the polynomial degree $m \in \mathbb{N}_0$. □

Exercise 23. Let $\mathcal{K}_F := \mathcal{K} \backslash \overline{\Gamma}_D$ denote the free nodes (where possibly $\mathcal{K}_F = \mathcal{K}$ if $\Gamma_D = \emptyset$). Define

$$J_h v := \sum_{z \in \mathcal{K}_F} v_z \zeta_z \quad \text{with} \quad v_z := \frac{1}{|\Omega_z|} \int_{\Omega_z} v \, dx, \tag{4.28}$$

where $\Omega_z \subseteq \overline{\Omega}$ denotes the patch of a node $z \in \mathcal{K}$. Prove that J_h satisfies the following properties:

(i) $J_h : L^2(\Omega) \to \mathcal{S}^1_D(\mathcal{T})$ is a well-defined linear operator.

(ii) J_h is locally L^2-stable, i.e., for all $T \in \mathcal{T}$, it holds that

$$\|(1 - J_h)v\|_{L^2(T)} \le C \, \|v\|_{L^2(\Omega_T)} \quad \text{for all } v \in L^2(\Omega).$$

(iii) J_h is locally H^1_D-stable, i.e., for all $T \in \mathcal{T}$, it holds that

$$\|\nabla(1 - J_h)v\|_{L^2(T)} \le C \, \|\nabla v\|_{L^2(\Omega_T)} \quad \text{for all } v \in H^1_D(\Omega).$$

(iv) J_h has a local first-order approximation property

$$\|(1 - J_h)v\|_{L^2(T)} \le C h_T \, \|\nabla v\|_{L^2(\Omega_T)} \quad \text{for all } v \in H^1_D(\Omega).$$

(v) $J_h : L^2(\Omega) \to L^2(\Omega)$ as well as $J_h : H^1_D(\Omega) \to H^1_D(\Omega)$ are bounded linear operators.

(vi) J_h is positivity preserving, i.e., $J_h v \ge 0$ for all $v \in L^2(\Omega)$ with $v \ge 0$.

(vii) With $\Pi_h : L^2(\Omega) \to \mathcal{P}^0(\mathcal{T})$ the L^2-orthogonal projection onto $\mathcal{P}^0(\mathcal{T})$, it holds that $J_h \Pi_h = J_h$.

The constant $C > 0$ depends only on γ-shape regularity of $\mathcal{T}$. □

Exercise 24. Find a counter example which shows that the operator J_h from Exercise 23 is no projection, i.e., it holds that $J_h v_h \neq v_h$ for some $v_h \in \mathcal{S}^1_D(\mathcal{T})$. □

4.3 Residual-Based Error Estimator

Residual-based a posteriori error estimates follow a general strategy. Recall that the weak solution $u \in H^1_D(\Omega)$ of (4.1) solves the variational form

$$(\nabla u\,;\,\nabla v)_{L^2(\Omega)} = (f\,;\,v)_{L^2(\Omega)} + (\phi\,;\,v)_{L^2(\Gamma_N)} \quad \text{for all } v \in H^1_D(\Omega). \tag{4.29}$$

For an approximation $u_h \in \mathcal{S}^1_D(\mathcal{T})$ it is thus natural to define the **residual** $R_h \in H^1_D(\Omega)^*$ by

$$R_h(v) := (f\,;\,v)_{L^2(\Omega)} + (\phi\,;\,v)_{L^2(\Gamma_N)} - (\nabla u_h\,;\,\nabla v)_{L^2(\Omega)}, \tag{4.30}$$

i.e., $R_h = 0$ if and only if $u_h = u$. Let $|\!|\!| w |\!|\!| := \|\nabla w\|_{L^2(\Omega)}$ denote the energy norm on $H^1_D(\Omega)$ and

$$|\!|\!| \Phi |\!|\!|_* := \sup_{w \in H^1_D(\Omega)\backslash\{0\}} \frac{\Phi(w)}{|\!|\!| w |\!|\!|}$$

the induced operator norm on $H^1_D(\Omega)^*$, where we stress that both are equivalent norms on $H^1_D(\Omega)$ and its dual space, respectively. Then, the Riesz theorem and $R_h(v) = (\nabla(u-u_h)\,;\,\nabla v)_{L^2(\Omega)}$ yield

$$|\!|\!| R_h |\!|\!|_* = |\!|\!| u - u_h |\!|\!|.$$

To derive a reliable error estimator η, we thus need to prove an estimate of the type

$$R_h(v) \le \widetilde{C_{\rm rel}}\,\eta\,|\!|\!| v |\!|\!| \quad \text{for all } v \in H^1_D(\Omega). \tag{4.31}$$

To derive an efficient error estimator η, we need to show

$$R_h(v) \ge \widetilde{C_{\rm eff}}\,\eta\,|\!|\!| v |\!|\!| \quad \text{for some } v \in H^1_D(\Omega)\backslash\{0\}, \tag{4.32}$$

where this $v \in H^1_D(\Omega)$ has to be constructed appropriately.

Exercise 25. Prove that reliability (4.2) of an error estimator η is, in fact, equivalent to (4.31). Prove that efficiency (4.3) of η holds if and only if (4.32) holds. □

So far, our observations did not use that we are dealing with Galerkin schemes. We stress that the Galerkin orthogonality reads

$$R_h(v_h) = 0 \quad \text{for all } v_h \in \mathcal{S}^1_D(\mathcal{T}) \tag{4.33}$$

with respect to the residual R_h. To provide a reliable (and residual-based) error estimator η, we will use some Clément-type operator $J_h : H^1(\Omega) \to \mathcal{S}^1_D(\Omega)$ in connection with the Galerkin orthogonality (4.33).

Before introducing a first a posteriori error estimator, we introduce the following notational conventions. We define the $\mathcal{T}$-piecewise resp. $\mathcal{E}$-piecewise constant mesh-width functions

$$h_{\mathcal{T}}|_T := h_T \quad \text{and} \quad h_{\mathcal{E}}|_E := h_E$$

for elements $T \in \mathcal{T}$ and edges $E \in \mathcal{E}$, respectively. Moreover, we write

$$\|h_{\mathcal{E}}^{1/2}\psi\|_{L^2(\mathcal{E}_*)} := \Big(\sum_{E\in\mathcal{E}_*} h_E\|\psi\|_{L^2(E)}^2\Big)^{1/2}$$

for any set $\mathcal{E}_* \subseteq \mathcal{E}$ of edges and any function ψ which belongs to $L^2(E)$ for all $E \in \mathcal{E}_*$. Recall that $\mathcal{E}_D$ and $\mathcal{E}_N$ denote the Dirichlet and Neumann edges of $\mathcal{T}$, respectively. Moreover, let $\mathcal{E}_\Omega$ denote the set of all **interior edges**, i.e., for $E \in \mathcal{E}_\Omega$, there are unique elements $T_E^+, T_E^- \in \mathcal{T}$ with $E = T_E^+ \cap T_E^-$. Finally, for $E \in \mathcal{E}_\Omega$, we define the **jump of the normal derivative** by

$$[\![\partial_n u_h]\!]_E := \frac{\partial u_h}{\partial n_E^+} + \frac{\partial u_h}{\partial n_E^-} \in \mathbb{R}, \tag{4.34}$$

where $n_E^\pm$ denote the outer normal vectors of the elements $T_E^\pm$ on the edge E. Note that $n_E^+ = -n_E^-$ so that the sum in the definition is, in fact, a difference.

Theorem 4.11. *The error estimator*

$$\eta := \big(\|h_{\mathcal{T}}f\|_{L^2(\Omega)}^2 + \|h_{\mathcal{E}}^{1/2}[\![\partial_n u_h]\!]\|_{L^2(\mathcal{E}_\Omega)}^2 + \|h_{\mathcal{E}}^{1/2}(\phi - \partial_n u_h)\|_{L^2(\mathcal{E}_N)}^2\big)^{1/2} \tag{4.35}$$

satisfies the reliability estimate

$$\|u - u_h\|_{H^1(\Omega)} \le C\,\eta, \tag{4.36}$$

where the constant $C > 0$ depends only on γ-shape regularity of $\mathcal{T}$.

Proof. For all $w \in H_D^1(\Omega)$, elementwise integration by parts proves

$$\begin{aligned}
R_h(w) &= (f\,;w)_{L^2(\Omega)} + (\phi\,;w)_{L^2(\Gamma_N)} - \sum_{T\in\mathcal{T}}(\nabla u_h\,;\nabla w)_{L^2(T)}\\
&= (f\,;w)_{L^2(\Omega)} + \sum_{E\in\mathcal{E}_N}(\phi\,;w)_{L^2(E)} - \sum_{T\in\mathcal{T}}(\partial_n u_h\,;w)_{L^2(\partial T)}\\
&= \sum_{T\in\mathcal{T}}(f\,;w)_{L^2(T)} + \sum_{E\in\mathcal{E}_N}(\phi - \partial_n u_h\,;w)_{L^2(E)} - \sum_{E\in\mathcal{E}_\Omega}([\![\partial_n u_h]\!]\,;w)_{L^2(E)}\\
&\le \sum_{T\in\mathcal{T}}\|f\|_{L^2(T)}\|w\|_{L^2(T)} + \sum_{E\in\mathcal{E}_N}\|\phi - \partial_n u_h\|_{L^2(E)}\|w\|_{L^2(E)} + \sum_{E\in\mathcal{E}_\Omega}\|[\![\partial_n u_h]\!]\|_{L^2(E)}\|w\|_{L^2(E)}.
\end{aligned}$$

For arbitrary $v \in H_D^1(\Omega)$, we now choose $w = v - J_h v$ and note that $R_h(v) = R_h(w)$ according to the Galerkin orthogonality. Then, we estimate the three sums separately. The approximation property of the Clément-type operator J_h and Lemma 4.3 imply

$$\begin{aligned}
\sum_{T\in\mathcal{T}}\|f\|_{L^2(T)}\|v - J_h v\|_{L^2(T)} &\lesssim \Big(\sum_{T\in\mathcal{T}}\|h_{\mathcal{T}}f\|_{L^2(T)}^2\Big)^{1/2}\Big(\sum_{T\in\mathcal{T}}\|\nabla v\|_{L^2(\Omega_T)}^2\Big)^{1/2}\\
&\lesssim \Big(\sum_{T\in\mathcal{T}}\|h_{\mathcal{T}}f\|_{L^2(T)}^2\Big)^{1/2}\Big(\sum_{T\in\mathcal{T}}\|\nabla v\|_{L^2(T)}^2\Big)^{1/2}\\
&= \|h_{\mathcal{T}}f\|_{L^2(\Omega)}\|\nabla v\|_{L^2(\Omega)}.
\end{aligned}$$

For each edge $E \in \mathcal{E}$, we choose an arbitrary element $T_E \in \mathcal{T}$ with $E \in \mathcal{E}_{T_E}$. Let $\mathcal{E}_* \subset \mathcal{E}$ and $\psi \in L^2(E)$ for all $E \in \mathcal{E}_*$. Recall that $\|(1-J_h)v\|_{L^2(E)} \lesssim h_E^{1/2}\,\|\nabla v\|_{L^2(\Omega_{T_E})}$. Therefore, the same arguments as before prove

$$\begin{aligned}\sum_{E\in\mathcal{E}_*}\|\psi\|_{L^2(E)}\|v-J_hv\|_{L^2(E)} &\lesssim \Big(\sum_{E\in\mathcal{E}_*}\|h_{\mathcal{E}}^{1/2}\psi\|_{L^2(E)}^2\Big)^{1/2}\Big(\sum_{E\in\mathcal{E}_*}\|\nabla v\|_{L^2(\Omega_{T_E})}^2\Big)^{1/2}\\ &\lesssim \|h_{\mathcal{E}}^{1/2}\psi\|_{L^2(\mathcal{E}_*)}\|\nabla v\|_{L^2(\Omega)},\end{aligned}$$

where we note that an element $T \in \mathcal{T}$ may satisfy $T = T_E$ for at most three edges. Altogether, we now see

$$\begin{aligned}R_h(v) &\lesssim \|\nabla v\|_{L^2(\Omega)}\big(\|h_{\mathcal{T}}f\|_{L^2(\Omega)} + \|h_{\mathcal{E}}^{1/2}[\![\partial_n u_h]\!]\|_{L^2(\mathcal{E}_\Omega)} + \|h_{\mathcal{E}}^{1/2}(\phi-\partial_n u_h)\|_{L^2(\mathcal{E}_N)}\big)\\ &\le \sqrt{3}\,\|\nabla v\|_{L^2(\Omega)}\,\eta.\end{aligned}$$

The hidden constant C depends only (on the Clément operator J_h and) on γ-shape regularity of $\mathcal{T}$. ■

Remark. Note that we have used $u_h \in \mathcal{S}^1(\mathcal{T})$ in the sense that the elementwise Laplacian satisfies $\Delta u_h|_T = 0$ for all $T \in \mathcal{T}$. For general $\mathcal{T}$-piecewise polynomials, the same proof applies with $\|h_{\mathcal{T}}f\|_{L^2(\Omega)}$ replaced by $\|h_{\mathcal{T}}(f+\Delta u_h)\|_{L^2(\Omega)}$. □

Exercise 26. We consider the mixed boundary value problem

$$\begin{aligned}-\Delta u &= f &&\text{in } \Omega,\\ u &= u_D &&\text{on } \Gamma_D,\\ \partial_n u &= \phi &&\text{on } \Gamma_N.\end{aligned}$$

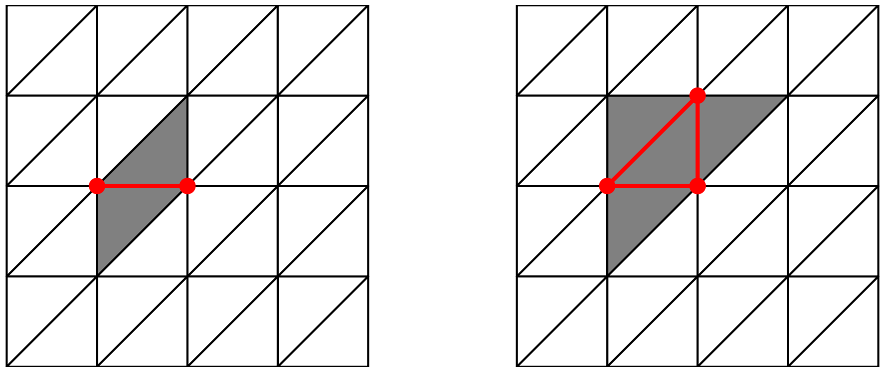
with inhomogeneous Dirichlet data $u_D \in H^{1/2}(\Gamma_D)$. Let $u \in H^1(\Omega)$ denote the weak solution and $u_h \in \mathcal{S}^1(\mathcal{T}_h)$ the P1-FEM solution for discrete Dirichlet data $u_{Dh} := \widehat{u}_{Dh}|_{\Gamma_D}$ with $\widehat{u}_{Dh} \in \mathcal{S}^1(\mathcal{T}_h)$. Use the additional problem

$$\begin{aligned}-\Delta w &= 0 &&\text{in } \Omega,\\ w &= u_D - u_{Dh} &&\text{on } \Gamma_D,\\ \partial_n u &= 0 &&\text{on } \Gamma_N.\end{aligned}$$

with weak solution $w \in H^1(\Omega)$ to derive a reliable error estimator for $\|u-u_h\|_{H^1(\Omega)}$.

Hint. Prove that $\|w\|_{H^1(\Omega)} \simeq \|u_D - u_{Dh}\|_{H^{1/2}(\Gamma_D)}$, where the right-hand side is already an a posteriori term. Then, consider the residual $\widetilde{R}_h \in H^1_D(\Omega)^*$ corresponding to the function $(u-u_h)-w \in H^1_D(\Omega)$. □

Next, we prove the efficiency of the residual-based error estimator η from (4.35) — at least up to terms of higher order. The efficiency estimate even holds locally with refined patches ω_E and ω_T shown in Figure 4.2: For an interior edge $E \in \mathcal{E}_\Omega$, let $T_E^+, T_E^- \in \mathcal{T}$ be the unique elements with



Patch ω_E of an edge $E \in \mathcal{E}$ *Patch ω_T of an element $T \in \mathcal{T}$*

FIGURE 4.2. *To prove the efficiency estimate, it suffices to consider smaller patches $\omega_E \subseteq \Omega_E$ and $\omega_T \subseteq \Omega_T$, for edges $E \in \mathcal{E}$ and elements $T \in \mathcal{T}$, respectively. For comparison with the larger patches Ω_E and Ω_T, the reader may consider Figure 4.1 on page 40.*

$E = T_E^+ \cap T_E^-$. For a boundary edge $E \in \mathcal{E}_\Gamma$, there is a unique element $T_E \in \mathcal{T}$ with $E \in \mathcal{E}_{T_E}$. We define the refined patch of an edge $E \in \mathcal{E}$ by

$$\omega_E := \begin{cases} T_E^+ \cup T_E^- & \text{for } E \in \mathcal{E}_\Omega, \\ T_E & \text{for } E \in \mathcal{E}_\Gamma. \end{cases} \tag{4.37}$$

Moreover, we define the refined patch of an element $T \in \mathcal{T}$ by

$$\omega_T := \bigcup \big\{ \omega_E \,\big|\, E \in \mathcal{E}_T \big\}. \tag{4.38}$$

Note that $\omega_E \subseteq \Omega_E$ and $\omega_T \subseteq \Omega_T$, so that Lemma 4.3 and Lemma 4.4 even hold for the refined patches.

Usually, one is interested in error estimators which are localized with respect to the elements or the edges of $\mathcal{T}$, respectively. For instance, one considers the **element-based residual error estimator**

$$\eta_{\mathcal{T}} := \Big(\sum_{T \in \mathcal{T}} \eta_T^2 \Big)^{1/2}, \tag{4.39}$$

where

$$\eta_T = \big(h_T^2 \|f\|_{L^2(T)}^2 + h_T \|[\![\partial_n u_h]\!]\|_{L^2(\partial T \cap \Omega)}^2 + h_T \|\phi - \partial_n u_h\|_{L^2(\partial T \cap \Gamma_N)}^2 \big)^{1/2} \tag{4.40}$$

or the **edge-based residual error estimator**

$$\eta_{\mathcal{E}} := \Big(\sum_{E \in \mathcal{E}} \eta_E^2 \Big)^{1/2}, \tag{4.41}$$

where

$$\eta_E = \begin{cases} \big(h_E^2 \|f\|_{L^2(\omega_E)}^2 + h_E \|[\![\partial_n u_h]\!]\|_{L^2(E)}^2 \big)^{1/2} & \text{for } E \in \mathcal{E}_\Omega, \\ \big(h_E^2 \|f\|_{L^2(\omega_E)}^2 + h_E \|\phi - \partial_n u_h\|_{L^2(E)}^2 \big)^{1/2} & \text{for } E \in \mathcal{E}_N, \\ 0 & \text{for } E \in \mathcal{E}_D. \end{cases} \tag{4.42}$$

Alternatively, one could also define

$$\eta_{\mathcal{T}\cup\mathcal{E}} := \Big(\sum_{T\in\mathcal{T}} \tilde{\eta}_T^2 + \sum_{E\in\mathcal{E}} \tilde{\eta}_E^2 \Big)^{1/2}, \tag{4.43}$$

where

$$\tilde{\eta}_T = h_T \|f\|_{L^2(T)}, \tag{4.44a}$$

$$\tilde{\eta}_E = \begin{cases} h_E^{1/2}\, \|[\![\partial_n u_h]\!]\|_{L^2(E)} & \text{for } E \in \mathcal{E}_\Omega, \\ h_E^{1/2}\, \|\phi - \partial_n u_h\|_{L^2(E)} & \text{for } E \in \mathcal{E}_N, \\ 0 & \text{for } E \in \mathcal{E}_D. \end{cases} \tag{4.44b}$$

Note that $\eta_\mathcal{T}$ as well as $\eta_\mathcal{E}$ are equivalent to the error estimator η from (4.35): There holds

$$\eta = \eta_{\mathcal{T}\cup\mathcal{E}} \le \eta_\mathcal{T} \le \sqrt{2}\,\sigma(\mathcal{T}_h)^{1/2}\,\eta \quad \text{as well as} \quad \sigma(\mathcal{T})^{-1}\,\eta \le \eta_\mathcal{E} \le \sqrt{3}\,\eta,$$

since $\eta_\mathcal{T}$ adds the contributions of interior edges twice and $h_E \le h_T \le \sigma(\mathcal{T}_h)\,h_E$ for each edge $E \in \mathcal{E}_T$, whereas $\eta_\mathcal{E}$ adds the element contribution at most three times. The local quantities η_T and η_E can be used to steer an adaptive mesh-refining algorithm. They are therefore called **refinement indicators**. We are going to discuss adaptive mesh-refinement below.

Theorem 4.12 (inverse estimate). *For all polynomial degrees $m \in \mathbb{N}$ and $k, r \in \mathbb{N}$ with $k > r$, there exists a constant $C > 0$ such that*

$$\|D^k v_h\|_{L^2(T)} \le C\sigma(\mathcal{T})\, h_T^{r-k} \|D^r v_h\|_{L^2(T)} \quad \textit{for all } v_h \in \mathcal{P}^m(\mathcal{T}) \textit{ and all } T \in \mathcal{T}, \tag{4.45}$$

where $\mathcal{P}^m(\mathcal{T}) := \big\{ v_h : \Omega \to \mathbb{R} \,\big|\, \forall T \in \mathcal{T} \quad v_h|_T \in \mathcal{P}^m(T) \big\}$.

Proof. The proof is done $\mathcal{T}$-elementwise and follows from a scaling argument. We start with the estimate on the reference element.

1. step. We prove

$$\|D^k w_h\|_{L^2(T_{\mathrm{ref}})} \le C_{\mathrm{ref}} \|D^r w_h\|_{L^2(T_{\mathrm{ref}})} \quad \text{for all } w_h \in \mathcal{P}^m(T_{\mathrm{ref}}). \tag{4.46}$$

To this end we introduce an interpolation operator $I : \mathcal{P}^m(T_{\mathrm{ref}}) \to \mathcal{P}^{r-1}(T_{\mathrm{ref}})$ via

$$(Iv)(z_j) = v(z_j), \qquad j = 1, \dots, N_r,$$

for $N_r = \dim \mathcal{P}^{r-1}(T_{\mathrm{ref}})$ piecewise different points $z_j \in T_{\mathrm{ref}}$. It is straightforward to see that

$$|\!|\!| v |\!|\!| := \|D^r v\|_{L^2(T_{\mathrm{ref}})} + \sum_{j=1}^{N_r} |v(z_j)|$$

is a norm on $\mathcal{P}^m(T_{\mathrm{ref}})$. Moreoever, since $\mathcal{P}^m(T_{\mathrm{ref}})$ is finite dimensional, it is equivalent to the norm $\|\cdot\|_{H^k(T_{\mathrm{ref}})}$. Due to $D^k v_h = D^r v_h = 0$ for all $v_h \in \mathcal{P}^{r-1}(T_{\mathrm{ref}})$ and $(Iw_h)(z_j) = w_h(z_j)$ for all $w_h \in \mathcal{P}^m(T_{\mathrm{ref}})$, we have

$$\|D^k w_h\|_{L^2(T_{\mathrm{ref}})} = \|D^k\,(w_h - Iw_h)\,\|_{L^2(T_{\mathrm{ref}})} \le \|w_h - Iw_h\|_{H^k(T_{\mathrm{ref}})} \le C_{\mathrm{ref}} |\!|\!| w_h - Iw_h |\!|\!| = C_{\mathrm{ref}} \|D^r w_h\|_{L^2(T_{\mathrm{ref}})}.$$

2. step. Proof of the statement: Let $\Phi : T_{\mathrm{ref}} \to T$ be an affine diffeomorphism and $B \in \mathbb{R}^{2\times 2}$ its linear part. We apply the transformation formula to Φ^{-1} to see that

$$\|D^k v_h\|_{L^2(T)} \le |\det B^{-1}|^{-1/2} \|B^{-1}\|_F^k \|D^k(v_h \circ \Phi)\|_{L^2(T_{\mathrm{ref}})}.$$

Note that the L^2-norm can be estimated by step 1 since $v_h \circ \Phi \in \mathcal{P}^m(T_{\mathrm{ref}})$. The application of the transformation formula to Φ proves that

$$\|D^r(v_h \circ \Phi)\|_{L^2(T_{\mathrm{ref}})} \le |\det B|^{-1/2} \|B\|_F^r \|D^r v_h\|_{L^2(T)}.$$

By definition of the shape regularity constant $\sigma(\mathcal{T})$, we obtain that

$$\begin{aligned}\|D^k v_h\|_{L^2(T)} &\le C_{\mathrm{ref}}\, \|B^{-1}\|_F^k \|B\|_F^r \|D^r v_h\|_{L^2(T)} \le \sqrt{2}\, C_{\mathrm{ref}}\, \varrho_T^{-k} h_T^r \|D^r v_h\|_{L^2(T)} \\ &\le \sqrt{2}\, C_{\mathrm{ref}}\, \sigma(\mathcal{T}) h_T^{r-k} \|D^r v_h\|_{L^2(T)},\end{aligned}$$

where we have used that $\|B^{-1}\|_F \le \sqrt{2}\, \varrho_T^{-1}$. This concludes the proof. ∎

Theorem 4.13. *We define $f_{\mathcal{T}} \in \mathcal{P}^0(\mathcal{T})$ by $f_{\mathcal{T}}|_T := |T|^{-1} \int_T f\, dx$ and $\phi_{\mathcal{E}} \in \mathcal{P}^0(\mathcal{E}_N)$ by $\phi_{\mathcal{E}}|_E := h_E^{-1} \int_E \phi\, ds$. For each element $T \in \mathcal{T}$, the refinement indicator η_T from* (4.40) *satisfies*

$$\eta_T \le C \left(\|\nabla(u - u_h)\|_{L^2(\omega_T)}^2 + \|h_{\mathcal{T}}(f - f_{\mathcal{T}})\|_{L^2(\omega_T)}^2 + \|h_{\mathcal{E}}^{1/2}(\phi - \phi_{\mathcal{E}})\|_{L^2(\partial T \cap \Gamma_N)}^2 \right)^{1/2}. \tag{4.47}$$

Moreover, the error estimator η from (4.35) *is efficient in the sense that*

$$\eta \le C \left(\|u - u_h\|_{H^1(\Omega)} + \|h_{\mathcal{T}}(f - f_{\mathcal{T}})\|_{L^2(\Omega)} + \|h_{\mathcal{E}}^{1/2}(\phi - \phi_{\mathcal{E}})\|_{L^2(\Gamma_N)} \right). \tag{4.48}$$

The constant $C > 0$ only depends on the shape regularity constant $\sigma(\mathcal{T})$.

Remark. For $f \in H^1(\mathcal{T})$ holds $\|h_{\mathcal{T}}(f - f_{\mathcal{T}})\|_{L^2(\Omega)} = \mathcal{O}(h^2)$. For $\phi \in C^1(\mathcal{E}_N)$ holds $\|h_{\mathcal{E}}^{1/2}(\phi - \phi_{\mathcal{E}})\|_{L^2(\Gamma_N)} = \mathcal{O}(h^{3/2})$. Even for $u \in H^2(\Omega)$, the error as well as the error estimator η only satisfy $\|u - u_h\|_{H^1(\Omega)} = \mathcal{O}(h) = \eta$. Therefore, the two terms on the right-hand side are of higher order. □

Proof of Theorem 4.13. **1. step.** Estimate (4.48) is a consequence of (4.47) since

$$\eta \le \eta_{\mathcal{T}} = \Big(\sum_{T \in \mathcal{T}} \eta_T^2 \Big)^{1/2} \le 2C \left(\|u - u_h\|_{H^1(\Omega)} + \|h_{\mathcal{T}}(f - f_{\mathcal{T}})\|_{L^2(\Omega)} + \|h_{\mathcal{E}}^{1/2}(\phi - \phi_{\mathcal{E}})\|_{L^2(\Gamma_N)} \right).$$

Here, the factor $2 = 4^{1/2}$ appears since each element $T \in \mathcal{T}$ belongs at most to four patches $\omega_{T'}$.

The proof of (4.47) is split into three steps, where we consider each of the three contributions of η_T separately.

2. step. There holds

$$\|h_{\mathcal{T}} f\|_{L^2(T)} \le C \left(\|\nabla(u - u_h)\|_{L^2(T)} + \|h_{\mathcal{T}}(f - f_{\mathcal{T}})\|_{L^2(T)} \right) : \tag{4.49}$$

For $T \in \mathcal{T}$, we define the **element bubble function**

$$b_T := \prod_{z \in \mathcal{K}_T} \zeta_z \in H^1_0(T) \cap \mathcal{P}^3(T)$$

as product of all three hat functions. It is essential to observe the following estimate

$$\|f_\mathcal{T} b_T\|_{L^2(T)} \le \|f_\mathcal{T} b_T^{1/2}\|_{L^2(T)} \le \|f_\mathcal{T}\|_{L^2(T)} \le C_{\rm ref}\|f_\mathcal{T} b_T\|_{L^2(T)}, \tag{4.50}$$

where the existence of an independent constant $C_{\rm ref} > 0$ follows from a scaling argument. We stress, however, that $\|f_\mathcal{T} b_T\|_{L^2(T)}$ and hence $C_{\rm ref}$ — since $f_\mathcal{T}$ is constant on T — can explicitly be computed. In the following, the main idea is to use integration by parts for $v := f_\mathcal{T} b_T \in H^1_0(T)$ to show

$$\begin{aligned} C_{\rm ref}^{-2}\|f_\mathcal{T}\|^2_{L^2(T)} \le \|f_\mathcal{T} b_T^{1/2}\|^2_{L^2(T)} = (f_\mathcal{T}\,;\,v)_{L^2(T)} &= (f_\mathcal{T} - f\,;\,v)_{L^2(T)} + (f - \Delta u_h\,;\,v)_{L^2(T)} \\ &= (f_\mathcal{T} - f\,;\,v)_{L^2(T)} + (\nabla(u-u_h)\,;\,\nabla v)_{L^2(T)}. \end{aligned}$$

Now, we estimate each of the two scalar products on the right-hand side by use of the Cauchy inequality. Together with $v = f_\mathcal{T} b_T \in \mathcal{P}^3(\mathcal{T})$ we observe

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

as well as

$$\begin{aligned} (\nabla(u-u_h)\,;\,\nabla v)_{L^2(T)} &\le \|\nabla(u-u_h)\|_{L^2(T)}\|\nabla(f_T b_T)\|_{L^2(T)} \\ &\le C_{\rm inv} h_T^{-1}\|\nabla(u-u_h)\|_{L^2(T)}\|f_\mathcal{T} b_T\|_{L^2(T)} \\ &\le C_{\rm inv} h_T^{-1}\|\nabla(u-u_h)\|_{L^2(T)}\|f_\mathcal{T}\|_{L^2(T)}. \end{aligned}$$

Altogether, we see

$$h_T\|f_\mathcal{T}\|_{L^2(T)} \le C_{\rm ref}^2\big(h_T\|f_\mathcal{T} - f\|_{L^2(T)} + C_{\rm inv}\|\nabla(u-u_h)\|_{L^2(T)}\big),$$

which finally results in

$$h_T\|f\|_{L^2(T)} \le (1 + C_{\rm ref}^2)\big(h_T\|f_\mathcal{T} - f\|_{L^2(T)} + C_{\rm inv}\|\nabla(u-u_h)\|_{L^2(T)}\big).$$

3. step. For an interior edge $E \in \mathcal{E}_\Omega$, there holds

$$h_E^{1/2}\|[\![\partial_n u_h]\!]\|_{L^2(E)} \le C\left(\|\nabla(u-u_h)\|_{L^2(\omega_E)} + \|h_\mathcal{T}(f - f_\mathcal{T})\|_{L^2(\omega_E)}\right): \tag{4.51}$$

To prove this estimate, we define the **edge bubble function**

$$b_E := \prod_{z \in \mathcal{K}_E} \zeta_z \in H^1_0(\omega_E) \cap \mathcal{P}^2(\mathcal{T}).$$

The essential estimate reads

$$\|b_E\|_{L^2(E)} \le \|b_E^{1/2}\|_{L^2(E)} \le h_E^{1/2} \le C_{\rm ref}\|b_E\|_{L^2(E)}, \tag{4.52}$$

where the constant $C_{\mathrm{ref}} > 0$ is independent of E. In particular, this provides

$$C_{\mathrm{ref}}^{-2}\|[\![\partial_n u_h]\!]\|_{L^2(E)}^2 \le \|[\![\partial_n u_h]\!] b_E^{1/2}\|_{L^2(E)}^2 = ([\![\partial_n u_h]\!] \,;\, [\![\partial_n u_h]\!] b_E)_{L^2(E)}.$$

Let $T_E^+, T_E^- \in \mathcal{T}$ be the unique elements with $T_E^+ \cap T_E^- = E$ and $\omega_E = T_E^+ \cup T_E^-$. Note that $v := [\![\partial_n u_h]\!] b_E \in \mathcal{P}^2(T_E^\pm)$ satisfies $v|_{\partial T_E^\pm \setminus E} = 0$. Therefore, integration by parts on $T_E^\pm$ proves

$$\begin{aligned}
([\![\partial_n u_h]\!] \,;\, v)_{L^2(E)} &= (\partial_n u_h \,;\, v)_{L^2(\partial T_E^+)} + (\partial_n u_h \,;\, v)_{L^2(\partial T_E^-)} \\
&= (\nabla u_h \,;\, \nabla v)_{L^2(\omega_E)} \\
&= (\nabla(u_h - u) \,;\, \nabla v)_{L^2(\omega_E)} + (f \,;\, v)_{L^2(\omega_E)} \\
&\le \big(C_{\mathrm{inv}}\|\nabla(u_h - u)\|_{L^2(\omega_E)} + \|h_{\mathcal{T}} f\|_{L^2(\omega_E)}\big)\|h_{\mathcal{T}}^{-1} v\|_{L^2(\omega_E)},
\end{aligned}$$

where we have applied the Cauchy inequality and an inverse estimate for $v \in \mathcal{P}^2(\mathcal{T})$. For $T \in \{T_E^+, T_E^-\}$ holds

$$\|v\|_{L^2(T)} = |[\![\partial_n u_h]\!]_E| \|b_E\|_{L^2(T)} \le |T|^{1/2} |[\![\partial_n u_h]\!]_E| \le \frac{h_T^{1/2}}{\sqrt{2}} \|[\![\partial_n u_h]\!]\|_{L^2(E)},$$

since $|T| \le \frac{1}{2} h_T h_E$. From this, we infer

$$h_E^{1/2}\|h_{\mathcal{T}}^{-1} v\|_{L^2(\omega_E)} \le \|h_{\mathcal{T}}^{-1/2} v\|_{L^2(\omega_E)} \le \|[\![\partial_n u_h]\!]\|_{L^2(E)}.$$

This finally proves

$$h_E^{1/2}\|[\![\partial_n u_h]\!]\|_{L^2(E)}^2 \le C_{\mathrm{ref}}^2\big(C_{\mathrm{inv}}\|\nabla(u_h - u)\|_{L^2(\omega_E)} + \|h_{\mathcal{T}} f\|_{L^2(\omega_E)}\big)\, \|[\![\partial_n u_h]\!]\|_{L^2(E)}$$

and we may conclude this step by use of step 2 to dominate $\|h_{\mathcal{T}} f\|_{L^2(\omega_E)}$.

4. step. For $T \in \mathcal{T}$ and a Neumann edge $E \in \mathcal{E}_N \cap \mathcal{E}_T$, it holds

$$\begin{aligned}
&h_E^{1/2}\|\phi - \partial_n u_h\|_{L^2(E)} \\
&\qquad \le C\big(\|h_{\mathcal{E}}^{1/2}(\phi - \phi_{\mathcal{E}})\|_{L^2(E)} + \|\nabla(u - u_h)\|_{L^2(T)} + \|h_{\mathcal{T}}(f - f_{\mathcal{T}})\|_{L^2(T)}\big) :
\end{aligned} \tag{4.53}$$

We consider again the edge bubble function $b_E \in \mathcal{P}^2(T)$ and note that $b_E|_{\partial T \setminus E} = 0$. With $v := (\phi_{\mathcal{E}} - \partial_n u_h) b_E \in \mathcal{P}^2(T)$, we proceed as in step 3 and obtain

$$C_{\mathrm{ref}}^{-2}\|\phi_{\mathcal{E}} - \partial_n u_h\|_{L^2(E)}^2 \le (\phi_{\mathcal{E}} - \partial_n u_h \,;\, v)_{L^2(E)} = (\phi_{\mathcal{E}} - \phi \,;\, v)_{L^2(E)} + (\phi - \partial_n u_h \,;\, v)_{L^2(E)}.$$

For the second term, we employ integration by parts to see

$$\begin{aligned}
(\phi - \partial_n u_h \,;\, v)_{L^2(E)} &= (\partial_n(u - u_h) \,;\, v)_{L^2(\partial T)} \\
&= (\nabla(u - u_h) \,;\, \nabla v)_{L^2(T)} - (f \,;\, v)_{L^2(T)} \\
&\le \big(C_{\mathrm{inv}}\|\nabla(u - u_h)\|_{L^2(T)} + \|h_{\mathcal{T}} f\|_{L^2(T)}\big) h_E^{-1/2}\|\phi_{\mathcal{E}} - \partial_n u_h\|_{L^2(E)}.
\end{aligned}$$

The first term is estimated by the Cauchy inequality directly

$$(\phi_{\mathcal{E}} - \phi \,;\, v)_{L^2(E)} \le \|h_{\mathcal{E}}^{1/2}(\phi_{\mathcal{E}} - \phi)\|_{L^2(E)}\|h_{\mathcal{E}}^{-1/2} v\|_{L^2(E)}.$$

There holds

$$\|v\|_{L^2(E)} = |(\phi_\mathcal{E} - \partial_n u_h)|_E| \|b_E\|_{L^2(E)} \leq h_E^{1/2} |(\phi_\mathcal{E} - \partial_n u_h)|_E| = \|\phi_\mathcal{E} - \partial_n u_h\|_{L^2(E)}.$$

Altogether, we thus have shown

$$\begin{aligned} &h_E^{1/2} \|\phi_\mathcal{E} - \partial_n u_h\|_{L^2(E)}^2 \\ &\qquad \leq C_{\rm ref}^2 \big(C_{\rm inv} \|\nabla(u - u_h)\|_{L^2(T)} + \|h_\mathcal{T} f\|_{L^2(T)} + \|h_\mathcal{E}^{1/2} (\phi_\mathcal{E} - \phi)\|_{L^2(E)} \big) \, \|\phi_\mathcal{E} - \partial_n u_h\|_{L^2(E)}. \end{aligned}$$

Here, $\|h_\mathcal{T} f\|_{L^2(T)}$ is estimated by step 2, and $\phi_\mathcal{E}$ on the left-hand side is replaced by ϕ with the help of the triangle inequality. ■

Exercise 27. Prove that $f_\mathcal{T}$ in Theorem 4.13 can be replaced by an arbitrary $\mathcal{T}$-elementwise polynomial $f_\mathcal{T} \in \mathcal{P}^m(\mathcal{T})$. The constant $C > 0$ in (4.47)–(4.48) then additionally depends on the polynomial degree $m \in \mathbb{N}_0$. □

Remark. With the help of a so-called extension operator that extends a polynomial $p : E \to \mathbb{R}$ to a polynomial $F_{\rm ext} p : T \to \mathbb{R}$, one can show that $\phi_\mathcal{E}$ in Theorem 4.13 can be replaced by an arbitrary $\mathcal{E}_N$-edgewise polynomial (with respect to the arclength). □

Actually, the volume residual contribution $\|h_\mathcal{T} f\|_{L^2(\Omega)} = \mathcal{O}(h)$ to η can be improved. This is done in the following exercise, where this term is replaced by some higher-order term $\mathcal{O}(h^2)$.

Exercise 28. Let $\Omega_z = \operatorname{supp}(\zeta_z)$ denote the node patch of $z \in \mathcal{K}$. For $f \in L^2(\Omega)$, let $f_z := |\Omega_z|^{-1} \int_{\Omega_z} f \, dx$ denote the corresponding integral mean. Prove the following claims:

(i) For all inner nodes $z \in \mathcal{K} \backslash \Gamma$, it holds

$$\int_{\Omega_z} f f_z \zeta_z \, dx \leq C \Big(\sum_{\substack{E \in \mathcal{E}_\Omega \\ z \in E}} \|[\![\partial_n u_h]\!]\|_{L^2(E)}^2 \Big)^{1/2} \|h_\mathcal{T}^{-1/2} f_z\|_{L^2(\Omega_z)}.$$

(ii) For all inner nodes $z \in \mathcal{K} \backslash \Gamma$ and elements $T \in \mathcal{T}$ with $z \in T$, it holds

$$C^{-1} \, \|h_\mathcal{T} f\|_{L^2(T)}^2 \leq \|h_\mathcal{T}(f - f_z)\|_{L^2(\Omega_z)}^2 + \sum_{\substack{E \in \mathcal{E}_\Omega \\ z \in E}} \|h_\mathcal{E}^{1/2} [\![\partial_n u_h]\!]\|_{L^2(E)}^2.$$

(iii) Derive the equivalence

$$\begin{aligned} C^{-1} \eta^2 \leq \widetilde{\eta}^2 :=& \|h_\mathcal{E}^{1/2} [\![\partial_n u_h]\!]\|_{L^2(\mathcal{E}_\Omega)}^2 + \|h_\mathcal{E}^{1/2} (\phi - \partial_n u_h)\|_{L^2(\mathcal{E}_N)}^2 \\ &+ \sum_{z \in \mathcal{K} \backslash \Omega} \|h_\mathcal{T}(f - f_z)\|_{L^2(\Omega_z)}^2 \leq C \, \eta^2. \end{aligned}$$

(iv) Conclude that the improved error estimator $\widetilde{\eta}$ is reliable and efficient.

(v) In what sense is the error estimator $\widetilde{\eta}$ improved when compared to η.

The constant $C > 0$ in (i)–(iii) depends only on γ-shape regularity of $\mathcal{T}$. □

4.4 Adaptive Mesh-Refining Algorithm

Usually, a posteriori error estimates are not only used to estimate the (unknown) error $\|\nabla(u - u_h)\|_{L^2(\Omega)}$ but even to steer the local mesh-refinement. Let

$$\eta := \Big(\sum_{T \in \mathcal{T}} \eta(T)^2 \Big)^{1/2}$$

be an a posteriori error estimator, where the quantities $\eta(T) := \eta_T$ reflect —at least heuristically— the (unknown) local error $\|\nabla(u - u_h)\|_{L^2(T)}$ for all $T \in \mathcal{T}$. We then aim to refine only the elements $T \in \mathcal{T}$, where $\eta(T)$ is large. Therefore, the quantities $\eta(T)$ are usually called **refinement indicators** (or error indicators). To state our version of an adaptive algorithm, we introduce some additional notation which will be used from now on.

- the index $\ell \in \mathbb{N}_0$ denotes the step of the adaptive algorithm,
- $\mathcal{T}_\ell$ is the mesh in the ℓ-th step of the adaptive algorithm.
- $\mathcal{N}_\ell$ and $\mathcal{E}_\ell$ denote the associated sets of nodes and edges, respectively.
- $U_\ell \in \mathcal{X}_\ell := \mathcal{S}^1_D(\mathcal{T}_\ell)$ denotes the Galerkin solution in the ℓ-th step.
- $h_\ell \in \mathcal{P}^0(\mathcal{T}_\ell)$, $h_\ell|_T := \mathrm{diam}(T)$ is the local mesh-side function.

With this notation, one common strategy is the following: Let $\theta \in (0, 1)$ be the parameter for the adaptive algorithm.

Algorithm 4.14 (Adaptive Mesh-Refinement). Input: *Initial triangulation $\mathcal{T}_0$, tolerance $\tau > 0$, adaptivity parameter $\theta \in (0, 1)$, counter $\ell := 0$.*

(i) *Compute discrete solution U_ℓ.*

(ii) *Compute refinement indicators $\eta_\ell(T)$ and error estimator $\eta_\ell = \big(\sum_{T \in \mathcal{T}_\ell} \eta_\ell(T)^2 \big)^{1/2}$.*

(iii) *Stop computation provided that $\eta_\ell \le \tau$*

(iv) *Choose the minimal set $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ of marked elements such that*

$$\theta\, \eta_\ell^2 = \theta \sum_{T \in \mathcal{T}_\ell} \eta_\ell(T)^2 \le \sum_{T \in \mathcal{M}_\ell} \eta_\ell(T)^2. \tag{4.54}$$

(v) *Generate a new regular mesh $\mathcal{T}_{\ell+1}$, where at least all marked elements have been refined.*

(vi) *Update $\ell \mapsto \ell + 1$ and goto* (i).

Output: *Finite sequence of discrete solutions U_ℓ and corresponding error estimators η_ℓ.*

Remark. Clearly, the stopping criterion (iii) is only meaningful if η_ℓ is reliable and if the reliability constant in $\|\nabla(u-U_\ell)\|_{L^2(\Omega)} \le C_{\rm rel}\,\eta_\ell$ is known. In practice, runtime and storage requirements are the limiting quantities for a numerical simulation. Usually, one thus uses rather a maximal runtime or a maximal storage requirement, e.g., the maximal number of elements, as a stopping criterion. Adaptivity is then used to obtain an —in some sense— optimal approximation with respect to these side constraints. □

Remark. The marking criterion (4.54) was introduced by DÖRFLER (1996). It will be crucial to prove convergence of U_ℓ to the exact solution $u \in H^1_D(\Omega)$ of (4.1). Note that the choice $\theta \to 0$ in (4.54) leads to highly adapted meshes, whereas $\theta \to 1$ corresponds to (almost) uniform mesh-refinement. However, for small θ, only a few elements are refined per step. This may result in too many steps in the sense that usually the assembly of the Galerkin data is the most time consuming part of the algorithm. In practice, a good compromise between sufficient mesh-adaption and as few steps in the loop as possible appears to be $\theta \approx 0.25$. □

Remark. In the beginning of the analysis of adaptive FEM, Babuška proposed the following marking criterion: An element $T \in \mathcal{T}$ is marked for refinement if and only if

$$\eta_T \ge \theta\, \max\big\{\eta_{T'}\,\big|\,T' \in \mathcal{T}\big\}, \tag{4.55}$$

which is called **bulk criterion** in the literature. Convergence (but *not* optimality) of this version of adaptive FEM was proven by MORIN, SIEBERT & VEESER (2008). Very recently, DIENING, KREUZER & STEVENSON (2014) proved the so-called *instance optimality* of the adaptive algorithm for some extended bulk criterion. □

4.4.1 Red-Green-Blue Refinement

It now remains to discuss the mesh-refinement. Recall that all error estimates are affected by the shape regularity $\sigma(\mathcal{T}_\ell)$ in the sense that the involved constants become unbounded for $\sigma(\mathcal{T}_\ell) \xrightarrow{\ell\to\infty} \infty$. Therefore, the mesh-refinement has to take care of the interior angles of the elements $T \in \mathcal{T}_\ell$ since $\sigma(\mathcal{T}_\ell)$ tends to infinity if and only if the minimal interior angle of the triangulation tends to zero. We follow the so-called **red-green-blue strategy** (or **RGB-refinement**): This refinement strategy is based on edge-refinement. First, we thus use the following marking rule:

- If an element $T \in \mathcal{T}_\ell$ is marked for refinement, we mark all edges $E \in \mathcal{E}_T$ for refinement.

We now proceed recursively as follows:

- For each element $T \in \mathcal{T}_\ell$, we mark its *longest* edge $E \in \mathcal{E}_T$ for refinement provided that $\mathcal{E}_T$ contains a marked edge.

Each marked edge will be halved, i.e., the midpoint m_E of a marked edge belongs to the new set $\mathcal{K}_{\ell+1}$ of nodes. Finally, we have the following refinement rules, for all $T \in \mathcal{T}_\ell$:

- If no edge in $\mathcal{E}_T$ is marked for refinement, T is not refined, i.e., $T \in \mathcal{T}_{\ell+1}$.
- If all edges in $\mathcal{E}_T$ are marked, we use a **red-refinement** of T, i.e., T is refined uniformly into four similar triangles, cf. Figure 4.3.
- If one edge in $\mathcal{E}_T$ is marked (and hence the longest edge), we use a **green-refinement**, i.e., T is refined into two triangles, cf. Figure 4.4.
- If two edges in $\mathcal{E}_T$ are marked — one of which is, according to the marking rule, the longest edge of T —, we use a **blue-refinement**, i.e., T is split into three triangles, cf. Figure 4.5.

In Figure 4.6, we visualize a simple example for an RGB-refined mesh.

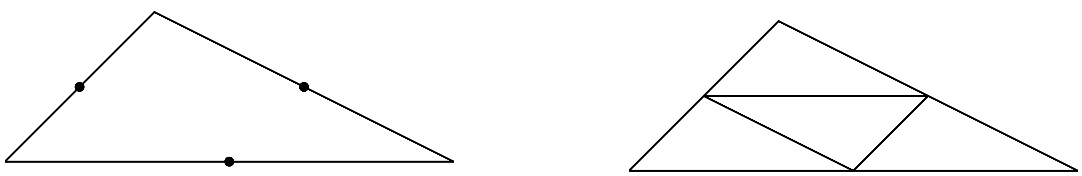


FIGURE 4.3. *Red-refinement: If all edges of a triangle $T \in \mathcal{T}_\ell$ are marked (left), T is refined into four similar triangles $T_1, T_2, T_3, T_4 \in \mathcal{T}_{\ell+1}$ (right).*

FIGURE 4.4. *Green-refinement: If only the longest edge of a triangle $T \in \mathcal{T}_\ell$ is marked (left), T is refined into two new triangles $T_1, T_2 \in \mathcal{T}_{\ell+1}$ (right).*

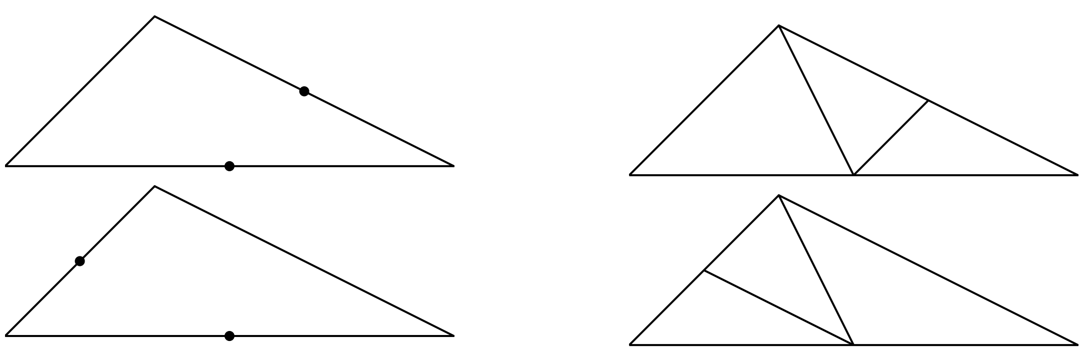


FIGURE 4.5. *Blue-refinement: If besides the longest edge of a triangle $T \in \mathcal{T}_\ell$ just one other edge is marked for refinement (left), T is refined into three new triangles $T_1, T_2, T_3 \in \mathcal{T}_{\ell+1}$ (right).*

We state the following elementary but important theorem without a proof.

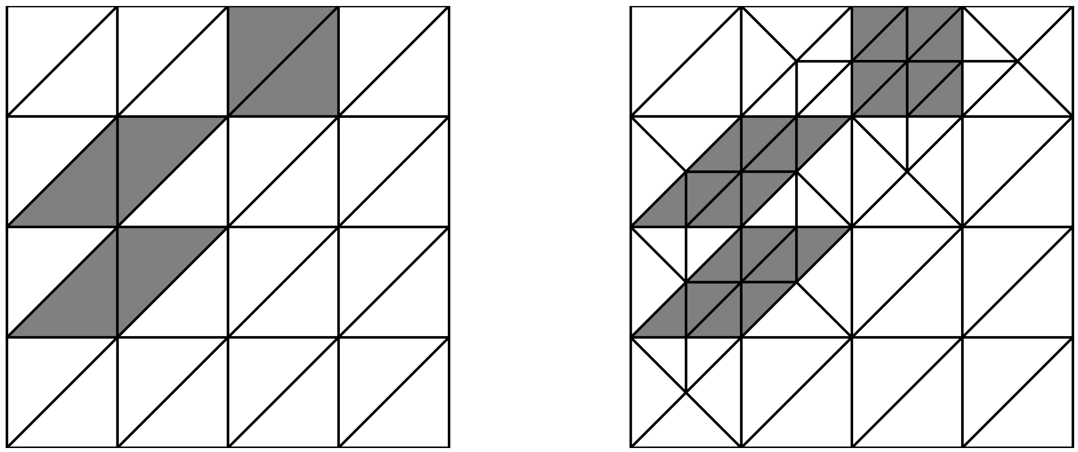


FIGURE 4.6. *The left plot shows an initial mesh $\mathcal{T}_\ell$ with marked elements coloured in grey. The right plot shows the mesh $\mathcal{T}_{\ell+1}$ obtained by RGB-refinement of the marked elements. The grey elements are obtained by uniform refinement of a marked element $T \in \mathcal{T}_0$.*

Theorem 4.15. *Let $\mathcal{T}_0$ be a regular triangulation such that $\varepsilon > 0$ is a lower bound for the smallest angle of a triangle $T \in \mathcal{T}_0$. Let $\mathcal{T}_\ell$ be a sequence of meshes, where $\mathcal{T}_\ell$ is obtained by RGB-refinement of the mesh $\mathcal{T}_{\ell-1}$ and where the set $\mathcal{M}_{\ell-1} \subseteq \mathcal{T}_{\ell-1}$ of marked elements is arbitrary. Then, $\mathcal{T}_\ell$ is regular and the smallest angle of all triangles $T \in \mathcal{T}_\ell$ is bounded from below by $\varepsilon/2$. In particular, there holds*

$$\sup_{\ell \in \mathbb{N}} \sigma(\mathcal{T}_\ell) < \infty, \tag{4.56}$$

which is an equivalent formulation for the fact that the smallest angles of the triangulations $\mathcal{T}_\ell$ do not tend to zero.

4.5 Convergence of Adaptive FEM

In the following, we aim to show that Algorithm 4.14 creates a sequence U_ℓ of discrete solutions which converges to the exact solution $u \in H := H^1_D(\Omega)$. The adaptive algorithm generates a sequence $\mathcal{X}_\ell = \mathcal{S}^1_D(\mathcal{T}_\ell)$ of finite dimensional nested subspaces of H, i.e., $\mathcal{X}_\ell \subsetneqq \mathcal{X}_{\ell+1}$ for all $\ell \in \mathbb{N}_0$, since $\mathcal{T}_{\ell+1}$ is some refinement of $\mathcal{T}_\ell$. We first stress that the sequence U_ℓ is always convergent to some limit $U_\infty \in H$. However, we even stress that one may in general expect that $U_\infty \neq u$.

Exercise 29. Let $\mathcal{X}_\ell$ be nested subspaces of a Hilbert space H, i.e., $\mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1}$ for all $\ell \in \mathbb{N}_0$. Let $\langle\!\langle \cdot\, ; \cdot \rangle\!\rangle$ be an elliptic and continuous bilinear form on H with corresponding Galerkin solutions $U_\ell \in \mathcal{X}_\ell$. Prove that the limit $U_\infty := \lim_{\ell\to\infty} U_\ell$ exists in H. **Hint:** Define $\mathcal{X}_\infty$ as the closure of $\bigcup_{\ell=0}^\infty \mathcal{X}_\ell$ in H. Let $U_\infty \in \mathcal{X}_\infty$ be the corresponding Galerkin solution, and prove that U_∞ is the limit of the sequence U_ℓ. $\square$

Exercise 30. Let $H = H^1_D(\Omega)$ and $\mathcal{X}_\ell = \mathcal{S}^1_D(\mathcal{T}_\ell)$, where the regular initial mesh $\mathcal{T}_0$ is given and where $\mathcal{T}_\ell$ is obtained iteratively by uniform refinement of $\mathcal{T}_{\ell-1}$. Prove that $\mathcal{X}_\infty = H$ for the space $\mathcal{X}_\infty$ from Exercise 29. $\square$

The interpretation of the last exercises is the following: For uniform mesh-refinement, there usually holds $\mathcal{X}_\infty = H$ and thus $u = U_\infty$, i.e., we have convergence of the sequence of discrete solutions U_ℓ towards the unique solution u. However, adaptive mesh-refinement may lead to $\mathcal{X}_\infty \subsetneqq H$. Consequently, the question arrises whether the adaptive algorithm guarantees $U_\infty = u$ or not. This will be discussed in the following sections.

Throughout the subsequent section, we use the following notation, which is now collected for the convenience of the reader:

- $U_\ell \in \mathcal{X}_\ell := \mathcal{S}^1_D(\mathcal{T}_\ell)$ denotes the Galerkin solution.
- For $T \in \mathcal{T}_\ell$ and some $V \in \mathcal{S}^1_D(\mathcal{T}_\ell)$, $\eta_\ell(T, V)$ denotes the associated refinement indicator, e.g.,

$$\eta_\ell(T,V)^2 = h_T^2\|f\|^2_{L^2(T)} + h_T\|[\![\partial_n V]\!]\|^2_{L^2(\partial T\cap\Omega)} + h_T\|\phi - \partial_n V\|^2_{L^2(\partial T\cap\Gamma_N)}. \tag{4.57}$$

- For some subset $\mathcal{M} \subseteq \mathcal{T}_\ell$ and $V \in \mathcal{S}^1_D(\mathcal{T}_\ell)$, let $\eta_\ell(\mathcal{M}, V) := \big(\sum_{T\in\mathcal{M}} \eta_\ell(T, V)^2\big)^{1/2}$.
- We abbreviate $\eta_\ell(\mathcal{M}) = \eta_\ell(\mathcal{M}, U_\ell)$ and $\eta_\ell = \eta_\ell(\mathcal{T}_\ell)$.

Note that in case of (4.57), η_ℓ is the residual a posteriori error estimator discussed in Section 4.3. We recall some technical terms, proven above for the residual error estimator η_ℓ.

- η_ℓ is **reliable** if

$$\|u - U_\ell\|_H \le C_{\rm rel}\eta_\ell. \tag{4.58}$$

- η_ℓ is **efficient** (up to oscillation terms which depend only on $\mathcal{T}_\ell$), if

$$\eta_\ell \le C_{\rm eff}\big(\|u - U_\ell\|_H + {\rm osc}_\ell\big), \tag{4.59}$$

  where ${\rm osc}_\ell := {\rm osc}_\ell(\mathcal{T}_\ell)$, ${\rm osc}_\ell(\mathcal{M}) := \big(\sum_{T\in\mathcal{M}} {\rm osc}_\ell(T)^2\big)^{1/2}$ for $\mathcal{M} \subseteq \mathcal{T}_\ell$, and

$${\rm osc}_\ell(T)^2 := h_T^2\|f - f_\mathcal{T}\|_{L^2(T)} + h_T\|\phi - \phi_\mathcal{E}\|^2_{L^2(\partial T\cap\Gamma_N)}. \tag{4.60}$$

- The set $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ of marked elements is usually assumed to satisfy the **Dörfler marking**

$$\theta\,\eta_\ell \le \eta_\ell(\mathcal{M}_\ell) \tag{4.61}$$

  for some fixed parameter $\theta \in (0, 1)$.

Exercise 31. Prove that $\|u - U_\ell\|_H$ as well as osc_ℓ are monotonously decreasing for $\ell \to \infty$. Prove that in case of the residual-based indicators (4.57), there holds $\mathrm{osc}_\ell(T) \le \eta_\ell(T)$ for all $T \in \mathcal{T}_\ell$, i.e., the error estimator dominates the oscillation terms. □

The following convergence theorem is a result of CASCÓN, KREUZER, NOCHETTO & SIEBERT from 2008, where it is proven that the combined error quantity, which consists of error and error estimator, has a contraction property. We stress two important observations:

- For their analysis, CASCÓN, KREUZER, NOCHETTO, and SIEBERT re-define the **mesh width**
$$h_T := |T|^{1/2} \quad \text{for } T \in \mathcal{T}_\ell, \tag{4.62}$$
whereas we considered $\mathrm{diam}(T)$ before. Note that, however, $|T| \le \mathrm{diam}(T)^2 \le 2\sigma(\mathcal{T})|T|$ so that both definition are equivalent for shape regular meshes, and we shall use the new definition in what follows.
- If $T \in \mathcal{T}_\ell$ is refined, each son $T' \in \mathcal{T}_{\ell+1}$ satisfies at least $|T'| \le |T|/2$, which now results in a strict reduction $h_{T'} \le h_T/\sqrt{2}$ of the local mesh-width (which fails, in general, for the usual definition $h_T = \mathrm{diam}(T)$). This observation is used in step 2 of the proof of the following theorem.

We note that the analysis holds for general symmetric problems. For non-symmetric problems, the correspoding result has been open until FEISCHL, FÜHRER & PRAETORIUS (2014).

Theorem 4.16 (Cascón, Kreuzer, Nochetto & Siebert '08). *Suppose that the set of marked elements $\mathcal{M}_\ell$ satisfies the Dörfler marking for some fixed $\theta \in (0,1)$. Then, there are constants $\kappa > 0$ and $q \in (0,1)$ which depend only on θ and uniform γ-shape regularity of $\mathcal{T}_\ell$ for all $\ell \in \mathbb{N}_0$, such that*
$$\big(\|\nabla(u - U_{\ell+1})\|_{L^2(\Omega)}^2 + \kappa\,\eta_{\ell+1}^2\big)^{1/2} \le q\,\big(\|\nabla(u - U_\ell)\|_{L^2(\Omega)}^2 + \kappa\,\eta_\ell^2\big)^{1/2} \text{ for all } \ell \in \mathbb{N}. \tag{4.63}$$
In particular, this implies convergence $\lim_{\ell\to\infty} \|\nabla(u - U_\ell)\|_{L^2(\Omega)} = 0 = \lim_{\ell\to\infty} \eta_\ell$.

Proof. **1. step.** There holds the following quasi-triangle inequality for the error estimator
$$\eta_\ell(V) \le \eta_\ell(W) + C_\Delta\,\|\nabla(V - W)\|_{L^2(\Omega)} \quad \text{for all } V, W \in \mathcal{S}^1_D(\mathcal{T}_\ell) \tag{4.64}$$
with some constant $C_\Delta > 0$ which depends only on $\sigma(\mathcal{T}_\ell)$: From the triangle inequalities in ℓ_2 and L^2, we infer
$$\begin{aligned}
\eta_\ell(V) &= \Big[\|h_\ell f\|_{L^2(\Omega)}^2 + \sum_{T\in\mathcal{T}_\ell} h_T\big(\|[\![\partial_n V]\!]\|_{L^2(\partial T\cap\Omega)}^2 + \|\phi - \partial_n V\|_{L^2(\partial T\cap\Gamma_N)}^2\big)\Big]^{1/2}\\
&\le \Big[\|h_\ell f\|_{L^2(\Omega)}^2 + \sum_{T\in\mathcal{T}_\ell} h_T\big(\|[\![\partial_n W]\!]\|_{L^2(\partial T\cap\Omega)}^2 + \|\phi - \partial_n W\|_{L^2(\partial T\cap\Gamma_N)}^2\big)\Big]^{1/2}\\
&\quad + \Big[\sum_{T\in\mathcal{T}_\ell} h_T\big(\|[\![\partial_n (V-W)]\!]\|_{L^2(\partial T\cap\Omega)}^2 + \|\partial_n (V-W)\|_{L^2(\partial T\cap\Gamma_N)}^2\big)\Big]^{1/2}.
\end{aligned}$$

For fixed $T \in \mathcal{T}_\ell$ and $E \in \mathcal{E}_T$, a scaling argument proves

$$h_T\big(\|[\![\partial_n(V-W)]\!]\|^2_{L^2(E\cap\Omega)} + \|\partial_n(V-W)\|^2_{L^2(E\cap\Gamma_N)}\big) \lesssim \|\nabla(V-W)\|^2_{L^2(\omega_E)},$$

where the constant depends only on $\sigma(\mathcal{T}_\ell)$. Consequently, we end up with (4.64).

2. step. There holds an estimator reduction in the sense that there is a constant $\varrho \in (0,1)$ with

$$\eta^2_{\ell+1} \le (1+\delta)\varrho\,\eta^2_\ell + C_\delta\|\nabla(U_{\ell+1}-U_\ell)\|^2_{L^2(\Omega)} \quad \text{for all } \delta > 0, \tag{4.65}$$

where $C_\delta > 0$ depends only on δ and $C_\Delta > 0$. The constant ϱ depends only on θ and the reduction of the mesh-side on marked elements: Let $\Omega_* := \bigcup_{T\in\mathcal{M}_\ell} T$ denote the subdomain of Ω, where the elements are marked. Recall that $h_{T'} \le h_T/\sqrt{2}$ for all sons $T' \in \mathcal{T}_{\ell+1}$ of a marked element $T \in \mathcal{M}_\ell$. The crucial step is to observe that the error indicators

$$\eta_\ell(T,V)^2 = h_T^2\|f\|^2_{L^2(T)} + h_T\|[\![\partial_n V]\!]\|^2_{L^2(\partial T\cap\Omega)} + h_T\|\phi - \partial_n V\|^2_{L^2(\partial T\cap\Gamma_N)}.$$

for $h_T = |T|^{1/2}$ lead to

$$\begin{aligned}
\eta_{\ell+1}(U_\ell)^2 &= \sum_{\substack{T'\in\mathcal{T}_{\ell+1}\\ T'\subseteq\Omega_*}} \eta_{\ell+1}(T',U_\ell)^2 + \sum_{\substack{T'\in\mathcal{T}_{\ell+1}\\ T'\subseteq\overline{\Omega\backslash\Omega_*}}} \eta_{\ell+1}(T',U_\ell)^2 \\
&\le \frac{1}{\sqrt{2}} \sum_{\substack{T\in\mathcal{T}_\ell\\ T\subseteq\Omega_*}} \eta_\ell(T,U_\ell)^2 + \sum_{\substack{T\in\mathcal{T}_\ell\\ T\subseteq\overline{\Omega\backslash\Omega_*}}} \eta_\ell(T,U_\ell)^2 \\
&= 2^{-1/2}\,\eta_\ell(\mathcal{M}_\ell)^2 + \eta_\ell(\mathcal{T}_\ell\backslash\mathcal{M}_\ell)^2 \\
&= (2^{-1/2}-1)\eta_\ell(\mathcal{M}_\ell)^2 + \eta_\ell^2.
\end{aligned}$$

By use of the Dörfler marking $\theta\eta_\ell^2 \le \eta_\ell(\mathcal{M}_\ell)^2$, we thus obtain

$$\eta^2_{\ell+1}(U_\ell) \le \eta_\ell^2 - (1-2^{-1/2})\eta_\ell(\mathcal{M}_\ell)^2 \le \varrho\,\eta_\ell^2 \quad \text{with} \quad \varrho := \big(1-\theta(1-2^{-1/2})\big).$$

Now, Young's inequality and step 1 conclude

$$\begin{aligned}
\eta^2_{\ell+1} &\le (1+\delta)\eta_{\ell+1}(U_\ell)^2 + (1+\delta^{-1})C_\Delta^2\,\|\nabla(U_{\ell+1}-U_\ell)\|^2_{L^2(\Omega)} \\
&\le (1+\delta)\varrho\,\eta_\ell^2 + (1+\delta^{-1})C_\Delta^2\,\|\nabla(U_{\ell+1}-U_\ell)\|^2_{L^2(\Omega)}.
\end{aligned}$$

3. step. Proof of contraction property (4.63): Let $\kappa, \delta, \beta > 0$ be constants which are fixed later. Let $\varrho \in (0,1)$ be the given constant from step 2. We recall the Galerkin orthogonality

$$\|\nabla(u-U_\ell)\|^2_{L^2(\Omega)} = \|\nabla(u-U_{\ell+1})\|^2_{L^2(\Omega)} + \|\nabla(U_{\ell+1}-U_\ell)\|^2_{L^2(\Omega)}$$

This and the estimator reduction imply

$$\begin{aligned}
\|\nabla(u-U_{\ell+1})\|^2_{L^2(\Omega)} + \kappa\,\eta^2_{\ell+1} &= \|\nabla(u-U_\ell)\|^2_{L^2(\Omega)} - \|\nabla(U_{\ell+1}-U_\ell)\|^2_{L^2(\Omega)} + \kappa\,\eta^2_{\ell+1} \\
&\le \|\nabla(u-U_\ell)\|^2_{L^2(\Omega)} + (\kappa C_\delta - 1)\,\|\nabla(U_{\ell+1}-U_\ell)\|^2_{L^2(\Omega)} + \kappa(1+\delta)\varrho\,\eta_\ell^2.
\end{aligned}$$

Provided that $\kappa C_\delta \le 1$, we infer

$$\begin{aligned}\|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}^2 + \kappa\,\eta_{\ell+1}^2 &\le \|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 + \kappa(1+\delta)\varrho\,\eta_\ell^2\\ &= \|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 - \kappa\beta\,\eta_\ell^2 + \kappa\big((1+\delta)\varrho+\beta\big)\,\eta_\ell^2.\end{aligned}$$

Reliability $\|\nabla(u-U_\ell)\|_{L^2(\Omega)} \le C_{\rm rel}\eta_\ell$ finally leads to

$$\begin{aligned}\|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}^2 + \kappa\,\eta_{\ell+1}^2 &\le (1-\kappa\beta C_{\rm rel}^{-2})\|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 + \kappa\big((1+\delta)\varrho+\beta\big)\,\eta_\ell^2\\ &\le \max\big\{1-\kappa\beta C_{\rm rel}^{-2}\,,\,(1+\delta)\varrho+\beta\big\}\,\big(\|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 + \kappa\,\eta_\ell^2\big).\end{aligned}$$

It remains to choose the constants κ,δ,β so that $q^2 := \max\big\{1-\kappa\beta C_{\rm rel}^{-2}\,,\,(1+\delta)\varrho+\beta\big\} \in (0,1)$:

- $\varrho := \big(1-\theta(1-2^{-1/2})\big) \in (0,1)$.
- Choose $\delta > 0$ such that $(1+\delta)\varrho < 1$.
- Choose $\kappa > 0$ such that $\kappa C_\delta \le 1$ with $C_\delta := (1+\delta^{-1})C_\Delta^2$.
- Choose $\beta > 0$ such that $(1+\delta)\varrho + \beta < 1$.

This implies $q \in (0,1)$ and concludes the proof. ■

Remark. We again collect the main arguments of the preceeding proof, namely a certain quasi-triangle inequality of the estimator (4.64) and a strict reduction $\eta_{\ell+1}(\texttt{sons}(\mathcal{M}_\ell),U_\ell) \le \kappa\,\eta_\ell(\mathcal{M}_\ell,U_\ell)$ with some $\kappa \in (0,1)$ based on the strict reduction of the local mesh-width for marked elements. Besides this, only Galerkin orthogonality and Dörfler marking are used. Therefore, the proof works for a quite general class of symmetric problems and a variety of error estimators. The original work of Cascón, Kreuzer, Nochetto & Siebert (2008) considers linear second order symmetric and elliptic problems in divergence form and H^1-conforming finite element spaces with fixed polynomial degree. Finally, we stress that the proof also works for higher dimensions $d \ge 2$, where $h_T = |T|^{-1/d}$. For 2D, the usual definition $h_T := \mathrm{diam}(T)$ is sufficient if marked elements are refined, e.g., by red-refinement or bisec(3), since then all edges are bisected. □

Exercise 32. Prove the following variants of Young's inequality, for all $a,b \in \mathbb{R}$ and $\delta > 0$,
- $ab \le \frac{a^2}{2\delta} + \frac{\delta b^2}{2}$.
- $(a+b)^2 \le (1+\delta^{-1})a^2 + (1+\delta)b^2$. □