ENSTA Paris

Lecture notes
Course ANN202

Analysis and approximation of partial differential equations by finite elements

Martin VOHRALÍK[§,¶]

March 8, 2024

[§]Inria, 2 rue Simone Iff, 75589 Paris, France.
[¶]CERMICS, Ecole des Ponts, 77455 Marne-la-Vallée, France.
martin.vohralik@inria.fr

Introduction

These notes were written for the lecture ANN202 "Analyse et approximation par éléments finis d'équations aux dérivées partielles", held at ENSTA Paris in academic year 2020/21. They detail the contents of the lecture catalogued at https://synapses.ensta-paris.fr/catalogue/2020-2021/ue/221/ANN202-analyse-et-approximation-par-elements-finis-d-edp, with description at https://who.rocq.inria.fr/Martin.Vohralik/Pages/FEM.html.

These lectures make a follow-up to the lecture ANN201 "Eléments finis", see https://synapses.ensta-paris.fr/catalogue/2020-2021/ue/212/ANN201-elements-finis. They are organized into preliminary reminder and 6 chapters, each corresponding to one lecture block.

Contents

Contents v

0 Preliminaries 7

0.1 Computational domain Ω 7
0.2 Lebesgue space $L^2(\Omega)$ 7
0.3 Sobolev space $H_0^1(\Omega)$ 8
0.4 Sobolev space $\boldsymbol{H}(\mathrm{div}, \Omega)$ 9
0.5 Computational mesh $\mathcal{T}_h$ 10
0.6 Broken Sobolev space $H^1(\mathcal{T}_h)$ 11
0.7 Poincaré–Friedrichs and trace inequalities 12
0.8 Piecewise polynomial space $\mathcal{P}_p(\mathcal{T}_h)$ 13

1 The nonconforming finite element method: a priori analysis 15

1.1 The Poisson equation 15
1.2 Weak formulation 15
1.3 The nonconforming finite element method 16
1.4 Weakly-continuous subspace of the broken Sobolev space $H^1(\mathcal{T}_h)$ and a broken Poincaré–Friedrichs inequality 17
1.5 Existence and uniqueness of the nonconforming finite element approximation 18
1.6 Nonconforming finite elements for $p = 1$ 19
1.6.1 A basis of V_{h1}^{nc} 19
1.6.2 An interpolation operator for V_{h1}^{nc} 20
1.7 A priori error estimates: best approximation up to the consistency error 22
1.8 A priori error estimates: rate of convergence for $p = 1$ 23
1.9 Complements 25
1.9.1 Proof of Theorem 1.4.4 25
1.9.2 Proof of Theorem 1.8.1 26
1.9.3 A priori error estimates: best approximation up to data oscillation 26
1.9.4 A priori error estimates: rate of convergence for $p \geq 1$ 28

2 Potential reconstruction in $H_0^1(\Omega)$ 31

2.1 Setting 31
2.2 Lagrangian nodes 31
2.3 Potential reconstruction by averaging 32
2.4 Approximation properties 32
2.5 Complements 36
2.5.1 Improved potential reconstruction by averaging 36
2.5.2 Potential reconstruction by solution of local Dirichlet problems and its approximation properties 36

3 Flux reconstruction in $\boldsymbol{H}(\mathrm{div}, \Omega)$ **39**
3.1 Setting and definition . . . 39
3.2 Properties . . . 39

4 The nonconforming finite element method: a posteriori analysis **41**
4.1 Setting . . . 41
4.2 Error characterization . . . 42
4.3 A posteriori error estimate for the nonconforming finite element method, $p = 1$. 44
4.4 Numerical illustration: a posteriori error localization and control . . . 46
4.5 Complements . . . 48
4.5.1 An abstract a posteriori error estimate . . . 48
4.5.2 A posteriori error estimate for the nonconforming finite element method, $p \geq 1$ . . . 49

5 Adaptive finite elements **51**
5.1 Adaptive finite elements: numerical illustration of mesh adaptivity . . . 51
5.2 Adaptive finite elements: some theoretical background . . . 52

6 *hp* finite elements **55**
6.1 Conforming finite element method . . . 55
6.2 *hp* approximation on a simplex . . . 55
6.3 A priori error estimate: *hp* rate of convergence . . . 56
6.4 A priori error estimate: exponential convergence rate with respect to the number of the degrees of freedom . . . 56

7 Finite elements for a nonlinear elliptic problem **59**
7.1 A nonlinear elliptic problem . . . 59
7.2 Strongly monotone and Lipschitz-continuous operators . . . 59
7.3 Weak formulation . . . 60
7.4 A contractive nonlinear operator . . . 60
7.5 Existence and uniqueness of a weak solution by the Banach fixed-point theorem . 61
7.6 The conforming finite element method for problem (7.1.1) . . . 63
7.7 A priori error estimates: rate of convergence by h-analysis . . . 63

8 Finite elements for the heat equation **65**
8.1 The heat equation . . . 65
8.2 Bochner function spaces . . . 65
8.3 Weak formulation . . . 66
8.4 Inf-sup condition . . . 66
8.5 Existence and uniqueness of a weak solution by the Banach closed range and open mapping theorems . . . 67
8.6 The conforming finite element method for problem (8.1.1) . . . 70

Bibliography **73**

Index **77**

Chapter 0

Preliminaries

We recall here some basic notions of numerical approximation of partial differential equetions, following the lecture ANN201 "Eléments finis" `https://synapses.ensta-paris.fr/catalogue/2020-2021/ue/212/ANN201-elements-finis`, the classical references Ciarlet [12] and Allaire [2], or the lecture notes [45] on a posteriori error estimates available at `https://who.rocq.inria.fr/Martin.Vohralik/Pages/APost.html`.

0.1 Computational domain Ω

In these lecture notes, $\Omega \subset \mathbb{R}^d$, $d \geq 1$, will be an open *polytope* (interval for $d = 1$, polygon for $d = 2$, polyhedron for $d = 3$); Ω is thus *open*, *bounded*, and *connected* domain of polytopal type in d space dimensions. Moreover, we suppose that the boundary $\partial\Omega$ is *Lipschitz*; recall that this covers most cases of practical interest but excludes, for example, the two-brick setting in three space dimensions. An example in two space dimensions, with $d = 2$, is given in Figure 1, left. Note that Ω in Figure 1 is not convex, which we do not request.

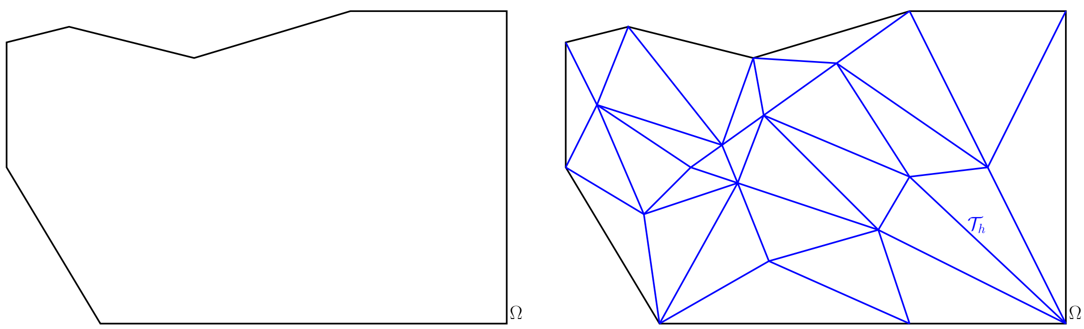

Figure 1: Example of a computational domain Ω (left) in two space dimensions ($d = 2$) and of a simplicial (triangular) mesh $\mathcal{T}_h$ (right)

0.2 Lebesgue space $L^2(\Omega)$

We let $L^2(\Omega)$ be the space of scalar-valued square-integrable functions defined on the polytope Ω, and $L^2(\omega)$ the space of scalar-valued square-integrable functions defined on a polytopal subdomain $\omega \subset \Omega$. We denote by $(v, w)_\omega$ the $L^2(\omega)$-scalar product of two functions $v, w \in L^2(\omega)$, given by

$$(v, w)_\omega := \int_\omega vw \,\mathrm{d}\boldsymbol{x}.$$

Then

$$\|v\|_\omega := (v, v)_\omega^{\frac{1}{2}}$$

denotes the associated norm on $L^2(\omega)$. We drop the index in both the scalar product and the norm notation when $\omega = \Omega$. A similar notation is used on $(d-1)$-dimensional manifolds, where we, however, rather employ $\langle \cdot, \cdot \rangle$ to denote the scalar product. We also employ this notation for vector-valued functions with each component in $L^2(\omega)$:

$$(\boldsymbol{v}, \boldsymbol{w})_\omega := \int_\omega \boldsymbol{v} \cdot \boldsymbol{w} \, \mathrm{d}\boldsymbol{x}, \quad \|\boldsymbol{v}\|_\omega := (\boldsymbol{v}, \boldsymbol{v})_\omega^{\frac{1}{2}} \tag{0.2.1}$$

for $\boldsymbol{v}, \boldsymbol{w} \in [L^2(\omega)]^d$, where $\boldsymbol{v} \cdot \boldsymbol{w}$ is the scalar product between the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$.

0.3 Sobolev space $H_0^1(\Omega)$

Let $\mathcal{D}(\Omega)$ be the space of functions from $C^\infty(\Omega)$ with a compact support in Ω. We first need to introduce the following concept:

Definition 0.3.1 (Weak partial derivative)**.** *Let a scalar-valued function $v : \Omega \to \mathbb{R}$ be given. We say that v admits a weak i-th partial derivative, $1 \le i \le d$, if*

1. $v \in L^2(\Omega)$;
2. *there exists a function $w_i : \Omega \to \mathbb{R}$ such that*
   (a) $w_i \in L^2(\Omega)$;
   (b) $(v, \partial_{\boldsymbol{x}_i} \varphi) = -(w_i, \varphi) \qquad \forall \varphi \in \mathcal{D}(\Omega)$.

We define the weak i-th partial derivative *of v, denoted by $\partial_{\boldsymbol{x}_i} v$, as*

$$\partial_{\boldsymbol{x}_i} v := w_i.$$

Weak partial derivative is a *generalization* of the concept of the partial derivative: whenever the function in question v admits the usual partial derivative, then the weak partial derivative equals the partial derivative. More functions, however, allow for a weak partial derivative. The prominent example is the absolute value function of Figure 2. Its weak (partial) derivative equals its derivative respectively on the left and right half-spaces. In general, in one space dimension, any function that is of class $C^0(\overline{\Omega})$ (continuous) and piecewise of class C^1 (admitting derivative on subintervals), admits a weak (partial) derivative, given by the derivative on subintervals.

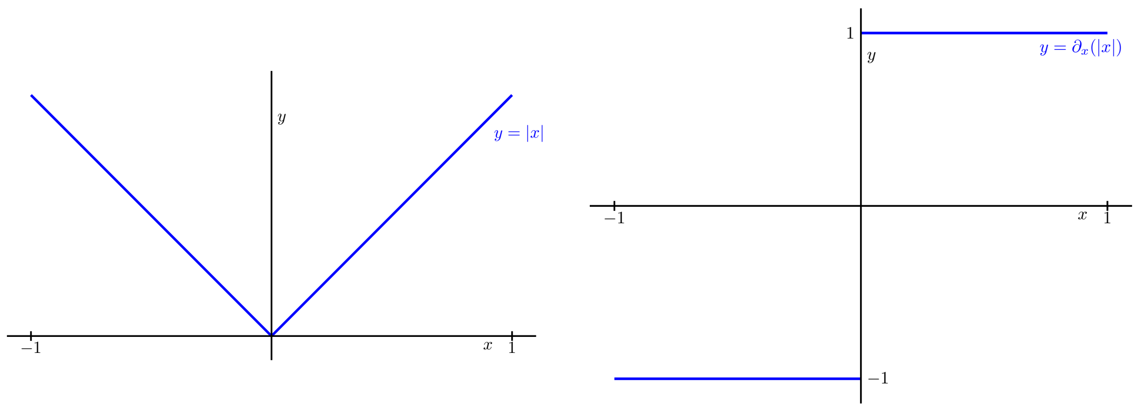

Figure 2: $\Omega = (-1, 1)$, $d = 1$. Example of a function $y(x) = |x|$ (left) which admits a weak partial derivative (right)

Definition 0.3.2 (Weak gradient)**.** *Let a scalar-valued function $v : \Omega \to \mathbb{R}$ be given. We say that v admits a weak gradient if v admits the weak i-th partial derivative for all $1 \le i \le d$. We define the* weak gradient *of v, denoted by $\nabla v \in [L^2(\Omega)]^d$, as*

$$\nabla v := (\partial_{\boldsymbol{x}_1} v, \ldots, \partial_{\boldsymbol{x}_d} v)^{\mathrm{t}}. \tag{0.3.1}$$

Definition 0.3.3 (Space $H^1(\Omega)$)**.** *The space $H^1(\Omega)$ is the space of all the functions which admit the weak gradient.*

Let us recall from [12, 2] that $H^1(\Omega)$ is a Hilbert space for the scalar product and the associated norm

$$(v, w)_{H^1(\Omega)} := (v, w) + (\nabla v, \nabla w) \qquad v, w \in H^1(\Omega), \tag{0.3.2a}$$

$$\|v\|^2_{H^1(\Omega)} := (v, v)_{H^1(\Omega)} = \|v\|^2 + \|\nabla v\|^2 \qquad v \in H^1(\Omega). \tag{0.3.2b}$$

A crucial property of functions from $H^1(\Omega)$ is that they are *continuous in the sense of traces*: for each $(d-1)$-dimensional interface inside Ω, the values from both sides coincide, see Figure 3 for an illustration, where the function in question is trace-continuous (actually, continuous) along all the depicted lines.

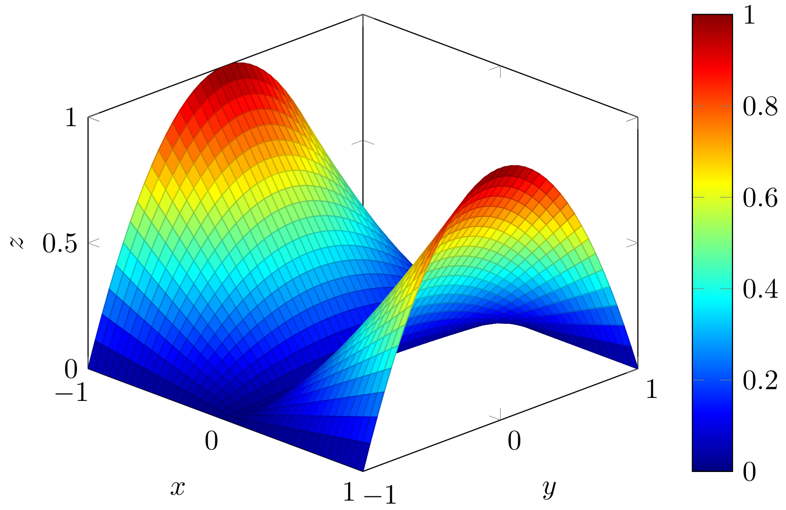

Figure 3: $\Omega = (-1, 1) \times (-1, 1)$, $d = 2$. Example of a function belonging to the Sobolev space $H^1(\Omega)$: note that $v|_{\Omega_1} = v|_{\Omega_2}$ on the interface between any two subdomains Ω_1 and Ω_2 of Ω

Definition 0.3.4 (Space $H^1_0(\Omega)$)**.** *The space $H^1_0(\Omega)$ is the space of functions $v \in H^1(\Omega)$ such that $v|_{\partial\Omega} = 0$, where $v|_{\partial\Omega}$ is the* trace *of the function v on the boundary $\partial\Omega$.*

0.4 Sobolev space $\boldsymbol{H}(\mathrm{div}, \Omega)$

The space $H^1_0(\Omega)$ from Section 0.3 is designed for scalar-valued functions. For vector-valued functions, we will sometimes need the following concepts:

Definition 0.4.1 (Weak divergence)**.** *Let a vector-valued function $\boldsymbol{v} : \Omega \to \mathbb{R}^d$ be given. We say that $\boldsymbol{v}$ admits a weak divergence if*

1. $\boldsymbol{v} \in [L^2(\Omega)]^d$;
2. *there exists a scalar-valued function $w : \Omega \to \mathbb{R}$ such that*
   (a) $w \in L^2(\Omega)$;
   (b) $(\boldsymbol{v}, \nabla\varphi) = -(w, \varphi) \qquad \forall \varphi \in \mathcal{D}(\Omega)$.

We define the weak divergence *of $\boldsymbol{v}$, denoted by $\nabla{\cdot}\boldsymbol{v}$, as*

$$\nabla{\cdot}\boldsymbol{v} := w.$$

Definition 0.4.2 (Space $\boldsymbol{H}(\mathrm{div}, \Omega)$)**.** *The space $\boldsymbol{H}(\mathrm{div}, \Omega)$ is the space of all the functions which admit the weak divergence.*

Let us recall from [2] that $\boldsymbol{H}(\mathrm{div}, \Omega)$ is a Hilbert space for the scalar product

$$(\boldsymbol{v}, \boldsymbol{w})_{\boldsymbol{H}(\mathrm{div},\Omega)} := (\boldsymbol{v}, \boldsymbol{w}) + (\nabla{\cdot}\boldsymbol{v}, \nabla{\cdot}\boldsymbol{w}) \qquad \boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{H}(\mathrm{div}, \Omega).$$

A crucial property of functions $\boldsymbol{v}$ from the Sobolev space $\boldsymbol{H}(\mathrm{div}, \Omega)$ is that their normal components are *continuous in the sense of normal traces*: for each $(d-1)$-dimensional interface dividing Ω into Ω_1 and Ω_2, the values of $\boldsymbol{v}|_{\Omega_i}{\cdot}\boldsymbol{n}$, where $\boldsymbol{n}$ is the unit normal vector of the interface between Ω_1 and Ω_2, are equal in appropriate normal trace sense. Figure 4 gives an illustration where $\boldsymbol{v}{\cdot}\boldsymbol{n}$ is well-defined everywhere in Ω, $\boldsymbol{v}{\cdot}\boldsymbol{n}$ on each interface Γ is an $L^2(\Gamma)$ function (actually piecewise constant), and $\boldsymbol{v}{\cdot}\boldsymbol{n}$ is continuous in the usual, strong, sense (and not just in the "appropriate normal trace" sense).

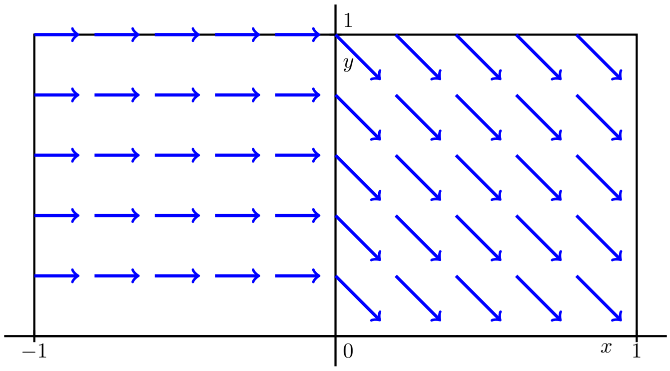

Figure 4: $\Omega = (-1, 1) \times (0, 1)$, $d = 2$. Example of a function belonging to the Sobolev space $\boldsymbol{H}(\mathrm{div}, \Omega)$: note that, on the interface $x = 0$ between the two subdomains $\Omega_1 = (-1, 0) \times (0, 1)$ and $\Omega_2 = (0, 1) \times (0, 1)$, $\boldsymbol{v}|_{\Omega_1}{\cdot}\boldsymbol{n} = \boldsymbol{v}|_{\Omega_2}{\cdot}\boldsymbol{n}$, with $\boldsymbol{n} = (1, 0)^{\mathrm{t}}$; the function $\boldsymbol{v}$ is not (trace-)continuous for each component (the y component of $\boldsymbol{v}$ is discontinuous, as it passes from value 0 in Ω_1 to a nonzero constant value in Ω_2), but $\boldsymbol{v}{\cdot}\boldsymbol{n}$ is (normal-trace) continuous (here $\boldsymbol{v}{\cdot}\boldsymbol{n}$ is simply the x component of $\boldsymbol{v}$, which has the same constant value in both Ω_1 and Ω_2)

0.5 Computational mesh $\mathcal{T}_h$

The computational mesh $\mathcal{T}_h$ is a partition of the closure $\overline{\Omega}$ of the polytope Ω such that

$$\cup_{K\in\mathcal{T}_h} K = \overline{\Omega},$$

where any element $K \in \mathcal{T}_h$ is a closed *simplex* (interval for $d = 1$, triangle for $d = 2$, tetrahedron for $d = 3$), and the intersection of two different simplices is either empty, a vertex, or an l-dimensional face, $1 \le l \le d-1$. An example for $d = 2$ is given in Figure 1, right. It is customary to define the *mesh size*

$$h := \max_{K\in\mathcal{T}_h} h_K, \tag{0.5.1}$$

where h_K is the diameter of the simplex K. The *shape-regularity* parameter of the mesh $\mathcal{T}_h$ is the positive real number

$$\kappa_{\mathcal{T}_h} := \max_{K\in\mathcal{T}_h} \frac{h_K}{\rho_K}, \tag{0.5.2}$$

where ρ_K is the diameter of the largest ball contained in K. We will only consider sequences of meshes where this number is uniformly bounded: this forbids "flat" simplices. Note, though, that highly graded meshes, containing a range of simplices from very small to very big, can be constructed for $\kappa_{\mathcal{T}_h}$ uniformly bounded, see Figure 5.1 below for an example.

We denote the set of vertices of the mesh $\mathcal{T}_h$ by $\mathcal{V}_h$; it is composed of vertices $\mathcal{V}_h^{\mathrm{int}}$ lying in the interior of Ω and of vertices $\mathcal{V}_h^{\mathrm{ext}}$ lying on the boundary $\partial\Omega$. Similarly, the set of $(d-1)$-dimensional faces of the mesh $\mathcal{T}_h$ is denoted by $\mathcal{F}_h$ and is again composed of interior faces $\mathcal{F}_h^{\mathrm{int}}$

and faces $\mathcal{F}_h^{\text{ext}}$ lying on the boundary $\partial\Omega$. For each face $F \in \mathcal{F}_h$, we fix a unique unit normal vector $\boldsymbol{n}_F$, arbitrary for $F \in \mathcal{F}_h^{\text{int}}$ and pointing outward of the domain Ω for $F \in \mathcal{F}_h^{\text{ext}}$. We will use $\mathcal{F}_K$ to denote the set of all the faces of the simplex K.

0.6 Broken Sobolev space $H^1(\mathcal{T}_h)$

The spaces $H_0^1(\Omega)$ from Section 0.3 and $\boldsymbol{H}(\text{div}, \Omega)$ from Section 0.4 are only related to the computational domain Ω. In these lecture notes, we will often use the space $H^1(\mathcal{T}_h)$ rather related to the computational mesh $\mathcal{T}_h$:

Definition 0.6.1 (Space $H^1(\mathcal{T}_h)$)**.** *The so-called* broken Sobolev space *is given by*

$$H^1(\mathcal{T}_h) := \{v \in L^2(\Omega);\, v|_K \in H^1(K) \quad \forall K \in \mathcal{T}_h\} = \prod_{K \in \mathcal{T}_h} H^1(K). \tag{0.6.1}$$

The space $H^1(\mathcal{T}_h)$ is thus a *collection of independent Sobolev spaces* $H^1(K)$ over the individual elements K of the mesh $\mathcal{T}_h$. In particular, functions from the space $H^1(\mathcal{T}_h)$ are possibly *completely discontinuous* from one mesh element to the other (they admit no trace continuity over the mesh interfaces) since $H^1(\mathcal{T}_h) \not\subset H^1(\Omega)$, see the illustration in Figure 5. Note, however, that $H^1(\Omega) \subset H^1(\mathcal{T}_h)$.

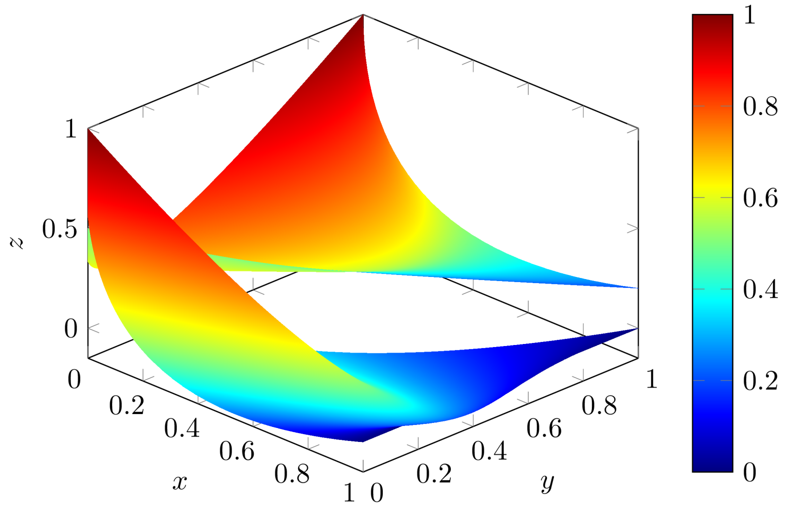

Figure 5: $\Omega = (0,1) \times (0,1)$, $d = 2$. Example of a function belonging to the broken Sobolev space $H^1(\mathcal{T}_h)$ for a mesh composed of two elements: note that there is no continuity on the interfaces between the two elements

An important concept related to the broken space $H^1(\mathcal{T}_h)$ is the concept of a jump:

Definition 0.6.2 (Jump of $v \in H^1(\mathcal{T}_h)$)**.** *Let $v \in H^1(\mathcal{T}_h)$. The* jump *over an* interior face *$F \in \mathcal{F}_h^{\text{int}}$ is given by*

$$[\![v]\!]_F := (v|_{K_1})|_F - (v|_{K_2})|_F, \tag{0.6.2a}$$

where K_1, K_2 are the two simplices from the mesh $\mathcal{T}_h$ that share the face F, enumerated such that the normal $\boldsymbol{n}_F$ points from K_1 to K_2. Here $(v|_{K_i})|_F$ is the trace *of the function $v|_{K_i} \in H^1(K_i)$ on the face F, which in particular belongs to $L^2(F)$. The* jump *over a* boundary face *$F \in \mathcal{F}_h^{\text{ext}}$ is then given by*

$$[\![v]\!]_F := (v|_K)|_F, \tag{0.6.2b}$$

where $K \in \mathcal{T}_h$ has F as face.

The following important relation holds between the spaces $H^1(\Omega)$ of Definition 0.3.3 and $H^1(\mathcal{T}_h)$ of Definition 0.6.1, see, e.g. [45, Theorems 4.4.1 and 4.4.3]:

Theorem 0.6.3 (Relations between the spaces $H^1(\Omega)$, $H^1_0(\Omega)$, and $H^1(\mathcal{T}_h)$)**.** *There holds*

$$v \in H^1(\Omega) \Longleftrightarrow v \in H^1(\mathcal{T}_h) \text{ and } [\![v]\!]_F = 0 \qquad \forall F \in \mathcal{F}_h^{\mathrm{int}}, \tag{0.6.3a}$$

$$v \in H^1_0(\Omega) \Longleftrightarrow v \in H^1(\mathcal{T}_h) \text{ and } [\![v]\!]_F = 0 \qquad \forall F \in \mathcal{F}_h. \tag{0.6.3b}$$

For a function $v \in H^1(\mathcal{T}_h)$, there in general does not exist a weak gradient in the sense of Definition 0.3.2. Since, of course, there exists a weak gradient for any function $v \in H^1(\mathcal{T}_h)$ when *restricted* to each individual mesh element $K \in \mathcal{T}_h$, we are lead to a *generalization* of the notion of the weak gradient:

Definition 0.6.4 (Broken weak gradient)**.** *Let a function $v \in H^1(\mathcal{T}_h)$ be given. The* broken weak gradient $\nabla_h v \in [L^2(\Omega)]^d$ *is on each element of the computational mesh $K \in \mathcal{T}_h$ given by*

$$(\nabla_h v)|_K := \nabla(v|_K). \tag{0.6.4}$$

With this notion, we then have:

Theorem 0.6.5 (Space $H^1(\mathcal{T}_h)$)**.** *The space $H^1(\mathcal{T}_h)$ is a Hilbert space for the scalar product*

$$\begin{aligned} (v,w)_{H^1(\mathcal{T}_h)} &:= (v,w) + (\nabla_h v, \nabla_h w) \\ &= \sum_{K \in \mathcal{T}_h} (v,w)_{H^1(K)} \\ &= \sum_{K \in \mathcal{T}_h} \left\{ (v,w)_K + (\nabla(v|_K), \nabla(w|_K))_K \right\} \qquad v,w \in H^1(\mathcal{T}_h) \end{aligned} \tag{0.6.5a}$$

and the associated norm

$$\|v\|^2_{H^1(\mathcal{T}_h)} := (v,v)_{H^1(\mathcal{T}_h)} = \|v\|^2 + \|\nabla_h v\|^2 = \sum_{K \in \mathcal{T}_h} \left\{ \|v\|_K^2 + \|\nabla(v|_K)\|_K^2 \right\} \qquad v \in H^1(\mathcal{T}_h). \tag{0.6.5b}$$

Proof. From (0.6.1),

$$H^1(\mathcal{T}_h) = \prod_{K \in \mathcal{T}_h} H^1(K),$$

so $H^1(\mathcal{T}_h)$ is a Hilbert space as a product of the Hilbert spaces $H^1(K)$. □

0.7 Poincaré–Friedrichs and trace inequalities

We recall here three basic inequalities that will be often used in the following chapters. We make a concise presentation; more details and numerous references can be found, e.g., in [45, Section 4.6]. Let $\omega \subset \Omega$ be an open polytope with Lipschitz boundary and let h_ω denote its diameter.

The Poincaré inequality states:

Theorem 0.7.1 (Poincaré inequality)**.** *There holds*

$$\|v - v_\omega\|_\omega \le C_{\mathrm{P},\omega} h_\omega \|\nabla v\|_\omega \qquad \forall v \in H^1(\omega), \tag{0.7.1}$$

where v_ω is the mean value of the function v over ω given by $v_\omega := (v,1)_\omega/|\omega|$.

The above generic constant $C_{\mathrm{P},\omega}$ only depends on the *shape* of the polytope ω but not on its size. It can be precisely estimated in many cases; in particular, whenever ω is convex, $C_{\mathrm{P},\omega}$ can be taken as $1/\pi$, cf. Payne and Weinberger [33] and Bebendorf [5].

Let $\partial\omega_{\mathrm{D}}$ be a subset of the boundary $\partial\omega$ with nonzero $(d-1)$-dimensional measure, $|\partial\omega_{\mathrm{D}}| \neq 0$. Then the Poincaré–Friedrichs inequality states:

Theorem 0.7.2 (Poincaré–Friedrichs inequality)**.** *There holds*

$$\|v\|_\omega \leq C_{\mathrm{PF},\omega,\partial\omega_{\mathrm{D}}} h_\omega \|\nabla v\|_\omega \qquad \forall v \in H^1(\omega) \text{ such that } v|_{\partial\omega_{\mathrm{D}}} = 0. \tag{0.7.2}$$

The generic constant $C_{\mathrm{PF},\omega,\partial\omega_{\mathrm{D}}}$ only depends on the shape of the polytope ω and on the boundary subset $\partial\omega_{\mathrm{D}}$. As long as ω and $\partial\omega_{\mathrm{D}}$ are such that there exists a vector $\boldsymbol{b} \in \mathbb{R}^d$ such that for almost all points $\boldsymbol{x} \in \omega$, the first intersection of the straight semi-line defined by the origin $\boldsymbol{x}$ and the vector $\boldsymbol{b}$ lies in $\partial\omega_{\mathrm{D}}$, the constant $C_{\mathrm{PF},\omega,\partial\omega_{\mathrm{D}}}$ can be taken equal to 1, cf. [43, Remark 5.8].

Let finally K be a simplex and let F be one of its faces. The trace inequalities state:

Theorem 0.7.3 (Trace inequalities)**.** *There holds*

$$\|v\|_F^2 \leq \widetilde{C}_{\mathrm{t},\kappa_K,d}^2 (h_K^{-1}\|v\|_K^2 + \|v\|_K\|\nabla v\|_K) \qquad \forall v \in H^1(K), \tag{0.7.3a}$$

$$\|v - v_F\|_F \leq \bar{C}_{\mathrm{t},\kappa_K,d} h_K^{\frac{1}{2}} \|\nabla v\|_K \qquad \forall v \in H^1(K), \tag{0.7.3b}$$

$$\|v - v_K\|_F \leq C_{\mathrm{t},\kappa_K,d} h_K^{\frac{1}{2}} \|\nabla v\|_K \qquad \forall v \in H^1(K), \tag{0.7.3c}$$

where the mean values over the element K and face F are respectively given by $v_K := (v,1)_K/|K|$ and $v_F := \langle v,1\rangle_F/|F|$.

The above three generic constants only depend on the shape-regularity parameter of the simplex K given by $\kappa_K := h_K/\rho_K$ (recall that h_K is the diameter of K and ρ_K is the diameter of the largest ball contained in K) and possibly on the space dimension d.

0.8 Piecewise polynomial space $\mathcal{P}_p(\mathcal{T}_h)$

Numerical methods typically rely on finite-dimensional spaces. Let $p \geq 1$ be a fixed polynomial degree and let $\mathcal{P}_p(K)$ be the space of *polynomials* of *total degree* less than or equal to p defined on one mesh element $K \in \mathcal{T}_h$. In these lecture notes, we will many times need:

Definition 0.8.1 (Piecewise polynomial space $\mathcal{P}_p(\mathcal{T}_h)$)**.** *The piecewise polynomial space $\mathcal{P}_p(\mathcal{T}_h)$ is given by*

$$\mathcal{P}_p(\mathcal{T}_h) := \{v_h \in L^2(\Omega);\, v_h|_K \in \mathcal{P}_p(K) \quad \forall K \in \mathcal{T}_h\} = \prod_{K\in\mathcal{T}_h} \mathcal{P}_p(K). \tag{0.8.1}$$

An illustration for $p = 2$ is given in Figure 6; clearly,

$$\mathcal{P}_p(\mathcal{T}_h) \subset H^1(\mathcal{T}_h) \text{ but } \mathcal{P}_p(\mathcal{T}_h) \not\subset H^1(\Omega).$$

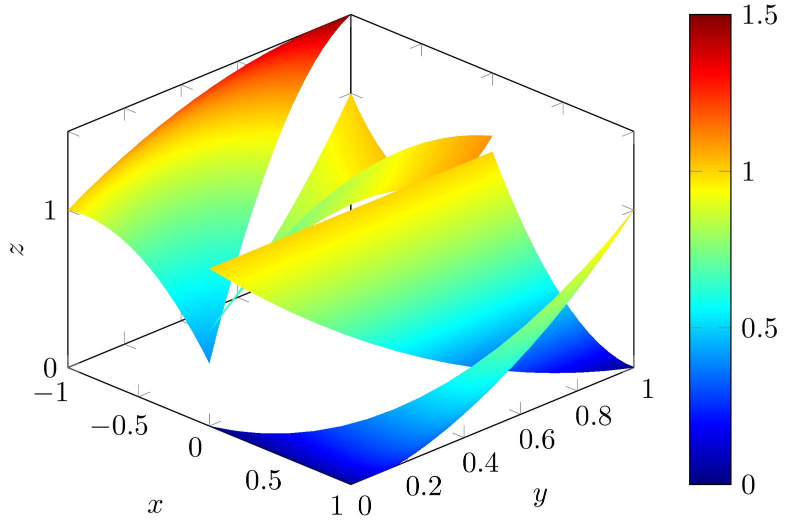

Figure 6: $\Omega = (-1,1) \times (0,1)$, $d = 2$. Example of a piecewise polynomial belonging to the piecewise polynomial space $\mathcal{P}_2(\mathcal{T}_h)$ with four mesh elements

Chapter 1

The nonconforming finite element method: a priori analysis

In this chapter, we are going to consider a finite element method that looks for an approximate solution of a model partial differential equation in a finite-dimensional space that is *not* a subspace of the space where the exact solution lies. This gives rise to the nomenclature *nonconforming* finite element method. We will in particular consider the so-called Crouzeix–Raviart nonconforming finite element method, named after Crouzeix and Raviart [13]. Our goal will be to justify the method and to derive *a priori* error estimates . We will partly follow the original analysis in Ciarlet [12], but we will also expose some recent achievements following, in particular, Gudi [26], Veeser [41], Ern *et al.* [19], and Ern and Guermond [20].

1.1 The Poisson equation

The Poisson equation with a homogeneous Dirichlet boundary condition reads: for a given source term $f \in L^2(\Omega)$, find a scalar-valued function $u : \Omega \to \mathbb{R}$ such that

$$-\Delta u = f \qquad \text{in } \Omega, \tag{1.1.1a}$$

$$u = 0 \qquad \text{on } \partial\Omega, \tag{1.1.1b}$$

where Δ stands for the Laplacian,

$$\Delta v := \sum_{i=1}^{d} \partial^2_{\boldsymbol{x}_i} v = \nabla{\cdot}(\nabla v), \tag{1.1.2}$$

a differential operator composed of the divergence and gradient.

1.2 Weak formulation

Problem (1.1.1), in general, does not have a classical solution $u \in C^2(\overline{\Omega})$. We are thus led to the weak formulation, relying on the Sobolev space $H^1_0(\Omega)$ recalled in Section 0.3:

Definition 1.2.1 (Weak formulation of problem (1.1.1))**.** *Find* $u \in H^1_0(\Omega)$ *such that*

$$(\nabla u, \nabla v) = (f, v) \qquad \forall v \in H^1_0(\Omega). \tag{1.2.1}$$

Recall that the existence and uniqueness of a solution of (1.2.1) can be shown as follows. The *Poincaré–Friedrichs inequality* (0.7.2) states

$$\|v\| \le C_{\mathrm{PF}} h_\Omega \|\nabla v\| \qquad \forall v \in H^1_0(\Omega), \tag{1.2.2}$$

where h_Ω is the diameter of the computational domain Ω. It follows from the discussion below Theorem 0.7.2, since here the zero trace boundary condition is imposed all along $\partial\Omega$, that $C_{\rm PF}$ is a generic constant at most equal to 1. As a consequence of (0.3.2b) and (1.2.2),

$$\|\nabla v\| \leq \|v\|_{H^1(\Omega)} \leq \left(1 + C_{\rm PF}^2 h_\Omega^2\right)^{\frac{1}{2}} \|\nabla v\| \qquad \forall v \in H_0^1(\Omega), \tag{1.2.3}$$

i.e., on the Sobolev space $H_0^1(\Omega)$, the $\|\nabla\cdot\|$ norm is equivalent to the $\|\cdot\|_{H^1(\Omega)}$ norm. Thus, $H_0^1(\Omega)$ equipped with the scalar product $(\nabla v, \nabla w)$ is also a Hilbert space, and the existence and uniqueness of a solution of (1.2.1) follows by the Riesz representation theorem. Note that one does not need to invoke the Lax–Milgram theorem [12, Theorem 1.1.3], since problem (1.2.1) is *symmetric*.

1.3 The nonconforming finite element method

Let $p \geq 1$ be a fixed polynomial degree. Recall the piecewise polynomial space $\mathcal{P}_p(\mathcal{T}_h)$ from Section 0.8, as well as the notion of the jump from Definition 0.6.2 (we will henceforth abbreviate $[\![v_h]\!]_F$ as $[\![v_h]\!]$; there will be no ambiguity.) We will employ the following weakly-continuous (Crouzeix–Raviart) space, with $\mathcal{P}_{p'}(F)$ denoting polynomials of total degree less than or equal to $p' \geq 0$ on the face F:

Definition 1.3.1 (Nonconforming finite element space)**.** *The weakly-continuous piecewise polynomial space $V_{hp}^{\rm nc}$ is given by*

$$V_{hp}^{\rm nc} := \{v_h \in \mathcal{P}_p(\mathcal{T}_h);\ \langle [\![v_h]\!], q_h\rangle_F = 0 \quad \forall q_h \in \mathcal{P}_{p-1}(F), \forall F \in \mathcal{F}_h\}. \tag{1.3.1}$$

By definition, $V_{hp}^{\rm nc}$ is a proper subspace of $\mathcal{P}_p(\mathcal{T}_h)$,

$$V_{hp}^{\rm nc} \subsetneq \mathcal{P}_p(\mathcal{T}_h),$$

since some conditions on the jumps over the faces are requested. This requested *weak continuity* (1.3.1) is that the jumps $[\![v_h]\!]$ are $L^2(F)$-orthogonal to all polynomials of degree $p-1$ on each mesh face F; since $v_h \in \mathcal{P}_p(\mathcal{T}_h)$, $[\![v_h]\!]|_F \in \mathcal{P}_p(F)$, the jumps have vanishing moments up to degree $p-1$, but are not entirely zero, $[\![v_h]\!] \neq 0$. Consequently, recalling Theorem 0.6.3, the functions in $V_{hp}^{\rm nc}$ are not (trace-)continuous, so that

$$V_{hp}^{\rm nc} \not\subset H_0^1(\Omega),$$

which is the *nonconformity*. An illustration for the lowest polynomial degree $p = 1$ is given in Figure 1.1. Here the zero moments, i.e., the *mean values* of the *jumps*, have to vanish on all mesh faces. Since for $p = 1$, the jumps are affine on each face F, this is equivalent to the condition of pointwise continuity (equality to zero on boundary faces) at face barycentres (see Section 1.6.1 below for more details).

Recalling the broken weak gradient from Definition 0.6.4, the finite element method we will study is:

Definition 1.3.2 (Nonconforming finite element method for problem (1.1.1))**.** *Find $u_h \in V_{hp}^{\rm nc}$ such that*

$$(\nabla_h u_h, \nabla_h v_h) = (f, v_h) \qquad \forall v_h \in V_{hp}^{\rm nc}. \tag{1.3.2}$$

Remark 1.3.3 (A first comparison with the conforming finite element method)**.** *The standard conforming finite element method employs the $H_0^1(\Omega)$-conforming piecewise polynomial space*

$$V_{hp} := \mathcal{P}_p(\mathcal{T}_h) \cap H_0^1(\Omega) \tag{1.3.3}$$

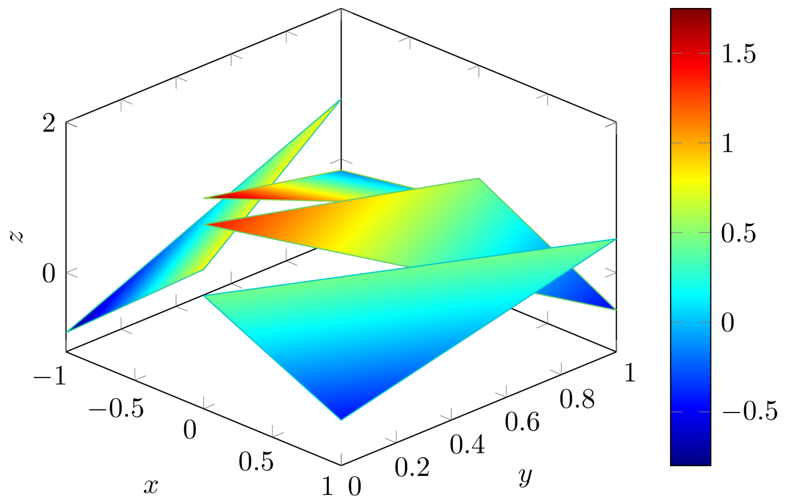

Figure 1.1: $\Omega = (-1,1) \times (0,1)$, $d = 2$. Example of a weakly-continuous piecewise polynomial belonging to the Crouzeix–Raviart space V_{hp}^{nc} with $p = 1$ (four mesh elements)

to look for $u_h \in V_{hp}$ such that

$$(\nabla u_h, \nabla v_h) = (f, v_h) \qquad \forall v_h \in V_{hp}. \tag{1.3.4}$$

Following Theorem 0.6.3, we could have also written, in place of (1.3.3),

$$\begin{aligned} V_{hp} := \{v_h \in \mathcal{P}_p(\mathcal{T}_h);\ [\![v_h]\!] = 0 \quad \forall F \in \mathcal{F}_h\} = \{v_h \in \mathcal{P}_p(\mathcal{T}_h);\ \langle [\![v_h]\!], q_h \rangle_F \\ = 0 \quad \forall q_h \in \mathcal{P}_p(F),\ \forall F \in \mathcal{F}_h\}. \end{aligned}$$

Relying on the broken weak gradient (0.6.4), (1.3.2) and (1.3.4) take the same form. Since it follows immediately from (1.3.1) that

$$V_{hp} \subset V_{hp}^{\mathrm{nc}},$$

it might come to one's mind that (1.3.2) simply needs to be "better". We will see below that such considerations are a bit precipitous and that the analysis of (1.3.2) is rather involved, because of the nonconformity $V_{hp}^{\mathrm{nc}} \not\subset H_0^1(\Omega)$. *Later, we will actually prove that the two methods are of the same quality in a quite sharp sense. Both of the approaches (1.3.2) and (1.3.4) have some (distinct) important advantages in (distinct) applications, which we will discuss below.*

Since $V_{hp}^{\mathrm{nc}} \not\subset H_0^1(\Omega)$, it is not obvious to show the existence and uniqueness of u_h from (1.3.2). It turns out that it would be good to have at our disposal an equivalent of the Poincaré–Friedrichs inequality (1.2.2) on the space V_{hp}^{nc}. This is indeed a central task that we will establish in a (much) larger setting of the following section.

1.4 Weakly-continuous subspace of the broken Sobolev space $H^1(\mathcal{T}_h)$ and a broken Poincaré–Friedrichs inequality

The space V_{hp}^{nc} from (1.3.1) is a finite-dimensional subspace of the broken Sobolev space $H^1(\mathcal{T}_h)$ of Definition 0.6.1. Its nonconforming nature can, however, be rather studied at the infinite-dimensional level. This better highlights the central issue. Let us for this purpose introduce:

Definition 1.4.1 (Weakly-continuous subspace $H^1_{[\![]\!]}(\mathcal{T}_h)$ of the broken Sobolev space $H^1(\mathcal{T}_h)$)**.** *Let*

$$H^1_{[\![]\!]}(\mathcal{T}_h) := \{v \in H^1(\mathcal{T}_h);\ \langle [\![v]\!], q_h \rangle_F = 0 \quad \forall q_h \in \mathcal{P}_0(F),\ \forall F \in \mathcal{F}_h\}. \tag{1.4.1}$$

Remark 1.4.2 (Higher vanishing moments of jumps). *We could have also defined spaces requesting orthogonality of jumps with respect to p'-degree polynomials on each face F, $p' \geq 0$, but the space $H^1_{[\![]\!]}(\mathcal{T}_h)$ of Definition 1.4.1 will be completely sufficient for our purposes.*

Remark 1.4.3 (Space $H^1_{[\![]\!]}(\mathcal{T}_h)$). *Clearly, $H^1_0(\Omega) \subset H^1_{[\![]\!]}(\mathcal{T}_h)$, as (0.6.3b) requests completely vanishing jumps as opposed to merely jumps with vanishing zero-order moments (i.e., mean values) in (1.4.1).*

The following theorem shows that Definition 1.4.1 is justified; the proof is given in Section 1.9.1 below.

Theorem 1.4.4 (Space $H^1_{[\![]\!]}(\mathcal{T}_h)$). *The space $H^1_{[\![]\!]}(\mathcal{T}_h)$ is a non-empty closed subspace of $H^1(\mathcal{T}_h)$, so that $H^1_{[\![]\!]}(\mathcal{T}_h)$ is a Hilbert space for the scalar product (0.6.5a). There holds $V^{\rm nc}_{hp} \subset H^1_{[\![]\!]}(\mathcal{T}_h)$.*

Recalling that h_Ω is the diameter of the computational domain Ω, a crucial property of the space $H^1_{[\![]\!]}(\mathcal{T}_h)$ is the following generalization of (1.2.2):

Theorem 1.4.5 (Broken Poincaré–Friedrichs inequality). *There exists a positive constant $C_{\rm bPF}$ only depending on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$ from (0.5.2) and the space dimension d such that*

$$\|v\| \leq C_{\rm bPF} h_\Omega \|\nabla_h v\| \qquad \forall v \in H^1_{[\![]\!]}(\mathcal{T}_h). \tag{1.4.2}$$

Proof. The proof of the result as stated can be found in [43], see also Temam [40] (for piecewise affine functions), Knobloch [29], or Brenner [9]. □

A crucial consequence of Theorem 1.4.5 is that $H^1_{[\![]\!]}(\mathcal{T}_h)$ is a Hilbert space for the scalar product given by the broken weak gradient of Definition 0.6.4:

Theorem 1.4.6 (Space $H^1_{[\![]\!]}(\mathcal{T}_h)$). *The space $H^1_{[\![]\!]}(\mathcal{T}_h)$ from Definition 1.4.1 is a Hilbert space for the scalar product*

$$(\nabla_h v, \nabla_h w) \qquad v, w \in H^1_{[\![]\!]}(\mathcal{T}_h).$$

Proof. The reasoning is as in (1.2.3). Definition (0.6.5b) and the broken Poincaré–Friedrichs inequality (1.4.2) immediately imply

$$\|\nabla_h v\| \leq \|v\|_{H^1(\mathcal{T}_h)} \leq \left(1 + C^2_{\rm bPF} h^2_\Omega\right)^{\frac{1}{2}} \|\nabla_h v\| \qquad \forall v \in H^1_{[\![]\!]}(\mathcal{T}_h),$$

so that the norms $\|\nabla_h \cdot\|$ and $\|\cdot\|_{H^1(\mathcal{T}_h)}$ are equivalent on the space $H^1_{[\![]\!]}(\mathcal{T}_h)$. Since, by Theorem 1.4.4, $H^1_{[\![]\!]}(\mathcal{T}_h)$ is a Hilbert space for the scalar product $(v, w)_{H^1(\mathcal{T}_h)}$, $H^1_{[\![]\!]}(\mathcal{T}_h)$ is also a Hilbert space for the scalar product $(\nabla_h v, \nabla_h w)$. □

1.5 Existence and uniqueness of the nonconforming finite element approximation

Armed with the results of the previous section, we now straightforwardly deduce:

Theorem 1.5.1 (Space $V^{\rm nc}_{hp}$). *The space $V^{\rm nc}_{hp}$ is a (finite-dimensional) Hilbert space for the scalar product*

$$(\nabla_h v_h, \nabla_h w_h) \qquad v_h, w_h \in V^{\rm nc}_{hp}. \tag{1.5.1}$$

Proof. In Theorem 1.4.4, we have seen that $V^{\rm nc}_{hp} \subset H^1_{[\![]\!]}(\mathcal{T}_h)$. $V^{\rm nc}_{hp}$ being piecewise polynomial on the mesh $\mathcal{T}_h$ and thus finite-dimensional, it is a closed subspace of $H^1_{[\![]\!]}(\mathcal{T}_h)$. Thus, the claim follows by virtue of Theorem 1.4.6. □

Consequently, there immediately holds:

Theorem 1.5.2 (Existence and uniqueness of u_h from (1.3.2))**.** *There is one and only one solution $u_h \in V_{hp}^{\mathrm{nc}}$ of the nonconforming finite element method of Definition 1.3.2.*

Proof. By virtue of the broken Poincaré–Friedrichs inequality (1.4.2),

$$v_h \to (f, v_h) \qquad v_h \in V_{hp}^{\mathrm{nc}},$$

is a bounded linear form on V_{hp}^{nc} equipped with the scalar product (1.5.1). Indeed,

$$|(f, v_h)| \leq \|f\| \|v_h\| \leq C_{\mathrm{bPF}} h_\Omega \|f\| \|\nabla_h v_h\| \qquad \forall v_h \in V_{hp}^{\mathrm{nc}}.$$

Thus in view of Theorem 1.5.1, the claim follows by the Riesz representation theorem. □

1.6 Nonconforming finite elements for $p = 1$

We now restrict our attention to the lowest-order setting $p = 1$ (see Figure 1.1) and discuss three important issues: we introduce a basis of the space V_{h1}^{nc} (handy for a computer implementation), an interpolation operator for the space V_{h1}^{nc} (useful for analysis), and a reconstruction of a flux in the $\boldsymbol{H}(\mathrm{div}, \Omega)$ space (which will be central later for a posteriori error estimates).

1.6.1 A basis of V_{h1}^{nc}

For a face $F \in \mathcal{F}_h$, denote by $\boldsymbol{x}_F$ its barycenter.

Definition 1.6.1 (Basis function ψ^F)**.** *For $F \in \mathcal{F}_h^{\mathrm{int}}$, let ψ^F, a priori defined in the space $\mathcal{P}_1(\mathcal{T}_h)$, be such that*

$$\psi^F(\boldsymbol{x}_F) := 1, \quad \psi^F(\boldsymbol{x}_{F'}) := 0 \quad \forall F' \in \mathcal{F}_h,\, F' \neq F.$$

Above, more precisely, we could have written $\psi^F|_K(\boldsymbol{x}_F)$ for all $K \in \mathcal{T}_h$ such that $F \in \mathcal{F}_K$. An illustration in space dimension $d = 2$ is given in Figure 1.2. In words, ψ^F is a piecewise affine function with respect to the mesh $\mathcal{T}_h$, taking value 1 in the barycenter of the face $F \in \mathcal{F}_h^{\mathrm{int}}$ and the value 0 in all other barycenters of mesh faces. We now confirm that ψ^F, $F \in \mathcal{F}_h^{\mathrm{int}}$, form a *basis* of V_{hp}^{nc} for $p = 1$. Thus, the dimension $|V_{h1}^{\mathrm{nc}}|$ of the space V_{h1}^{nc} will be $|\mathcal{F}_h^{\mathrm{int}}|$, the number of mesh interior faces:

Lemma 1.6.2 (ψ^F)**.** *There holds*

$$V_{h1}^{\mathrm{nc}} = \mathrm{span}_{F \in \mathcal{F}_h^{\mathrm{int}}} \psi^F \tag{1.6.1}$$

Proof. We proceed in two steps.

1) We first show that

$$\psi^F \in V_{h1}^{\mathrm{nc}} \qquad \forall F \in \mathcal{F}_h^{\mathrm{int}}.$$

Recall for this purpose that for any $v_h \in V_{h1}^{\mathrm{nc}}$, the jump on all faces $F' \in \mathcal{F}_h$ has to be of mean value zero, since $p - 1 = 0$ in (1.3.1). Notice now that the trace of any $v_h \in V_{h1}^{\mathrm{nc}}$ on each mesh face $F' \in \mathcal{F}_h$ is an affine function (since $p = 1$). Consequently, as a mean value of an affine function on F' (which is a $(d-1)$-simplex) is equal to the value in the barycenter, the condition of the jump of mean value zero on F' is equivalent to the condition of the jump pointwise zero in the barycenter $\boldsymbol{x}_{F'}$ of F'. Thus, $\psi^F \in V_{h1}^{\mathrm{nc}}$.

2) We now prove (1.6.1). Let $v_h \in V_{h1}^{\mathrm{nc}}$ be given. Restricted to a mesh element $K \in \mathcal{T}_h$, $v_h|_K \in \mathcal{P}_1(K)$, with possibly mean value zero on boundary faces, and thus $v_h|_K$ can be combined from $\psi^F|_K$, $F \in \mathcal{F}_K \cap \mathcal{F}_h^{\mathrm{int}}$. Now on any element $K' \in \mathcal{T}_h$ sharing a face F' with K, again $v_h|_{K'}$ can be combined from $\psi^F|_{K'}$, $F \in \mathcal{F}_{K'} \cap \mathcal{F}_h^{\mathrm{int}}$, and the coefficients for $\psi^{F'}|_K$ and $\psi^{F'}|_{K'}$ are the same, given by the point value (from either K or K') of v_h in the barycenter $\boldsymbol{x}_{F'}$ of F'. □

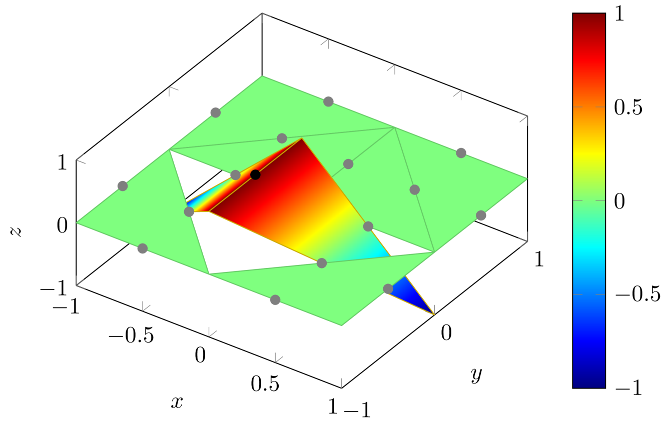

Figure 1.2: Basis function ψ^F of the space V_{h1}^{nc} for $p = 1$

1.6.2 An interpolation operator for V_{h1}^{nc}

We will need below the following interpolation operator, taking values in the weakly-continuous space $H^1_{[\![]\!]}(\mathcal{T}_h)$ of Definition 1.4.1 and producing a piecewise polynomial in the V_{h1}^{nc} space of Definition 1.3.1.

Definition 1.6.3 (Interpolation operator I_1^{nc})**.** *Let the* global interpolation operator

$$I_1^{\mathrm{nc}} : H^1_{[\![]\!]}(\mathcal{T}_h) \to V_{h1}^{\mathrm{nc}}, \quad I_1^{\mathrm{nc}} v := \sum_{F \in \mathcal{F}_h^{\mathrm{int}}} \frac{\langle v, 1 \rangle_F}{|F|} \psi^F \quad \textit{for} \quad v \in H^1_{[\![]\!]}(\mathcal{T}_h). \tag{1.6.2}$$

For $K \in \mathcal{T}_h$, let the local interpolation operator

$$I_{1,K}^{\mathrm{nc}} : H^1(K) \to \mathcal{P}_1(K), \quad I_{1,K}^{\mathrm{nc}} v := \sum_{F \in \mathcal{F}_K \cap \mathcal{F}_h^{\mathrm{int}}} \frac{\langle v, 1 \rangle_F}{|F|} \psi^F|_K \quad \textit{for} \quad v \in H^1(K). \tag{1.6.3}$$

Actually, (1.6.2) *is equivalent to*

$$(I_1^{\mathrm{nc}} v)|_K := I_{1,K}^{\mathrm{nc}}(v|_K) \qquad \forall v \in H^1_{[\![]\!]}(\mathcal{T}_h), \, \forall K \in \mathcal{T}_h. \tag{1.6.4}$$

Several remarks are in order:

- The interpolation operator I_1^{nc} is well defined on the weakly-continuous space $H^1_{[\![]\!]}(\mathcal{T}_h)$; indeed, for both elements $K_1, K_2 \in \mathcal{T}_h$ sharing a face $F \in \mathcal{F}_h^{\mathrm{int}}$, the mean value $\langle v|_{K_i}, 1 \rangle_F / |F|$ is the same by (1.4.1), giving a meaning to (1.6.2).
- Remarkably, as $H_0^1(\Omega) \subset H^1_{[\![]\!]}(\mathcal{T}_h)$, see Remark 1.4.3, I_1^{nc} is well defined over the *whole Sobolev space* $H_0^1(\Omega)$, in contrast to common interpolation operators that take some *point values* (recall that traces are well defined on $H_0^1(\Omega)$ but not point evaluations).
- One calls I_1^{nc} the *global* interpolation operator since it is defined over the whole computational domain Ω and its action concerns functions from $H^1_{[\![]\!]}(\mathcal{T}_h)$. In turn, $I_{1,K}^{\mathrm{nc}}$ is the *local* interpolation operator since it is connected with one mesh element $K \in \mathcal{T}_h$ and it acts on all $v \in H^1(K)$. Remark that the equivalent definition (1.6.4) again crucially hinges upon the weak jump continuity of any $v \in H^1_{[\![]\!]}(\mathcal{T}_h)$.

The following is a simple but important result:

Theorem 1.6.4 (Projection of I_1^{nc})**.** *The operators I_1^{nc} and $I_{1,K}^{\mathrm{nc}}$ from Definition 1.6.3 are* projections *in that*

$$I_1^{\mathrm{nc}} v_h = v_h \qquad \forall v_h \in V_{h1}^{\mathrm{nc}}, \tag{1.6.5a}$$
$$I_{1,K}^{\mathrm{nc}} v_h = v_h \qquad \forall v_h \in \mathcal{P}_1(K),\, \forall K \in \mathcal{T}_h. \tag{1.6.5b}$$

Proof. The second statement (1.6.5b) is trivial. As for the first one (1.6.5a), any $v_h \in V_{h1}^{\mathrm{nc}}$ can be written as

$$v_h = \sum_{F \in \mathcal{F}_h^{\mathrm{int}}} \frac{\langle v_h, 1\rangle_F}{|F|} \psi^F$$

by the reasoning of Lemma 1.6.2, which is exactly the from in (1.6.2). □

We finish this section by proving that I_1^{nc} is a *globally* as well as *locally stable* operator:

Theorem 1.6.5 (Stability of I_1^{nc})**.** *The global interpolation operator I_1^{nc} is stable in that*

$$\|\nabla_h (I_1^{\mathrm{nc}} v)\| \le C \|\nabla_h v\| \qquad \forall v \in H^1_{[\![\,]\!]}(\mathcal{T}_h), \tag{1.6.6}$$

where the constant C only depends on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$ and on the space dimension d. With the same constant, the local interpolation operator $I_{1,K}^{\mathrm{nc}}$ is also stable as

$$\|\nabla (I_{1,K}^{\mathrm{nc}} v)\|_K \le C \|\nabla v\|_K \qquad \forall v \in H^1(K),\, \forall K \in \mathcal{T}_h. \tag{1.6.7}$$

Proof. We only prove (1.6.7) which implies (1.6.6) in view of the link (1.6.4); in this proof, $a \lesssim b$ means that $a \le cb$, where c only depends on $\kappa_{\mathcal{T}_h}$ and d. Let $K \in \mathcal{T}_h$ and $v \in H^1(K)$ be fixed. We will rely on definition (1.6.3). First, we remark that if we shift v by a constant, both the gradient and the gradient of the interpolant stay intact, $\nabla v = \nabla(v + c)$ and $\nabla(I_{1,K}^{\mathrm{nc}} v) = \nabla(I_{1,K}^{\mathrm{nc}}(v + c))$. Thus, we can shift v (still denoted by v with an abuse of notation) such that the mean value $v_K := (v, 1)_K / |K| = 0$. Using (1.6.3), we see that

$$\|\nabla (I_{1,K}^{\mathrm{nc}} v)\|_K = \left\| \nabla \left(\sum_{F \in \mathcal{F}_K \cap \mathcal{F}_h^{\mathrm{int}}} \frac{\langle v, 1\rangle_F}{|F|} \psi^F \right) \right\|_K \le \sum_{F \in \mathcal{F}_K \cap \mathcal{F}_h^{\mathrm{int}}} \frac{|\langle v, 1\rangle_F|}{|F|} \|\nabla \psi^F\|_K.$$

Now, since $|\nabla \psi^F| \lesssim h_K^{-1}$ (note that ψ^F is affine on K and varies between 0 and 1 in faces barycentres),

$$\|\nabla \psi^F\|_K \lesssim h_K^{-1} |K|^{\frac{1}{2}}, \tag{1.6.8}$$

where $|K|$ is the d-volume (length/surface/volume) of the simplex K. Moreover, the Cauchy–Schwarz inequality together with the trace inequality (0.7.3c) yield

$$|\langle v, 1\rangle_F| \le \|v\|_F |F|^{\frac{1}{2}} = \|v - v_K\|_F |F|^{\frac{1}{2}} \lesssim h_K^{\frac{1}{2}} \|\nabla v\|_K |F|^{\frac{1}{2}}.$$

Finally, since the mesh shape-regularity definition (0.5.2) implies

$$|K| \lesssim h_K |F| \qquad \forall F \in \mathcal{F}_K,$$

we conclude that

$$\|\nabla (I_{1,K}^{\mathrm{nc}} v)\|_K \lesssim \sum_{F \in \mathcal{F}_K \cap \mathcal{F}_h^{\mathrm{int}}} \underbrace{h_K^{\frac{1}{2}} |F|^{\frac{1}{2}} |F|^{-1} h_K^{-1} h_K^{\frac{1}{2}} |F|^{\frac{1}{2}}}_{=1} \|\nabla v\|_K \le (d+1) \|\nabla v\|_K.$$

□

1.7 A priori error estimates: best approximation up to the consistency error

We now move towards giving estimates on the error $\|\nabla_h(u-u_h)\|$. The following is a crucial result in this direction (below and in the sequel, we use the shorthand notation $0/0=0$):

Theorem 1.7.1 (Second Strang lemma). *Let $u \in H_0^1(\Omega)$ be the weak solution of Definition 1.2.1 and $u_h \in V_{hp}^{\mathrm{nc}}$ its nonconforming finite element approximation of Definition 1.3.2. Then*

$$\|\nabla_h(u-u_h)\|^2 = \min_{v_h \in V_{hp}^{\mathrm{nc}}} \|\nabla_h(u-v_h)\|^2 + \left(\max_{w_h \in V_{hp}^{\mathrm{nc}}} \frac{(f, w_h) - (\nabla u, \nabla_h w_h)}{\|\nabla_h w_h\|} \right)^2. \tag{1.7.1}$$

Proof. Let $z_h \in V_{hp}^{\mathrm{nc}}$ be such that

$$(\nabla_h z_h, \nabla_h v_h) = (\nabla u, \nabla_h v_h) \qquad \forall v_h \in V_{hp}^{\mathrm{nc}}. \tag{1.7.2}$$

Since the linear form on the right-hand side of (1.7.2) is bounded, $|(\nabla u, \nabla_h v_h)| \le \|\nabla u\| \|\nabla_h v_h\|$, and recalling Theorem 1.5.1, there exists a unique solution z_h to (1.7.2) by the Riesz representation theorem. For an arbitrary $v_h \in V_{hp}^{\mathrm{nc}}$,

$$\begin{aligned}
\|\nabla_h(u-v_h)\|^2 &= \|\nabla_h(u-z_h+z_h-v_h)\|^2 \\
&= \|\nabla_h(u-z_h)\|^2 + \|\nabla_h(z_h-v_h)\|^2 + 2(\nabla_h(u-z_h), \nabla_h(z_h-v_h)) \\
&= \|\nabla_h(u-z_h)\|^2 + \|\nabla_h(z_h-v_h)\|^2,
\end{aligned} \tag{1.7.3}$$

where the second equality follows from (1.7.2) as $z_h - v_h \in V_{hp}^{\mathrm{nc}}$. Consequently,

$$\|\nabla_h(u-z_h)\|^2 = \|\nabla_h(u-v_h)\|^2 - \|\nabla_h(z_h-v_h)\|^2 \le \|\nabla_h(u-v_h)\|^2 \qquad \forall v_h \in V_{hp}^{\mathrm{nc}}$$

and

$$\|\nabla_h(u-z_h)\| = \min_{v_h \in V_{hp}^{\mathrm{nc}}} \|\nabla_h(u-v_h)\|. \tag{1.7.4}$$

The piecewise polynomial $z_h \in V_{hp}^{\mathrm{nc}}$ is simply the orthogonal projection of u onto the space V_{hp}^{nc} (note, however, that $V_{hp}^{\mathrm{nc}} \not\subset H_0^1(\Omega)$, so that one needs to consider u as an element of $H_{[\![]\!]}^1(\mathcal{T}_h)$ rather than as an element of $H_0^1(\Omega)$).

We now apply (1.7.3) to $v_h = u_h$. This gives

$$\|\nabla_h(u-u_h)\|^2 = \|\nabla_h(u-z_h)\|^2 + \|\nabla_h(z_h-u_h)\|^2.$$

Moreover, the second term above becomes

$$\|\nabla_h(z_h-u_h)\| = \max_{w_h \in V_{hp}^{\mathrm{nc}}} \frac{(\nabla_h(z_h-u_h), \nabla_h w_h)}{\|\nabla_h w_h\|} = \max_{w_h \in V_{hp}^{\mathrm{nc}}} \frac{(\nabla u, \nabla_h w_h) - (f, w_h)}{\|\nabla_h w_h\|}$$

by (1.7.2) and (1.3.2). Thus (1.7.1) follows. □

Remark 1.7.2 (Comparison with the conforming finite element method). *In the weak formulation (1.2.1), we can take test functions $v_h \in V_{hp} \subset H_0^1(\Omega)$. Thus, subtracting with (1.3.4), in the conforming finite element method, there immediately holds*

$$(\nabla(u-u_h), \nabla v_h) = 0 \qquad \forall v_h \in V_{hp}. \tag{1.7.5}$$

This is called Galerkin orthogonality*. Consequently, the conforming finite element approximation u_h is nothing but the* orthogonal projection *of the weak solution u to the space V_{hp}, so that*

$$\|\nabla(u-u_h)\| = \min_{v_h \in V_{hp}} \|\nabla(u-v_h)\|. \tag{1.7.6}$$

The first term in (1.7.1) is similar. It is the nonconformity *that gives rise to the second term in (1.7.1), often called* consistency error *term. Indeed, if the function space V_{hp}^{nc} was conforming, $V_{hp}^{\mathrm{nc}} \subset H_0^1(\Omega)$, then this second term would disappear by virtue of (1.2.1).*

1.8 A priori error estimates: rate of convergence for $p = 1$

We will now study how fast the error $\|\nabla_h(u - u_h)\|$ goes to zero with respect to the maximal mesh size h, which is, recall, defined by (0.5.1). We will see that the speed in the lowest-order case $p = 1$ will be of order h^1 if the exact solution u is sufficiently regular.

Let a subdomain $\omega \subset \Omega$ be given. Then the Sobolev space of functions with second-order weak derivatives $H^2(\omega)$ can simply be described as

$$H^2(\omega) = \{v \in H^1(\omega);\, \partial_{\boldsymbol{x}_i} v \in H^1(\omega),\, 1 \le i \le d\}. \tag{1.8.1}$$

Note that, consequently, the weak derivatives $\partial_{\boldsymbol{x}_j}\partial_{\boldsymbol{x}_i} v$, $1 \le j, i \le d$, are well defined. Let $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_d\} \in \mathbb{N}^d$ be a multi-index, a d-tuple of non-negative integers, with $|\boldsymbol{\alpha}| := \alpha_1 + \ldots + \alpha_d \le 2$. With the notation $\partial_{\boldsymbol{\alpha}} v := \partial_{\boldsymbol{x}_1}^{\alpha_1} \ldots \partial_{\boldsymbol{x}_d}^{\alpha_d} v$ and $\partial_{\boldsymbol{\alpha}} v := v$ if $\boldsymbol{\alpha} = \{0, \ldots, 0\}$, the associated semi-norm and norm on $H^2(\omega)$ are then respectively

$$|v|^2_{H^2(\omega)} := \sum_{|\boldsymbol{\alpha}|=2} \|\partial_{\boldsymbol{\alpha}} v\|^2_\omega = \sum_{i=1}^{d}\sum_{j=i}^{d} \|\partial_{\boldsymbol{x}_i}\partial_{\boldsymbol{x}_j} v\|^2_\omega, \tag{1.8.2a}$$

$$\|v\|^2_{H^2(\omega)} := \sum_{|\boldsymbol{\alpha}|\le 2} \|\partial_{\boldsymbol{\alpha}} v\|^2_\omega = \sum_{|\boldsymbol{\alpha}|=0} \|\partial_{\boldsymbol{\alpha}} v\|^2_\omega + \sum_{|\boldsymbol{\alpha}|=1} \|\partial_{\boldsymbol{\alpha}} v\|^2_\omega + \sum_{|\boldsymbol{\alpha}|=2} \|\partial_{\boldsymbol{\alpha}} v\|^2_\omega \tag{1.8.2b}$$

$$= \|v\|^2_\omega + \sum_{i=1}^{d} \|\partial_{\boldsymbol{x}_i} v\|^2_\omega + |v|^2_{H^2(\omega)} \tag{1.8.2c}$$

$$= \|v\|^2_\omega + \|\nabla v\|^2_\omega + |v|^2_{H^2(\omega)} = \|v\|^2_{H^1(\omega)} + |v|^2_{H^2(\omega)}. \tag{1.8.2d}$$

Following Deny and Lions [15] or Bramble and Hilbert [8], we first have the following crucial approximation result (the proof is given in Section 1.9.2 below):

Theorem 1.8.1 (Deny–Lions/Bramble–Hilbert lemma for $p = 1$)**.** *There holds*

$$\min_{v_h \in \mathcal{P}_1(K)} \|\nabla(v - v_h)\|_K \le \sqrt{2}\frac{h_K}{\pi}|v|_{H^2(K)} \qquad \forall v \in H^2(K),\, \forall K \in \mathcal{T}_h. \tag{1.8.3}$$

We now present a "classical" proof of a priori rate of convergence of the nonconforming finite element method (1.3.2) for $p = 1$, in the spirit of Ciarlet [12]:

Theorem 1.8.2 (A priori rate of convergence, $p = 1$)**.** *Let $u \in H^1_0(\Omega)$ be the weak solution of Definition 1.2.1 and $u_h \in V^{\mathrm{nc}}_{hp}$ its nonconforming finite element approximation of Definition 1.3.2 with the polynomial degree $p = 1$. Let additionally*

$$u|_K \in H^2(K) \qquad \forall K \in \mathcal{T}_h. \tag{1.8.4}$$

Then there exists a constant C only depending on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$ and the space dimension d such that

$$\|\nabla_h(u - u_h)\| \le C\left\{\sum_{K \in \mathcal{T}_h} h_K^2 |u|^2_{H^2(K)}\right\}^{\frac{1}{2}} \le Ch\left\{\sum_{K \in \mathcal{T}_h} |u|^2_{H^2(K)}\right\}^{\frac{1}{2}}. \tag{1.8.5}$$

Proof. From the characterization (1.7.1), we see that we need to bound two terms, so we proceed in two steps. Here again, $a \lesssim b$ means that $a \le cb$, where c only depends on $\kappa_{\mathcal{T}_h}$ and d.

1) Recall the interpolation operator I^{nc}_1 from Definition 1.6.3. There immediately holds

$$\min_{v_h \in V^{\mathrm{nc}}_{hp}} \|\nabla_h(u - v_h)\|^2 \le \|\nabla_h(u - I^{\mathrm{nc}}_1 u)\|^2 = \sum_{K \in \mathcal{T}_h} \|\nabla(u - I^{\mathrm{nc}}_{1,K} u)\|^2_K,$$

so that we can henceforth only study the *interpolation error* $\|\nabla(u - I_{1,K}^{\mathrm{nc}}u)\|_K$ on each *single mesh element* $K \in \mathcal{T}_h$. For this purpose, we will use the orthogonal projector P_1 onto $\mathcal{P}_1(K)$: for each $v \in H^1(K)$, $P_1 v \in \mathcal{P}_1(K)$ is such that

$$(\nabla P_1 v, \nabla v_h)_K = (\nabla v, \nabla v_h)_K \qquad \forall v_h \in \mathcal{P}_1(K); \tag{1.8.6}$$

to make $P_1 v$ unique, it is convenient to fix

$$(P_1 v, 1)_K = (v, 1)_K,$$

i.e., having the same mean value on K as v. The orthogonal projection property (1.8.6) immediately gives

$$\|\nabla(v - P_1 v)\|_K = \min_{v_h \in \mathcal{P}_1(K)} \|\nabla(v - v_h)\|_K, \tag{1.8.7}$$

We will also use the projection property (1.6.5b) of the interpolation operator I_1^{nc}, namely $I_{1,K}^{\mathrm{nc}} P_1 u = P_1 u$. This gives, on the element K,

$$u - I_{1,K}^{\mathrm{nc}}u = u - I_{1,K}^{\mathrm{nc}}u - P_1 u + P_1 u = u - P_1 u + I_{1,K}^{\mathrm{nc}}(P_1 u - u),$$

so that

$$\|\nabla(u - I_{1,K}^{\mathrm{nc}}u)\|_K \le \|\nabla(u - P_1 u)\|_K + \|\nabla(I_{1,K}^{\mathrm{nc}}(P_1 u - u))\|_K \le (1 + C)\|\nabla(u - P_1 u)\|_K,$$

where C is the constant from the crucial local stability bound (1.6.7). Thus, (1.8.7) and (1.8.3) allow us to conclude

$$\min_{v_h \in V_{hp}^{\mathrm{nc}}} \|\nabla_h(u - v_h)\| \lesssim \left\{ \sum_{K \in \mathcal{T}_h} h_K^2 |u|_{H^2(K)}^2 \right\}^{\frac{1}{2}}.$$

2) Let $w_h \in V_{hp}^{\mathrm{nc}}$ be fixed. We manipulate, using the assumption $u|_K \in H^2(K)$ for all $K \in \mathcal{T}_h$, the Green theorem on each mesh element, (1.1.1a), the fact that $w_h \in V_{hp}^{\mathrm{nc}}$, and the Cauchy–Schwarz inequality

$$\begin{aligned}
(f, w_h) - (\nabla u, \nabla_h w_h) &= \sum_{K \in \mathcal{T}_h} \{(f, w_h)_K - (\nabla u, \nabla_h w_h)_K\} \\
&= \sum_{K \in \mathcal{T}_h} \{(f + \Delta u, w_h)_K - (\nabla u {\cdot} \boldsymbol{n}_K, w_h)_{\partial K}\} \\
&\overset{(1.1.1a)}{=} - \sum_{K \in \mathcal{T}_h} (\nabla u {\cdot} \boldsymbol{n}_K, w_h)_{\partial K} \\
&= - \sum_{F \in \mathcal{F}_h} (\nabla u {\cdot} \boldsymbol{n}_F, [\![w_h]\!])_F \\
&\overset{(1.3.1)}{=} - \sum_{F \in \mathcal{F}_h} ([\nabla u - (\nabla u)_F] {\cdot} \boldsymbol{n}_F, [\![w_h]\!])_F \\
&\le \sum_{F \in \mathcal{F}_h} \|\nabla u - (\nabla u)_F\|_F \|[\![w_h]\!]\|_F,
\end{aligned}$$

where $(\nabla u)_F$ is the componentwise mean value of ∇u on the face F (recall that $\partial_{\boldsymbol{x}_i} u \in H^1(K)$ for all $1 \le i \le d$ from (1.8.1), so that $\partial_{\boldsymbol{x}_i} u$ on the faces F exists in the trace sense). We have also crucially used here that

$$[\![\nabla u {\cdot} \boldsymbol{n}_F]\!] = 0 \qquad \forall F \in \mathcal{F}_h^{\mathrm{int}},$$

which is a consequence of assumption (1.8.4) together with the fact that $\nabla u \in \boldsymbol{H}(\mathrm{div}, \Omega)$ by Definitions 1.2.1 and 0.4.2.

For a fixed face $F \in \mathcal{F}_h^{\text{int}}$ shared by two mesh elements K_1 and K_2, we now on the one hand have, using (0.6.2a) and the trace inequality (0.7.3b),

$$\begin{aligned}
\|[\![w_h]\!]\|_F = \|w_h|_{K_1} - w_h|_{K_2}\|_F = \|w_h|_{K_1} - (w_h)_F + (w_h)_F - w_h|_{K_2}\|_F \\
\leq \sum_{K \in \mathcal{T}_h,\, F \in \mathcal{F}_K} \|w_h|_K - (w_h)_F\|_F \overset{(0.7.3b)}{\lesssim} \sum_{K \in \mathcal{T}_h,\, F \in \mathcal{F}_K} h_K^{\frac{1}{2}} \|\nabla w_h\|_K,
\end{aligned}$$

and the same bound holds for any $F \in \mathcal{F}_h^{\text{ext}}$ as well. On the other hand, the trace inequality (0.7.3b) also gives

$$\|\nabla u - (\nabla u)_F\|_F^2 = \sum_{i=1}^d \|\partial_{\boldsymbol{x}_i} u - (\partial_{\boldsymbol{x}_i} u)_F\|_F^2 \lesssim \sum_{i=1}^d h_K \|\nabla \partial_{\boldsymbol{x}_i} u\|_K^2 \leq 2 h_K |u|_{H^2(K)}^2$$

for any of the (two for $F \in \mathcal{F}_h^{\text{int}}$ and one for $F \in \mathcal{F}_h^{\text{ext}}$) elements K that share the face F, where we have proceeded as in (1.9.4). Collecting the above bounds, the Cauchy–Schwarz inequality gives

$$\begin{aligned}
|(f, w_h) - (\nabla u, \nabla_h w_h)| &\lesssim \sum_{F \in \mathcal{F}_h} \sum_{K \in \mathcal{T}_h,\, F \in \mathcal{F}_K} h_K \|\nabla w_h\|_K |u|_{H^2(K)} \\
&\leq 2 \sum_{K \in \mathcal{T}_h} h_K \|\nabla w_h\|_K |u|_{H^2(K)} \\
&\leq 2 \left\{ \sum_{K \in \mathcal{T}_h} h_K^2 |u|_{H^2(K)}^2 \right\}^{\frac{1}{2}} \|\nabla w_h\|,
\end{aligned}$$

which allows to conclude the proof. □

1.9 Complements

We collect here some additional material, developing further the above contents.

1.9.1 Proof of Theorem 1.4.4

Proof of Theorem 1.4.4. From Remark 1.4.3, $H^1_{[\![]\!]}(\mathcal{T}_h)$ is non-empty. The inclusion $V_{hp}^{\text{nc}} \subset H^1_{[\![]\!]}(\mathcal{T}_h)$ follows immediately from definitions (1.3.1) and (1.4.1) since $\mathcal{P}_p(\mathcal{T}_h) \subset H^1(\mathcal{T}_h)$ and the jumps orthogonality constraint in (1.3.1) (requesting jump orthogonality with respect to $(p-1)$-degree polynomials) is stronger than that in (1.4.1) (requesting jump orthogonality with respect to 0-degree polynomials).

To show that $H^1_{[\![]\!]}(\mathcal{T}_h)$ is a closed subspace of $H^1(\mathcal{T}_h)$, we will rely on the second characterization in (0.6.1). Consider a sequence of functions $v_i \in H^1_{[\![]\!]}(\mathcal{T}_h)$. We need to show that whenever $v \in H^1(\mathcal{T}_h)$ is its limit, i.e., v is such that

$$\lim_{i \to \infty} \|v_i - v\|_{H^1(\mathcal{T}_h)} = 0, \tag{1.9.1}$$

where $\|\cdot\|_{H^1(\mathcal{T}_h)}$ is the norm on the Hilbert space $H^1(\mathcal{T}_h)$ from (0.6.5b), then $v \in H^1_{[\![]\!]}(\mathcal{T}_h)$, i.e.,

$$\langle [\![v]\!], q_h \rangle_F = 0 \qquad \forall q_h \in \mathcal{P}_0(F),\, \forall F \in \mathcal{F}_h. \tag{1.9.2}$$

Consider an arbitrary face $F \in \mathcal{F}_h$ and a constant function $q_h \in \mathcal{P}_0(F)$. Adding and subtracting v_i, and since $v_i \in H^1_{[\![]\!]}(\mathcal{T}_h)$, we have

$$\langle [\![v]\!], q_h \rangle_F = \langle [\![v - v_i]\!], q_h \rangle_F + \langle [\![v_i]\!], q_h \rangle_F = \langle [\![v - v_i]\!], q_h \rangle_F.$$

Now Definition 0.6.2 of the jump, the triangle inequality, the Cauchy–Schwarz inequality, and the trace inequality (0.7.3a) imply

$$\begin{aligned}\langle [\![v-v_i]\!], q_h\rangle_F &\le \sum_{K\in\mathcal{T}_h;F\in\mathcal{F}_K} |\langle (v-v_i)|_K, q_h\rangle_F| \le \sum_{K\in\mathcal{T}_h;F\in\mathcal{F}_K} \|(v-v_i)|_K\|_F\|q_h\|_F\\ &\le \sum_{K\in\mathcal{T}_h;F\in\mathcal{F}_K} \widetilde{C}_{\mathrm{t},\kappa_K,d}(h_K^{-1}\|v-v_i\|_K^2 + \|v-v_i\|_K\|\nabla(v-v_i)\|_K)^{\frac{1}{2}}\|q_h\|_F,\end{aligned}$$

where, recall, $\mathcal{F}_K$ denotes the set of faces of the simplex K (the above sum has two summands for interior faces F and only one for boundary faces F). Now assumption (1.9.1) and definition (0.6.5b) imply, for all mesh elements $K\in\mathcal{T}_h$,

$$\begin{aligned}\lim_{i\to\infty}\|v-v_i\|_K &= 0,\\ \lim_{i\to\infty}\|\nabla(v-v_i)\|_K &= 0,\end{aligned}$$

so that $|\langle [\![v-v_i]\!], q_h\rangle_F|$ tends to zero and (1.9.2) follows. □

1.9.2 Proof of Theorem 1.8.1

Proof of Theorem 1.8.1. We present a short direct proof, as in Ern and Guermond [20, Corollary 9.13]. Let $K\in\mathcal{T}_h$ and $v\in H^2(K)$ be fixed. Recall that P_1v is the orthogonal projection onto $\mathcal{P}_1(K)$ given by (1.8.6). Choosing $v_h = \boldsymbol{x}_i|_K$, it follows from (1.8.6) that

$$(\partial_{\boldsymbol{x}_i}(v-P_1v),1)_K = 0 \qquad \forall 1\le i\le d, \tag{1.9.3}$$

i.e., the difference $v-P_1v$ has vanishing first order weak partial derivatives on K. We see, using (0.3.1), (0.2.1), (1.9.3), (0.7.1) (recall that K is a convex, so that $C_{\mathrm{P},K}\le 1/\pi$)

$$\begin{aligned}\|\nabla(v-P_1v)\|_K^2 &= \sum_{i=1}^d \|\partial_{\boldsymbol{x}_i}(v-P_1v)\|_K^2 \overset{(1.9.3)}{=} \sum_{i=1}^d \|\partial_{\boldsymbol{x}_i}(v-P_1v) - (\partial_{\boldsymbol{x}_i}(v-P_1v))_K\|_K^2\\ &\overset{(0.7.1)}{\le} \frac{h_K^2}{\pi^2}\sum_{i=1}^d \|\nabla\partial_{\boldsymbol{x}_i}(v-P_1v)\|_K^2 = \frac{h_K^2}{\pi^2}\sum_{i=1}^d\|\nabla\partial_{\boldsymbol{x}_i}v\|_K^2\\ &= \frac{h_K^2}{\pi^2}\sum_{i=1}^d\sum_{j=1}^d \|\partial_{\boldsymbol{x}_j}\partial_{\boldsymbol{x}_i}v\|_K^2 \le 2\frac{h_K^2}{\pi^2}\sum_{i=1}^d\sum_{j=i}^d\|\partial_{\boldsymbol{x}_i}\partial_{\boldsymbol{x}_j}v\|_K^2 = 2\frac{h_K^2}{\pi^2}|v|_{H^2(K)}^2;\end{aligned} \tag{1.9.4}$$

crucially, we also use that $P_1v\in\mathcal{P}_1(K)$, so that $\nabla\partial_{\boldsymbol{x}_i}P_1v = 0$ for all $1\le i\le d$, and compensate by the factor 2 the fact that the cross derivatives are not repeated in (1.8.2a). Now, using (1.8.7), (1.8.3) follows. □

1.9.3 A priori error estimates: best approximation up to data oscillation

We now present some recent extensions of Theorem 1.7.1.

For $p'\ge 0$, let $\Pi_{p'}$ be the $L^2(\Omega)$ orthogonal projection onto the space $\mathcal{P}_{p'}(\mathcal{T}_h)$ of Definition 0.8.1, i.e., for $v\in L^2(\Omega)$,

$$\Pi_{p'}v\in\mathcal{P}_{p'}(\mathcal{T}_h), \quad (v-\Pi_{p'}v, v_h) = 0 \qquad \forall v_h\in\mathcal{P}_{p'}(\mathcal{T}_h). \tag{1.9.5}$$

Remark that this can be equivalently defined elementwise as

$$(\Pi_{p'}v)|_K\in\mathcal{P}_{p'}(K), \quad (v-\Pi_{p'}v, v_h)_K = 0 \qquad \forall v_h\in\mathcal{P}_{p'}(K),\ \forall K\in\mathcal{T}_h.$$

With this notation, let us now define the following computable quantity η_{osc}. We note straight away that for smooth right-hand sides f, η_{osc} goes to zero by one order faster than the error $\|\nabla_h(u-u_h)\|$, which we will discuss in details in Section 1.9.4 below, see (1.9.16).

Definition 1.9.1 (Data oscillation). *Let*

$$\eta_{\mathrm{osc}}^2 := \sum_{K\in\mathcal{T}_h} \frac{h_K^2}{\pi^2}\|f - \Pi_{p-1} f\|_K^2. \tag{1.9.6}$$

Theorem 1.9.2 (Best approximation in V_{hp}^{nc}). *Let $u \in H_0^1(\Omega)$ be the weak solution of Definition 1.2.1 and $u_h \in V_{hp}^{\mathrm{nc}}$ its nonconforming finite element approximation of Definition 1.3.2. Then*

$$\|\nabla_h(u-u_h)\| \le C\left(\min_{v_h\in V_{hp}^{\mathrm{nc}}}\|\nabla_h(u-v_h)\| + \eta_{\mathrm{osc}}\right), \tag{1.9.7}$$

where the generic constant C only depends on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$, the space dimension d, and the polynomial degree p.

Proof. We need to bound the second term in (1.7.1). On each mesh element $K \in \mathcal{T}_h$, let the specific space of vector-valued polynomials of Raviart and Thomas [6, 36] of degree $p' \ge 0$ be given by

$$\boldsymbol{\mathcal{RT}}_{p'}(K) := [\mathcal{P}_{p'}(K)]^3 + \mathcal{P}_{p'}(K)\boldsymbol{x}. \tag{1.9.8}$$

We then define

$$\boldsymbol{V}_{hp'} := \{\boldsymbol{v}_h \in \boldsymbol{H}(\mathrm{div},\Omega);\, \boldsymbol{v}_h|_K \in \boldsymbol{\mathcal{RT}}_{p'}(K) \quad \forall K\in\mathcal{T}_h\}, \tag{1.9.9}$$

where, recall, $\boldsymbol{H}(\mathrm{div},\Omega)$ is introduced in Definition 0.4.2. Two particular features of $\boldsymbol{v}_h \in \boldsymbol{V}_{h(p-1)}$ are that 1) $(\nabla{\cdot}\boldsymbol{v}_h)|_K \in \mathcal{P}_{p-1}(K)$ for all mesh elements $K\in\mathcal{T}_h$; 2) $(\boldsymbol{v}_h{\cdot}\boldsymbol{n}_F)|_F \in \mathcal{P}_{p-1}(F)$ for all mesh faces $F\in\mathcal{F}_h$. This second property, together with the jump orthogonality assumption (1.3.1), implies the *discrete Green formula*

$$(\boldsymbol{v}_h, \nabla_h w_h) + (\nabla{\cdot}\boldsymbol{v}_h, w_h) = 0 \qquad \forall \boldsymbol{v}_h \in \boldsymbol{V}_{h(p-1)},\ \forall w_h \in V_{hp}^{\mathrm{nc}}.$$

Let now $w_h \in V_{hp}^{\mathrm{nc}}$ be fixed and consider any $\boldsymbol{v}_h \in \boldsymbol{V}_{h(p-1)}$ such that $\nabla{\cdot}\boldsymbol{v}_h = \Pi_{p-1}f$. Then the Poincaré inequality (0.7.1) applied on convex simplices K (so that $C_{\mathrm{P},K} \le 1/\pi$) together with the Cauchy–Schwarz inequality give

$$\begin{aligned}
(f,w_h) - (\nabla u, \nabla_h w_h) &= (f-\Pi_{p-1}f, w_h) + (\Pi_{p-1}f, w_h) - (\nabla u, \nabla_h w_h)\\
&= (f-\Pi_{p-1}f, w_h) + (\nabla{\cdot}\boldsymbol{v}_h, w_h) - (\nabla u, \nabla_h w_h)\\
&= \sum_{K\in\mathcal{T}_h}(f-\Pi_{p-1}f, w_h - \Pi_0 w_h)_K - (\boldsymbol{v}_h + \nabla u, \nabla_h w_h)\\
&\le \sum_{K\in\mathcal{T}_h}\|f-\Pi_{p-1}f\|_K\frac{h_K}{\pi}\|\nabla_h w_h\|_K + \|\boldsymbol{v}_h+\nabla u\|\|\nabla_h w_h\|\\
&\le (\eta_{\mathrm{osc}} + \|\boldsymbol{v}_h + \nabla u\|)\|\nabla_h w_h\|.
\end{aligned}$$

Consequently, the second term in (1.7.1) (without square) can be bounded by

$$\eta_{\mathrm{osc}} + \min_{\substack{\boldsymbol{v}_h\in\boldsymbol{V}_{h(p-1)}\\ \nabla\cdot\boldsymbol{v}_h = \Pi_{p-1}f}} \|\boldsymbol{v}_h + \nabla u\|.$$

Thus the *consistency error* related to the nonconformity $u_h \notin H_0^1(\Omega)$ has been changed into a *best-approximation* in the discrete subspace $\boldsymbol{V}_{h(p-1)}$ of the $\boldsymbol{H}(\mathrm{div},\Omega)$ space plus data oscillation. The decisive advantage is that this can now be manipulated without any complications related to the treatment of the *nonconformity*. In particular, the results from [19] imply

$$\min_{\substack{\boldsymbol{v}_h\in\boldsymbol{V}_{h(p-1)}\\ \nabla\cdot\boldsymbol{v}_h = \Pi_{p-1}f}} \|\boldsymbol{v}_h + \nabla u\| \le C\left(\min_{v_h\in\mathcal{P}_p(\mathcal{T}_h)}\|\nabla_h(u-v_h)\| + \eta_{\mathrm{osc}}\right)$$

with a generic constant C with same dependencies as in (1.9.7). Since

$$\min_{v_h \in \mathcal{P}_p(\mathcal{T}_h)} \|\nabla_h(u - v_h)\| \leq \min_{v_h \in V_{hp}^{\mathrm{nc}}} \|\nabla_h(u - v_h)\|,$$

as $V_{hp}^{\mathrm{nc}} \subset \mathcal{P}_p(\mathcal{T}_h)$, this finishes the proof.

$\square$

Theorem 1.9.3 (Best approximation in $\mathcal{P}_p(\mathcal{T}_h)$)**.** *Let $u \in H_0^1(\Omega)$ be the weak solution of Definition 1.2.1 and $u_h \in V_{hp}^{\mathrm{nc}}$ its nonconforming finite element approximation of Definition 1.3.2. Then*

$$\|\nabla_h(u - u_h)\| \leq C\left(\min_{v_h \in \mathcal{P}_p(\mathcal{T}_h)} \|\nabla_h(u - v_h)\| + \eta_{\mathrm{osc}}\right), \tag{1.9.10}$$

where the generic constant C only depends on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$, the space dimension d, and the polynomial degree p.

Proof. The seminal result of Veeser [41] shows that

$$\min_{v_h \in V_{hp}} \|\nabla_h(v - v_h)\| \leq C \min_{v_h \in \mathcal{P}_p(\mathcal{T}_h)} \|\nabla_h(v - v_h)\| \qquad \forall v \in H_0^1(\Omega), \tag{1.9.11}$$

i.e., up to a generic constant C, the capability of the (much) smaller space V_{hp} to approximate a function from the space $H_0^1(\Omega)$ is comparable to that of the entire space $\mathcal{P}_p(\mathcal{T}_h)$. Recall that $V_{hp} \subset \mathcal{P}_p(\mathcal{T}_h)$, which in particular immediately implies the converse inequality

$$\min_{v_h \in \mathcal{P}_p(\mathcal{T}_h)} \|\nabla_h(v - v_h)\| \leq \min_{v_h \in V_{hp}} \|\nabla_h(v - v_h)\| \qquad \forall v \in H_0^1(\Omega).$$

Since

$$\min_{v_h \in V_{hp}^{\mathrm{nc}}} \|\nabla_h(u - v_h)\| \leq \min_{v_h \in V_{hp}} \|\nabla_h(u - v_h)\|, \tag{1.9.12}$$

as $V_{hp} \subset V_{hp}^{\mathrm{nc}}$, one concludes (1.9.10) from (1.9.7), (1.9.12), and (1.9.11). $\square$

Remark 1.9.4 (Theorems 1.7.1, 1.9.2, and 1.9.3)**.** *In Theorem 1.7.1, the bound (1.7.1) consists of the best approximation of u in the nonconforming finite element space V_{hp}^{nc} plus the consistency error. In Theorem 1.9.2, first shown in Gudi [26] for $p = 1$, the consistency term is absorbed in the best approximation in V_{hp}^{nc} plus the data oscillation. The result (1.9.7) makes, up to the data oscillation term, a parallel with the characterization (1.7.6) in the $H_0^1(\Omega)$-conforming space V_{hp}. Finally, Theorem 1.9.3 is a nonconforming extension of Veeser [41]. The bound (1.9.10) might seem quite surprising at a first sight, since the best approximation in $\mathcal{P}_p(\mathcal{T}_h)$ on the right-hand side might initially seem to be (much) smaller than the best approximation in V_{hp}^{nc} on the left-hand side. This is structurally the sharpest result.*

1.9.4 A priori error estimates: rate of convergence for $p \geq 1$

In extension of Section 1.8, we prove here that the speed of convergence of $\|\nabla_h(u - u_h)\|$ for $p \geq 1$ is of order h^p. We first need to extend Definition 0.3.1 to higher-order derivatives:

Definition 1.9.5 (Weak multi-index derivative)**.** *Let a scalar-valued function $v : \Omega \to \mathbb{R}$ be given. Let $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_d\} \in \mathbb{N}^d$ be a multi-index, i.e., a d-tuple of non-negative integers. We say that v admits a weak $\boldsymbol{\alpha}$-th partial derivative if*

1. *$v \in L^2(\Omega)$;*
2. *there exists a function $w_{\boldsymbol{\alpha}} : \Omega \to \mathbb{R}$ such that*
   - *(a) $w_{\boldsymbol{\alpha}} \in L^2(\Omega)$;*
   - *(b) $(v, \partial_{\boldsymbol{\alpha}} \varphi) = (-1)^{|\boldsymbol{\alpha}|}(w_{\boldsymbol{\alpha}}, \varphi) \qquad \forall \varphi \in \mathcal{D}(\Omega)$,*

where $|\boldsymbol{\alpha}| := \alpha_1 + \ldots + \alpha_d$ *and* $\partial_{\boldsymbol{\alpha}}\varphi := \partial_{\boldsymbol{x}_1}^{\alpha_1} \ldots \partial_{\boldsymbol{x}_d}^{\alpha_d}\varphi$.

We define the weak $\boldsymbol{\alpha}$-th partial derivative *of* v, *denoted by* $\partial_{\boldsymbol{\alpha}}v$, *as*

$$\partial_{\boldsymbol{\alpha}}v := \partial_{\boldsymbol{x}_1}^{\alpha_1} \ldots \partial_{\boldsymbol{x}_d}^{\alpha_d}v := w_{\boldsymbol{\alpha}}. \tag{1.9.13}$$

We use the notation $\partial_{\boldsymbol{\alpha}}v := v$ *if* $\boldsymbol{\alpha} = \{0, \ldots, 0\}$.

With the notation of Definition 1.9.5, the Sobolev space with m-th order weak partial derivatives is defined as:

Definition 1.9.6 (Space $H^m(\Omega)$)**.** *For an integer* $m \geq 0$, *the space* $H^m(\Omega)$ *is the space of all functions which admit the weak multi-index derivatives of order* $|\boldsymbol{\alpha}| \leq m$. *This is often abbreviated as*

$$H^m(\Omega) = \{v \in L^2(\Omega);\, \partial_{\boldsymbol{\alpha}}v \in L^2(\Omega) \quad \forall \boldsymbol{\alpha} \in \mathbb{N}^d \text{ with } |\boldsymbol{\alpha}| \leq m\}.$$

The m*-th order semi-norm is defined as*

$$|u|^2_{H^m(\omega)} = \sum_{|\boldsymbol{\alpha}|=m} \|\partial_{\boldsymbol{\alpha}}v\|^2_\omega. \tag{1.9.14}$$

Equipped with the previous definitions, one can extend Theorem 1.8.1 to any order:

Theorem 1.9.7 (Deny–Lions/Bramble–Hilbert lemma)**.** *There holds*

$$\min_{v_h \in \mathcal{P}_p(K)} \|\nabla(v - v_h)\|_K \leq \sqrt{(p+1)!}\,\frac{h_K^p}{\pi^p}|v|_{H^{p+1}(K)} \qquad \forall v \in H^{p+1}(K),\, \forall K \in \mathcal{T}_h. \tag{1.9.15}$$

Finally, Theorems 1.9.3 and 1.9.7 immediately imply:

Theorem 1.9.8 (A priori rate of convergence)**.** *Let* $u \in H^1_0(\Omega)$ *be the weak solution of Definition 1.2.1 and* $u_h \in V^{\mathrm{nc}}_{hp}$ *its nonconforming finite element approximation of Definition 1.3.2. Let additionally*

$$\begin{aligned} u|_K &\in H^{p+1}(K) \qquad && \forall K \in \mathcal{T}_h, \\ f|_K &\in H^p(K) \qquad && \forall K \in \mathcal{T}_h. \end{aligned}$$

Then there exists a constant C *only depending on the mesh shape-regularity parameter* $\kappa_{\mathcal{T}_h}$, *the space dimension* d, *and the polynomial degree* p *such that*

$$\begin{aligned} \|\nabla_h(u - u_h)\| &\leq C\left(\left\{\sum_{K \in \mathcal{T}_h} h_K^{2p}|u|^2_{H^{p+1}(K)}\right\}^{\frac{1}{2}} + \eta_{\mathrm{osc}}\right) \\ &\leq Ch^p\left\{\sum_{K \in \mathcal{T}_h} |u|^2_{H^{p+1}(K)}\right\}^{\frac{1}{2}} + Ch^{p+1}\left\{\sum_{K \in \mathcal{T}_h} |f|^2_{H^p(K)}\right\}^{\frac{1}{2}}. \end{aligned} \tag{1.9.16}$$

Chapter 2

Potential reconstruction in $H_0^1(\Omega)$

In this chapter, we show how to, from a possibly completely *discontinuous* piecewise polynomial, efficiently reconstruct a *continuous* piecewise polynomial. This will turn as an important tool to obtain reliable and efficient a posteriori error estimates in Chapter 4. As a theoretical device, though, it also serves in a priori error analysis: was have already used in the background of the proof of Theorem 1.9.3 in Chapter 1 (it serves to demonstrate the result of [41]), and we will also use it in a priori error estimates in Chapter 6. Below, we follow the results in Karakashian and Pascal [27], Burman and Ern [10], and [24, 25].

2.1 Setting

Let $p \geq 1$ be a given polynomial degree. Recall Definition 0.8.1 of the piecewise polynomial space

$$\mathcal{P}_p(\mathcal{T}_h) = \{v_h \in L^2(\Omega);\ v_h|_K \in \mathcal{P}_p(K) \quad \forall K \in \mathcal{T}_h\} \tag{2.1.1}$$

and Definition 1.3.1 of the weakly-continuous (Crouzeix–Raviart) nonconforming finite element space

$$V_{hp}^{\mathrm{nc}} = \{v_h \in \mathcal{P}_p(\mathcal{T}_h);\ \langle [\![v_h]\!], q_h \rangle_F = 0 \quad \forall q_h \in \mathcal{P}_{p-1}(F), \forall F \in \mathcal{F}_h\}. \tag{2.1.2}$$

Let $p' \geq 1$ be a polynomial degree, possibly different from p. The standard finite element space from Remark 1.3.3 is then

$$V_{hp'} = \mathcal{P}_{p'}(\mathcal{T}_h) \cap H_0^1(\Omega). \tag{2.1.3}$$

2.2 Lagrangian nodes

Lagrangian nodes for a mesh element $K \in \mathcal{T}_h$ are points $\boldsymbol{x}_i \in K$ (recall that K is supposed closed) that enable a unique prescription of a polynomial $v_h \in \mathcal{P}_p(K)$ by its values in $\boldsymbol{x}_i$. The typical (but not mandatory) choice is depicted in Figure 2.1 for $p = 1$ and $p = 2$ in two or three space dimensions. Similarly, Lagrangian nodes for the entire computational domain Ω are points $\boldsymbol{x}_i \in \overline{\Omega}$ that enable a unique prescription of a continuous piecewise polynomial $v_h \in V_{hp}$. The typical (but again not mandatory) choice is depicted in Figure 2.2; this choice matches with that in Figure 2.1, left. Note in particular that prescribing the point values in the "free" nodes located inside Ω ensures the $H^1(\Omega)$ trace-continuity, which is equivalent to the $C^0(\overline{\Omega})$-continuity for piecewise polynomials, see the exercices. Imposing zero point values in the "frozen" nodes located at the boundary $\partial\Omega$ then in turn ensures the homogeneous Dirichlet boundary condition. We refer for details to Ciarlet [12, Section 2.2.2].

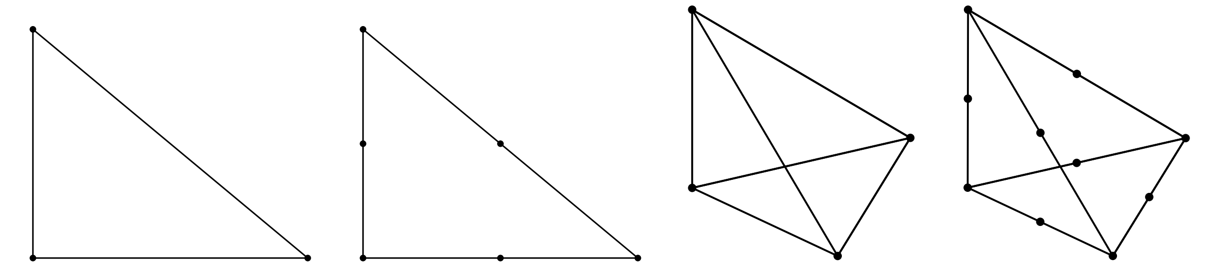

Figure 2.1: Lagrangian nodes for a mesh element $K \in \mathcal{T}_h$, $p = 1$ and $p = 2$; $d = 2$ (left) and $d = 3$ (right)

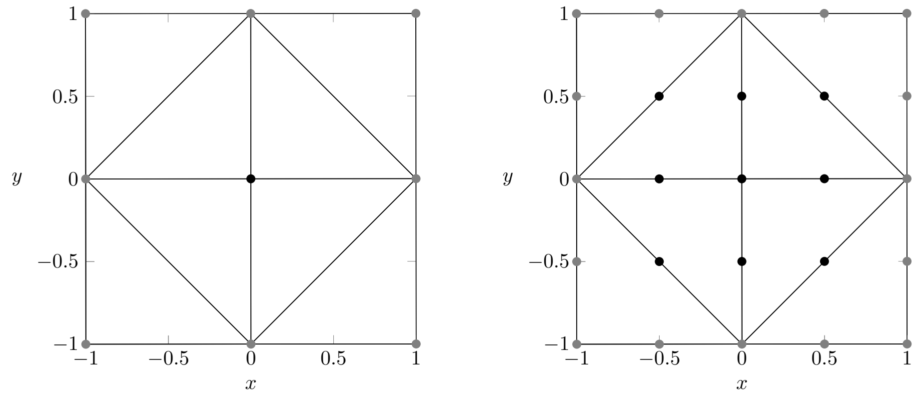

Figure 2.2: $\Omega = (-1, 1) \times (-1, 1)$, $d = 2$. Lagrangian nodes for the space V_{hp} with $p = 1$ (left) and $p = 2$ (right); "free" nodes in black, "frozen" nodes where values are fixed to zero in grey

2.3 Potential reconstruction by averaging

The following is a simple *locally-defined* operator from $\mathcal{P}_p(\mathcal{T}_h)$ to V_{hp}. We illustrate it in Figure 2.3 when applied to $v_h \in V_{h1}^{\mathrm{nc}} \subset \mathcal{P}_1(\mathcal{T}_h)$ given by the basis function ψ^F from Definition 1.6.1. A more general example is given in Figure 2.4.

Definition 2.3.1 (Potential reconstruction by averaging)**.** *Let $v_h \in \mathcal{P}_p(\mathcal{T}_h)$ be given. Then the* potential reconstruction by averaging*, $s_h \in V_{hp}$, is prescribed by its values in the Lagrange points of the space V_{hp} by*

$$s_h(\boldsymbol{x}) := \frac{1}{|\mathcal{T}_{\boldsymbol{x}}|} \sum_{K \in \mathcal{T}_{\boldsymbol{x}}} v_h|_K(\boldsymbol{x}) \qquad \boldsymbol{x} \text{ is a Lagrange point of } V_{hp} \text{ included in } \Omega, \tag{2.3.1a}$$

$$s_h(\boldsymbol{x}) := 0 \qquad \boldsymbol{x} \text{ is a Lagrange point of } V_{hp} \text{ included in } \partial\Omega, \tag{2.3.1b}$$

where $\mathcal{T}_{\boldsymbol{x}}$ denotes the set of elements of the mesh $\mathcal{T}_h$ that contain the point $\boldsymbol{x}$ and $|\mathcal{T}_{\boldsymbol{x}}|$ is the cardinality (number of elements) of this set.

2.4 Approximation properties

The procedure of Definition 2.3.1 has been analyzed in [27, 10, 1] to yield the following crucial result. It states that, up to a generic constant, no function from the *infinite-dimensional* space $H_0^1(\Omega)$, $s \in H_0^1(\Omega)$, can be made closer to v_h than s_h in the $[L^2(\Omega)]^d$ norm of the broken weak gradient when $v_h \in V_{hp}^{\mathrm{nc}}$. This is truly remarkable, since s_h is only a *piecewise polynomial* that has been *constructed locally*. Below, $\mathcal{T}_K$ is the set of mesh elements that share at least a vertex with $K \in \mathcal{T}_h$ and ω_K is the corresponding subdomain of Ω.

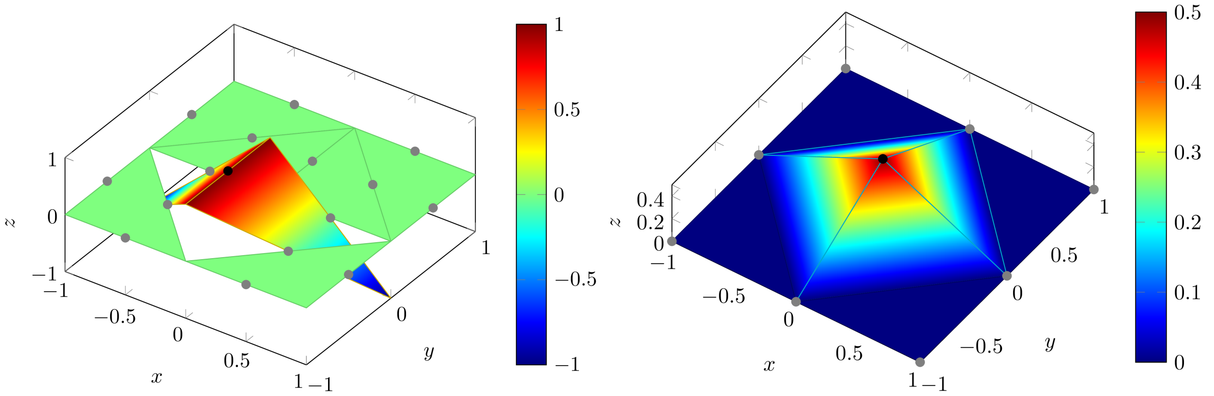

Figure 2.3: $d = 2$, $p = 1$. The function $v_h = \psi^F \in V_{h1}^{\mathrm{nc}}$ from Definition 1.6.1 and the corresponding potential reconstruction $s_h \in V_{h1}$ from Definition 2.3.1

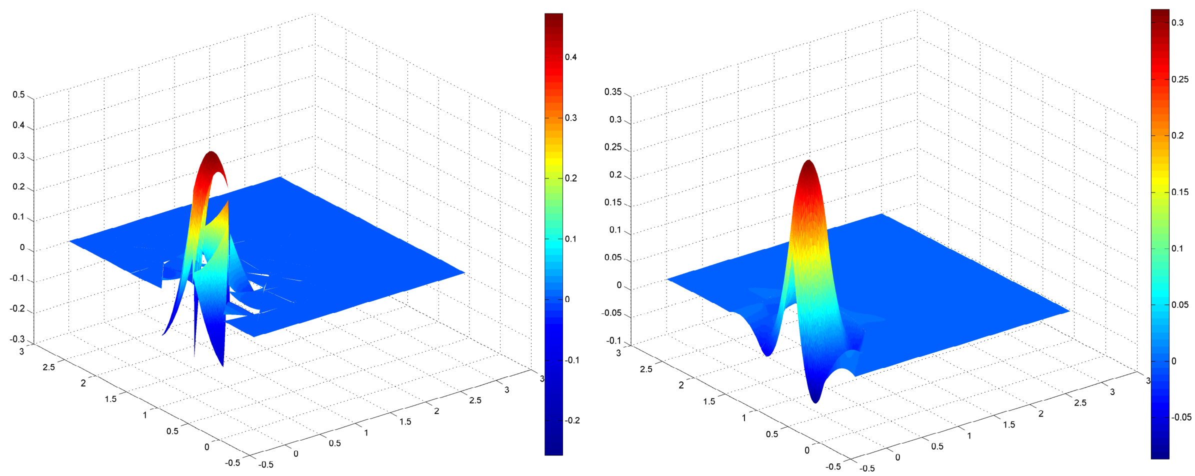

Figure 2.4: $d = 2$, $p = 2$. A function $v_h \in V_{h2}^{\mathrm{nc}}$ and the corresponding potential reconstruction $s_h \in V_{h2}$ from Definition 2.3.1

Theorem 2.4.1 (Approximation of the potential reconstruction of Definition 2.3.1 on V_{hp}^{nc})**.** *Let $v_h \in V_{hp}^{\mathrm{nc}}$ be arbitrary and let $s_h \in V_{hp}$ be given by Definition 2.3.1. Then, for all $s \in H_0^1(\Omega)$, there holds the* local approximation property

$$\|\nabla_h(v_h - s_h)\|_K \leq C\|\nabla_h(v_h - s)\|_{\omega_K} \qquad \forall K \in \mathcal{T}_h, \tag{2.4.1}$$

where the constant C only depends (unfavorably) on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$, on the space dimension d, and on the polynomial degree p. With a constant with same dependencies, there also holds the global approximation property

$$\|\nabla_h(v_h - s_h)\| \leq C\|\nabla_h(v_h - s)\|. \tag{2.4.2}$$

Proof. Each mesh element $K' \in \mathcal{T}_h$ is only included a given number of times in $\mathcal{T}_K$ for some $K \in \mathcal{T}_h$, where this number solely depends on d and $\kappa_{\mathcal{T}_h}$. Thus, upon squaring and summing (2.4.1) over all mesh elements $K \in \mathcal{T}_h$, property (2.4.2) follows. Consequently, we only show (2.4.1). For this purpose, we proceed in two steps. Let $a \lesssim b$ stand for $a \leq cb$, where c only depends on $\kappa_{\mathcal{T}_h}$, d, and p. Fix $v_h \in V_{hp}^{\mathrm{nc}}$ and $K \in \mathcal{T}_h$.

1) We first show that

$$\|\nabla_h(v_h - s_h)\|_K \lesssim \left\{ \sum_{F\in\mathfrak{F}_K} h_F^{-1}\|[\![v_h]\!]\|_F^2 \right\}^{\frac{1}{2}}, \tag{2.4.3}$$

where $\mathfrak{F}_K$ denotes all the faces of the mesh $\mathcal{T}_h$ that share at least a vertex with the simplex K. We start preparing some concepts.

a) Recall that $\mathcal{P}_p(K)$ denotes the space of polynomials on the simplex K of total degree at most p. Its Lagrangian basis is given by functions $\psi^i \in \mathcal{P}_p(K)$ such that

$$\psi^i(\boldsymbol{x}_j) = \delta_{ij}, \tag{2.4.4}$$

where $\boldsymbol{x}_j$ are the corresponding Lagrangian nodes.

b) Since ψ^i is polynomial on K and varies between 0 and 1 in the Lagrangian nodes, $|\nabla\psi^i| \lesssim h_K^{-1}$, so that

$$\|\nabla\psi^i\|_K \lesssim h_K^{-1}|K|^{\frac{1}{2}}, \tag{2.4.5}$$

where $|K|$ is the d-volume of the simplex K (cf. the similar reasoning in (1.6.8)).

c) Let $F \in \mathcal{F}_K \cap \mathcal{F}_h^{\rm int}$ be an interior face shared by the simplex K together with simplex $K' \in \mathcal{T}_h$, with $\boldsymbol{n}_F$ pointing from, say, K to K'. We distinguish four cases. Let i) $\boldsymbol{x}_i$ be a Lagrangian node in K located in the interior of the face F, so that $\mathcal{T}_{\boldsymbol{x}_i} = \{K, K'\}$ and $|\mathcal{T}_{\boldsymbol{x}_i}| = 2$ in (2.3.1a). Then, there holds

$$\begin{aligned}(v_h - s_h)|_K(\boldsymbol{x}_i) &= v_h|_K(\boldsymbol{x}_i) - \tfrac{1}{2}(v_h|_K(\boldsymbol{x}_i) + v_h|_{K'}(\boldsymbol{x}_i))\\ &= \tfrac{1}{2}(v_h|_K(\boldsymbol{x}_i) - v_h|_{K'}(\boldsymbol{x}_i)) = \tfrac{1}{2}[\![v_h]\!]_F(\boldsymbol{x}_i),\end{aligned} \tag{2.4.6}$$

where we have also used Definition 0.6.2 of the jump. For ii) a Lagrangian node $\boldsymbol{x}_i$ in K located at the boundary of the above face F but not on $\partial\Omega$, which in particular includes the mesh interior vertices $\boldsymbol{a}$, (2.3.1a) similarly gives

$$\begin{aligned}(v_h - s_h)|_K(\boldsymbol{x}_i) &= v_h|_K(\boldsymbol{x}_i) - \frac{1}{|\mathcal{T}_{\boldsymbol{x}_i}|}\sum_{K'\in\mathcal{T}_{\boldsymbol{x}_i}} v_h|_{K'}(\boldsymbol{x}_i)\\ &= \frac{1}{|\mathcal{T}_{\boldsymbol{x}_i}|}\sum_{K'\in\mathcal{T}_{\boldsymbol{x}_i}} \big(v_h|_K(\boldsymbol{x}_i) - v_h|_{K'}(\boldsymbol{x}_i)\big)\\ &= \sum_{F'\in\mathcal{F}_i} \alpha_{F',i}[\![v_h]\!]_{F'}(\boldsymbol{x}_i),\end{aligned} \tag{2.4.7}$$

where $\mathcal{F}_i \subset \mathfrak{F}_K$ is a set of mesh faces located in the element patch $\mathcal{T}_K$ specific for $\boldsymbol{x}_i$ and $\alpha_{F,i}$ are some positive coefficients only depending on $|\mathcal{T}_{\boldsymbol{x}_i}|$, which itself only depends on d and $\kappa_{\mathcal{T}_h}$. Remark that formula (2.4.6) is a specific instance of (2.4.7) with $\mathcal{F}_i = \{F\}$ and $\alpha_{F,i} = \frac{1}{2}$. Finally, for iii) a Lagrangian node $\boldsymbol{x}_i$ located in the interior of the element K, (2.3.1a) trivially gives

$$(v_h - s_h)|_K(\boldsymbol{x}_i) = v_h|_K(\boldsymbol{x}_i) - v_h|_K(\boldsymbol{x}_i) = 0, \tag{2.4.8}$$

since there no averaging was performed, whereas for iv) any Lagrangian node $\boldsymbol{x}_i$ in K located on $F \in \mathcal{F}_K$, $F \subset \partial\Omega$, (2.3.1b) gives

$$(v_h - s_h)|_K(\boldsymbol{x}_i) = v_h|_K(\boldsymbol{x}_i) = [\![v_h]\!]_F(\boldsymbol{x}_i) \tag{2.4.9}$$

in view of Definition 0.6.2 of the jump. From (2.4.6)–(2.4.9), we observe that, for any Lagrangian node $\boldsymbol{x}_i$ in K, there are non-negative coefficients $\alpha_{F,i}$ only depending on d and $\kappa_{\mathcal{T}_h}$ such that

$$(v_h - s_h)|_K(\boldsymbol{x}_i) = \sum_{F\in\mathfrak{F}_K} \alpha_{F,i}[\![v_h]\!]_F(\boldsymbol{x}_i). \tag{2.4.10}$$

d) Using (2.4.4) gives

$$(v_h - s_h)|_K = \sum_{i=1}^{|\mathcal{P}_p(K)|} (v_h - s_h)|_K(\boldsymbol{x}_i)\psi^i, \tag{2.4.11}$$

so that

$$\begin{aligned}
\|\nabla(v_h - s_h)\|_K &\overset{(2.4.11)}{=} \left\| \sum_{i=1}^{|\mathcal{P}_p(K)|} (v_h - s_h)|_K(\boldsymbol{x}_i)\nabla\psi^i \right\|_K \\
&\overset{(2.4.10)}{=} \left\| \sum_{i=1}^{|\mathcal{P}_p(K)|} \sum_{F\in\mathfrak{F}_K} \alpha_{F,i}[\![v_h]\!]_F(\boldsymbol{x}_i)\nabla\psi^i \right\|_K \\
&\leq \sum_{i=1}^{|\mathcal{P}_p(K)|} \sum_{F\in\mathfrak{F}_K} \left|\alpha_{F,i}[\![v_h]\!]_F(\boldsymbol{x}_i)\right| \left\|\nabla\psi^i\right\|_K \\
&\overset{(2.4.5)}{\lesssim} h_K^{-1}|K|^{\frac{1}{2}} \sum_{i=1}^{|\mathcal{P}_p(K)|} \sum_{F\in\mathfrak{F}_K} \left|\alpha_{F,i}[\![v_h]\!]_F(\boldsymbol{x}_i)\right| \\
&\lesssim h_K^{-1}|K|^{\frac{1}{2}}|\mathcal{P}_p(K)| \sum_{F\in\mathfrak{F}_K} \|[\![v_h]\!]\|_{\infty,F} \\
&\lesssim h_K^{-1} \sum_{F\in\mathfrak{F}_K} |F|^{\frac{1}{2}} h_F^{\frac{1}{2}} |F|^{-\frac{1}{2}} \|[\![v_h]\!]\|_F \\
&\lesssim \sum_{F\in\mathfrak{F}_K} h_F^{-\frac{1}{2}} \|[\![v_h]\!]\|_F \\
&\lesssim \left\{ \sum_{F\in\mathfrak{F}_K} h_F^{-1} \|[\![v_h]\!]\|_F^2 \right\}^{\frac{1}{2}},
\end{aligned}$$

where we have also used the norm equivalence

$$\|[\![v_h]\!]\|_{\infty,F} \lesssim |F|^{-\frac{1}{2}} \|[\![v_h]\!]\|_F,$$

the mesh shape-regularity consequences

$$|K|^{\frac{1}{2}} \lesssim |F|^{\frac{1}{2}} h_F^{\frac{1}{2}}, \quad h_F \lesssim h_K,$$

the facts that the dimension $|\mathcal{P}_p(K)|$ only depends on p and d and that the number of faces $|\mathfrak{F}_K|$ only depends on d and $\kappa_{\mathcal{T}_h}$, and the Cauchy–Schwarz inequality. This establishes (2.4.3).

2) Let $s \in H_0^1(\Omega)$ be arbitrary. We now crucially use that fact that $[\![s]\!] = 0$ on all faces $F \in \mathcal{F}_h$ by Theorem 0.6.3. Let F be a face from the element patch $\mathcal{T}_K$ not lying on the domain boundary $\partial\Omega$, $F \in \mathfrak{F}_K \cap \mathcal{F}_h^{\mathrm{int}}$. Let be F shared by two mesh elements K_1 and K_2. Then, using the jump definition (0.6.2a) and the trace inequality (0.7.3b), there holds

$$\begin{aligned}
\|[\![v_h]\!]\|_F &= \|[\![v_h - s]\!]\|_F = \|(v_h - s)|_{K_1} - (v_h - s)|_{K_2}\|_F \\
&= \|(v_h - s)|_{K_1} - (v_h - s)_F + (v_h - s)_F - (v_h - s)|_{K_2}\|_F \\
&\leq \sum_{K\in\mathcal{T}_h,\, F\in\mathcal{F}_K} \|(v_h - s)|_K - (v_h - s)_F\|_F \overset{(0.7.3b)}{\lesssim} \sum_{K\in\mathcal{T}_h,\, F\in\mathcal{F}_K} h_K^{\frac{1}{2}} \|\nabla(v_h - s)\|_K,
\end{aligned} \tag{2.4.12}$$

where $(v_h - s)_F$ is the mean value of the function $v_h - s$ on the face F. We recall that $(v_h - s)_F$ is crucially uniquely defined from (1.3.1), independently of whether $(v_h)_F$ means $(v_h|_{K_1})_F$ or

$(v_h|_{K_2})_F$; it is at this place that we use the assumption $v_h \in V_{hp}^{\mathrm{nc}}$ and where we could not proceed with a general $v_h \in \mathcal{P}_p(\mathcal{T}_h)$. Since the inequality (2.4.12) also holds for boundary faces $F \in \mathfrak{F}_K \cap \mathcal{F}_h^{\mathrm{ext}}$, we deduce

$$h_F^{-\frac{1}{2}} \|[\![v_h]\!]\|_F \lesssim \sum_{K \in \mathcal{T}_h,\, F \in \mathcal{F}_K} \|\nabla(v_h - s)\|_K \tag{2.4.13}$$

by the mesh shape-regularity. Now combining (2.4.3) and (2.4.13), (2.4.1) follows. □

2.5 Complements

The potential reconstruction by averaging of Definition 2.3.1 is extremely simple to use in practice and, as per Theorem 2.4.1, is a powerful tool. However, some structural limits can be observed. We discuss some of them here and show possible improvements.

2.5.1 Improved potential reconstruction by averaging

In the eye-ball measure, the reconstruction by averaging of Definition 2.3.1 does not produce a spectacular result in Figure 2.3, which, again in the eye-ball measure, seems to be well improved in Figure 2.5. In particular when to be applied on the weakly-continuos space V_{hp}^{nc}, this more precise alternative is:

Remark 2.5.1 (Potential reconstruction by averaging, preserving the trace moments)**.** *Since $v_h \in V_{hp}^{\mathrm{nc}}$ preserve polynomial moments of the traces up to degree $p-1$, it is interesting to maintain such a property, i.e.,*

$$\langle s_h, q_h \rangle_F = \langle v_h, q_h \rangle_F \qquad \forall q_h \in \mathcal{P}_{p-1}(F),\ \forall F \in \mathcal{F}_h.$$

Example of such a potential reconstruction for $d = 2$ and $p = 1$ is given Figure 2.5. Note, however, that one has to increase the polynomial degree of the reconstruction s_h; in Figure 2.5, one defines $s_h \in V_{h2}$ and not $s_h \in V_{h1}$.

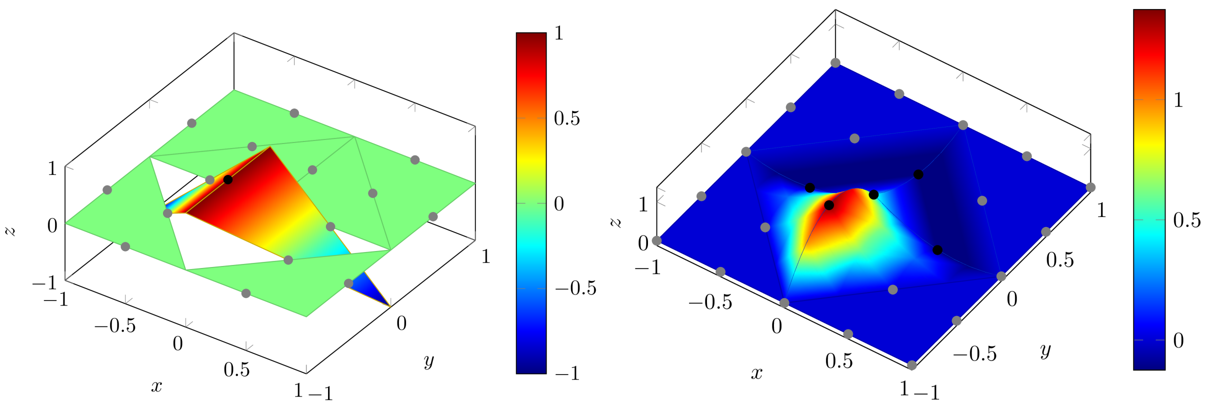

Figure 2.5: $d = 2$, $p = 1$. The function $v_h = \psi^F \in V_{h1}^{\mathrm{nc}}$ from Definition 1.6.1 and the corresponding potential reconstruction $s_h \in V_{h2}$ from Remark 2.5.1

2.5.2 Potential reconstruction by solution of local Dirichlet problems and its approximation properties

The potential reconstruction by averaging of Definition 2.3.1 only allows to bring to the space V_{hp} functions v_h that enable pointwise evaluations in each mesh element. Thus, one cannot apply it to functions v_h from the spaces $H^1_{[\![]\!]}(\mathcal{T}_h)$ from Definition 1.4.1 or $H^1(\mathcal{T}_h)$ from Definition 0.6.1. This comes in concordance with the fact that the generic constant C in Theorem 2.4.1 unfortunately (unfavorably) depends on the polynomial degree p. These two issues are rectified upon proceeding by following [24, 25].

Let $\mathcal{T}_{\boldsymbol{a}}$ be the set of mesh elements that share the given vertex $\boldsymbol{a} \in \mathcal{V}_h$ and $\omega_{\boldsymbol{a}}$ the corresponding subdomain of Ω. Let also $\psi^{\boldsymbol{a}} \in V_{h1}$ be such that $\psi^{\boldsymbol{a}}(\boldsymbol{a}) = 1$ and $\psi^{\boldsymbol{a}}(\boldsymbol{a}') = 0$ for all other vertices $\boldsymbol{a}' \in \mathcal{V}_h$ different from $\boldsymbol{a}$ ($\psi^{\boldsymbol{a}}$ are the Lagrange basis functions of the space V_{h1}, supported on $\omega_{\boldsymbol{a}}$). Let:

Definition 2.5.2 (Potential reconstruction by solution of local Dirichlet problems)**.** *Let $v_h \in H^1(\mathcal{T}_h)$ and a polynomial degree $p' \geq 1$ be given. For each vertex $\boldsymbol{a} \in \mathcal{V}_h$, define $s_h^{\boldsymbol{a}} \in \mathcal{P}_{p'}(\mathcal{T}_{\boldsymbol{a}}) \cap H^1_0(\omega_{\boldsymbol{a}})$ by*

$$(\nabla s_h^{\boldsymbol{a}}, \nabla w_h)_{\omega_{\boldsymbol{a}}} = (\nabla_h(\psi^{\boldsymbol{a}} v_h), \nabla w_h)_{\omega_{\boldsymbol{a}}} \qquad \forall w_h \in \mathcal{P}_{p'}(\mathcal{T}_{\boldsymbol{a}}) \cap H^1_0(\omega_{\boldsymbol{a}}). \tag{2.5.1}$$

Then, extending $s_h^{\boldsymbol{a}}$ by zero outside of the vertex patch subdomain $\omega_{\boldsymbol{a}}$, set

$$s_h := \sum_{\boldsymbol{a} \in \mathcal{V}_h} s_h^{\boldsymbol{a}} \in V_{hp'}. \tag{2.5.2}$$

Remark that the existence and uniqueness of each $s_h^{\boldsymbol{a}}$ in (2.5.1) is straightforward from the Riesz representation theorem; an illustration of the construction of $s_h^{\boldsymbol{a}}$ for $v_h \in \mathcal{P}_1(\mathcal{T}_h)$ and $p' = 2$ in one space dimension is provided in Figure 2.6. The key ideas of the procedure of Definition 2.5.2 are the following:

- One performs a *local minimization* in the patch subdomain $\omega_{\boldsymbol{a}}$ around each vertex $\boldsymbol{a} \in \mathcal{V}_h$ to find the *best-possible* local reconstruction contribution $s_h^{\boldsymbol{a}}$. Indeed, each $s_h^{\boldsymbol{a}}$ of (2.5.1) is equivalently given by

$$s_h^{\boldsymbol{a}} := \arg \min_{w_h \in \mathcal{P}_{p'}(\mathcal{T}_{\boldsymbol{a}}) \cap H^1_0(\omega_{\boldsymbol{a}})} \|\nabla_h(\psi^{\boldsymbol{a}} v_h - w_h)\|_{\omega_{\boldsymbol{a}}}. \tag{2.5.3}$$

  Thus, (2.5.1) are localized version of

$$\tilde{s}_h := \arg \min_{w_h \in \mathcal{P}_{p'}(\mathcal{T}_h) \cap H^1_0(\Omega)} \|\nabla_h(v_h - w_h)\| \tag{2.5.4}$$

  which would give the *best choice* for the potential reconstruction in terms of its distance to v_h. Since (2.5.4) has a practically unacceptable solution price of a global minimization problem, (2.5.1)/(2.5.3) is tempting.

- One applies a *cut-off* by hat basis functions $\psi^{\boldsymbol{a}}$ to the datum v_h, so that $\psi^{\boldsymbol{a}} v_h$ takes the value zero on $\partial\omega_{\boldsymbol{a}}$ for $\boldsymbol{a} \in \mathcal{V}_h^{\mathrm{int}}$. This is compatible with the zero trace condition in $H^1_0(\omega_{\boldsymbol{a}})$ appearing in (2.5.1) and (2.5.3).

- The contributions $s_h^{\boldsymbol{a}}$ satisfy a *homogeneous Dirichlet boundary condition* on $\partial\omega_{\boldsymbol{a}}$, so that they can indeed be combined to $s_h \in V_{hp'} \subset H^1_0(\Omega)$ in (2.5.2).

- Behind the construction, there is the *partition of unity*

$$\sum_{\boldsymbol{a} \in \mathcal{V}_h} \psi^{\boldsymbol{a}} = 1,$$

  which is then crucial for the analysis.

In practice, we can apply Definition 2.5.2 to $v_h \in V_{hp}^{\mathrm{nc}}$ and choose the reconstruction polynomial degree $p' = p + 1$ in order to obtain $s_h \in V_{h(p+1)}$. Then, the following improvement of the dependencies of the generic constant C of Theorem 2.4.1 is achieved; note that the constant C below is *independent* of the polynomial degree p:

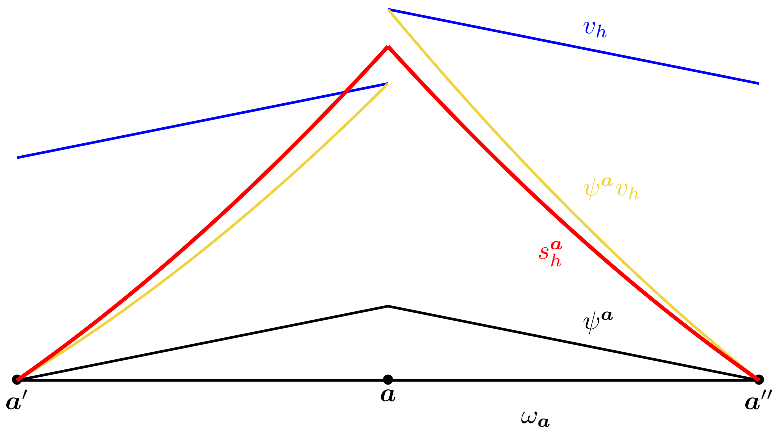

Figure 2.6: $d = 1$, $p = 1$, $p' = 2$. Illustration of the local projection $s_h^{\boldsymbol{a}}$ from Definition 2.5.2

Theorem 2.5.3 (Approximation of the potential reconstruction of Definition 2.5.2 on V_{hp}^{nc})**.** *Let $v_h \in V_{hp}^{\mathrm{nc}}$ be arbitrary and let $s_h \in V_{h(p+1)}$ be given by Definition 2.5.2 for $p' = p + 1$. Then, for all $s \in H_0^1(\Omega)$, there holds the* local approximation property

$$\|\nabla_h(v_h - s_h)\|_K \leq C\|\nabla_h(v_h - s)\|_{\omega_K} \qquad \forall K \in \mathcal{T}_h, \tag{2.5.5}$$

where the constant C only depends on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$ and on the space dimension d when $1 \leq d \leq 3$. With a constant with same dependencies, there also holds the global approximation property

$$\|\nabla_h(v_h - s_h)\| \leq C\|\nabla_h(v_h - s)\|. \tag{2.5.6}$$

Chapter 3

Flux reconstruction in $\boldsymbol{H}(\mathrm{div}, \Omega)$

In this short chapter, we show that the noncoforming finite element method allows for a straightforward elementwise reconstruction a *normal-trace continuous* vector-valued piecewise polynomial lying in the $\boldsymbol{H}(\mathrm{div}, \Omega)$ space introduced in Section 0.4. This will turn as in important tool to obtain a posteriori error estimates in Chapter 4 below. This remarkable property was probably first noted by Marini [30]. We consider u_h of (1.3.2) for simplicity for the polynomial degree $p = 1$. Again for simplicity, let the source term f be piecewise constant,

$$f \in \mathcal{P}_0(\mathcal{T}_h). \tag{3.0.1}$$

3.1 Setting and definition

For a simplex $K \in \mathcal{T}_h$, let $\boldsymbol{x}_K$ denote its barycenter and define

$$\boldsymbol{f}_h|_K(\boldsymbol{x}) := \frac{f|_K}{d}(\boldsymbol{x} - \boldsymbol{x}_K) \qquad \forall K \in \mathcal{T}_h. \tag{3.1.1}$$

Note that, on each mesh element $K \in \mathcal{T}_h$, this is an affine vector-valued function only depending on the barycenter $\boldsymbol{x}_K$, the constant value $f|_K$, and the space dimension d.

Then we can state:

Definition 3.1.1 (Elementwise flux prescription for the nonconforming finite element method)**.** *Let u_h be given by* (1.3.2) *for $p = 1$, with f satisfying* (3.0.1)*. Then prescribe*

$$\boldsymbol{\sigma}_h|_K := -\nabla_h u_h|_K + \boldsymbol{f}_h|_K \qquad \forall K \in \mathcal{T}_h. \tag{3.1.2}$$

3.2 Properties

The function $\boldsymbol{\sigma}_h$ from Definition 3.1.1 satisfies:

Theorem 3.2.1 (Equilibrated flux reconstruction)**.** *Let $\boldsymbol{\sigma}_h$ be given by Definition 3.1.1. Then it is an* equilibrated flux reconstruction *in that*

$$\boldsymbol{\sigma}_h \in \boldsymbol{H}(\mathrm{div}, \Omega), \tag{3.2.1a}$$

$$\nabla{\cdot}\boldsymbol{\sigma}_h = f. \tag{3.2.1b}$$

More precisely, $\boldsymbol{\sigma}_h \in \boldsymbol{V}_{h0}$, where the vector-valued piecewise polynomial space $\boldsymbol{V}_{h0}$ is introduced in (1.9.9) *above.*

Proof. (3.2.1b) is obvious, taking into account that $-\nabla_h u_h$ is piecewise constant and thus its elementwise divergence is zero, whereas the piecewise divergence of $\boldsymbol{f}_h$ prescribed by (3.1.1) is precisely f (supposing f piecewise constant as per (3.0.1)). So we are left with verifying that $\boldsymbol{\sigma}_h$

given by (3.1.2) belongs to the space $\boldsymbol{H}(\mathrm{div},\Omega)$. As discussed in Section 0.4, we need to show that

$$[\![\boldsymbol{\sigma}_h\cdot\boldsymbol{n}_F]\!] = 0$$

for all interior mesh faces $F\in\mathcal{F}_h^{\mathrm{int}}$.

Let $F\in\mathcal{F}_h^{\mathrm{int}}$ be given, and let K_1 and K_2 be the two elements that share the face F. Now consider the test function $v_h=\psi^F$ in (1.3.2), where ψ^F is the basis function associated with the face F of Definition 1.6.1. Taking into account that ψ^F is only supported on $K_1\cup K_2$, that $\nabla_h u_h$ is piecewise constant (so that its piecewise divergence vanishes and $\nabla_h u_h\cdot\boldsymbol{n}_{K_i}$ is constant on each face), the properties of the basis function ψ^F (recall Figure 1.2), and the Green theorem gives

$$\begin{aligned}(\nabla_h u_h,\nabla_h\psi^F)_{K_1\cup K_2} &= \langle\nabla_h u_h\cdot\boldsymbol{n}_{K_1},\psi^F\rangle_{\partial K_1}+\langle\nabla_h u_h\cdot\boldsymbol{n}_{K_2},\psi^F\rangle_{\partial K_2}\\ &= \langle\nabla_h u_h\cdot\boldsymbol{n}_{K_1},1\rangle_F+\langle\nabla_h u_h\cdot\boldsymbol{n}_{K_2},1\rangle_F.\end{aligned}$$

Note the simple property of the barycentre $\boldsymbol{x}_K$ of the simplex $K\in\mathcal{T}_h$

$$(\boldsymbol{x},\nabla_h\psi^F)_K=(\boldsymbol{x}_K,\nabla_h\psi^F)_K,$$

which implies

$$(\boldsymbol{f}_h,\nabla_h\psi^F)_K=0$$

for all $K\in\mathcal{T}_h$, and recall that $\nabla\cdot\boldsymbol{f}_h|_K=f|_K$ on all $K\in\mathcal{T}_h$. Thus, the Green theorem, the fact that the normal component $\boldsymbol{f}_h\cdot\boldsymbol{n}_{K_i}$ is facewise constant (please verify!), and the properties of ψ^F imply

$$\begin{aligned}(f,\psi^F)_{K_1\cup K_2} &= (\boldsymbol{f}_h,\nabla_h\psi^F)_{K_1}+(\boldsymbol{f}_h,\nabla_h\psi^F)_{K_2}+(\nabla\cdot\boldsymbol{f}_h,\psi^F)_{K_1}+(\nabla\cdot\boldsymbol{f}_h,\psi^F)_{K_2}\\ &= \langle\boldsymbol{f}_h\cdot\boldsymbol{n}_{K_1},\psi^F\rangle_{\partial K_1}+\langle\boldsymbol{f}_h\cdot\boldsymbol{n}_{K_2},\psi^F\rangle_{\partial K_2}\\ &= \langle\boldsymbol{f}_h\cdot\boldsymbol{n}_{K_1},1\rangle_F+\langle\boldsymbol{f}_h\cdot\boldsymbol{n}_{K_2},1\rangle_F.\end{aligned}$$

The assertion follows by combining the two above identities with (1.3.2) for $v_h=\psi^F$ and the definition (3.1.2) of $\boldsymbol{\sigma}_h$. □

Chapter 4

The nonconforming finite element method: a posteriori analysis

Theorems 1.8.2 and 1.9.8 ensure that for a fixed polynomial degree p and a *piecewise smooth* weak solution u, $u|_K \in H^{p+1}(K)$ for all $K \in \mathcal{T}_h$, the error by the nonconforming finite element approximation $u_h \in V_{hp}^{\mathrm{nc}}$ of Definition 1.3.2 decreases as h^p. These bounds *justify* the nonconforming finite element method in that a higher effort (finer mesh with smaller mesh size h or a higher polynomial degree p) will make u_h approach u. Unfortunately, these bounds but do not allow to assess the *actual distance* of u_h to u measured as $\|\nabla_h(u-u_h)\|$.

This is the subject of the present chapter, where we will in particular design a procedure of obtaining a quantity $\eta(u_h)$, *fully* and *locally computable* from u_h, which structurally gives

$$\|\nabla_h(u-u_h)\| \le \eta(u_h), \quad \eta(u_h) \le C\|\nabla_h(u-u_h)\|, \tag{4.0.1}$$

where C is a generic constant only depending (unfavorably) on the shape-regularity parameter $\kappa_{\mathcal{T}_h}$, the space dimension d, and (typically) on the polynomial degree p (actually we will show how to avoid this last dependency at least for $1 \le d \le 3$). The quantity $\eta(u_h)$ is called an *a posteriori error estimate*, since it can only be obtained a posteriori, once u_h has been computed. This is another difference with the *a priori error estimates* from Theorems 1.8.2 and 1.9.8 which can be obtained a priori, before u_h has been computed. We will in this chapter mainly follow the expositions in Prager and Synge [34], Destuynder and Métivet, [16, 17], Karakashian and Pascal [27], Repin [37], Braess and Schöberl [7], and [24, 25], with some technical tools from Verfürth [42].

4.1 Setting

We adopt here the setting of Chapter 1. We consider the Poisson equation with a homogeneous Dirichlet boundary condition, where, for $f \in L^2(\Omega)$, we look for $u : \Omega \to \mathbb{R}$ such that

$$-\Delta u = f \qquad \text{in } \Omega, \tag{4.1.1a}$$
$$u = 0 \qquad \text{on } \partial\Omega; \tag{4.1.1b}$$

recall that the Laplace differential operator Δ is defined by (1.1.2). The weak formulation of problem (4.1.1) consists in finding $u \in H_0^1(\Omega)$ such that

$$(\nabla u, \nabla v) = (f, v) \qquad \forall v \in H_0^1(\Omega). \tag{4.1.2}$$

Relying on Definition 1.3.1, the nonconforming finite element method for problem (4.1.1) looks for $u_h \in V_{hp}^{\mathrm{nc}}$ such that

$$(\nabla_h u_h, \nabla_h v_h) = (f, v_h) \qquad \forall v_h \in V_{hp}^{\mathrm{nc}}. \tag{4.1.3}$$

4.2 Error characterization

To start with, we have the following characterization of the error which reveals that $\|\nabla_h(u-u_h)\|$ is given by the *distance* of u_h to the *correct space* for the *primal variable* (potential) u, which is $H_0^1(\Omega)$, plus the *distance* of $\nabla_h u_h$ to the *correct space* for the *dual variable* (flux) $-\nabla u$, which is $\boldsymbol{H}(\mathrm{div}, \Omega)$, subject to a divergence constraint (recall at this occasion Definitions 0.3.4 and 0.4.2).

Theorem 4.2.1 (Error characterization). *Let $u \in H_0^1(\Omega)$ be the weak solution of (4.1.2) and let $u_h \in H^1(\mathcal{T}_h)$ be arbitrary. Then*

$$\|\nabla_h(u-u_h)\|^2 = \min_{\substack{\boldsymbol{v}\in \boldsymbol{H}(\mathrm{div},\Omega)\\ \nabla\cdot\boldsymbol{v}=f}} \|\nabla_h u_h + \boldsymbol{v}\|^2 + \min_{v\in H_0^1(\Omega)} \|\nabla_h(u_h - v)\|^2. \tag{4.2.1}$$

Proof. Let us define a function $s \in H_0^1(\Omega)$ by

$$(\nabla s, \nabla v) = (\nabla_h u_h, \nabla v) \qquad \forall v \in H_0^1(\Omega). \tag{4.2.2}$$

There exists one and only one s by the Riesz representation theorem. The function s is the orthogonal projection of the approximate solution u_h (when seen as $u_h \in H^1_{[\![]\!]}(\mathcal{T}_h)$) onto the space $H_0^1(\Omega)$ with respect to the scalar product $(\nabla_h \cdot, \nabla_h \cdot)$; recall in this respect from Theorem 1.4.6 that $H^1_{[\![]\!]}(\mathcal{T}_h)$ is a Hilbert space for this scalar product. With the aid of s, we can thus write the Pythagorean equality

$$\|\nabla_h(u-u_h)\|^2 = \|\nabla(u-s)\|^2 + \|\nabla_h(s-u_h)\|^2. \tag{4.2.3}$$

To see this shortly,

$$\|\nabla_h(u-u_h)\|^2 = \|\nabla_h(u-s+s-u_h)\|^2 = \|\nabla(u-s)\|^2 + \|\nabla_h(s-u_h)\|^2 + 2(\nabla(u-s), \nabla_h(s-u_h)),$$

and the last term in the above expression vanishes in view of the orthogonality (4.2.2), since $u-s$ can be taken as a test function $v \in H_0^1(\Omega)$ in (4.2.2). We now continue in two steps.

1) Since s is a projection of u_h,

$$\|\nabla_h(s-u_h)\|^2 = \min_{v\in H_0^1(\Omega)} \|\nabla_h(v-u_h)\|^2. \tag{4.2.4}$$

This can again be proven directly from (4.2.3) used for any function $v \in H_0^1(\Omega)$ in place of $u \in H_0^1(\Omega)$,

$$\|\nabla_h(v-u_h)\|^2 = \|\nabla(v-s)\|^2 + \|\nabla_h(s-u_h)\|^2,$$

from where we get

$$\|\nabla_h(s-u_h)\|^2 = \|\nabla_h(v-u_h)\|^2 - \|\nabla(v-s)\|^2 \le \|\nabla_h(v-u_h)\|^2 \qquad \forall v \in H_0^1(\Omega).$$

This handles the second term in (4.2.3) in the form needed in (4.2.1).

2) As for the first term in (4.2.3), we first notice that $u - s \in H_0^1(\Omega)$. Thus, (4.2.2) gives

$$\|\nabla(u-s)\| = \max_{\substack{\varphi\in H_0^1(\Omega)\\ \|\nabla\varphi\|=1}} (\nabla(u-s), \nabla\varphi) = \max_{\substack{\varphi\in H_0^1(\Omega)\\ \|\nabla\varphi\|=1}} (\nabla_h(u-u_h), \nabla\varphi). \tag{4.2.5}$$

Let now $\varphi \in H_0^1(\Omega)$ with $\|\nabla\varphi\| = 1$ be fixed. Using the characterization (4.1.2) of the weak solution, we have

$$(\nabla_h(u-u_h), \nabla\varphi) = (f, \varphi) - (\nabla_h u_h, \nabla\varphi). \tag{4.2.6}$$

Finally, for an arbitrary $\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ such that $\nabla\cdot\boldsymbol{v} = f$, the Green theorem gives

$$(f, \varphi) - (\nabla_h u_h, \nabla\varphi) = (\nabla\cdot\boldsymbol{v}, \varphi) - (\nabla_h u_h, \nabla\varphi) = -(\nabla_h u_h + \boldsymbol{v}, \nabla\varphi).$$

Consequently, by the Cauchy–Schwarz inequality,

$$\|\nabla(u-s)\| \le \min_{\substack{\boldsymbol{v}\in \boldsymbol{H}(\mathrm{div},\Omega)\\ \nabla\cdot\boldsymbol{v}=f}} \|\nabla u_h + \boldsymbol{v}\|. \tag{4.2.7}$$

In the rest of the proof, we show that actually

$$\|\nabla(u-s)\| = \min_{\substack{\boldsymbol{v}\in \boldsymbol{H}(\mathrm{div},\Omega)\\ \nabla\cdot\boldsymbol{v}=f}} \|\nabla_h u_h + \boldsymbol{v}\|, \tag{4.2.8}$$

which handles the first term in (4.2.3) in the form needed in (4.2.1).

The argument of the minimum in (4.2.7) is

$$\boldsymbol{\sigma} := \arg \min_{\substack{\boldsymbol{v}\in \boldsymbol{H}(\mathrm{div},\Omega)\\ \nabla\cdot\boldsymbol{v}=f}} \|\nabla_h u_h + \boldsymbol{v}\|$$

and is characterized by the Euler–Lagrange conditions as a function $\boldsymbol{\sigma} \in \boldsymbol{H}(\mathrm{div},\Omega)$ with $\nabla\cdot\boldsymbol{\sigma} = f$ such that

$$(\boldsymbol{\sigma},\boldsymbol{v}) = -(\nabla_h u_h, \boldsymbol{v}) \qquad \forall \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div},\Omega) \text{ with } \nabla\cdot\boldsymbol{v} = 0.$$

This problem is in turn equivalent to finding $\boldsymbol{\sigma} \in \boldsymbol{H}(\mathrm{div},\Omega)$ and $r \in L^2(\Omega)$ such that

$$(\boldsymbol{\sigma},\boldsymbol{v}) - (r, \nabla\cdot\boldsymbol{v}) = -(\nabla_h u_h, \boldsymbol{v}) \qquad \forall \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div},\Omega), \tag{4.2.9}$$

$$(\nabla\cdot\boldsymbol{\sigma}, q) = (f, q) \qquad \forall q \in L^2(\Omega). \tag{4.2.10}$$

Now, (4.2.9) implies, see, e.g., [45, Theorem 6.3.1], that $r \in H_0^1(\Omega)$ with $\nabla r = -\boldsymbol{\sigma} - \nabla_h u_h$. Consequently, by the Green theorem,

$$\begin{aligned} \min_{\substack{\boldsymbol{v}\in \boldsymbol{H}(\mathrm{div},\Omega)\\ \nabla\cdot\boldsymbol{v}=f}} \|\nabla_h u_h + \boldsymbol{v}\| &= \|\nabla_h u_h + \boldsymbol{\sigma}\| = \|\nabla r\| = \max_{\substack{\varphi\in H_0^1(\Omega)\\ \|\nabla\varphi\|=1}} (\nabla r, \nabla\varphi) \\ &= \max_{\substack{\varphi\in H_0^1(\Omega)\\ \|\nabla\varphi\|=1}} (-\boldsymbol{\sigma} - \nabla_h u_h, \nabla\varphi) = \max_{\substack{\varphi\in H_0^1(\Omega)\\ \|\nabla\varphi\|=1}} \{(f,\varphi) - (\nabla_h u_h, \nabla\varphi)\}, \end{aligned} \tag{4.2.11}$$

and (4.2.5)–(4.2.6) show that (4.2.8) holds true. Thus (4.2.3), (4.2.4), and (4.2.8) imply the claim (4.2.1). □

The heart of the above result can be traced back to at least Prager and Synge [34] who showed the following result. Here u_h is supposed conforming, $u_h \in H_0^1(\Omega)$.

Theorem 4.2.2 (Prager–Synge equality)**.** *Let $u \in H_0^1(\Omega)$ be the weak solution of (4.1.2) and let $u_h \in H_0^1(\Omega)$ and $\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div},\Omega)$ with $\nabla\cdot\boldsymbol{v} = f$ be arbitrary. Then*

$$\|\nabla(u-u_h)\|^2 + \|\nabla u + \boldsymbol{v}\|^2 = \|\nabla u_h + \boldsymbol{v}\|^2. \tag{4.2.12}$$

Proof. Adding and subtracting ∇u, we develop

$$\|\nabla u_h + \boldsymbol{v}\|^2 = \|\nabla(u_h - u)\|^2 + \|\nabla u + \boldsymbol{v}\|^2 + 2(\nabla(u_h - u), \nabla u + \boldsymbol{v}).$$

It follows from (4.1.2) that $\nabla u \in \boldsymbol{H}(\mathrm{div},\Omega)$ with $\nabla\cdot(\nabla u) = -f$, see the weak divergence Definition 0.4.1. Thus $(\nabla u + \boldsymbol{v}) \in \boldsymbol{H}(\mathrm{div},\Omega)$, and, importantly, $\nabla\cdot(\nabla u + \boldsymbol{v}) = 0$. Consequently, using that $u_h - u \in H_0^1(\Omega)$, the Green theorem gives

$$(\nabla u + \boldsymbol{v}, \nabla(u_h - u)) = -(\nabla\cdot(\nabla u + \boldsymbol{v}), u_h - u) = 0,$$

whence the assertion follows. □

4.3 A posteriori error estimate for the nonconforming finite element method, $p = 1$

In order to arrive at our first goal in (4.0.1), we can use the characterization of Theorem 4.2.1 which gives

$$\|\nabla_h(u - u_h)\|^2 \le \|\nabla_h u_h + \boldsymbol{\sigma}_h\|^2 + \|\nabla_h(u_h - s_h)\|^2$$

for any

$$s_h \in H_0^1(\Omega) \quad \text{and} \quad \boldsymbol{\sigma}_h \in \boldsymbol{H}(\mathrm{div}, \Omega) \text{ such that } \nabla\cdot\boldsymbol{\sigma}_h = f.$$

For nonconforming finite elements with $p = 1$, suitable s_h and $\boldsymbol{\sigma}_h$ are:

Theorem 4.3.1 (A posteriori error estimate for the nonconforming finite element method, $p = 1$). *Let $f \in \mathcal{P}_0(\mathcal{T}_h)$ for simplicity, let $u \in H_0^1(\Omega)$ be the weak solution of (4.1.2), and let $u_h \in V_{hp}^{\mathrm{nc}}$ from (4.1.3) be its nonconforming finite element approximation for $p = 1$. Let $s_h \in V_{h1}$ be the potential reconstruction given by Definition 2.3.1 for $v_h = u_h$, and let $\boldsymbol{\sigma}_h \in \boldsymbol{V}_{h0} \subset \boldsymbol{H}(\mathrm{div}, \Omega)$ be the flux reconstruction given by Definition 3.1.1. Then*

$$\|\nabla_h(u - u_h)\|^2 \le \underbrace{\sum_{K \in \mathcal{T}_h} \left(\|\nabla_h u_h + \boldsymbol{\sigma}_h\|_K^2 + \|\nabla_h(u_h - s_h)\|_K^2 \right)}_{\eta^2(u_h)}. \tag{4.3.1}$$

Proof. Immediate from Theorem 4.2.1, relying on Theorem 3.2.1. □

A few remarks are in order:

- Define
$$\eta_K^2(u_h) := \|\nabla_h u_h + \boldsymbol{\sigma}_h\|_K^2 + \|\nabla_h(u_h - s_h)\|_K^2. \tag{4.3.2}$$
This is often called an *element estimator*. It is remarkably calculable from the nonconforming finite element approximation u_h, locally on the element K and in the elements sharing a vertex with K.

- The estimator η_K is extremely cheap to compute: actually
$$\|\nabla_h u_h + \boldsymbol{\sigma}_h\|_K = \|\boldsymbol{f}_h\|_K \tag{4.3.3}$$
from (3.1.2), where, recall, $\boldsymbol{f}_h$ is given by (3.1.1).

- There is no hidden unknown constant: (4.3.1) allows to *control the error* and thus assess the quality of the numerical solution.

The second goal in (4.0.1) aims at ensuring the quality of the a posteriori error estimate $\eta(u_h)$: whatever happens (nice or ugly domain Ω, fine or coarse mesh $\mathcal{T}_h$, regular or singular solution u), $\eta(u_h)$ needs to behave as the error $\|\nabla_h(u - u_h)\|$. Only then the known a posteriori error *estimate* $\eta(u_h)$ is sound, mathematically *equivalent* to the unknown *error*. Actually, better than the second property in (4.0.1), we even achieve a *local* bound on ω_K, the subdomain corresponding to $\mathcal{T}_K$, all mesh elements that share at least a vertex with $K \in \mathcal{T}_h$. The following is a useful result in this direction, following Verfürth [42]:

Lemma 4.3.2 (Element residual). *For $K \in \mathcal{T}_h$, let $u_h \in \mathcal{P}_p(K)$ and $f \in \mathcal{P}_p(K)$. Then*

$$h_K \|f + \Delta u_h\|_K \le C \|\nabla_h(u - u_h)\|_K, \tag{4.3.4}$$

where the constant C only depends on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$, on the space dimension d, and on the polynomial degree p.

Proof. Let $a \lesssim b$ stand for $a \le cb$, where c only depends on $\kappa_{\mathcal{T}_h}$, d, and p. Set

$$v_K := (f + \Delta u_h)|_K. \tag{4.3.5}$$

Let ψ_K be the bubble function on K given by the product of the $d+1$ Lagrange hat basis functions $\psi^{\boldsymbol{a}}$, $\boldsymbol{a} \in \mathcal{V}_K$, on the element K. Note that $\psi_K|_{\partial K} = 0$. Also note that both v_K and ψ_K are polynomials. By equivalence of norms on finite-dimensional spaces, there consequently holds

$$(v_K, v_K)_K \lesssim (v_K, \psi_K v_K)_K. \tag{4.3.6}$$

Using the inverse inequality (cf. Quarteroni and Valli [35, Proposition 6.3.2]), we obtain

$$h_K \|\nabla(\psi_K v_K)\|_K \lesssim \|\psi_K v_K\|_K. \tag{4.3.7}$$

Finally, from the definition of the bubble function ψ_K, there holds

$$\|\psi_K v_K\|_K \le \|\psi_K\|_{\infty,K} \|v_K\|_K \le \|v_K\|_K. \tag{4.3.8}$$

Thus, using these properties and noting that $\psi_K v_K \in H_0^1(K)$, together with (4.1.2), the Green theorem, and the Cauchy–Schwarz inequality, we see

$$\begin{aligned}
\|v_K\|^2 &\overset{(4.3.6)}{\lesssim} (v_K, \psi_K v_K)_K \overset{(4.3.5)}{=} (f + \Delta u_h, \psi_K v_K)_K \overset{\substack{(4.1.2)\\ \text{Green}}}{=} (\nabla(u - u_h), \nabla(\psi_K v_K))_K \\
&\le \|\nabla(u - u_h)\|_K \|\nabla(\psi_K v_K)\|_K \overset{\substack{(4.3.7)\\(4.3.8)}}{\lesssim} \|\nabla(u - u_h)\|_K h_K^{-1} \|v_K\|_K.
\end{aligned}$$

Therefrom, the assertion of the lemma follows. □

We are now ready to assess the quality of the estimates of Theorem 4.3.1:

Theorem 4.3.3 (Efficiency of the a posteriori error estimate for the nonconforming finite element method, $p = 1$)**.** *Let the assumptions of Theorem 4.3.1 be verified. Then, with the notation* (4.3.2)*, the* local efficiency *holds,*

$$\eta_K(u_h) \le C \|\nabla_h(u - u_h)\|_{\omega_K}, \tag{4.3.9}$$

where the constant C only depends on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$ and on the space dimension d. With a constant with same dependencies, there also holds the global efficiency

$$\eta(u_h) \le C \|\nabla_h(u - u_h)\|. \tag{4.3.10}$$

Proof. The property (4.3.10) is clearly a consequence of (4.3.9), just as in the proof of Theorem 2.4.1. Actually, choosing $v_h = u_h$ and $s = u$ in Theorem 2.4.1, we immediately have

$$\|\nabla_h(u_h - s_h)\|_K \le C \|\nabla_h(u - u_h)\|_{\omega_K}.$$

For the other term in (4.3.2), we use (4.3.3), which together with the definition (3.1.1) of $\boldsymbol{f}_h$ gives

$$\|\nabla_h u_h + \boldsymbol{\sigma}_h\|_K = \|\boldsymbol{f}_h\|_K = \left\| \frac{f|_K}{d} (\boldsymbol{x} - \boldsymbol{x}_K) \right\|_K \le \frac{h_K}{d} \|f + \Delta u_h\|_K;$$

here, we have used the fact that, since $p = 1$, $\Delta(u_h|_K) = 0$, as well as the estimate $|\boldsymbol{x} - \boldsymbol{x}_K| \le h_K$. Thus, Lemma 4.3.2 allows us to conclude

$$\|\nabla_h u_h + \boldsymbol{\sigma}_h\|_K \le C \|\nabla_h(u - u_h)\|_K,$$

and consequently (4.3.9) follows. □

4.4 Numerical illustration: a posteriori error localization and control

Consider a problem of the form (4.1.1), featuring additionally a diffusion coefficient $\boldsymbol{S}$:

$$-\nabla\cdot(\boldsymbol{S}\nabla u) = 0 \quad \text{in} \quad \Omega = (-1,1)\times(-1,1). \tag{4.4.1}$$

Let more precisely $\boldsymbol{S}|_{\Omega_i} = s_i \boldsymbol{I}$ on the four subdomains Ω_i of Ω, as illustrated in the two settings of Figure 4.1. Here $\boldsymbol{I}$ is the identity matrix and s_i are positive constants.

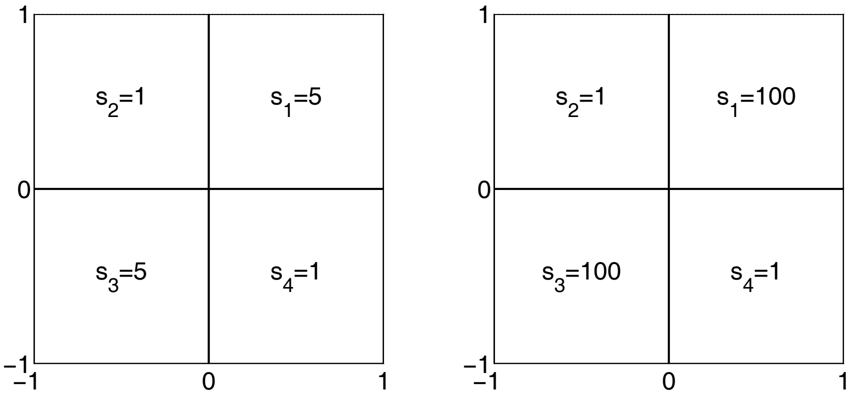

Figure 4.1: Subdomains Ω_i, $1 \le i \le 4$, and the diffusion tensor $\boldsymbol{S}$

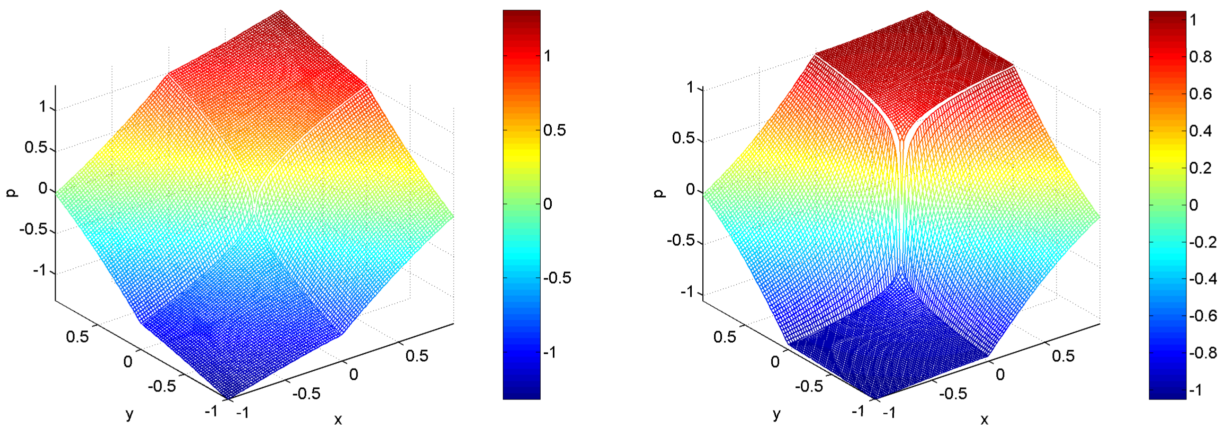

Figure 4.2: Exact solutions of the from (4.4.2) for $\alpha \approx 0.54$ (left) and $\alpha \approx 0.13$ (right)

The considered setting leads to exact solutions featuring the singularity at the origin $(0,0)$ which take the form, following Kellogg [28],

$$u(r,\theta) = r^{\alpha}(a_i \sin(\alpha\theta) + b_i \cos(\alpha\theta)), \tag{4.4.2}$$

where

- (r,θ) are the polar coordinates in Ω,
- a_i, b_i are constants depending on Ω_i,
- α is the regularity of the exact solution; $u \in H^{1+\alpha}(\Omega)$, where $\alpha \approx 0.54$ for the left setting of Figure 4.1 and $\alpha \approx 0.13$ for the right setting of Figure 4.1.

These exact solutions for the two settings of Figure 4.1 are depicted in Figure 4.2. We also use them to impose an inhomogeneous Dirichlet boundary condition that complements (4.4.1).

We now apply the nonconforming finite element method with the polynomial degree $p = 1$ to problem (4.4.1) and evaluate the a posteriori error estimators $\eta_K(u_h)$ following Theorem 4.3.1. The forthcoming results are taken from [44]. The estimators $\eta_K(u_h)$ are depicted in the left part of Figure 4.3, whereas the errors $\|\nabla(u-u_h)\|_K$ (using the knowledge of the exact solution u given by (4.4.2) in this model test case) ($\alpha \approx 0.54$) are depicted in the right part of Figure 4.3. We clearly see that the estimators $\eta_K(u_h)$ allow to predict where the *error is localized*, without the knowledge of the exact solution u. This is the practical consequence of the local efficiency (4.3.9): if we predict the error to be of some size in some element, then it is the case (possibly also in the neighborhood of the element).

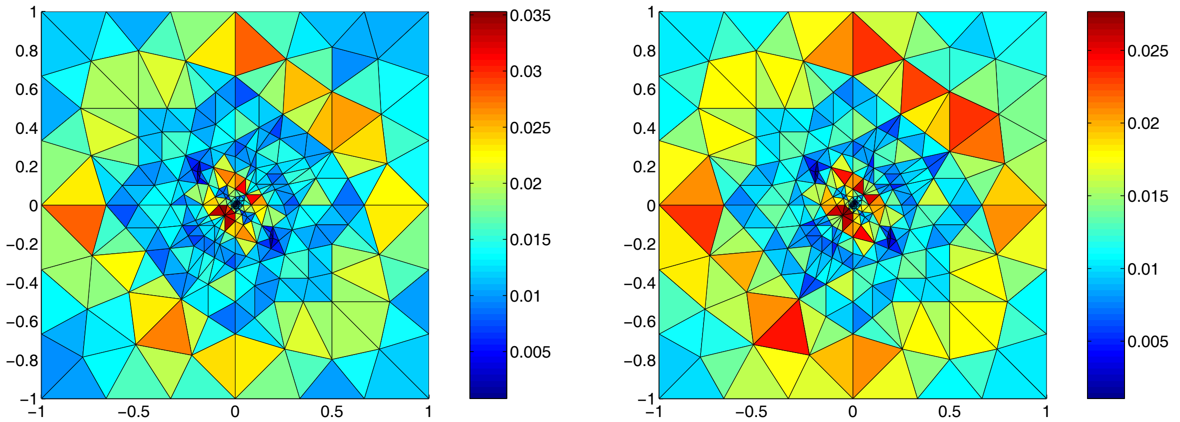

Figure 4.3: Estimated error distribution $\eta_K(u_h)$ (left) and exact error distribution $\|\nabla(u-u_h)\|_K$ (right), $\alpha \approx 0.54$

We finally assess the overall quality of the a posteriori error estimators of Theorem 4.3.1. This is best done in terms of the *effectivity index*

$$I_{\text{eff}} := \frac{\eta(u_h)}{\|\nabla_h(u-u_h)\|}.$$

From (4.3.1), we know that $I_{\text{eff}} \geq 1$, and from (4.3.10) of Theorem 4.3.3, we know that I_{eff} is bounded from above by a generic constant. Figure 4.4 shows that I_{eff} is actually quite close to the optimal value of 1.

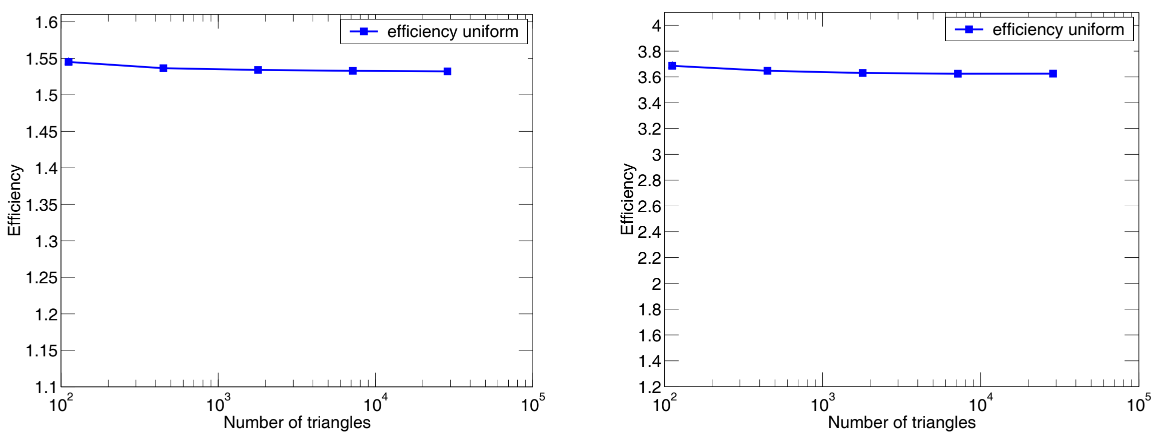

Figure 4.4: Effectivity indices, $\alpha \approx 0.54$ (left) and $\alpha \approx 0.13$ (right)

4.5 Complements

We collect here some additional material, developing further the above contents.

4.5.1 An abstract a posteriori error estimate

Theorem 4.2.1 is the basis for the following more general result:

Theorem 4.5.1 (Abstract a posteriori error estimate)**.** *Let $u \in H_0^1(\Omega)$ be the weak solution of (4.1.2) and let $u_h \in H^1(\mathcal{T}_h)$ be arbitrary. Let*

$$s_h \in H_0^1(\Omega) \quad \textit{and} \quad \boldsymbol{\sigma}_h \in \boldsymbol{H}(\mathrm{div}, \Omega) \textit{ such that } (\nabla\cdot\boldsymbol{\sigma}_h, 1)_K = (f, 1)_K \qquad \forall K \in \mathcal{T}_h \tag{4.5.1}$$

be arbitrary. Then

$$\|\nabla_h(u - u_h)\|^2 \le \sum_{K\in\mathcal{T}_h} \Big(\|\nabla_h u_h + \boldsymbol{\sigma}_h\|_K + \frac{h_K}{\pi}\|f - \nabla\cdot\boldsymbol{\sigma}_h\|_K\Big)^2 + \sum_{K\in\mathcal{T}_h} \|\nabla_h(u_h - s_h)\|_K^2. \tag{4.5.2}$$

Proof. We build on (4.2.1). The second term is trivially bounded by $\|\nabla_h(u_h - s_h)\|^2$, so we are left with the first one. From (4.2.11), we need to work with

$$(f, \varphi) - (\nabla_h u_h, \nabla\varphi)$$

for a given $\varphi \in H_0^1(\Omega)$ with $\|\nabla\varphi\| = 1$. Adding and subtracting $(\boldsymbol{\sigma}_h, \nabla\varphi)$ and using the Green theorem, we infer

$$(f, \varphi) - (\nabla_h u_h, \nabla\varphi) = (f - \nabla\cdot\boldsymbol{\sigma}_h, \varphi) - (\nabla_h u_h + \boldsymbol{\sigma}_h, \nabla\varphi).$$

The Cauchy–Schwarz inequality gives for the second term above

$$-(\nabla_h u_h + \boldsymbol{\sigma}_h, \nabla\varphi) = -\sum_{K\in\mathcal{T}_h} (\nabla_h u_h + \boldsymbol{\sigma}_h, \nabla\varphi)_K \le \sum_{K\in\mathcal{T}_h} \|\nabla_h u_h + \boldsymbol{\sigma}_h\|_K \|\nabla\varphi\|_K,$$

whereas the approximate equilibrium property (4.5.1), the Poincaré inequality (0.7.1) applied on convex mesh elements K, and the Cauchy–Schwarz inequality give for the first term above

$$\begin{aligned}(f - \nabla\cdot\boldsymbol{\sigma}_h, \varphi) &= \sum_{K\in\mathcal{T}_h} (f - \nabla\cdot\boldsymbol{\sigma}_h, \varphi)_K = \sum_{K\in\mathcal{T}_h} (f - \nabla\cdot\boldsymbol{\sigma}_h, \varphi - \varphi_K)_K \\ &\le \sum_{K\in\mathcal{T}_h} \frac{h_K}{\pi}\|f - \nabla\cdot\boldsymbol{\sigma}_h\|_K \|\nabla\varphi\|_K.\end{aligned}$$

The Cauchy–Schwarz inequality concludes the proof in that

$$\begin{aligned}\min_{\substack{\boldsymbol{v}\in\boldsymbol{H}(\mathrm{div},\Omega)\\ \nabla\cdot\boldsymbol{v}=f}} \|\nabla_h u_h + \boldsymbol{v}\| &= \max_{\substack{\varphi\in H_0^1(\Omega)\\ \|\nabla\varphi\|=1}} \{(f, \varphi) - (\nabla_h u_h, \nabla\varphi)\} \\ &\le \Big\{\sum_{K\in\mathcal{T}_h} \Big(\|\nabla_h u_h + \boldsymbol{\sigma}_h\|_K + \frac{h_K}{\pi}\|f - \nabla\cdot\boldsymbol{\sigma}_h\|_K\Big)^2\Big\}^{\frac{1}{2}}.\end{aligned}$$

□

Remark 4.5.2 (Theorem 4.5.1)**.** *The three estimators of Theorem 4.5.1 reflect respectively the three possible violations of physical properties of an approximate solution u_h. Note that whenever $u_h \in H_0^1(\Omega)$, we can set $s_h = u_h$, and the corresponding estimator vanishes. Similarly, shall it happen that $-\nabla_h u_h \in \boldsymbol{H}(\mathrm{div}, \Omega)$ and $\nabla\cdot(-\nabla_h u_h) = f$, we can set $\boldsymbol{\sigma}_h = -\nabla_h u_h$, and the two other estimators vanish.*

4.5.2 A posteriori error estimate for the nonconforming finite element method, $p \geq 1$

We will now finally indicate how to perform an a posteriori error analysis in the spirit of (4.0.1) for the nonconforming finite element method (4.1.3) for any polynomial degree $p \geq 1$.

Recall the definition $\Pi_{p'}$ of the $L^2(\Omega)$ orthogonal projection onto the piecewise polynomial space $\mathcal{P}_{p'}(\mathcal{T}_h)$ from (1.9.5). We will adjust Definition 1.9.1 here to:

Definition 4.5.3 (Data oscillation)**.** *For $p' \geq 0$, let*

$$\eta_{\mathrm{osc},K,p'} := \frac{h_K}{\pi}\|f - \Pi_{p'} f\|_K, \tag{4.5.3}$$

$$\eta^2_{\mathrm{osc},p'} := \sum_{K\in\mathcal{T}_h} \eta^2_{\mathrm{osc},K,p'}. \tag{4.5.4}$$

Recall next the Raviart–Thomas space (1.9.8) as well as the notation $\mathcal{T}_{\boldsymbol{a}}$ for the set of mesh elements that share the given vertex $\boldsymbol{a} \in \mathcal{V}_h$ and $\omega_{\boldsymbol{a}}$ for the corresponding patch subdomain. Then we define

$$\boldsymbol{\mathcal{RT}}_{p'}(\mathcal{T}_{\boldsymbol{a}}) := \{\boldsymbol{v}_h \in [L^2(\omega_{\boldsymbol{a}})]^d; \boldsymbol{v}_h|_K \in \boldsymbol{\mathcal{RT}}_{p'}(K) \quad \forall K \in \mathcal{T}_{\boldsymbol{a}}\}. \tag{4.5.5}$$

The following is the counterpart of Definition 2.5.2 for the vector-valued dual variable; an illustration is given in Figure 4.5.

Definition 4.5.4 (Flux reconstruction by solution of local Neumann problems)**.** *Let $v_h \in H^1(\mathcal{T}_h)$ such that*

$$(\nabla_h v_h, \nabla\psi^{\boldsymbol{a}}) = (f, \psi^{\boldsymbol{a}}) \qquad \forall \boldsymbol{a} \in \mathcal{V}_h^{\mathrm{int}} \tag{4.5.6}$$

and a polynomial degree $p' \geq 0$ be arbitrary. For each vertex $\boldsymbol{a} \in \mathcal{V}_h$, set up the local space

$$\boldsymbol{V}^{\boldsymbol{a}}_{hp'} := \left\{ \boldsymbol{v}_h \in \boldsymbol{\mathcal{RT}}_{p'}(\mathcal{T}_{\boldsymbol{a}}) \cap \boldsymbol{H}(\mathrm{div}, \omega_{\boldsymbol{a}});\ \begin{array}{l} \boldsymbol{v}_h{\cdot}\boldsymbol{n} = 0 \text{ on } \partial\omega_{\boldsymbol{a}} \text{ for } \boldsymbol{a} \in \mathcal{V}_h^{\mathrm{int}}, \\ \boldsymbol{v}_h{\cdot}\boldsymbol{n} = 0 \text{ on } \partial\omega_{\boldsymbol{a}} \setminus \text{ faces sharing } \boldsymbol{a} \text{ for } \boldsymbol{a} \in \mathcal{V}_h^{\mathrm{ext}} \end{array} \right\}.$$

Define

$$\boldsymbol{\sigma}^{\boldsymbol{a}}_h := \arg \min_{\substack{\boldsymbol{v}_h \in \boldsymbol{V}^{\boldsymbol{a}}_{hp'} \\ \nabla\cdot\boldsymbol{v}_h = \Pi_{p'}(f\psi^{\boldsymbol{a}} - \nabla_h v_h \cdot \nabla\psi^{\boldsymbol{a}})}} \|\psi^{\boldsymbol{a}}\nabla_h v_h + \boldsymbol{v}_h\|_{\omega_{\boldsymbol{a}}}. \tag{4.5.7}$$

Then, extending $\boldsymbol{\sigma}^{\boldsymbol{a}}_h$ by zero outside of $\omega_{\boldsymbol{a}}$, set

$$\boldsymbol{\sigma}_h := \sum_{\boldsymbol{a}\in\mathcal{V}_h} \boldsymbol{\sigma}^{\boldsymbol{a}}_h \in \boldsymbol{V}_{hp'}. \tag{4.5.8}$$

Following [7, 24, 25], we can now identify a practical variant of Theorem 4.5.1, extending Theorems 4.3.1 and 4.3.3 to an arbitrary polynomial degree $p \geq 1$. Crucially, this result is *robust* with respect to the polynomial degree p, at least for $1 \leq d \leq 3$ (with a constant C independent of p, not deteriorating with p). This is typically very useful when working with hp finite elements that we will discuss in the following Chapter 6.

Theorem 4.5.5 (A posteriori error estimate for the nonconforming finite element method and its efficiency)**.** *Let $u \in H^1_0(\Omega)$ be the weak solution of (4.1.2) and let $u_h \in V^{\mathrm{nc}}_{hp}$ from (4.1.3) be its nonconforming finite element approximation for $p \geq 1$. Let $s_h \in V_{h(p+1)}$ be given by Definition 2.5.2 for $v_h = u_h$ and $p' = p+1$ and let $\boldsymbol{\sigma}_h \in \boldsymbol{V}_{hp} \subset \boldsymbol{H}(\mathrm{div}, \Omega)$ be given by Definition 4.5.4 for $v_h = u_h$ and $p' = p$. Then*

$$\|\nabla_h(u - u_h)\|^2 \leq \underbrace{\sum_{K\in\mathcal{T}_h} \left(\|\nabla_h u_h + \boldsymbol{\sigma}_h\|_K + \eta_{\mathrm{osc},K,p}\right)^2 + \sum_{K\in\mathcal{T}_h} \|\nabla_h(u_h - s_h)\|^2_K}_{\eta^2(u_h)}. \tag{4.5.9}$$

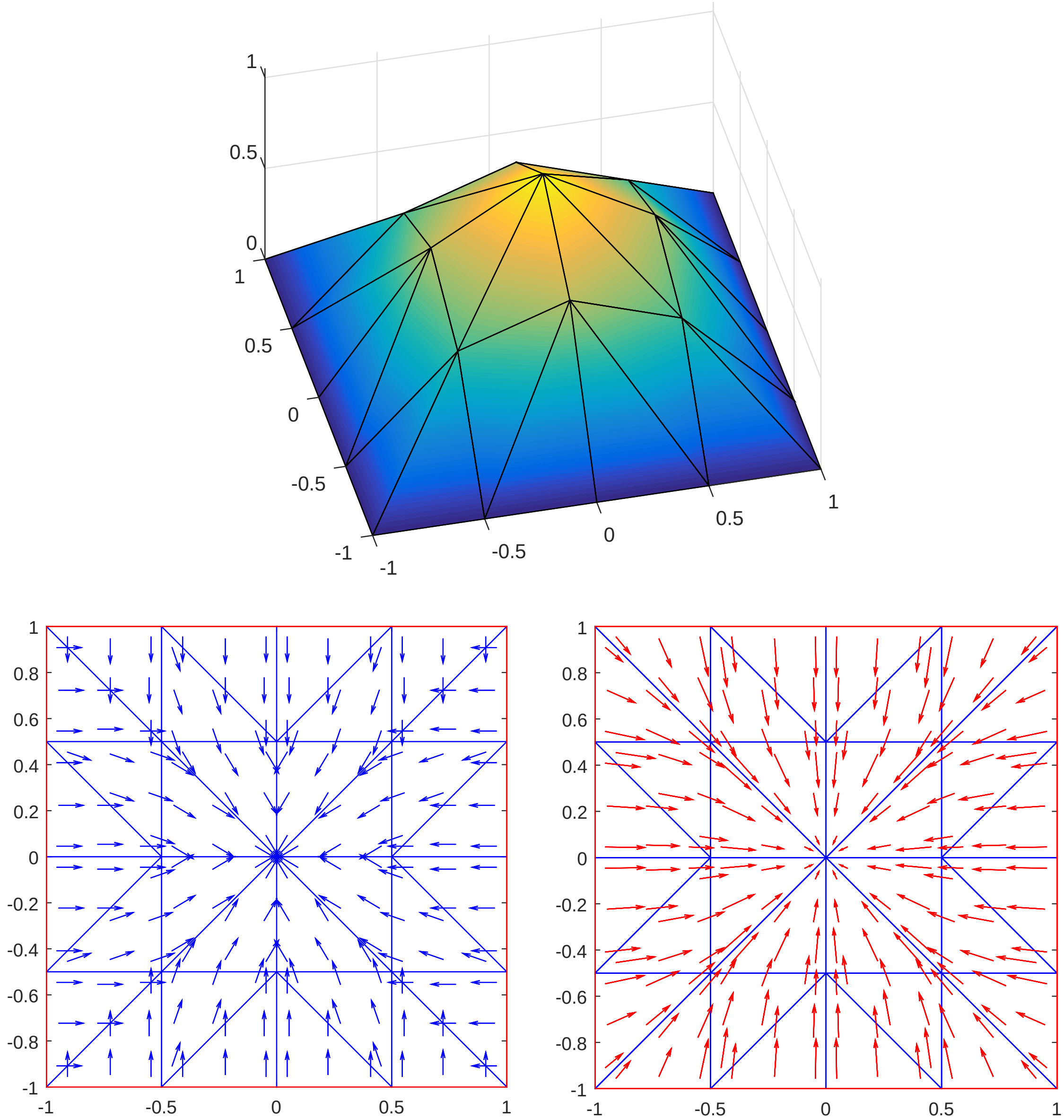

Figure 4.5: $\Omega = (-1, 1) \times (-1, 1)$, $d = 2$, $p = 1$. A function $v_h \in V_{h1}$ (top), its broken weak gradient $\nabla_h v_h \in [\mathcal{P}_0(\mathcal{T}_h)]^d$ (bottom left), and the corresponding equilibrated flux reconstruction $-\boldsymbol{\sigma}_h \in \boldsymbol{V}_{h1}$ from Definition 4.5.4 (bottom right)

Moreover, there holds the local efficiency

$$\|\nabla_h u_h + \boldsymbol{\sigma}_h\|_K + \|\nabla_h(u_h - s_h)\|_K \leq C\left(\|\nabla_h(u - u_h)\|_{\omega_K} + \left\{\sum_{K' \in \mathcal{T}_K} \eta^2_{\mathrm{osc},K',p-1}\right\}^{\frac{1}{2}}\right), \quad (4.5.10)$$

as well as the global efficiency

$$\eta(u_h) \leq C\|\nabla_h(u - u_h)\| + \eta_{\mathrm{osc},p-1}, \quad (4.5.11)$$

where the constant C only depends on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$ and on the space dimension d when $1 \leq d \leq 3$.

Chapter 5

Adaptive finite elements

In this rather exploratory chapter, we first continue the numerical example of Section 4.4. We then quickly present some important theoretical results.

5.1 Adaptive finite elements: numerical illustration of mesh adaptivity

We now continue further with the numerical example from Section 4.4.

From (4.3.9) and (4.5.10), illustrated in Figure 4.3, we know where the *error* is *localized*: if we predict that an error of some size in mesh element $K \in \mathcal{T}_h$, then, up to a generic constant, error of this size is indeed present in the element K and possibly in its neighbors. We can then try to adapt the mesh $\mathcal{T}_h$ by refining it in the problematic parts only, there where the values of the estimators $\eta_K(u_h)$ are increased, to hopefully better approximate the exact solution u. The outcome for the example from Section 4.4 is visualized in Figure 5.1 and is clearly beneficial visually.

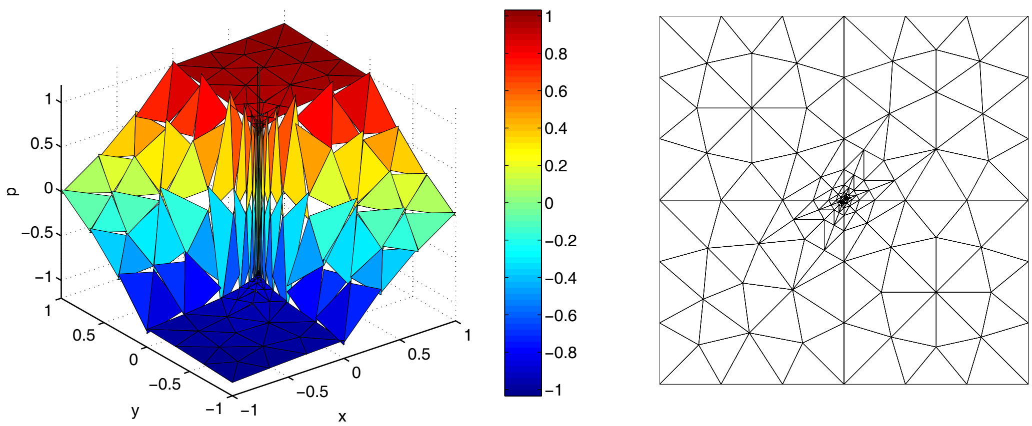

Figure 5.1: Nonconforming finite element approximate solution (left) on an adaptively refined mesh (right), $\alpha \approx 0.13$

We can now ask the following important question: does the adaptive mesh refinement (driven by the estimators $\eta_K(u_h)$) from Theorem 4.3.1 or 4.5.5 lead to a faster error decrease than the uniform mesh refinement, where each triangle is always cut into four congruent triangles (so that all mesh elements are of comparable size h, in the spirit of the a priori bound of Theorem 1.8.2)? The answer, presented in Figure 5.2, is decisively yes. Actually, for the uniform mesh refinement, the error decrease is of order h^α, i.e., it depends (and deteriorates) with the exact solution regularity. In contrast, the speed of the error decrease (rate of convergence) with respect to the number of mesh elements/degrees of freedom in the space V_{hp}^{nc} (these two numbers are similar, up to a constant) is *independent* of the exact solution *regularity* α. Actually, the speed of the error decrease (convergence rate) is the best-possible by piecewise first-order polynomials, the same as the order h from Theorem 1.8.2 for regular solution and uniform mesh refinement. Note that the maximal mesh size h does not give much sense anymore whereas the number of degrees of freedom always gives a good sense.

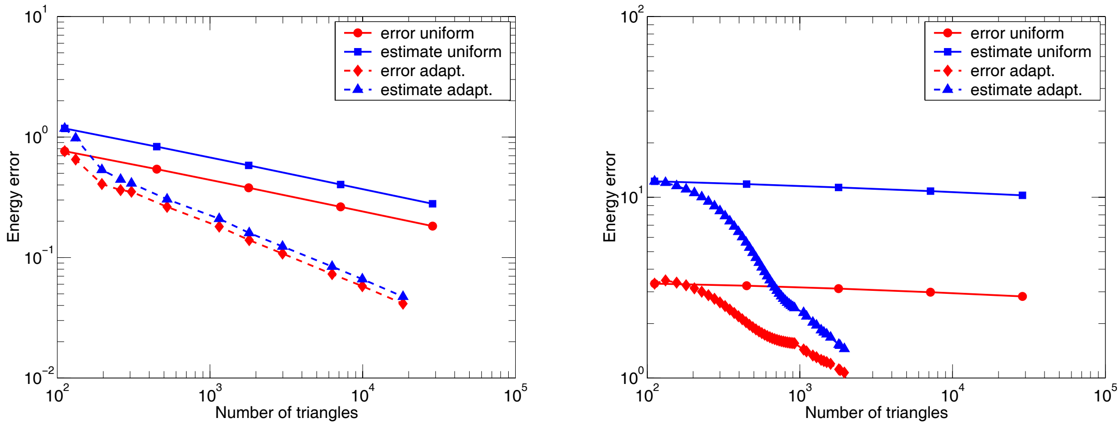

Figure 5.2: Errors and estimators for uniform and adaptive mesh refinement, $\alpha \approx 0.54$ (left), $\alpha \approx 0.13$ (right)

We finally again assess the quality of the a posteriori error estimators in terms of the *effectivity index*

$$I_{\mathrm{eff}} := \frac{\eta(u_h)}{\|\nabla_h(u-u_h)\|},$$

now both for uniform and adaptive mesh refinement. Still, from (4.3.1) of Theorem 4.3.1, we know that $I_{\mathrm{eff}} \geq 1$, and from (4.3.10) of Theorem 4.3.3, we know that I_{eff} is bounded from above by a generic constant; this holds on any sequence of meshes, uniformly or adaptively refined. Figure 5.3 shows that I_{eff} is still quite close to the optimal value of 1, and, additionally, improves with adaptive mesh refinement.

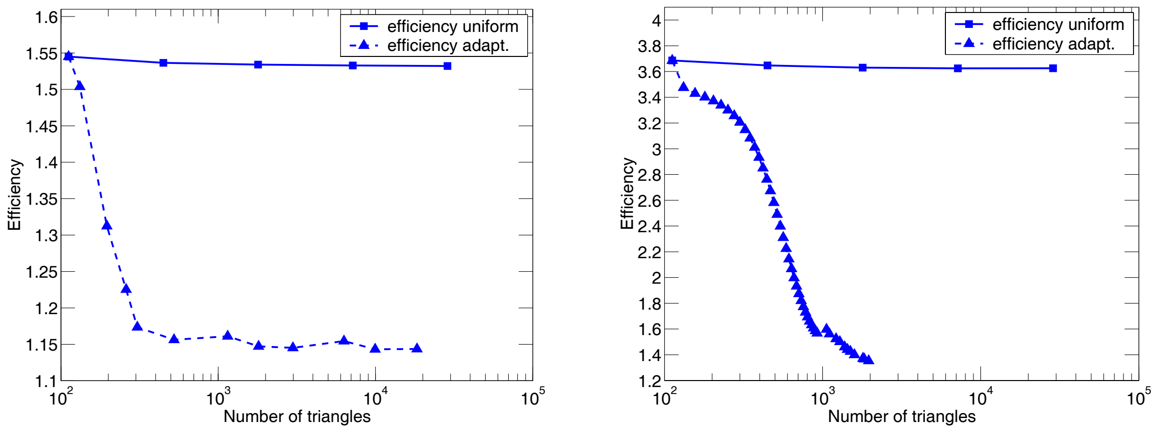

Figure 5.3: Effectivity indices, $\alpha \approx 0.54$ (left) and $\alpha \approx 0.13$ (right)

5.2 Adaptive finite elements: some theoretical background

Let us now quickly present some important theoretical results that are beyond what can be observed in the numerical experiment above in Section 5.1. We do so in the context of the conforming finite element method of Remark 1.3.3, following the seminal contributions in [18, 31, 39, 11].

The central point is to select visely the mesh elements to be refined. This is done by identifying the (minimal) subset $\mathcal{M}_\ell$ of all mesh elements in $\mathcal{T}_\ell$ containing the *θ-fraction* of the whole estimated error:

$$\sum_{K\in\mathcal{M}_\ell} \eta_K(u_\ell)^2 \geq \theta^2 \sum_{K\in\mathcal{T}_\ell} \eta_K(u_\ell)^2.$$

Here, we have quitted the notation $\mathcal{T}_h$, u_h that gives no much sense anymore, since not all mesh elements will have a comparable diameter h. Instead, we give a number ℓ to each mesh (and we will form a sequence od meshes). Now, we do not refine all mesh elements in $\mathcal{T}_\ell$ but only those contained in the set $\mathcal{M}_\ell$; we call the elements in $\mathcal{M}_\ell$ *marked* for refinement. In practice, one needs to also refine some neighbors of the elements $\mathcal{M}_\ell$.

The first important (and maybe a little surprising, since in adaptive mesh refinement, some mesh elements may not be refined at all, so that the maximal element diameter $h_\ell := \max_{K\in\mathcal{T}_\ell} h_K$ may stay constant) result shows that

$$\|\nabla(u-u_\ell)\| \to 0 \text{ for } \ell\to\infty, \tag{5.2.1}$$

i.e., the sequence of the adaptive finite element approximations u_ℓ *converges* towards u.

The second, and crucial, result states that

$$\|\nabla(u-u_\ell)\| \leq C|\mathrm{DoF}_\ell|^{-p/d}, \tag{5.2.2}$$

where DoF_ℓ stands for the number of degrees of freedom in the finite element space $V_{\ell p} := \mathcal{P}_p(\mathcal{T}_\ell)\cap H^1_0(\Omega)$. To apprehend formula (5.2.2), recall the standard a priori error decay result (1.9.16) of Theorem 1.9.8: it's essentially like if we had $\|\nabla(u-u_h)\| \leq Ch^p$ but all the time, for both *smooth* and *singular* solutions, and for both uniform and adaptive mesh refinement. In other words, the standard claim that convergence rates are limited by the exact solution regularity, so that higher-order finite elements only pay-off for smooth solutions, when $u|_K \in H^{p+1}(K)$ for all $K\in\mathcal{T}_h$, is only valid on uniformly refined meshes; on adaptively refined meshes, this is, fortunately, not true anymore. Moreover, the result in (5.2.2) actually shows that the error $\|\nabla(u-u_\ell)\|$ decays to zero as fast as in a best-possible way, on a *best-possible* sequence of meshes. This is termed *optimality* with respect to the number of degrees of freedom.

Chapter 6

hp finite elements

In Chapter 1, we have performed the so-called h-analysis of the Crouzeix–Raviart nonconforming finite element method of Definition 1.3.2. Namely, Theorems 1.8.2 and 1.9.8 assert hat for each *fixed polynomial degree* p, the error $\|\nabla(u - u_h)\|$ goes to zero as h^p when the mesh size h goes to zero. The deficiency of this traditional analysis is that it does not allow one to assess what happens if the role of the two discretization parameters is flipped: h is fixed and p goes to infinity, neither what happens if simultaneously h goes to zero and p to infinity. This is the purpose of the so-called hp-analysis that we briefly outline here. To simplify ideas, we will work in this chapter in the context of the conforming finite element method of Remark 1.3.3. We follow the developments in Babuška and Suri [4] and Schwab [38].

6.1 Conforming finite element method

We start with the conforming finite element space that we have already seen in Remark 1.3.3:

Definition 6.1.1 (Conforming finite element space)**.** *The continuous piecewise polynomial space V_{hp} is given by*

$$V_{hp} := \mathcal{P}_p(\mathcal{T}_h) \cap H^1_0(\Omega). \tag{6.1.1}$$

The conforming finite element method is the same as in Remark 1.3.3; it is the analysis that will done differently than in Chapter 1:

Definition 6.1.2 (Conforming finite element method for problem (1.1.1))**.** *Find $u_h \in V_{hp}$ such that*

$$(\nabla u_h, \nabla v_h) = (f, v_h) \qquad \forall v_h \in V_{hp}. \tag{6.1.2}$$

6.2 *hp* approximation on a simplex

The first central result is the following structural improvement of the Deny–Lions/Bramble–Hilbert Theorem 1.9.7, following Babuška and Suri [4, Lemma 4.1]:

Theorem 6.2.1 (*hp* approximation on a simplex)**.** *There holds*

$$\min_{v_h \in \mathcal{P}_p(K)} \|\nabla(v - v_h)\|_K \le C \frac{h_K^{\min\{s-1,p\}}}{p^{s-1}} \|v\|_{H^s(K)} \qquad \forall v \in H^s(K),\ s \ge 1,\ \forall K \in \mathcal{T}_h, \tag{6.2.1}$$

where the constant C only depends on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$, the space dimension d, and the regularity exponent s.

Remark 6.2.2 (*hp* approximation on a simplex)**.** *If the function v above has at least $H^{p+1}(K)$-regularity, i.e., $s \geq p+1$, than the above bound gives*

$$\min_{v_h \in \mathcal{P}_p(K)} \|\nabla(v - v_h)\|_K \leq C\Big(\frac{h_K}{p}\Big)^p \|v\|_{H^{p+1}(K)} \qquad \forall v \in H^{p+1}(K), \forall K \in \mathcal{T}_h. \tag{6.2.2}$$

Thus, in comparison with Theorem 1.9.7, we gain approximation properties in the polynomial degree p: the right-hand side in (6.2.1) goes to zero both *when the polynomial degree p is fixed and the mesh size h goes to zero and when h is fixed and p goes to infinity (for sufficiently smooth v).*

6.3 A priori error estimate: *hp* rate of convergence

The elementwise interpolation result of Theorem 6.2.1 together with further important (involved) developments which we do not detail enable to give the following result in the spirit of Theorem 1.9.8:

Theorem 6.3.1 (A priori *hp* rate of convergence)**.** *Let $u \in H_0^1(\Omega)$ be the weak solution of Definition 1.2.1 and $u_h \in V_{hp}$ its conforming finite element approximation of Definition 6.1.2. Let additionally*

$$u|_K \in H^s(K), s \geq 1, \qquad \forall K \in \mathcal{T}_h.$$

Then there exists a constant C only depending on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$, the space dimension d, and the regularity exponent s such that

$$\|\nabla(u - u_h)\| \leq C\frac{h^{\min\{s-1,p\}}}{p^{s-1}}\left\{\sum_{K\in\mathcal{T}_h} \|u\|^2_{H^s(K)}\right\}^{\frac{1}{2}}. \tag{6.3.1}$$

Remark 6.3.2 (A priori *hp* rate of convergence)**.** *The result of Theorem 1.9.8 can be in a simplified way (for sufficiently smooth u and f) described as*

$$\|\nabla(u - u_h)\| \leq Ch^p, \tag{6.3.2}$$

where the constant C only depends on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$, the space dimension d, the weak solution u, and additionally (in an unfavorable way) on the polynomial degree p. In contrast, a simplified writing of (6.3.1) (for sufficiently smooth u) gives

$$\|\nabla(u - u_h)\| \leq C\Big(\frac{h}{p}\Big)^p, \tag{6.3.3}$$

where the constant C only depends on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$, the space dimension d, and the weak solution u, but is independent of the polynomial degree p. In comparison with (6.3.2), (6.3.3) allows to asses the convergence rate when the mesh size h goes to zero or *the polynomial degree p goes to infinity. The hp-analysis of Theorem 6.3.1 thus improves over the h-analysis of Theorem 1.9.8.*

6.4 A priori error estimate: exponential convergence rate with respect to the number of the degrees of freedom

We finish this chapter by the following claim that we again state without a proof:

Theorem 6.4.1 (Exponential convergence rate)**.** *Let $u \in H_0^1(\Omega)$ be the weak solution of Definition 1.2.1 and $u_h \in V_{hp}$ its conforming finite element approximation of Definition 6.1.2. Let the space dimension $d = 2$. Then, for sufficiently smooth u, only increasing the polynomial degree p*

while keeping the mesh size h constant, there exist constants C_1 and C_2 only depending on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$ and the weak solution u such that

$$\|\nabla(u - u_h)\| \leq C_1 \frac{1}{e^{C_2 |V_{hp}|^{1/3}}}, \tag{6.4.1}$$

where $|V_{hp}|$ denotes the dimension of the space V_{hp}, i.e., the number of the degrees of freedom in V_{hp}.

The strength of Theorem 6.4.1 is best seen in a comparative example:

Remark 6.4.2 (Exponential convergence rate)**.** *Let the space dimension $d = 2$, let the polynomial degree $p = 1$, let the computational domain Ω be a unit square, and let the mesh $\mathcal{T}_h$ be composed of uniform isosceles triangles of diameter h that we decrease. One easily sees that there are roughly $1/h^2$ degrees of freedom in the space V_{hp}^{nc} of Definition 1.3.1, $|V_{hp}^{\mathrm{nc}}| \approx 1/h^2$. Then, Theorem 1.8.2 gives*

$$\|\nabla(u - u_h)\| \leq Ch \approx C \frac{1}{|V_{hp}|^{1/2}}, \tag{6.4.2}$$

which is a rate of convergence algebraic *with respect to $|V_{hp}|$, in contrast to (6.4.1), which is* exponential*.*

Remark 6.4.3 (Exponential convergence rate, singular solutions)**.** *The power of the result of Theorem 6.4.1 is further enhanced in that it actually holds also for a vast class of* singular *solutions.* Adapting *the* mesh *$\mathcal{T}_h$ and the* distribution *of the* polynomial degree *in a* non-uniform *way to the exact solution u (in an a priori way, using the knowledge of u (impractical) or relying on a posteriori error estimates of the form of Theorem 4.5.5 (practical)) (simultaneous mesh and polynomial degree adaptivity), a bound of the form (6.4.1) can still be obtained. Typically, one refines the mesh $\mathcal{T}_h$ towards the singularities of u and increases the polynomial degree in the smooth regions of u. The hp (adaptive) method and analysis then spectacularly outperform the traditional uniform h-refinement and analysis, see Figure 6.1 for an example.*

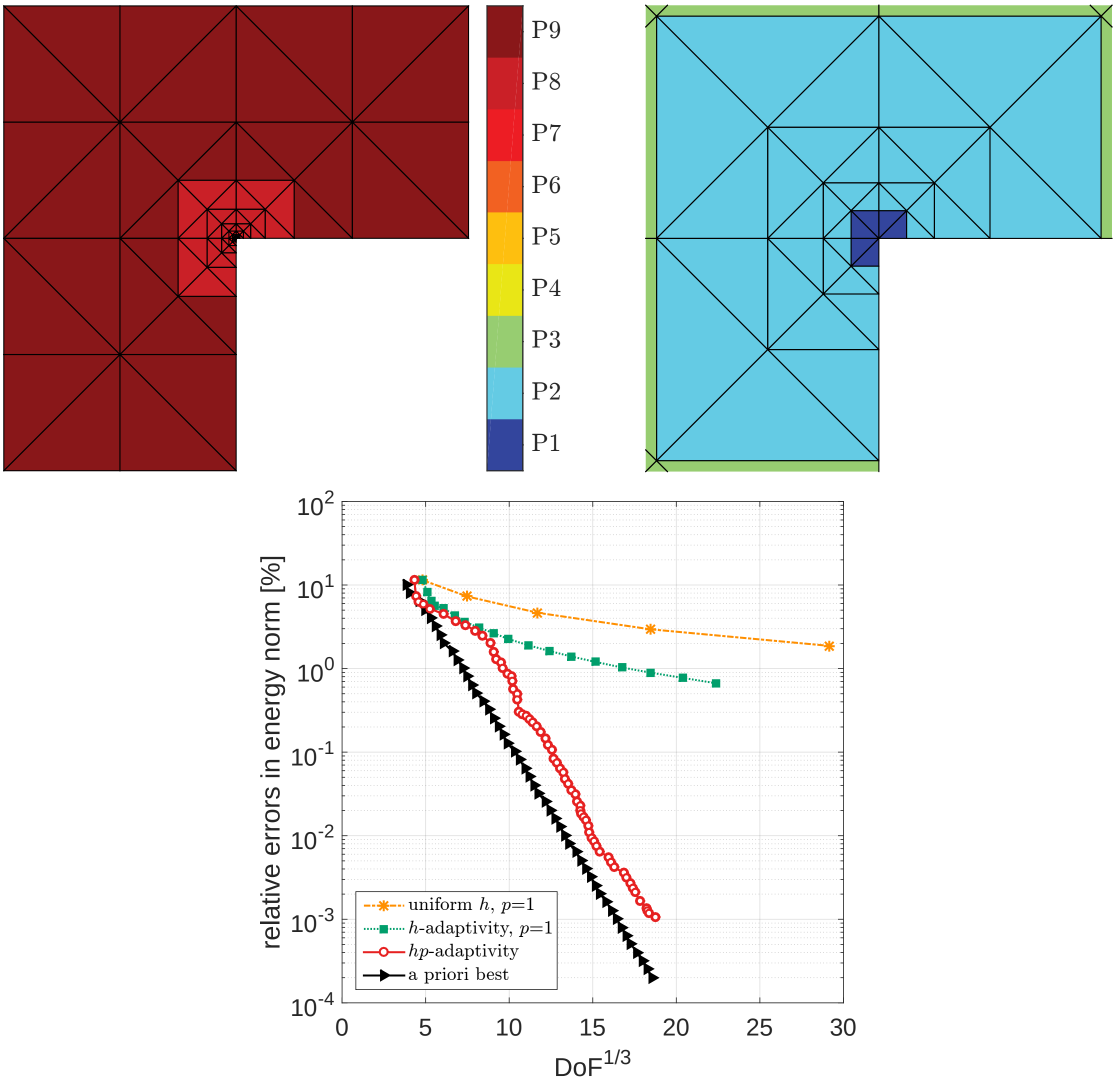

Figure 6.1: L-shaped domain $\Omega = (-1, 1) \times (-1, 1) \setminus [0, 1] \times [-1, 0]$, $d = 2$. hp finite element approximation of a singular solution $u \in H^{5/3}(\Omega)$. Non-uniformly refined mesh $\mathcal{T}_h$ and polynomial degree p (top left) and detail around the origin (top right): mesh refinement around the singularity in the origin $(0, 0)$ with $p = 1$ and increase of the polynomial degree p away from the origin, where u is smooth. Decrease of the relative energy error $\|\nabla(u - u_h)\|/\|\nabla u\|$ (log scale) in function of the number of the degrees of freedom in the finite element space V_{hp} (to the power 1/3) (bottom): a straight line represents a behavior of the form (6.4.1). A priori best: mesh and polynomial degree adapted in a best-possible way knowing the exact solution u (impractical but illustrative). hp-adaptivity: mesh and polynomial degree adapted on the basis of a posteriori error estimates of the form of Theorem 4.5.5. h-adaptivity, $p = 1$: mesh adapted on the basis of a posteriori error estimates of the form of Theorem 4.5.5, polynomial degree fixed to 1. Uniform h, $p = 1$: mesh uniformly refined (all triangles cut into four congruent triangles), polynomial degree fixed to $p = 1$; this curve corresponds to the relative error decrease as $Ch^{2/3} \approx C/|V_{hp}|^{1/3}$, in the sprit of a priori estimates such as that of Theorem 1.8.2. The results taken from [14].

Chapter 7

Finite elements for a nonlinear elliptic problem

We generalize in this chapter the linear differential operator of the Poisson equation (1.1.1) into a *nonlinear* one. We then define its conforming finite element approximation and perform its a priori h-convergence analysis. We will follow Ciarlet [12] and Zeidler [46].

7.1 A nonlinear elliptic problem

Let $f \in L^2(\Omega)$ be given. We consider in this chapter a nonlinear function $\boldsymbol{A} : \mathbb{R}^d \to \mathbb{R}^d$ giving rise to the *nonlinear* elliptic boundary value problem: find a scalar-valued function $u : \Omega \to \mathbb{R}$ such that

$$-\nabla{\cdot}\boldsymbol{A}(\nabla u) = f \qquad \text{in } \Omega, \tag{7.1.1a}$$

$$u = 0 \qquad \text{on } \partial\Omega. \tag{7.1.1b}$$

This generalizes the Poisson equation (1.1.1), where $\boldsymbol{A}(\boldsymbol{v}) = \boldsymbol{v}$ for all $\boldsymbol{v} \in \mathbb{R}^d$, i.e., $\boldsymbol{A}$ is the identity. We will establish the existence and uniqueness of a weak solution u of (7.1.1) under appropriate assumptions on the nonlinear function $\boldsymbol{A}$, relying on the Banach fixed-point theorem that we recall.

7.2 Strongly monotone and Lipschitz-continuous operators

We will focus on the following generalization of the weak gradient of Definition 0.3.2, which was a linear differential operator $H_0^1(\Omega) \to [L^2(\Omega)]^d$, $v \to \nabla v$:

Definition 7.2.1 (Strongly monotone and Lipschitz-continuous differential operator)**.** *We say that a differential operator $H_0^1(\Omega) \to [L^2(\Omega)]^d$, $v \to \boldsymbol{A}(\nabla v)$ is* strongly monotone *if there exists a positive constant α such that*

$$\alpha\|\nabla(v-w)\|^2 \le (\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w), \nabla(v-w)) \qquad \forall v, w \in H_0^1(\Omega). \tag{7.2.1}$$

We say that it is Lipschitz-continuous *if there exists a positive constant L such that*

$$\|\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w)\| \le L\|\nabla(v-w)\| \qquad \forall v, w \in H_0^1(\Omega). \tag{7.2.2}$$

Remark 7.2.2 (Constants α and L)**.** *There necessarily holds $\alpha \le L$. Indeed,* (7.2.1) *together with* (7.2.2) *and the Cauchy–Schwarz inequality yield*

$$\begin{aligned}
\alpha\|\nabla(v-w)\|^2 &\le (\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w), \nabla(v-w)) \\
&\le \|\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w)\|\|\nabla(v-w)\| \\
&\le L\|\nabla(v-w)\|^2 \qquad \forall v, w \in H_0^1(\Omega).
\end{aligned}$$

Note that when $\boldsymbol{A}$ is the identity, then $\alpha = L = 1$.

Example 7.2.3 (Mean-curvature flow). *A prototypical example of a nonlinear operator $\boldsymbol{A}$ is*

$$\boldsymbol{A}(\nabla v) := A(|\nabla v|)\nabla v \qquad \text{with } A(x) = \alpha + (L-\alpha)\sqrt{1+x^2}. \tag{7.2.3}$$

Note that when $L = \alpha$, $\boldsymbol{A}$ becomes linear, $\boldsymbol{A}(\nabla v) = \alpha \nabla v$.

7.3 Weak formulation

Generalizing the concepts of Section 1.2, the weak formulation of problem (7.1.1) is:

Definition 7.3.1 (Weak formulation of problem (7.1.1)). *Find $u \in H_0^1(\Omega)$ such that*

$$(\boldsymbol{A}(\nabla u), \nabla v) = (f, v) \qquad \forall v \in H_0^1(\Omega). \tag{7.3.1}$$

7.4 A contractive nonlinear operator

The key for the study of problem (7.3.1) is the following nonlinear operator:

Definition 7.4.1 (Nonlinear operator Φ). *Define the operator $\Phi : H_0^1(\Omega) \to H_0^1(\Omega)$, $v \to \Phi(v)$ by*

$$(\nabla\Phi(v), \nabla z) = (\nabla v, \nabla z) + \frac{\alpha}{L^2}\big[(f, z) - (\boldsymbol{A}(\nabla v), \nabla z)\big] \qquad \forall z \in H_0^1(\Omega). \tag{7.4.1}$$

Note that the above operator Φ is nonlinear whenever $\boldsymbol{A}$ is a nonlinear function. In contrast, problem (7.4.1) that defines $\Phi(v) \in H_0^1(\Omega)$ is a *linear Poisson* problem of the same form as problem (1.2.1), so that $\Phi(v)$ is uniquely defined by the Riesz representation theorem. Indeed, for the right-hand side, there holds

$$\left|(\nabla v, \nabla z) + \frac{\alpha}{L^2}\big[(f, z) - (\boldsymbol{A}(\nabla v), \nabla z)\big]\right| \le \|\nabla v\|\|\nabla z\| + \frac{\alpha}{L^2}\big[\|f\| C_{\mathrm{PF}} h_\Omega \|\nabla z\| + \|\boldsymbol{A}(\nabla v)\|\|\nabla z\|\big]$$

by the Cauchy–Schwarz inequality and the Poincaré–Friedrichs inequality (1.2.2), so that it is a bounded linear form on $H_0^1(\Omega)$; $\boldsymbol{A}(\nabla v)$ is a datum here whose action is well-defined in that $\|\boldsymbol{A}(\nabla v)\| = \|\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla 0) + \boldsymbol{A}(\nabla 0)\| \le L\|\nabla v\| + \|\boldsymbol{A}(\nabla 0)\|$ by (7.2.2).

The crucial property of operator Φ is that it is a *contraction* on the Sobolev space $H_0^1(\Omega)$:

Theorem 7.4.2 (Contraction of Φ). *There holds*

$$\|\nabla(\Phi(v) - \Phi(w))\| \le \underbrace{\left(1 - \frac{\alpha^2}{L^2}\right)^{\frac{1}{2}}}_{<1} \|\nabla(v - w)\| \qquad \forall v, w \in H_0^1(\Omega). \tag{7.4.2}$$

Proof. Let $v, w \in H_0^1(\Omega)$ be fixed. Introduce the following Riesz representers $\tilde v, \tilde w \in H_0^1(\Omega)$:

$$(\nabla \tilde v, \nabla z) = (\boldsymbol{A}(\nabla v), \nabla z) \qquad \forall z \in H_0^1(\Omega), \tag{7.4.3a}$$

$$(\nabla \tilde w, \nabla z) = (\boldsymbol{A}(\nabla w), \nabla z) \qquad \forall z \in H_0^1(\Omega). \tag{7.4.3b}$$

Note that $\|\boldsymbol{A}(\nabla v)\| \le L\|\nabla v\| + \|\boldsymbol{A}(\nabla 0)\|$ by (7.2.2) and similarly for w, so that the right-hand sides in (7.4.3) are again well-defined bounded forms and $\boldsymbol{A}(\nabla v)$ is fixed, so that these forms are indeed linear. It follows immediately by subtracting the lines in (7.4.3) and the Cauchy–Schwarz inequality that

$$\begin{aligned}
\|\nabla(\tilde v - \tilde w)\| &= \sup_{\substack{z \in H_0^1(\Omega)\\ \|\nabla z\|=1}} (\nabla(\tilde v - \tilde w), \nabla z)\\
&= \sup_{\substack{z \in H_0^1(\Omega)\\ \|\nabla z\|=1}} (\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w), \nabla z)\\
&\le \|\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w)\|.
\end{aligned} \tag{7.4.4}$$

To conclude, we carefully develop, noting that all $\Phi(v) - \Phi(w), v - w$, and $\tilde{v} - \tilde{w} \in H_0^1(\Omega)$ can be taken as the test function z in (7.4.1) and $\Phi(v) - \Phi(w)$ and $v - w \in H_0^1(\Omega)$ can be taken as the test function z in (7.4.3),

$$
\begin{aligned}
&\|\nabla(\Phi(v) - \Phi(w))\|^2 = (\nabla(\Phi(v) - \Phi(w)), \nabla(\Phi(v) - \Phi(w)))\\
&\overset{(7.4.1)}{=} (\nabla(v - w), \nabla(\Phi(v) - \Phi(w))) - \frac{\alpha}{L^2}\big[(\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w), \nabla(\Phi(v) - \Phi(w)))\big]\\
&\overset{(7.4.1)}{=} \|\nabla(v - w)\|^2 - \frac{\alpha}{L^2}\big[(\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w), \nabla(v - w))\big]\\
&\quad - \frac{\alpha}{L^2}\big[(\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w), \nabla(\Phi(v) - \Phi(w)))\big]\\
&\overset{(7.4.3)}{=} \|\nabla(v - w)\|^2 - \frac{\alpha}{L^2}\big[(\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w), \nabla(v - w))\big]\\
&\quad - \frac{\alpha}{L^2}\big[(\nabla(\tilde{v} - \tilde{w}), \nabla(\Phi(v) - \Phi(w)))\big]\\
&\overset{(7.4.1)}{=} \|\nabla(v - w)\|^2 - \frac{\alpha}{L^2}\big[(\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w), \nabla(v - w))\big]\\
&\quad - \frac{\alpha}{L^2}\Big[(\nabla(\tilde{v} - \tilde{w}), \nabla(v - w)) - \frac{\alpha}{L^2}\big[(\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w), \nabla(\tilde{v} - \tilde{w}))\big]\Big]\\
&\overset{(7.4.3)}{=} \|\nabla(v - w)\|^2 - 2\frac{\alpha}{L^2}\big[(\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w), \nabla(v - w))\big]\\
&\quad + \frac{\alpha^2}{L^4}\big[(\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w), \nabla(\tilde{v} - \tilde{w}))\big]\\
&\overset{(7.2.1)}{\leq} \|\nabla(v - w)\|^2 - 2\frac{\alpha^2}{L^2}\|\nabla(v - w)\|^2 + \frac{\alpha^2}{L^4}\big[(\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w), \nabla(\tilde{v} - \tilde{w}))\big]\\
&\overset{(7.4.4)}{\leq} \|\nabla(v - w)\|^2 - 2\frac{\alpha^2}{L^2}\|\nabla(v - w)\|^2 + \frac{\alpha^2}{L^4}\|\boldsymbol{A}(\nabla v) - \boldsymbol{A}(\nabla w)\|^2\\
&\overset{(7.2.2)}{\leq} \|\nabla(v - w)\|^2 - 2\frac{\alpha^2}{L^2}\|\nabla(v - w)\|^2 + \frac{\alpha^2}{L^2}\|\nabla(v - w)\|^2.
\end{aligned}
$$

□

7.5 Existence and uniqueness of a weak solution by the Banach fixed-point theorem

The following lemma is instrumental in understanding the role of the operator Φ of Definition 7.4.1 in the proof of existence and uniqueness of u from (7.3.1):

Lemma 7.5.1 (Fixed point of Φ and existence and uniqueness of u from (7.3.1))**.** *The following equivalence holds: $u \in H_0^1(\Omega)$ is a fixed point of the operator Φ if and only if u solves (7.3.1), i.e.,*

$$\Phi(u) = u \Longleftrightarrow (7.3.1) \text{ holds.} \tag{7.5.1}$$

Proof. Definition (7.4.1) applied to $u \in H_0^1(\Omega)$ gives

$$(\nabla(\Phi(u) - u), \nabla z) = \frac{\alpha}{L^2}\big[(f, z) - (\boldsymbol{A}(\nabla u), \nabla z)\big] \qquad \forall z \in H_0^1(\Omega),$$

which readily implies the assertion. □

Theorem 7.5.2 (Existence and uniqueness of $u \in H_0^1(\Omega)$ from Definition 7.3.1, Banach fixed–point theorem)**.** *There exists a unique solution $u \in H_0^1(\Omega)$ from Definition 7.3.1.*

Proof. The space $H_0^1(\Omega)$ is Hilbert for the scalar product $(\nabla v, \nabla w)$, $v, w \in H_0^1(\Omega)$, and the associated norm given by the $[L^2(\Omega)]^d$-norm of the weak gradient, $\|\nabla v\|$. Since the operator

Φ is contractive in this setting by Theorem 7.4.2, the Banach fixed-point theorem, see, e.g., Zeidler [46, Section 25.4], implies that Φ has a unique fixed point. Thus the conclusion follows by Lemma 7.5.1.

For illustration and to make the text self-contained, we include the proof of the Banach fixed-point theorem directly in our setting. Denote by q the contraction factor from (7.4.2),

$$q := \left(1 - \frac{\alpha^2}{L^2}\right)^{\frac{1}{2}} < 1.$$

Consider an arbitrary $u_0 \in H_0^1(\Omega)$ and define a sequence

$$u_k := \Phi(u_{k-1}), \quad k \geq 1.$$

Then (7.4.2) immediately implies

$$\|\nabla(u_{k+1} - u_k)\| = \|\nabla(\Phi(u_k) - \Phi(u_{k-1}))\| \overset{(7.4.2)}{\leq} q\|\nabla(u_k - u_{k-1})\| \leq \ldots \leq q^k\|\nabla(u_1 - u_0)\|.$$

We will use this property to show that the sequence u_k is Cauchy. Consider integer indices $k > 0$ and $m > 0$. We have, by the triangle inequality, the above property, and a geometric series sum

$$\begin{aligned}
\|\nabla(u_{k+m} - u_k)\| &\leq \|\nabla(u_{k+m} - u_{k+m-1})\| + \|\nabla(u_{k+m-1} - u_{k+m-2})\| + \ldots + \|\nabla(u_{k+1} - u_k)\| \\
&\leq (q^{k+m-1} + q^{k+m-2} + \ldots + q^k)\|\nabla(u_1 - u_0)\| \\
&= q^k(q^{m-1} + q^{m-2} + \ldots + 1)\|\nabla(u_1 - u_0)\| \\
&\leq q^k \sum_{i=0}^{\infty} q^i \|\nabla(u_1 - u_0)\| \\
&\leq q^k \frac{1}{1-q}\|\nabla(u_1 - u_0)\|.
\end{aligned}$$

Thus, for any real $\epsilon > 0$, we can take $k > 0$ large enough such that

$$q^k \frac{1}{1-q}\|\nabla(u_1 - u_0)\| < \epsilon,$$

so that

$$\|\nabla(u_{k+m} - u_k)\| < \epsilon$$

for any $m > 0$. The space $H_0^1(\Omega)$ being complete, we conclude that there exists $u \in H_0^1(\Omega)$ such that

$$u = \lim_{k\to\infty} u_k \quad \Longleftrightarrow \quad \lim_{k\to\infty} \|\nabla(u - u_k)\| = 0.$$

Finally, the triangle inequality and (7.4.2) show that for any $k > 0$

$$\begin{aligned}
\|\nabla(\Phi(u) - u)\| &\leq \|\nabla(\Phi(u) - u_k)\| + \|\nabla(u_k - u)\| \\
&= \|\nabla(\Phi(u) - \Phi(u_{k-1}))\| + \|\nabla(u_k - u)\| \\
&\leq \underbrace{q\|\nabla(u - u_{k-1})\| + \|\nabla(u - u_k)\|}_{\to 0 \text{ for } k\to\infty}.
\end{aligned}$$

Thus, we conclude that u is a fixed point of Φ, $\Phi(u) = u$. Moreover, Φ cannot have two distinct fixed points u and $\tilde{u}$, since

$$\|\nabla(\Phi(u) - \Phi(\tilde{u}))\| = \|\nabla(u - \tilde{u})\| \not\leq q\|\nabla(u - \tilde{u})\|.$$

□

7.6 The conforming finite element method for problem (7.1.1)

The discretization of the weak formulation (7.3.1) by conforming finite elements is straightforward:

Definition 7.6.1 (Conforming finite element method for problem (7.1.1))**.** *Find $u_h \in V_{hp}$ such that*

$$(\boldsymbol{A}(\nabla u_h), \nabla v_h) = (f, v_h) \qquad \forall v_h \in V_{hp}. \tag{7.6.1}$$

We also readily obtain:

Theorem 7.6.2 (Existence and uniqueness of $u_h \in V_{hp}$ from Definition 7.6.1)**.** *There exists a unique solution $u_h \in V_{hp}$ from Definition 7.6.1.*

Proof. One merely replaces the Hilbert space $H_0^1(\Omega)$ by the Hilbert space V_{hp} and proceeds as above for u. More concretely, one defines a nonlinear operator $\Phi_h : V_{hp} \to V_{hp}$, $v_h \to \Phi_h(v_h)$ by

$$(\nabla \Phi_h(v_h), \nabla z_h) = (\nabla v_h, \nabla z_h) + \frac{\alpha}{L^2}\big[(f, z_h) - (\boldsymbol{A}(\nabla v_h), \nabla z_h)\big] \qquad \forall z_h \in V_{hp}, \tag{7.6.2}$$

as a finite-dimensional equivalent of (7.4.1). One then finds easily that it is contractive just as Φ in Theorem 7.4.2,

$$\|\nabla(\Phi_h(v_h) - \Phi_h(w_h))\| \le \Big(1 - \frac{\alpha^2}{L^2}\Big)^{\frac{1}{2}} \|\nabla(v_h - w_h)\| \qquad \forall v_h, w_h \in V_{hp}. \tag{7.6.3}$$

Thus, it has a fixed point in the space V_{hp}, and one concludes by the equivalent of (7.5.1), stating that

$$\Phi_h(u_h) = u_h \iff (7.6.1) \text{ holds}, \tag{7.6.4}$$

which yields the assertion. $\square$

7.7 A priori error estimates: rate of convergence by h-analysis

Recall the monotonicity and Lipschitz-continuity constants α and L from Definition 7.2.1. We now have all the ingredients to rather quickly perform an a priori h-convergence analysis:

Theorem 7.7.1 (A priori rate of convergence)**.** *Let $u \in H_0^1(\Omega)$ be the weak solution of Definition 7.3.1 and $u_h \in V_{hp}$ its conforming finite element approximation of Definition 7.6.1. Let additionally*

$$u|_K \in H^{p+1}(K) \qquad \forall K \in \mathcal{T}_h.$$

Then there exists a constant C only depending on the mesh shape-regularity parameter $\kappa_{\mathcal{T}_h}$, the space dimension d, and the polynomial degree p such that

$$\|\nabla(u - u_h)\| \le \frac{L}{\alpha} C h^p \Bigg\{ \sum_{K \in \mathcal{T}_h} |u|^2_{H^{p+1}(K)} \Bigg\}^{\frac{1}{2}}. \tag{7.7.1}$$

Proof. By subtracting (7.6.1) from (7.3.1), we obtain the Galerkin orthogonality

$$(\boldsymbol{A}(\nabla u) - \boldsymbol{A}(\nabla u_h), \nabla v_h) = 0 \qquad \forall v_h \in V_{hp}. \tag{7.7.2}$$

Consequently, there holds, for an arbitrary $v_h \in V_{hp}$,

$$\begin{aligned}
\|\nabla(u - u_h)\|^2 &\overset{(7.2.1)}{\le} \frac{1}{\alpha}(\boldsymbol{A}(\nabla u) - \boldsymbol{A}(\nabla u_h), \nabla(u - u_h)) \\
&\overset{(7.7.2)}{=} \frac{1}{\alpha}(\boldsymbol{A}(\nabla u) - \boldsymbol{A}(\nabla u_h), \nabla(u - v_h)) \\
&\overset{(7.2.2)}{\le} \frac{L}{\alpha}\|\nabla(u - u_h)\| \|\nabla(u - v_h)\|.
\end{aligned}$$

Thus, we see that

$$\|\nabla(u - u_h)\| \leq \frac{L}{\alpha} \min_{v_h \in V_{hp}} \|\nabla(u - v_h)\|, \tag{7.7.3}$$

which is, up to the ratio α/L, a best-approximation result as in (1.7.6). Consequently, (1.9.11) allows to reduce the question to elementwise best-approximation, and (7.7.1) follows by the Deny–Lions/Bramble–Hilbert Theorem 1.9.7. $\square$

From Theorem 7.7.1, we see that the rate of convergence for a finite element approximation of the nonlinear problem (7.1.1) is the same as that for the linear problem (1.1.1).

Chapter 8

Finite elements for the heat equation

The purpose of this last chapter is to present a second important extension of the linear elliptic Poisson problem (1.1.1), this time into a time-dependent partial differential equation. We investigate the existence and uniqueness of a weak solution and an appropriate finite element discretization. We will follow Ern and Guermond [20, 21] and [23, 22].

8.1 The heat equation

The heat equation reads as follows: for a final time $T > 0$ and source term $f \in L^2(0, T; L^2(\Omega))$, find a scalar-valued function $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t u - \Delta u &= f && \text{in } \Omega \times (0, T), && (8.1.1a) \\ u &= 0 && \text{on } \partial\Omega \times (0, T), && (8.1.1b) \\ u(0) &= 0 && \text{in } \Omega. && (8.1.1c) \end{aligned}$$

8.2 Bochner function spaces

Bochner function spaces are a generalization of Lebesgue spaces to functions whose values lie in a Banach space, instead of real or complex numbers, cf. Ern and Guermond [21, Section 56.1]. We will in particular need the function space with weak partial derivatives with respect to the spatial variables to belong to L^2 in both space and time

$$X := L^2(0, T; H_0^1(\Omega)). \tag{8.2.1}$$

We will also need its subspace additionally requesting the weak partial derivative with respect to the time variable to belong to H^{-1} in space and L^2 in time,

$$Y := \{v \in X;\ \partial_t v \in L^2(0, T; H^{-1}(\Omega))\} = L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)). \tag{8.2.2}$$

We will also impose the zero initial condition in the subspace

$$Y_0 := \{v \in Y;\ v(0) = 0\}, \tag{8.2.3}$$

where we note that $Y \subset C(0, T; L^2(\Omega))$. We equip the spaces X and Y with the following norms:

$$\begin{aligned} \|v\|_X^2 &:= \int_0^T \|\nabla v\|^2 \,\mathrm{d}t && v \in X, && (8.2.4a) \\ \|v\|_Y^2 &:= \int_0^T \|\nabla v\|^2 + \|\partial_t v\|_{H^{-1}(\Omega)}^2 \,\mathrm{d}t + \|v(T)\|^2 && v \in Y_0, && (8.2.4b) \end{aligned}$$

where $H^{-1}(\Omega)$ is the dual space to $H_0^1(\Omega)$. With $\langle\cdot,\cdot\rangle$ the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$,

$$\|v\|_{H^{-1}(\Omega)} = \max_{\substack{\varphi\in H_0^1(\Omega)\\ \|\nabla\varphi\|=1}} \langle v,\varphi\rangle. \tag{8.2.5}$$

8.3 Weak formulation

The weak formulation for the heat equation is:

Definition 8.3.1 (Weak formulation of problem (8.1.1))**.** *Find $u \in Y_0$ such that*

$$\int_0^T \langle\partial_t u, v\rangle + (\nabla u,\nabla v)\,\mathrm{d}t = \int_0^T (f,v)\,\mathrm{d}t \qquad \forall v\in X. \tag{8.3.1}$$

One remarks that in contrast to the developments of the previous chapters, one looks for the weak solution in the *trial space* Y which is *different* from the *test space* X. This is in line with the nonsymmetry between space and time in (8.1.1). This is also the origin of the fact that the analysis of (8.3.1) will be more involved. We also remark that formulation (8.3.1) is equivalent to finding $u\in Y_0$ such that

$$\langle\partial_t u(t), v\rangle + (\nabla u(t),\nabla v) = (f(t),v) \qquad \forall v\in H_0^1(\Omega),\ \text{ for a.e. } t\in(0,T). \tag{8.3.2}$$

8.4 Inf-sup condition

Recall that

$$\|\nabla\varphi\| = \max_{\substack{v\in H_0^1(\Omega)\\ \|\nabla v\|=1}} (\nabla\varphi,\nabla v) = \max_{v\in H_0^1(\Omega)} \frac{(\nabla\varphi,\nabla v)}{\|\nabla v\|} \qquad \forall\varphi\in H_0^1(\Omega), \tag{8.4.1}$$

where we use the shorthand notation $0/0 = 0$. One finds similarly

$$\|\varphi\|_X = \max_{v\in X} \frac{\int_0^T (\nabla\varphi,\nabla v)\,\mathrm{d}t}{\|v\|_X} \qquad \forall\varphi\in X. \tag{8.4.2}$$

Indeed, on the one hand,

$$\max_{v\in X} \frac{\int_0^T (\nabla\varphi,\nabla v)\,\mathrm{d}t}{\|v\|_X} \le \|\varphi\|_X \max_{v\in X}\frac{\|v\|_X}{\|v\|_X} = \|\varphi\|_X$$

by virtue of (8.2.4a) and the Cauchy–Schwarz inequality. On the other hand,

$$\max_{v\in X} \frac{\int_0^T (\nabla\varphi,\nabla v)\,\mathrm{d}t}{\|v\|_X} \ge \frac{\int_0^T (\nabla\varphi,\nabla\varphi)\,\mathrm{d}t}{\|\varphi\|_X} = \frac{\|\varphi\|_X^2}{\|\varphi\|_X} = \|\varphi\|_X,$$

where the lower bound follows by picking $\varphi\in X$ in the max.

The following is a central result for problem (8.1.1), identifying a suitable extension of the property (8.4.2):

Theorem 8.4.1 (Inf-sup identity)**.** *For every $\varphi\in Y_0$, there holds*

$$\|\varphi\|_Y = \max_{v\in X} \frac{\int_0^T \langle\partial_t\varphi, v\rangle + (\nabla\varphi,\nabla v)\,\mathrm{d}t}{\|v\|_X}. \tag{8.4.3}$$

Proof. For a fixed $\varphi \in Y_0$, let $w_* \in X$ be defined by, a.e. in $(0,T)$,

$$(\nabla w_*, \nabla v) = \langle \partial_t \varphi, v\rangle \qquad \forall v \in H_0^1(\Omega). \tag{8.4.4}$$

Using (8.4.1) and (8.2.5), this implies the identity

$$\|\nabla w_*\| = \|\partial_t \varphi\|_{H^{-1}(\Omega)} \tag{8.4.5}$$

a.e. in $(0,T)$, as well as

$$\int_0^T \langle \partial_t \varphi, v\rangle + (\nabla \varphi, \nabla v)\,\mathrm{d}t = \int_0^T (\nabla(w_* + \varphi), \nabla v)\,\mathrm{d}t \qquad \forall v \in X.$$

Consequently, using (8.4.2),

$$\|w_* + \varphi\|_X = \max_{v\in X} \frac{\int_0^T \langle \partial_t \varphi, v\rangle + (\nabla \varphi, \nabla v)\,\mathrm{d}t}{\|v\|_X}. \tag{8.4.6}$$

Moreover, the following useful identity holds true on the space Y_0:

$$2\int_0^T \langle \partial_t \varphi, \varphi\rangle\,\mathrm{d}t = \int_0^T \frac{d}{dt}\|\varphi\|^2\,\mathrm{d}t = \|\varphi(T)\|^2 - \|\varphi(0)\|^2 = \|\varphi(T)\|^2. \tag{8.4.7}$$

Consequently,

$$\begin{aligned}
\|w_* + \varphi\|_X^2 &\overset{(8.2.4a)}{=} \int_0^T \|\nabla(w_* + \varphi)\|^2\,\mathrm{d}t \\
&= \int_0^T \|\nabla w_*\|^2 + 2(\nabla w_*, \nabla\varphi) + \|\nabla\varphi\|^2\,\mathrm{d}t \\
&\overset{(8.4.4)\,(8.4.5)}{=} \int_0^T \|\partial_t\varphi\|_{H^{-1}(\Omega)}^2 + 2\langle \partial_t\varphi, \varphi\rangle + \|\nabla\varphi\|^2\,\mathrm{d}t \\
&\overset{(8.4.7)\,(8.2.4b)}{=} \|\varphi\|_Y^2,
\end{aligned}$$

so that the claim (8.4.3) follows from (8.4.6). □

Remark 8.4.2 (Inf-sup condition)**.** *One remarks easily that* (8.4.3) *in particular implies*

$$\inf_{\varphi\in Y_0} \sup_{v\in X} \frac{\int_0^T \langle \partial_t \varphi, v\rangle + (\nabla \varphi, \nabla v)\,\mathrm{d}t}{\|v\|_X \|\varphi\|_Y} \geq C. \tag{8.4.8}$$

We from (8.4.3) *actually have an equality with* $C = 1$ *and* min *and* max *in place of* inf *and* sup*. The writing* (8.4.8) *gives rise to the nomenclature "inf-sup condition", which is central in analysis of partial differential equations and finite element methods of nonsymmetric problems, see Nečas [32], Babuška [3], or the summary in the form of the Banach–Nečas–Babuška (also called inf-sup or Babuška–Brezzi–Ladyzhenskaya) theorem in, e.g., Ern and Guermond [20, Theorem 22.8].*

8.5 Existence and uniqueness of a weak solution by the Banach closed range and open mapping theorems

Let X' be the dual of X, $X' = L^2(0,T;H^{-1}(\Omega))$, and let $\langle\cdot,\cdot\rangle_{X',X}$ denote the corresponding duality pairing. Define the operator $B_Y : Y_0 \to X'$ by

$$\langle B_Y(\varphi), v\rangle_{X',X} := \int_0^T \langle \partial_t \varphi, v\rangle + (\nabla \varphi, \nabla v)\,\mathrm{d}t \qquad v \in X,\ \varphi \in Y_0. \tag{8.5.1}$$

This operator is clearly linear and bounded as, for all $\varphi \in Y_0$,

$$\begin{aligned} \|B_Y(\varphi)\|_{X'} &= \sup_{v\in X} \frac{\langle B_Y(\varphi), v\rangle_{X',X}}{\|v\|_X} \\ &\overset{(8.5.1)}{=} \sup_{v\in X} \frac{\int_0^T \langle \partial_t\varphi, v\rangle + (\nabla\varphi, \nabla v)\,\mathrm{d}t}{\|v\|_X} \\ &\overset{(8.4.3)}{=} \|\varphi\|_Y; \end{aligned} \tag{8.5.2}$$

B_Y is actually an isometry. The weak formulation (8.3.1) can then be equivalently rewritten as: find $u \in Y_0$ such that

$$B_Y(\varphi) = f \qquad \text{in } X'. \tag{8.5.3}$$

Let Y_0' be the dual of Y_0 and $\langle\cdot,\cdot\rangle_{Y_0',Y_0}$ the corresponding duality pairing. We will also need below the adjoint operator $B_Y^* : X \to Y_0'$ defined by

$$\langle B_Y^*(v), \varphi\rangle_{Y_0',Y_0} := \langle B_Y(\varphi), v\rangle_{X',X} \qquad v \in X,\ \varphi \in Y_0. \tag{8.5.4}$$

We can now state and prove the central result of this chapter:

Theorem 8.5.1 (Existence and uniqueness of $u \in Y_0$ from Definition 8.3.1 by the Banach closed range and open mapping theorems)**.** *There exists a unique solution $u \in Y_0$ from Definition 8.3.1.*

Proof. Since (8.3.1) is equivalent to (8.5.3), the existence and uniqueness of a weak solution $u \in Y_0$ follows when the operator B_Y from (8.5.1) is bijective. By the Banach closed range and open mapping theorems, this is in turn equivalent to showing that

$$(1)\ B_Y \text{ is injective}, \quad (2)\ B_Y \text{ is surjective}, \tag{8.5.5a}$$

$$(A)\ B_Y^* \text{ is surjective}, \quad (B)\ B_Y^* \text{ is injective}, \tag{8.5.5b}$$

$$(i)\ B_Y \text{ is injective}, \quad (ii) \text{ range of } B_Y \text{ is closed in } X', \quad (iii)\ B_Y^* \text{ is injective}, \tag{8.5.5c}$$

see, e.g., Ern and Guermond [21, Section 22.1] and [21, Lemmas A.43 and A.44]. We will prove the three properties in (8.5.5c), as in [21, Theorem 56.21].

(i) Let $B_Y(\varphi) = 0$ for some $\varphi \in Y_0$. Then, by (8.5.1),

$$\int_0^T \langle \partial_t\varphi, v\rangle + (\nabla\varphi, \nabla v)\,\mathrm{d}t = 0 \qquad v \in X.$$

By virtue of Theorem 8.4.1, this implies $\varphi = 0$, i.e., injectivity (actually, B_Y is automatically injective, since it is an isometry).

(ii) Consider a sequence $\varphi_i \in Y_0$ such that $B_Y(\varphi_i)$ is a Cauchy sequence in X'. Thus, for any real $\epsilon > 0$, there exists $k > 0$ such that for all $m > 0$,

$$\|B_Y(\varphi_{k+m}) - B_Y(\varphi_k)\|_{X'} \le \epsilon.$$

This, however, immediately implies that φ_i is a Cauchy sequence in Y_0, since

$$\|B_Y(\varphi_{k+m}) - B_Y(\varphi_k)\|_{X'} = \|B_Y(\varphi_{k+m} - \varphi_k)\|_{X'} \overset{(8.5.2)}{=} \|\varphi_{k+m} - \varphi_k\|_Y.$$

Taking its limit $\varphi \in Y_0$, we obtain that

$$\lim_{k\to\infty} \|B_Y(\varphi) - B_Y(\varphi_k)\|_{X'} \overset{(8.5.2)}{=} \lim_{k\to\infty} \|\varphi - \varphi_k\|_Y = 0,$$

so that the range of B_Y is closed in X'.

(iii) Finally, let $B_Y^*(v) = 0$ for some $v \in X$. Then, by (8.5.4) and (8.5.1),

$$\langle B_Y^*(v), \varphi\rangle_{Y_0', Y_0} = \langle B_Y(\varphi), v\rangle_{X',X} = \int_0^T \langle \partial_t \varphi, v\rangle + (\nabla\varphi, \nabla v)\,\mathrm{d}t = 0 \qquad \forall \varphi \in Y_0 \tag{8.5.6}$$

Define $\xi \in L^2(0,T;H^{-1}(\Omega))$ by, a.e. in $(0,T)$,

$$\langle \xi, w\rangle = (\nabla v, \nabla w) \qquad \forall w \in H_0^1(\Omega). \tag{8.5.7}$$

Then (8.5.6) in particular implies

$$\int_0^T \langle \partial_t\varphi, v\rangle\,\mathrm{d}t = -\int_0^T \langle \xi, \varphi\rangle\,\mathrm{d}t \qquad \forall \varphi \in \mathcal{D}((0,T)\times\Omega) \subset Y_0,$$

which is (compare with Definition 0.3.1) the meaning of

$$\partial_t v = \xi \tag{8.5.8}$$

and in particular shows that, actually, $v \in Y$.

Consider now an arbitrary function $w \in H_0^1(\Omega)$, so that when multiplied by the time variable t, $tw \in Y_0$. Taking tw as a test function φ in (8.5.6) and using the integration by parts in time formula, we see

$$\begin{aligned}
0 &\overset{(8.5.7)}{=} \int_0^T \langle \partial_t(tw), v\rangle + \langle \xi, tw\rangle\,\mathrm{d}t \\
&\overset{(8.5.8)}{=} \int_0^T \langle \partial_t(tw), v\rangle + \langle \partial_t v, tw\rangle\,\mathrm{d}t \\
&= \; T(w, v(T)) - 0 = T(w, v(T)).
\end{aligned}$$

Since $w \in H_0^1(\Omega)$ was arbitrary and since $H_0^1(\Omega)$ is dense in $L^2(\Omega)$, we infer that

$$v(T) = 0. \tag{8.5.9}$$

We finally use tv as a test function $\varphi \in Y_0$ in (8.5.6). This gives, similarly as in (8.4.7),

$$\begin{aligned}
\int_0^T \langle \partial_t(tv), v\rangle\,\mathrm{d}t &\overset{(8.5.7)}{=} -\int_0^T \langle \xi, tv\rangle\,\mathrm{d}t \overset{(8.5.8)}{=} -\int_0^T \langle \partial_t v, tv\rangle\,\mathrm{d}t \\
&= \; -\int_0^T t\langle \partial_t v, v\rangle\,\mathrm{d}t = -\frac{1}{2}\int_0^T t\frac{d}{dt}\|v\|^2\,\mathrm{d}t \\
&= \; -\frac{1}{2}\left[T\|v(T)\|^2 - 0\|v(0)\|^2 - \int_0^T \|v\|^2\,\mathrm{d}t\right] \\
&\overset{(8.5.9)}{=} \frac{1}{2}\int_0^T \|v\|^2\,\mathrm{d}t.
\end{aligned}$$

Consequently, still taking tv as a test function $\varphi \in Y_0$ in (8.5.6), and using this result,

$$\begin{aligned}
0 &= \int_0^T \langle \partial_t(tv), v\rangle + (\nabla(tv), \nabla v)\,\mathrm{d}t \\
&= \frac{1}{2}\int_0^T \|v\|^2\,\mathrm{d}t + \int_0^T t\|\nabla v\|^2\,\mathrm{d}t.
\end{aligned}$$

From here, we conclude $v = 0$, i.e., the injectivity of B_Y^*. □

8.6 The conforming finite element method for problem (8.1.1)

In all the previous chapters, the application of a finite element method to a weak formulation was straightforward. Moreover, for a conforming discretization, the finite element method turned out to give a projection of the weak solution, as in (1.7.5)–(1.7.6), or at least a best approximation up to a constant as in (7.7.3). Unfortunately, both conception of the finite element approximation and analysis and more involved in the present nonsymmetric, time-dependent setting. We will thus only describe a possible discretization with the simplest approximation of the time derivative and we do not perform analysis here.

Let $N > 1$ be the number of time steps and let $0 = t_0 < t_1 < \ldots < t_n < \ldots < t_N = T$ be the discrete times; we will denote by I_n the n-th time interval, $[t_{n-1}, t_n]$ and τ_n the length of the n-th time step, $\tau_n := t_n - t_{n-1} = |I_n|$, $1 \le n \le N$. As in the previous chapters, we let $\mathcal{T}_h$ be a simplicial mesh of the closure of the computational domain Ω. Recall from (1.3.3) that $V_{hp} = \mathcal{P}_p(\mathcal{T}_h) \cap H_0^1(\Omega)$.

Definition 8.6.1 (Conforming finite elements in space and backward Euler in time for problem (8.1.1))**.** *Let* $u_0 = 0$. *For all* $1 \le n \le N$, *find* $u_h^n \in V_{hp}$ *such that*

$$\left(\frac{u_h^n - u_h^{n-1}}{\tau_n}, v_h\right) + (\nabla u_h^n, \nabla v_h) = \frac{1}{\tau_n}\int_{I_n} (f, v_h)\,\mathrm{d}t \qquad \forall v_h \in V_{hp}. \tag{8.6.1}$$

In contrast to the (rather involved) proof of Theorem 8.5.1, existence and uniqueness of all u_h^n from (8.6.1) is straightforward:

Theorem 8.6.2 (Existence and uniqueness of $u_h^n \in V_{hp}$ from Definition 8.6.1)**.** *There exists a unique solution* $u_h^n \in V_{hp}$ *from Definition 8.6.1 for all* $1 \le n \le N$.

Proof. Let $1 \le n \le N$ be fixed. Definition 8.6.1 can be equivalently rewritten as: find $u_h^n \in V_{hp}$ such that

$$(u_h^n, v_h) + \tau_n(\nabla u_h^n, \nabla v_h) = \int_{I_n} (f, v_h)\,\mathrm{d}t + (u_h^{n-1}, v_h) \qquad \forall v_h \in V_{hp}. \tag{8.6.2}$$

The claim is thus a consequence of the Riesz representation theorem, since the left-hand side of (8.6.2) is equivalent to the $H^1(\Omega)$ scalar product (0.3.2a) and the right-hand side of (8.6.2) is a bounded linear form. The twist is that in contrast to (8.3.1), (8.6.2) is discrete; in particular when the time step size τ_n goes to zero, the left-hand side of (8.6.2) stops to be equivalent to the $H^1(\Omega)$ scalar product (0.3.2a). □

The outcome of Definition 8.6.1 is a collection of piecewise polynomial functions in V_{hp}. The procedure is *sequential*, where one first poses $u_0 = 0$, then one computes u_1 from u_0, then u_2 from u_1, …, and finally u_N from u_{N-1}. This is *advantageous for computation*, since the size of the arising linear systems is still as in the steady cases (but one has to solve N such systems).

By the finite element approximation of the heat equation, one usually understands:

Definition 8.6.3 (Space-time finite element approximation)**.** *A finite element approximation of* (8.1.1) *is the space-time function* $u_{h\tau} : \Omega \times (0, T) \to \mathbb{R}$

$$u_{h\tau}|_{I_n} := u_h^n \qquad \forall 1 \le n \le N.$$

Recalling the function spaces X and Y from (8.2.1)–(8.2.2), it unfortunately turns our that the finite element approximation $u_{h\tau}$ is *nonconforming* in the sense that $u_{h\tau} \in X$ but $u_{h\tau} \notin Y$. Indeed, $u_{h\tau}$ is piecewise constant, discontinuous in time, so that $\partial_t u_{h\tau}$ does not exist. This can be remedied partly in using a reconstruction $I u_{h\tau}$:

Definition 8.6.4 (Conforming reconstruction of the finite element approximation)**.** *Define* $Iu_{h\tau}$*, piecewise affine and continuous in time, such that* $Iu_{h\tau}(t_n) := u_h^n$, $Iu_{h\tau}(t_{n-1}) := u_h^{n-1}$, $1 \leq n \leq N$*, and* $Iu_{h\tau}$ *varies affinely inbetween, that is*

$$Iu_{h\tau}(t) := u_h^{n-1} + \frac{u_h^n - u_h^{n-1}}{\tau_n}(t - t_{n-1}),\ t \in I_n \qquad \forall 1 \leq n \leq N.$$

Using Definition 8.6.4, two important properties arise. First, there holds

$$Iu_{h\tau} \in Y_0,$$

so that $Iu_{h\tau}$ is now a *conforming* approximation of the weak solution $u \in Y_0$. Second, there holds

$$\partial_t Iu_{h\tau}|_{I_n} = \frac{u_h^n - u_h^{n-1}}{\tau_n} \qquad \forall 1 \leq n \leq N,$$

so that (8.6.1) yields

$$(\partial_t Iu_{h\tau}, v_h) + (\nabla u_{h\tau}, \nabla v_h) = \frac{1}{\tau_n}\int_{I_n} (f, v_h)\,\mathrm{d}t \qquad \forall v_h \in V_{hp}, \quad \forall 1 \leq n \leq N. \tag{8.6.3}$$

The link of (8.6.3) to (8.3.2) may at a first sight seem much more in line with, e.g., the link between (1.2.1) and (1.3.4). The nonconformity bottleneck has, however, not been resolved but merely reshaped, since in (8.6.3), two different objects appear, namely the space-time finite element approximation $u_{h\tau}$ and its conforming reconstruction $Iu_{h\tau}$. This is yet another obstacle that one has to overcome if one wishes to accomplish a convergence analysis of the finite element approximation of the heat equation.

Bibliography

[1] Achdou, Y., Bernardi, C., and Coquel, F. A priori and a posteriori analysis of finite volume discretizations of Darcy's equations. *Numer. Math.* **96** (2003), 17–42. http://dx.doi.org/10.1007/s00211-002-0436-7.

[2] Allaire, G. *Analyse numérique et optimisation.* Editions Ecole Polytechnique, Palaiseau, France, 2005.

[3] Babuška, I. Error-bounds for finite element method. *Numer. Math.* **16** (1971), 322–333.

[4] Babuška, I., and Suri, M. The *h-p* version of the finite element method with quasi-uniform meshes. *RAIRO Modél. Math. Anal. Numér.* **21** (1987), 199–238. http://dx.doi.org/10.1051/m2an/1987210201991.

[5] Bebendorf, M. A note on the Poincaré inequality for convex domains. *Z. Anal. Anwendungen* **22** (2003), 751–756. http://dx.doi.org/10.4171/ZAA/1170.

[6] Boffi, D., Brezzi, F., and Fortin, M. *Mixed finite element methods and applications*, vol. **44** of *Springer Series in Computational Mathematics.* Springer, Heidelberg, 2013. https://doi.org/10.1007/978-3-642-36519-5.

[7] Braess, D., and Schöberl, J. Equilibrated residual error estimator for edge elements. *Math. Comp.* **77** (2008), 651–672. http://dx.doi.org/10.1090/S0025-5718-07-02080-7.

[8] Bramble, J. H., and Hilbert, S. R. Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation. *SIAM J. Numer. Anal.* **7** (1970), 112–124. https://doi.org/10.1137/0707006.

[9] Brenner, S. C. Poincaré-Friedrichs inequalities for piecewise H^1 functions. *SIAM J. Numer. Anal.* **41** (2003), 306–324. https://doi.org/10.1137/S0036142902401311.

[10] Burman, E., and Ern, A. Continuous interior penalty *hp*-finite element methods for advection and advection-diffusion equations. *Math. Comp.* **76** (2007), 1119–1140.

[11] Cascón, J. M., Kreuzer, C., Nochetto, R. H., and Siebert, K. G. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.* **46** (2008), 2524–2550. http://dx.doi.org/10.1137/07069047X.

[12] Ciarlet, P. G. *The Finite Element Method for Elliptic Problems*, vol. **4** of *Studies in Mathematics and its Applications.* North-Holland, Amsterdam, 1978.

[13] Crouzeix, M., and Raviart, P.-A. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge* **7** (1973), 33–75.

[14] Daniel, P., Ern, A., Smears, I., and Vohralík, M. An adaptive *hp*-refinement strategy with computable guaranteed bound on the error reduction factor. *Comput. Math. Appl.* **76** (2018), 967–983. https://doi.org/10.1016/j.camwa.2018.05.034.

[15] Deny, J., and Lions, J. L. Les espaces du type de Beppo Levi. *Ann. Inst. Fourier (Grenoble)* **5** (1954), 305–370. http://aif.cedram.org/item?id=AIF_1953__5__305_0.

[16] Destuynder, P., and Métivet, B. Explicit error bounds for a nonconforming finite element method. *SIAM J. Numer. Anal.* **35** (1998), 2099–2115. https://doi.org/10.1137/S0036142996300191.

[17] Destuynder, P., and Métivet, B. Explicit error bounds in a conforming finite element method. *Math. Comp.* **68** (1999), 1379–1396. http://dx.doi.org/10.1090/S0025-5718-99-01093-5.

[18] Dörfler, W. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.* **33** (1996), 1106–1124. http://dx.doi.org/10.1137/0733054.

[19] Ern, A., Gudi, T., Smears, I., and Vohralík, M. Equivalence of local- and global-best approximations, a simple stable local commuting projector, and optimal *hp* approximation estimates in $\boldsymbol{H}(\mathrm{div})$. *IMA J. Numer. Anal.* (2021). DOI 10.1093/imanum/draa103, http://dx.doi.org/10.1093/imanum/draa103.

[20] Ern, A., and Guermond, J.-L. *Finite Elements I. Approximation and Interpolation*, vol. **72** of *Texts in Applied Mathematics*. Springer International Publishing, Springer Nature Switzerland AG, 2021. https://doi-org.ezproxy.is.cuni.cz/10.1007/978-3-030-56341-7.

[21] Ern, A., and Guermond, J.-L. *Finite Elements II. Galerkin Approximation, Elliptic and Mixed PDEs*, vol. **73** of *Texts in Applied Mathematics*. Springer International Publishing, Springer Nature Switzerland AG, 2021. https://doi-org.ezproxy.is.cuni.cz/10.1007/978-3-030-56923-5.

[22] Ern, A., Smears, I., and Vohralík, M. Guaranteed, locally space-time efficient, and polynomial-degree robust a posteriori error estimates for high-order discretizations of parabolic problems. *SIAM J. Numer. Anal.* **55** (2017), 2811–2834. https://doi.org/10.1137/16M1097626.

[23] Ern, A., Smears, I., and Vohralík, M. Equilibrated flux a posteriori error estimates in $L^2(H^1)$-norms for high-order discretizations of parabolic problems. *IMA J. Numer. Anal.* **39** (2019), 1158–1179. https://doi.org/10.1093/imanum/dry035.

[24] Ern, A., and Vohralík, M. Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations. *SIAM J. Numer. Anal.* **53** (2015), 1058–1081. http://dx.doi.org/10.1137/130950100.

[25] Ern, A., and Vohralík, M. Stable broken H^1 and $\boldsymbol{H}(\mathrm{div})$ polynomial extensions for polynomial-degree-robust potential and flux reconstruction in three space dimensions. *Math. Comp.* **89** (2020), 551–594. http://dx.doi.org/10.1090/mcom/3482.

[26] Gudi, T. A new error analysis for discontinuous finite element methods for linear elliptic problems. *Math. Comp.* **79** (2010), 2169–2189. http://dx.doi.org/10.1090/S0025-5718-10-02360-4.

[27] Karakashian, O. A., and Pascal, F. A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems. *SIAM J. Numer. Anal.* **41** (2003), 2374–2399. http://dx.doi.org/10.1137/S0036142902405217.

[28] Kellogg, R. B. On the Poisson equation with intersecting interfaces. *Applicable Anal.* **4** (1974/75), 101–129. Collection of articles dedicated to Nikolai Ivanovich Muskhelishvili, http://dx.doi.org/10.1080/00036817408839086.

[29] Knobloch, P. Uniform validity of discrete Friedrichs' inequality for general nonconforming finite element spaces. *Numer. Funct. Anal. Optim.* **22** (2001), 107–126.

[30] Marini, L. D. An inexpensive method for the evaluation of the solution of the lowest order Raviart–Thomas mixed method. *SIAM J. Numer. Anal.* **22** (1985), 493–496.

[31] Morin, P., Nochetto, R. H., and Siebert, K. G. Convergence of adaptive finite element methods. *SIAM Rev.* **44** (2002), 631–658. Revised reprint of "Data oscillation and convergence of adaptive FEM" [SIAM J. Numer. Anal. **38** (2000), no. 2, 466–488; MR1770058 (2001g:65157)].

[32] Nečas, J. Sur une méthode pour résoudre les équations aux dérivées partielles du type elliptique, voisine de la variationnelle. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)* **16** (1962), 305–326.

[33] Payne, L. E., and Weinberger, H. F. An optimal Poincaré inequality for convex domains. *Arch. Rational Mech. Anal.* **5** (1960), 286–292.

[34] Prager, W., and Synge, J. L. Approximations in elasticity based on the concept of function space. *Quart. Appl. Math.* **5** (1947), 241–269.

[35] Quarteroni, A., and Valli, A. *Numerical approximation of partial differential equations*, vol. **23** of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1994.

[36] Raviart, P.-A., and Thomas, J.-M. A mixed finite element method for 2nd order elliptic problems. In *Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975)*. Springer, Berlin, 1977, pp. 292–315. Lecture Notes in Math., Vol. 606.

[37] Repin, S. *A posteriori estimates for partial differential equations*, vol. **4** of *Radon Series on Computational and Applied Mathematics*. Walter de Gruyter GmbH & Co. KG, Berlin, 2008. http://dx.doi.org/10.1515/9783110203042.

[38] Schwab, C. Theory and applications in solid and fluid mechanics. *p- and hp-finite element methods*. Numerical Mathematics and Scientific Computation. The Clarendon Press, Oxford University Press, New York, 1998.

[39] Stevenson, R. An optimal adaptive finite element method. *SIAM J. Numer. Anal.* **42** (2005), 2188–2217. http://dx.doi.org/10.1137/S0036142903425082.

[40] Temam, R. *Navier-Stokes equations*, revised ed., vol. **2** of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1979.

[41] Veeser, A. Approximating gradients with continuous piecewise polynomial functions. *Found. Comput. Math.* **16** (2016), 723–750. http://dx.doi.org/10.1007/s10208-015-9262-z.

[42] Verfürth, R. *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Teubner-Wiley, Stuttgart, 1996.

[43] Vohralík, M. On the discrete Poincaré–Friedrichs inequalities for nonconforming approximations of the Sobolev space H^1. *Numer. Funct. Anal. Optim.* **26** (2005), 925–952. http://dx.doi.org/10.1080/01630560500444533.

[44] Vohralík, M. A posteriori error estimates for lowest-order mixed finite element discretizations of convection-diffusion-reaction equations. *SIAM J. Numer. Anal.* **45** (2007), 1570–1599. http://dx.doi.org/10.1137/060653184.

[45] Vohralík, M. Lecture notes. *A posteriori error estimates for efficiency and error control in numerical simulations.* Sorbonne University, Paris & Charles University, Prague, 2018. https://who.rocq.inria.fr/Martin.Vohralik/Enseig/APost/a_posteriori.pdf.

[46] Zeidler, E. *Nonlinear functional analysis and its applications. II/B.* Springer-Verlag, New York, 1990. https://doi.org/10.1007/978-1-4612-0985-0.

Index

a posteriori
error estimate, 31, 39, 41, 44, 45, 47
a priori
error estimate, 15, 23, 29, 31, 41, 55, 56, 59, 63, 65
adaptive
error convergence, 53
finite element method, 53

best approximation, 27, 28
broken
Poincaré–Friedrichs inequality, 18, 19
Sobolev space $H^1(\mathcal{T}_h)$, 11, 17
weak gradient, 12, 16, 18, 41
weakly-continuous Sobolev space $H^1_{[\![\,]\!]}(\mathcal{T}_h)$, 17, 18
bubble function
element, 45

conforming
finite element method, 16, 22, 55, 63, 70
consistency
error, 22, 27
convergence
adaptive, 53
rate, 23, 29, 56, 63

data
oscillation, 26, 29, 47
differential
operator, 15, 41, 59, 65
divergence
weak, 9, 39, 47

efficiency
global, 46
local, 46
equation
Poisson, 15
equilibrated
flux reconstruction, 39, 47
error
consistency, 22, 27
estimate a posteriori, 31, 39, 41, 44, 45, 47
estimate a priori, 15, 23, 29, 31, 41, 55, 56, 59, 63, 65

finite element method
adaptive, 53
conforming, 16, 22, 55, 63, 70
hp, 55
nonconforming, 15, 16, 41
flux
equilibrated reconstruction, 39, 47
reconstruction, 39

gradient
broken weak, 12, 16, 18, 41, 63, 70
weak, 8, 15, 41, 55, 60, 66

hp
finite element method, 55

inequality
broken Poincaré–Friedrichs, 18, 19
inverse, 46
Poincaré, 12, 27, 44
Poincaré–Friedrichs, 13, 15
trace, 13, 21, 24, 25, 36

jump, 11, 16

nonconforming
finite element method, 15, 16, 41
nonconformity, 16
nonlinear
operator, 59

operator
differential, 15, 41, 59, 65
nonlinear, 59
oscillation
data, 26, 29, 47

partial derivative
weak, 8, 28
Poincaré
inequality, 12, 27
Poincaré–Friedrichs
broken inequality, 18, 19

inequality, 13, 15
Poisson
equation, 15
potential
reconstruction, 31, 32, 36

rate
convergence, 23, 29, 56, 63
reconstruction
equilibrated flux, 39, 47
flux, 39
potential, 31, 32, 36

solution
approximate, 44
exact, 15, 41, 42, 60, 66
space
$\mathcal{D}(\Omega)$, 8
$\boldsymbol{H}(\mathrm{div}, \Omega)$, 9, 39
broken Sobolev $H^1(\mathcal{T}_h)$, 11, 17
conforming V_{hp}, 55
Crouzeix–Raviart nonconforming V_{hp}^{nc}, 16
piecwise polynomial $\mathcal{P}_p(\mathcal{T}_h)$, 13, 16, 55
Sobolev $H^m(\Omega)$, 29
Sobolev $H^1(\Omega)$, 9
Sobolev $H_0^1(\Omega)$, 9, 15, 41, 42, 60, 66
weakly-continuous broken Sobolev $H^1_{[\![\,]\!]}(\mathcal{T}_h)$, 17, 18
Strang
second lemma, 22

theorem
Lax–Milgram, 16
Riesz, 16, 42, 60, 70
second Strang lemma, 22
trace, 9
inequality, 13, 21, 24, 25, 36

weak
divergence, 9, 39, 47
gradient, 8, 15, 41, 60, 66
partial derivative, 8, 28