Technische Universität Chemnitz

Sonderforschungsbereich 393

Numerische Simulation auf massiv parallelen Rechnern

Gerd Kunert

Error estimation for anisotropic tetrahedral and triangular finite element meshes

Preprint SFB393/97_16

Abstract

Some boundary value problems yield anisotropic solutions, e.g. solutions with boundary layers. If such problems are to be solved with the finite element method (FEM), anisotropically refined meshes can be advantageous.

In order to construct these meshes or to control the error one aims at reliable error estimators. For *isotropic* meshes many estimators are known, but they either fail when used on *anisotropic* meshes, or they were not applied yet. For rectangular (or cuboidal) anisotropic meshes a modified error estimator had already been found.

We are investigating error estimators on anisotropic tetrahedral or triangular meshes because such grids offer greater geometrical flexibility. For the Poisson equation a residual error estimator, a local Dirichlet problem error estimator, and an L_2 error estimator are derived, respectively. Additionally a residual error estimator is presented for a singularly perturbed reaction diffusion equation.

It is important that the anisotropic mesh corresponds to the anisotropic solution. Provided that a certain condition is satisfied, we have proven that all estimators bound the error reliably.

Keywords: finite elements, error estimator, anisotropic solution, stretched elements, tetrahedral mesh, singularly perturbed problem

AMS(MOS): 65N30, 65N15, 35B25

Preprint-Reihe des Chemnitzer SFB 393

SFB393/97_16 **August 1997**

Contents

1 Introduction **4**

2 Preliminaries **9**

2.1 Notation 9

2.1.1 General notation 9

2.1.2 Notation of the tetrahedron 9

2.1.3 Transformations and their properties 11

2.1.4 The directional derivative $\tilde{D}_i$ 13

2.1.5 Auxiliary subdomains 14

2.2 Requirements on the mesh 15

2.3 Basic tools 15

2.3.1 Anisotropic trace inequalities 15

2.3.2 Inverse inequalities for finite element functions 17

2.3.3 Bubble functions and their inverse inequalities 18

3 The Poisson equation **21**

3.1 Analytical Background 21

3.2 Residual error estimator 23

3.2.1 The set $H^1_{\mathcal{T}}(\Omega)$ of adapted functions 23

3.2.2 Anisotropic interpolation estimates 24

3.2.3 Anisotropic residual error estimator 31

3.3 A local Dirichlet problem error estimator 37

3.3.1 Introduction and definition 37

3.3.2 Equivalence and bounds of the estimator 38

3.3.3 Condition number of the FEM matrix 43

3.4 L_2 error estimator 45

3.4.1 Special L_2 bubble functions 45

3.4.2 Additional mesh requirement and inverse inequalities 47

3.4.3 The set $H^2_{\mathcal{T}}(\Omega)$ of L_2 adapted functions 52

3.4.4 Anisotropic interpolation estimates 52

3.4.5 Anisotropic L_2 error estimator 54

3.5 An anisotropic Zienkiewicz-Zhu like error estimator 57

3.5.1 Cuboidal or rectangular mesh 57

3.5.2 Tetrahedral mesh of tensor product type 58

4 A reaction–diffusion equation **59**

4.1 Analytical Background 59

4.2 Residual error estimator 61

4.2.1 Special face bubble functions 61

4.2.2 Anisotropic interpolation estimates 62
4.2.3 Anisotropic residual error estimator 63

5 Numerical examples **67**
5.1 Scope of and introduction to the numerical experiments 67
5.2 Two numerical examples 69
5.2.1 The condition on $u - u_h$ 70
5.2.2 Error bounds in the energy norm 70
5.2.3 Error bounds in the L_2 norm 71
5.3 Conclusions 72

6 Summary **73**
6.1 Open points and future work 73

Bibliography **74**

Author's address:

Gerd Kunert
TU Chemnitz
Fakultät Mathematik
09107 Chemnitz

Germany

Email: *gerd.kunert@mathematik.tu-chemnitz.de*
Phone: +49 - 0371 - 531 - 2603

Chapter 1

Introduction

Many models in science and engineering lead to partial differential equations. Some of them, the boundary value problems, are written (in the so-called classical formulation) as

$$\text{Find } u: \qquad \begin{aligned} Lu &= f \qquad \text{in } \Omega \qquad , \\ lu &= g \qquad \text{on } \Gamma_D = \partial\Omega \qquad . \end{aligned}$$

Here L is the differential operator which is supposed to be self-adjoint and elliptic, and l represents the operator of the boundary conditions. We restrict ourselves to bounded, polygonal, three-dimensional or two-dimensional domains $\Omega \subset \mathbb{R}^d$, i.e. $d = 3$ or 2. Variational analysis leads to the so-called variational (or weak) formulation

$$\text{Find } u \in V: \qquad a(u, v) = (f, v) \qquad \forall\, v \in V$$

with a symmetric, elliptic and continuous bilinear form $a(\cdot\,,\,\cdot)$ and a functional $(f,\,\cdot)$. In chapters 3 and 4 (where we are dealing with the Poisson problem and a singularly perturbed reaction diffusion equation, both with homogeneous Dirichlet boundary conditions) the space $V = H^1_o(\Omega)$ is appropriate. Then the *finite element method (FEM)* can be employed to solve this problem numerically. An approximate space $V_{o,h} \subset V$ yields the FEM formulation

$$\text{Find } u_h \in V_{o,h}: \qquad a(u_h, v_h) = (f, v_h) \qquad \forall\, v_h \in V_{o,h} \qquad .$$

To obtain the approximate space $V_{o,h}$ assume a family $\mathcal{F} = \{\mathcal{T}_h\}$ of triangulation $\mathcal{T}_h$ of Ω. Then let $V_{o,h}$ be the space of continuous, piecewise linear functions over $\mathcal{T}_h$ that satisfy the homogeneous Dirichlet boundary conditions.

The finite element method shall be *accurate* and *efficient*. The accuracy is assessed by the error $u - u_h$ in some suitable norm. The efficiency is, roughly speaking, related to the number of elements, the degree of the basis functions, the solution method etc.

Usually the search for an accurate (approximate) solution u_h is an iterative procedure. One constructs a sequence $\{u_h\}$ of FEM solutions whose error $\|u - u_h\|_*$ decreases until a prescribed accuracy is obtained. The well known adaptive process has the form:

1. Estimate the error locally for a solution on a given mesh.

2. Based on this information, construct a new mesh or perform a mesh refinement.

3. Solve the arising finite element system.

The topic of our work is a special class of problems which can be solved very efficiently by a non-classical finite element method. Some boundary value problems (arising e.g. from fluid dynamics, weather simulation etc.) yield a solution which exhibits an almost one-dimensional behaviour, i.e. the solution varies significantly only in one direction but remains almost constant in other directions. Such solutions are called *anisotropic*. Examples include solutions with a boundary or an interior layer.

One feature of the classical finite element method is that the ratio of the diameters of the circumscribed and inscribed spheres of a finite element (e.g. rectangle, tetrahedron, or cube) is bounded. Such meshes are referred to as *isotropic* meshes. But when an anisotropic solution as mentioned above occurs it is sensible to violate this condition and to use highly stretched elements instead. One hopes to capture in this way the important features of the solution with much less elements. Numerical evidence confirms that problems with anisotropic solutions can indeed be solved much more efficiently on anisotropic meshes.

A (certainly incomplete) list of engineers and scientists dealing with such anisotropic problems include Beinert and Kröner [10], Fröhlich, Lang and Roitzsch [16], Kornhuber and Roitzsch [17], Nochetto [20], Peraire et al.[21], Rachowicz [22], Rick, Greza and Koschel [24], Siebert [28], Vilsmeier, Hänel et al. [33], Zienkiewicz and Wu [35]. Anisotropic finite element methods with emphasis on *a priori* error estimation have been considered for example by Apel and Dobrowolski [3], Apel and Lube [4], Apel and Nicaise [5], Miller, O'Riordan and Shishkin [18], Roos [26], Zhou and Rannacher [34]. But although anisotropic finite elements are used, its theoretical foundation is much weaker than for isotropic elements.

An adaptive strategy that takes account of an *anisotropic* solution clearly involves the following tasks.

1. Estimate the error for a solution on a given mesh.
2. Obtain information for a new, better mesh. This includes:
   - Detect regions of anisotropic behaviour of the solution.
   - Determine a (quasi) optimal aspect ratio and stretching direction of the finite elements.
   - Determine the element size.
3. Based on this information, construct a new mesh or perform a mesh refinement.
4. Solve the arising finite element system.

Obviously, every adaptive process has to answer questions 1 and 2 in some form. Yet explicit and analytically based error estimators or indicators (like in [28]) are rather rare; often estimators/indicators are hidden behind some refinement criterion or are derived by heuristic considerations [10, 21, 22, 24].

Similarly, information of the anisotropic solution is often drawn from heuristic arguments. This includes the analysis of the partial second derivatives [21, 24, 35], of the level lines [17] or of the gradient (or gradient jump) of some values [10, 22, 28].

The next step, namely the remeshing, is done either by mesh refinement and adjustment (see, e.g. [10, 11, 16, 17, 22, 24]), or a new mesh generation (for example coupled with a background mesh, or by means of a virtual transformation, [12, 24, 33]).

Finally, the solution of the resulting system does not seem to be too difficult compared with these first three steps.

Our work focuses mainly on the first task, the *a posteriori* error estimators, although some aspects related to the other steps are discussed occasionally.

On isotropic meshes the theory of error estimators is fairly well established (see e.g. [1, 7, 8, 9, 31, 36, 37]. On anisotropic meshes such estimators cannot be applied (e.g. because an anisotropic element T cannot be described by a single element size h_T), or they were not investigated yet. To our knowledge, the only mathematically exact estimator is due to Siebert [28]. He considers the Poisson equation, utilizes cuboidal, rectangular or prismatic grids (triangulations) and modifies the well known residual error estimator.

The aim of this work is threefold. Firstly we want to derive estimators for tetrahedral and triangular grids because of their greater geometrical flexibility. Understandably this requires more effort than for cuboidal grids since tetrahedra do not have three natural directions, and since they can not be aligned with the coordinate axes (in general).

Secondly several estimators (residual error estimators, local problem error estimators, Zienkiewicz-Zhu like error estimators) and different norms (energy norm, L_2 norm) are considered.

Thirdly, several differential equations are investigated into. The Poisson equation, being one of the simplest boundary value problems, is chosen to identify and study the effects of anisotropic finite elements. The singularly perturbed reaction diffusion equation in chapter 4 further reveals properties due to the anisotropy but also features that are related to the governing equation. Furthermore, this example shall show (or at least indicate) that an anisotropic theory can be applied to (almost) real life problems.

The paper is organized as follows. At the end of this introduction a list of commonly used symbols is given. In chapter 2 the notation is introduced and basic relations are derived. Chapters 3 and 4 are devoted to the Poisson equation and a singularly perturbed reaction diffusion equation, respectively. Numerical examples are briefly discussed in chapter 5. A summary completes this paper.

List of Symbols

The list below comprises important notation accompanied by a brief explanation and (where possible) the page number of its definition or first occurrence. A unified notation for the $\mathbb{R}^3$ and the $\mathbb{R}^2$ is used as far as it is unambiguous. Note however that some meanings are different (e.g. T denotes either a tetrahedron or a triangle).

Symbol		**page**
Ω	bounded polygonal domain in the $\mathbb{R}^3$ (or the $\mathbb{R}^2$)	
d	dimension of $\Omega (d = 2, 3)$	
$\Gamma_D = \partial\Omega$	Dirichlet boundary of Ω	
$\mathbf{e}_i$	unitary vectors of $\mathbb{R}^d$	9
L_2, H^1, H^1_o	usual Sobolev spaces over Ω	
$(\cdot,\cdot)$, $(\cdot,\cdot)_\omega$	$L_2(\Omega)$ scalar product or $L_2(\omega)$ scalar product	
$a(\cdot\,,\,\cdot)$	bilinear form	21, 59
$\Vert\cdot\Vert$	L_2 norm over Ω	
$\Vert\cdot\Vert_\omega$, $\Vert\cdot\Vert_E$	L_2 norm over a domain ω or a face E	
$\Vert\cdot\Vert_{\mathbb{R}^{3\times 3}}$	spectral norm of a matrix	12
$\mathcal{T}_h$	triangulation of Ω	15
$V_h, V_{o,h}$	finite element spaces over $\mathcal{T}_h$	21
$\mathbb{P}^m(\omega)$	space of polynomials of degree $\leq m$ over domain ω	17
$T \in \mathcal{T}_h$	tetrahedron (or triangle)	
$\vert T\vert$	volume $\mathrm{meas}_d(T)$ of T	
$\mathbf{p}_i$	special vectors of T, $i = 1 \dots d$	10
$h_i = h_{i,T}$	length $\vert\mathbf{p}_i\vert$ of vector $\mathbf{p}_i$	10
$h_{min,T}$	$:= \min\{h_{i,T}\} = h_{d,T}$	10
$h_i(\mathbf{x})$	global function that has value $h_{i,T}$ over T	10
E, E_T	arbitrary face of T (or edge of a triangle T)	10
$\vert E\vert$	$\mathrm{meas}_{d-1}(E)$ of E	10
$h_E, h_{E,T}$	length of the height over E in a tetrahedron T	10
ω_T, ω_E	auxiliary local subdomains	14
$\lambda_{T,i}$	barycentric coordinates of T	18
b_T	element bubble function (related to T)	18
	special L_2 element bubble function	45
b_E	face bubble function (related to E)	18
	special L_2 face bubble function	46
F_{ext}	extension operator $F_{ext} : \mathbb{P}^0(E) \to \mathbb{P}^0(T)$	18
$\bar{T}$	standard tetrahedron (see definition)	11
$\hat{T}$	reference tetrahedron	11
F_A	affine linear mapping from standard tetrahedron $\bar{T}$ onto T	11
F_C	affine linear mapping from reference tetrahedron $\hat{T}$ onto T	11
A_T, C_T	transformation matrices of the maps F_A and F_C	11
$C(\mathbf{x})$	global matrix function that coincides with C_T over T	11
$\tilde{D}_i$	(unitary) directional derivative in the direction $\mathbf{p}_i$	13

Symbol		**page**
$H^1_{\mathcal{T}}(\Omega), H^1_{o,\mathcal{T}}(\Omega)$	sets of adapted functions	23
φ_j	linear basis function of node a_j, $\varphi_j(a_i) = \delta_i^j$ (Kronecker symbol)	27
R_o	Clément interpolation operator $H^1_{o,\mathcal{T}}(\Omega) \mapsto V_{o,h}$	30
P_{L_2}	L_2 projection onto piecewise constant functions	31
$r_E(v_h)$	(scalar) gradient jump of a function $v_h \in V_h$ across a face E	32
$r_T(v_h)$	element residual (problem dependent)	31, 63
$D_{h,m}(v_h)$	discrete, mesh dependent norm representing the jump residuals	32
$\partial/\partial n$	directional derivative with respect to the outer normal unit vector n	33
$\eta_{R,T}$	residual error estimator (energy norm), problem dependent	32, 63
$\eta_{D,T}$	local Dirichlet problem error estimator (energy norm)	37
$\eta_{R,L_2,T}$	L_2 residual error estimator	54
$\vert\vert\vert\cdot\vert\vert\vert$	energy norm $\vert\vert\vert\cdot\vert\vert\vert^2 = a(\cdot,\cdot)$	60

Constants are denoted by c. They are generic constants, i.e. always independent of the underlying triangulation or the function in question, and may have different values at different occurrences. We write

$$\begin{array}{lcll} x \lesssim y & \Longleftrightarrow & x \le c \cdot y & \\ x \gtrsim y & \Longleftrightarrow & x \ge c \cdot y, & c > 0 \\ x \sim y & \Longleftrightarrow & \underline{c} \cdot x \le y \le \overline{c} \cdot x, & \underline{c} > 0 \end{array} .$$

The notation with an explicit constant c is used only when a dependence on some other values is expressed (e.g. c_Ω) or when further details are thus revealed. Also, the (sharper) notation $\le$ is used instead of $\lesssim$ wherever possible.

Chapter 2

Preliminaries

2.1 Notation

2.1.1 General notation

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ be an open, bounded, polygonal domain over which the differential equation is posed. The following spaces are frequently employed (with ω being an arbitrary open domain).

$C^k(\overline{\omega})$	space of k times continuously differentiable functions
$L_2(\omega)$	space of square integrable functions
$H^k(\omega)$	Sobolev space of functions whose k^{th} derivative is in $L_2(\omega)$
$H^k_o(\omega)$	Sobolev space of functions of $H^k(\omega)$ satisfying the corresponding homogeneous Dirichlet boundary conditions
$\mathbb{P}^k(\omega)$	space of polynomials of order k or less.

All norms of functions are L_2 norms unless otherwise stated. A norm without subscript denotes $\|\cdot\| = \|\cdot\|_{L_2(\Omega)}$, i.e. the L_2 norm over the whole domain Ω. All vectors norms are Euclidean norms, and norms of matrices are spectral norms. The unitary vectors of $\mathbb{R}^d$ are denoted by $\mathbf{e}_i$, $i = 1 \dots d$.

All considerations are made for the three-dimensional case. The application to the easier two-dimensional case is readily possible.

2.1.2 Notation of the tetrahedron

Assume that a triangulation $\mathcal{T}_h$ (also called a mesh or a grid) is given which satisfies the usual conformity conditions (see Ciarlet [13], Chapter 2). Let T be an arbitrary tetrahedron thereof. For this tetrahedron the following notation is introduced. The four vertices of T are denoted by $P_0, \dots, P_3$ according to these three conditions:

- Let P_0P_1 be the longest edge of T.
- There exist two triangles that contain the edge P_0P_1. The one with largest area is denoted by $\triangle P_0P_1P_2$.
- Let P_0P_2 be the shortest edge of $\triangle P_0P_1P_2$. This determines which vertex is P_0 and P_1, respectively. Let P_3 be the remaining vertex.

This notation is not uniquely determined if, for instance, T has two edges which are simultaneously the longest ones. However, it turns out that then either choice of the notation fits into the theory. Additionally we define three vectors:

- $\mathbf{p}_1 := \overrightarrow{P_0 P_1}$.
- Let $\mathbf{p}_2$ be that vector in the plane of $P_0P_1P_2$ that points to P_2 and that is perpendicular to $\mathbf{p}_1$
- Let $\mathbf{p}_3$ be that vector to P_3 that is perpendicular to $\triangle P_0P_1P_2$.

Hence $\mathbf{p}_1 \ldots \mathbf{p}_3$ are mutually orthogonal. Figure 2.1 visualizes this notation.

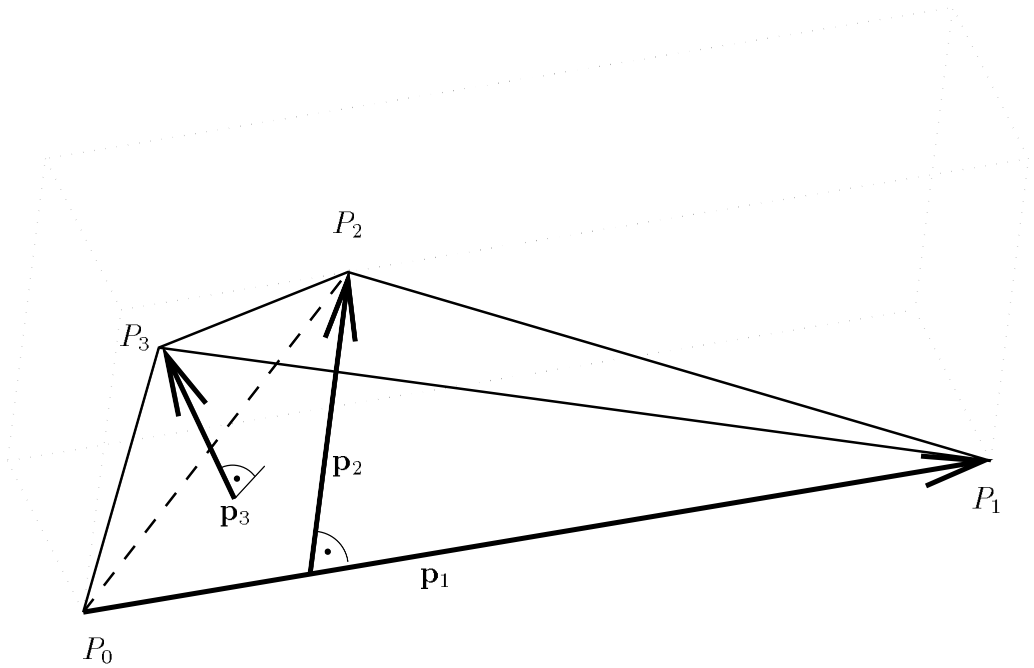


Figure 2.1: Notation of tetrahedron T

The length of the vectors $\mathbf{p}_i$ is denoted by $h_i = h_{i,T} := |\mathbf{p}_i|$, $i = 1, 2, 3$. Because of the definition of the P_i we conclude immediately $h_1 > h_2 \geq h_3$. We further define

$$h_{min,T} \ := \ \min_{i=1\ldots d} \{h_{i,T}\} \ = \ h_{d,T}$$

(in $\mathbb{R}^3$ thus $h_{min,T} = h_{3,T}$ holds). Furthermore a piecewise constant function $h_i(\mathbf{x})$ is defined for almost all $\mathbf{x} \in \Omega$ according to

$$h_i(\mathbf{x}) := h_{i,T} \qquad \text{for } \mathbf{x} \in T\,,\ i = 1 \ldots d.$$

Analogously $h_{min}(\mathbf{x})$ is defined.

The boundary of a tetrahedron T consists of four faces (i.e. triangles). Such a face is denoted by E, and its $(d-1)$ dimensional content is expressed by $|E| := \text{meas}_{d-1}(E)$. The length of the height over such a face E will be denoted by $h_{E,T} = h_E = 3|T|/|E|$.

2.1.3 Standard and reference tetrahedron, transformations and their properties

Let T be an arbitrary but fixed tetrahedron. Mainly we will employ two affine linear mappings F_A and F_C which will be defined as follows.

Let $\vec{P_0}$ be the (column) vector from the origin of the coordinate system to P_0, and let $\overrightarrow{P_0P_i}$ be the (column) vectors from P_0 to P_i, $i = 1, 2, 3$. We define the matrices $A_T, C_T \in \mathbb{R}^{3\times 3}$ by

$$A_T := \left(\overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \overrightarrow{P_0P_3} \right) \qquad \text{and} \qquad C_T := \left(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\right) \qquad . \tag{2.1}$$

Sometimes we want to refer to the matrix C_T not only on an actual tetrahedron T but on a larger domain. Thus we introduce a matrix (or more precisely a matrix function) $C(\mathbf{x})$ which is defined globally for almost all $\mathbf{x} \in \Omega$ and which coincides with C_T on a tetrahedron T:

$$C(\mathbf{x}) := C_T \qquad \text{for } \mathbf{x} \in T \qquad . \tag{2.2}$$

Additionally a matrix H_T is defined by

$$H_T := \operatorname{diag}(h_1, h_2, h_3) \qquad .$$

Let now the affine linear mappings be

$$\begin{aligned} &F_A : \bar{T} \mapsto T && \text{and} && F_C : \hat{T} \mapsto T \\ &F_A : \mathbf{x}(\boldsymbol{\mu}) = A_T \cdot \boldsymbol{\mu} + \vec{P_0} && && F_C : \mathbf{x}(\boldsymbol{\mu}) = C_T \cdot \boldsymbol{\mu} + \vec{P_0} \end{aligned}$$

with $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)^T$.

Definition 2.1 (Standard tetrahedron and reference tetrahedron) *The* standard tetrahedron $\bar{T}$ *has vertices* $\bar{P}_0 = (0,0,0)^T$ *and* $\bar{P}_i = \mathbf{e}_i^T \quad, i = 1 \ldots d$. *Enumerate the faces* $\bar{E}_i$ *of* $\bar{T}$ *such that*

$$\bar{E}_i := \bar{T} \cap \{x_i = 0\} \quad , i = 1 \ldots d \qquad \textit{and} \qquad \bar{E}_0 := \bar{T} \cap \{|\mathbf{x}|_1 = 1\} \qquad ,$$

i.e. face $\bar{E}_i$ *is opposite the vertex* $\bar{P}_i$.

The reference tetrahedron $\hat{T}$ *is defined implicitly by the mapping* F_C, *i.e.* $\hat{T} = F_C^{-1}(T)$.

The vertices of $\hat{T}$ are $\hat{P}_0 = (0,0,0)^T$, $\hat{P}_1 = (1,0,0)^T$, $\hat{P}_2 = (\hat{x}_2, 1, 0)^T$ and $\hat{P}_3 = (\hat{x}_3, \hat{y}_3, 1)^T$ because of the definition of F_C. The conditions on the P_i yield immediately $0 < \hat{x}_2, \hat{x}_3 < 1$ and $-1 < \hat{y}_3 < 1$. Figures 2.1 and 2.2 may illustrate this definition (the circumscribed rectangular prisms shall facilitate the visualization).

Variables that are related to the standard tetrahedron $\bar{T}$ and the reference tetrahedron $\hat{T}$ are referred to with a bar and a hat, respectively (e.g. $\bar{\nabla}$, $\hat{v}$).

The determinants of both mappings are

$$|\det(A_T)| = |\det(C_T)| = h_1 \cdot h_2 \cdot h_3 = 6 \cdot |T| \qquad .$$

The transformed derivatives satisfy

$$\bar{\nabla}\bar{v} = A_T^T \nabla v \qquad \text{and} \qquad \hat{\nabla}\hat{v} = C_T^T \nabla v \qquad .$$

In order to bound the norms of some transformation matrix we state the following simple lemma (see also [13]).

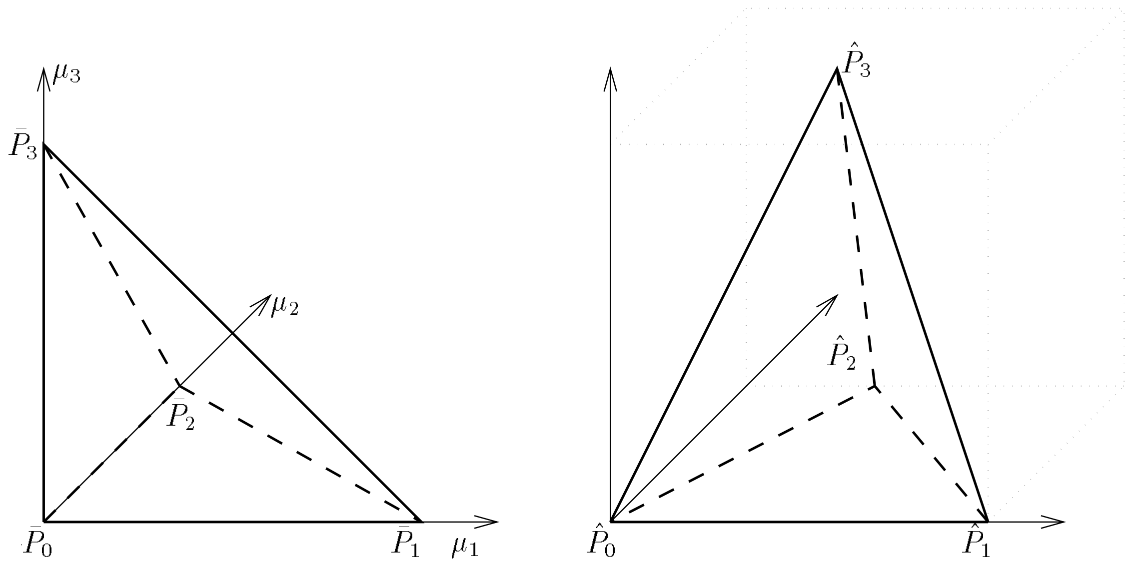


Figure 2.2: Standard tetrahedron $\bar{T}$ and reference tetrahedron $\hat{T}$

Lemma 2.1 (Bound of the norm of a transformation matrix) *Let A be a linear transformation that maps the (closed) domain $\hat{G} \subset \mathbb{R}^d$ onto G. The spectral norm of the corresponding transformation matrix satisfies*

$$\|A\|_{\mathbb{R}^{d\times d}} \leq d(G)/\varrho(\hat{G})$$

with $\quad d(G) := \max\limits_{x,y\in G} \|x-y\|_{\mathbb{R}^3} \quad$ *and* $\quad \varrho(\hat{G}) :=$ *diameter of the largest sphere* $S \subset \hat{G}$*.*

Lemma 2.2 (Norms of some matrices) *The following relations hold.*

$$\|A_T^T C_T^{-T}\|_{\mathbb{R}^{3\times3}} = \|C_T^{-1}A_T\|_{\mathbb{R}^{3\times3}} \lesssim 1 \tag{2.3}$$
$$\|C_T^T A_T^{-T}\|_{\mathbb{R}^{3\times3}} = \|A_T^{-1}C_T\|_{\mathbb{R}^{3\times3}} \lesssim 1 \tag{2.4}$$
$$\|C_T H_T^{-1}\|_{\mathbb{R}^{3\times3}} = \|H_T C_T^{-1}\|_{\mathbb{R}^{3\times3}} = 1 \tag{2.5}$$
$$\|H_T^{-1}\|_{\mathbb{R}^{3\times3}} = \|C_T^{-1}\|_{\mathbb{R}^{3\times3}} = h_{min,T}^{-1} \tag{2.6}$$
$$\|A_T^{-1}\|_{\mathbb{R}^{3\times3}} \sim h_{min,T}^{-1} \tag{2.7}$$

Proof: Let $T-\vec{P_0}$ be the tetrahedron T shifted by $-\vec{P_0}$. The mappings A_T, C_T^{-1} and $C_T^{-1}A_T$ act as follows:

$$\bar{T} \stackrel{A_T}{\longmapsto} (T-\vec{P_0}) \stackrel{C_T^{-1}}{\longmapsto} \hat{T} \qquad \text{and thus} \qquad \bar{T} \stackrel{C_T^{-1}A_T}{\longmapsto} \hat{T},$$

i.e. $C_T^{-1}A_T$ maps the standard tetrahedron $\bar{T}$ onto the reference tetrahedron $\hat{T}$. The lemma from above implies immediately (2.3) and analogously (2.4).

Because of $C_T^T \cdot C_T = H_T^2$ from (2.1) we conclude $(H_T C_T^{-1})^T \cdot H_T C_T^{-1} = I$ and $(C_T H_T^{-1})^T \cdot C_T H_T^{-1} = I$. Thus (2.5) is derived. Note that $\|C_T^{-1}H_T\|_{\mathbb{R}^{3\times3}} \neq 1$.

The equality $\|H_T^{-1}\|_{\mathbb{R}^{3\times3}} = h_{min,T}^{-1}$ is obvious. Equality (2.6) follows immediately from $\|C_T^{-1}\|_{\mathbb{R}^{3\times3}}^2 = \lambda_{\max}(C_T^{-1}C_T^{-T}) = \lambda_{\max}(H_T^{-2})$.

The inequalities $\|C_T^{-1}\|_{\mathbb{R}^{3\times3}} = \|C_T^{-1}A_T \cdot A_T^{-1}\|_{\mathbb{R}^{3\times3}} \leq \|C_T^{-1}A_T\|_{\mathbb{R}^{3\times3}} \cdot \|A_T^{-1}\|_{\mathbb{R}^{3\times3}}$ and $\|A_T^{-1}\|_{\mathbb{R}^{3\times3}} = \|A_T^{-1}C_T \cdot C_T^{-1}\|_{\mathbb{R}^{3\times3}} \leq \|A_T^{-1}C_T\|_{\mathbb{R}^{3\times3}} \cdot \|C_T^{-1}\|_{\mathbb{R}^{3\times3}}$ and (2.3)–(2.6) finally imply (2.7). ∎

Finally, a norm $\|\cdot\|_T$ over an actual tetrahedron T is often transformed into a norm over the standard tetrahedron $\bar{T}$ or the reference tetrahedron $\hat{T}$. The following relations hold. Let $v \in L_2(T)$ and $T \subset \mathbb{R}^3$. For a mapping $F_A(\boldsymbol{\mu}) = A_T \cdot \boldsymbol{\mu} + \vec{P_0}$ one obtains

$$\begin{aligned} \int_T v^2(\mathbf{x}) d\mathbf{x} &= \int_{\bar{T}} \bar{v}^2(\boldsymbol{\mu}) \cdot |\det\, A_T|\, d\boldsymbol{\mu} = 6|T| \cdot \int_{\bar{T}} \bar{v}^2(\boldsymbol{\mu}) d\boldsymbol{\mu} \\ \text{or} \qquad \|v\|_T &= \sqrt{6|T|} \cdot \|\bar{v}\|_{\bar{T}} \\ \text{and similarly} \qquad \|v\|_T &= \sqrt{6|T|} \cdot \|\hat{v}\|_{\hat{T}} \\ \text{and} \qquad \|v\|_E &= \sqrt{|E|/|\bar{E}|} \cdot \|\bar{v}\|_{\bar{E}} \qquad . \end{aligned}$$

2.1.4 The directional derivative $\tilde{D}_i$

In order to motivate the derivatives $\tilde{D}_i$ consider rectangular or cuboidal finite elements (cf. [28]). There are three (or two) natural directions that correspond to the coordinate axes. The partial derivatives that correspond to these axes too are thus sufficient for an error analysis.

In contrast to this a tetrahedron or a triangle does not possess these natural directions. However the (normalized) directions $\mathbf{p}_1, \mathbf{p}_2$, and $\mathbf{p}_3$ that correspond to C_T will prove to be useful. This leads to the following definition.

Definition 2.2 (Directional derivative) *Let v be a function in $H^1(T)$. The* directional derivative $\tilde{D}_{i,T}$ *is defined by*

$$\begin{pmatrix} \tilde{D}_{1,T}\, v \\ \tilde{D}_{2,T}\, v \\ \tilde{D}_{3,T}\, v \end{pmatrix} := H_T^{-1} C_T^T \cdot \nabla v \qquad , v \in H^1(T). \tag{2.8}$$

Here this derivative $\tilde{D}_{i,T}$ is defined for a fixed tetrahedron T. Hence we introduce a derivative $\tilde{D}_i$ which is defined globally for almost all $\mathbf{x} \in \Omega$, and which coincides with $\tilde{D}_{i,T}$ on a tetrahedron T:

$$\tilde{D}_i\, v(\mathbf{x}) := \tilde{D}_{i,T}\, v(\mathbf{x}) \qquad \text{for} \quad \mathbf{x} \in T \qquad .$$

Note that this derivative $\tilde{D}_i$ depends on the triangulation $\mathcal{T}_h$, and it is defined separately over each tetrahedron T.

When considering each component in the definition above, the directional derivative is equivalent to

$$\tilde{D}_{i,T} v = h_i^{-1} \cdot (\mathbf{p}_i, \nabla v) \qquad i = 1 \ldots d$$

i.e. $\tilde{D}_{i,T}$ is the (unitary) directional derivative along the direction $\mathbf{p}_i$.

The orthogonality of the vectors $\mathbf{p}_i$ and the definition $h_i = |\mathbf{p}_i|$ implies that $H_T^{-1} C_T^T$ is an orthogonal matrix. Thus

$$\sum_{i=1}^{d} (\tilde{D}_i v)^2 = |\nabla v|^2 \qquad \text{or} \qquad \|H_T^{-1} C_T^T \nabla v\|_T = \|\nabla v\|_T \tag{2.9}$$

and
$$\sum_{i=1}^{d} h_i^2 (\tilde{D}_i v)^2 = |C_T^T \nabla v|^2 \qquad \sum_{i=1}^{d} h_i^2 \|\tilde{D}_{i,T} v\|_T^2 = \|C^T \nabla v\|_T^2. \tag{2.10}$$

The last equations indicate that derivatives can be written either component-wise in terms of $\tilde{D}_i$, or they can be written in the compact form of $C_T^T \nabla$.

In this work all results and proofs are given in this compact form since $C_T^T \nabla$ on the actual tetrahedron T is related (via F_C) directly to $\hat{\nabla}$ on the reference tetrahedron $\hat{T}$. Main results however are also given in terms of $\tilde{D}_i$ for two reasons. Firstly this might facilitate the understanding of the underlying principles, and secondly an extension to rectangular or cuboidal finite elements is then readily possible.

With the help of (2.9) and (2.10) any term involving derivatives can be expressed easily in either form.

2.1.5 Auxiliary subdomains

Two auxiliary subdomains that occur in many estimates are defined now. Let $T \in \mathcal{T}_h$ be an arbitrary tetrahedron. Let ω_T be that domain that is formed by T and all (at most four) adjacent tetrahedra that have a common face with T:

$$\omega_T := \bigcup_{T' \cap T = E} T' \quad .$$

Let E be an inner face (triangle) of $\mathcal{T}_h$, i.e. there are two tetrahedra T_1 and T_2 having the common face E. Let the domain $\omega_E := T_1 \cup T_2$. Figure 2.3 depicts both domains for the two-dimensional and the three-dimensional case.

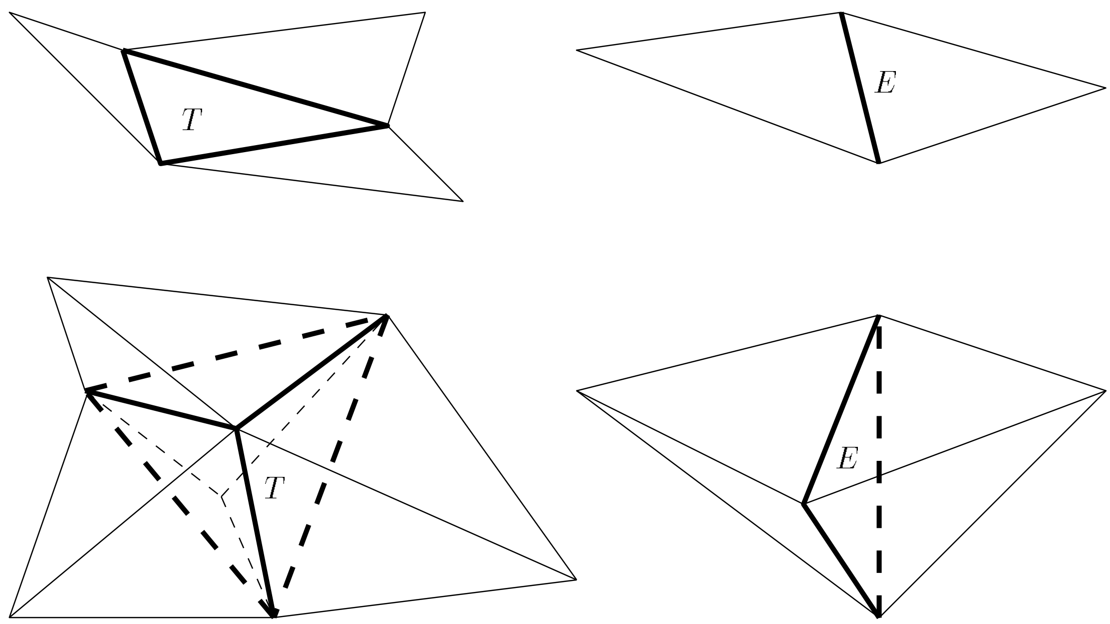


Figure 2.3: Auxiliary domains ω_T (left) and ω_E (right) for $\Omega \subset \mathbb{R}^2$ and $\Omega \subset \mathbb{R}^3$

2.2 Requirements on the mesh

Let $a_1, \ldots, a_N$ be the nodes of the triangulation $\mathcal{T}_h$. In addition to the usual conformity conditions of the mesh (see Ciarlet [13], Chapter 2) we demand the following assumptions.

1. The number of tetrahedra that contain the node a_j is bounded.

2. The dimensions of adjacent tetrahedra must not change rapidly, i.e.

$$h_{i,T'} \sim h_{i,T} \qquad \forall\, T, T' \text{ with } T \cap T' \neq \emptyset\,,\; i = 1 \ldots d \quad . \tag{2.11}$$

Remark 2.1 Assume that T and T' are adjacent tetrahedra. If in any inequality the terms $h_{i,T}$ or $h_{i,T'}$ occur then assumption (2.11) implies that both terms can be exchanged mutually. The inequality constants are then still independent of T or $\mathcal{T}_h$. This feature is exploited to some extend, e.g. for inequalities written component-wise.

Assumption (2.11) implies in particular that we can use a term $h_{i,T} = h_i$ for describing the dimension h_i of a local subdomain like ω_T or ω_E. Additionally, for $\omega_E = T_1 \cup T_2$ we can simply write h_E instead of h_{E,T_1} or h_{E,T_2}. □

2.3 Basic tools

In this section some basic tools and inequalities are listed.

2.3.1 Anisotropic trace inequalities

The first trace inequality is readily obtained by standard scaling techniques.

Lemma 2.3 (First trace inequality) *Let T be an arbitrary tetrahedron and E be a face of it. For $v \in H^1(T)$ the trace inequality*

$$\|v\|_E^2 \;\lesssim\; h_E^{-1}\, \left(\|v\|_T^2 + \|C_T^T \nabla v\|_T^2\right) \tag{2.12}$$

holds. The component-wise form is

$$\|v\|_E^2 \;\lesssim\; h_E^{-1} \cdot \left(\|v\|_T^2 + \sum_{i=1}^{d} h_{i,T}^2\, \|\tilde{D}_i v\|_T^2 \right) \qquad .$$

Proof: Consider the transformation F_A, the standard tetrahedron $\bar{T}$, the face $\bar{E}$ of $\bar{T}$, and the function $\bar{v} := v \circ F_A \in H^1(\bar{T})$. The trace theorem gives

$$\|\bar{v}\|_{\bar{E}}^2 \lesssim \|\bar{v}\|_{H^1(\bar{T})}^2 = \|\bar{v}\|_{\bar{T}}^2 + \|\bar{\nabla}\bar{v}\|_{\bar{T}}^2 \qquad .$$

The transformation into the actual tetrahedron (via F_A) yields

$$|E|^{-1} \cdot \|v\|_E^2 \lesssim |T|^{-1} \left(\|v\|_T^2 + \|A_T^T \nabla v\|_T^2\right) \qquad .$$

From (2.3) and (2.10)

$$\|A_T^T \nabla v\|_T = \|A_T^T C_T^{-T} \cdot C_T^T \nabla v\|_T \leq \|A_T^T C_T^{-T}\|_{\mathbb{R}^{3\times 3}} \cdot \|C_T^T \nabla v\|_T \lesssim \|C_T^T \nabla v\|_T$$

can be derived. Utilizing $6\,|T| = |E| \cdot h_E$ results in the trace inequality (2.12). ■

The second, improved trace inequality in the isotropic version (i.e. on the standard tetrahedron) is, to our knowledge, due to Verfürth [32]. We state this inequality and, for self-containment, repeat the proof.

Lemma 2.4 (Second trace inequality) *Let T be an arbitrary tetrahedron and E be a face of it. For $v \in H^1(T)$ the trace inequality*

$$\|v\|_E^2 \;\lesssim\; h_E^{-1} \cdot \|v\|_T \cdot \left(\|v\|_T + \|C_T^T \nabla v\|_T\right) \tag{2.13}$$

holds. The component-wise form is

$$\|v\|_E^2 \;\lesssim\; h_E^{-1} \cdot \|v\|_T \cdot \left(\|v\|_T + \sum_{i=1}^{d} h_{i,T}\, \|\tilde{D}_i v\|_T\right) \qquad .$$

Proof: Again standard scaling arguments will be used. Therefore consider first the standard tetrahedron $\bar{T}$.

Let $\bar{v} \in H^1(\bar{T})$ vanish on $\bar{E}_0$. Then

$$\|\bar{v}\|_{\bar{E}_k}^2 \;\leq\; 2 \cdot \|\bar{v}\|_{\bar{T}} \cdot \|\partial \bar{v}/\partial x_k\|_{\bar{T}}$$

holds for $k = 1 \ldots d$. To derive this consider a fixed index k. Since $\bar{v}$ vanishes on $\bar{E}_0$ we obtain for all $\mathbf{y} \in \bar{E}_k$

$$|\bar{v}(\mathbf{y})|^2 \;=\; |\bar{v}(\mathbf{y})|^2 - |\bar{v}(\mathbf{y} + (1 - |\mathbf{y}|_1)e_k)|^2 \;=\; -2 \int_0^{1-|\mathbf{y}|_1} \bar{v}(\mathbf{y} + te_k) \cdot \frac{\partial}{\partial x_k} \bar{v}(\mathbf{y} + te_k)\mathrm{d}t$$

since $\mathbf{y} + (1 - |\mathbf{y}|_1)e_k \in \bar{E}_0$. Integrating over $\bar{E}_k$, invoking Fubini's theorem and the Cauchy-Schwarz inequality establishes the desired estimate.

Consider now a function $v \in H^1(\bar{T})$ that vanishes on an arbitrary face $\bar{E}_i \quad , 0 \leq i \leq d$. Let $\bar{E}$ be a face of $\bar{T}$. Then

$$\|v\|_{\bar{E}}^2 \;\lesssim\; \|v\|_{\bar{T}} \cdot \|\nabla v\|_{\bar{T}} \qquad .$$

To prove this assume $\bar{E} \neq \bar{E}_i$ since otherwise the inequality is trivial. We employ an affine linear mapping F_i which satisfies

$$\begin{array}{llll} F_i & : & \mathbf{x}(\boldsymbol{\mu}) = F \cdot \boldsymbol{\mu} + \boldsymbol{\mu}_0 & \qquad F \in \mathbb{R}^{d \times d} \\ F_i & : & \bar{T} \mapsto \bar{T} \qquad \text{and} \qquad \bar{E}_0 \mapsto \bar{E}_i & \qquad . \end{array}$$

Assume that the face $\bar{E}_k$ is mapped onto $\bar{E}$, with $k \neq 0$. The function $\bar{v} := v \circ F_i$ vanishes on $\bar{E}_0$ and thus the previous inequality implies

$$\|\bar{v}\|_{\bar{E}_k}^2 \;\leq\; 2 \cdot \|\bar{v}\|_{\bar{T}} \cdot \|\partial \bar{v}/\partial x_k\|_{\bar{T}} \qquad .$$

Lemma 2.1 yields readily $\|F\|_{\mathbb{R}^{d\times d}} \lesssim 1$, and $|\bar{E}|/|\bar{E}_k| \lesssim 1$ is obvious. The transformation back to v results in the desired inequality

$$\|v\|_{\bar{E}}^2 \lesssim \|v\|_{\bar{T}} \cdot \|e_k^T F^T \nabla v\|_{\bar{T}} \leq \|v\|_{\bar{T}} \cdot \|e_k^T F^T\|_{\mathbb{R}^d} \cdot \|\nabla v\|_{\bar{T}} \lesssim \|v\|_{\bar{T}} \cdot \|\nabla v\|_{\bar{T}} \quad .$$

Consider finally an arbitrary function $v \in H^1(\bar{T})$. Let $\bar{E}$ be any of the faces of $\bar{T}$, and enumerate the vertices of $\bar{T}$ such that the vertices of $\bar{E}$ are numbered first. Denote by $\lambda_1 \cdots \lambda_{d+1}$ the barycentric coordinates of $\bar{T}$. Then $\lambda_1 + \cdots + \lambda_d = 1$ on $\bar{E}$, and thus

$$\|v\|_{\bar{E}} \leq \sum_{i=1}^{d} \|\lambda_i \cdot v\|_{\bar{E}} \lesssim \sum_{i=1}^{d} \|\lambda_i \cdot v\|_{\bar{T}}^{1/2} \cdot \|\nabla(\lambda_i \cdot v)\|_{\bar{T}}^{1/2}$$

since $\lambda_i \cdot v$ vanishes on $\bar{E}_i$. The chain rule, the Cauchy-Schwarz inequality, the actual representation of λ_i, and $|\lambda_i| \leq 1$ imply

$$\|\nabla(\lambda_i \cdot v)\|_{\bar{T}}^2 = \sum_{j=1}^{d} \left\| v \cdot \frac{\partial \lambda_i}{\partial x_j} + \lambda_i \cdot \frac{\partial v}{\partial x_j} \right\|_{\bar{T}}^2 \leq 4 \cdot \|v\|_{\bar{T}}^2 + 2 \cdot \|\nabla v\|_{\bar{T}}^2$$

yielding $$\|v\|_{\bar{E}}^2 \lesssim \|v\|_{\bar{T}} \cdot (\|v\|_{\bar{T}} + \|\nabla v\|_{\bar{T}}) \quad .$$

This constitutes the trace inequality on the standard tetrahedron $\bar{T}$. The transformation onto the actual tetrahedron T is completely analogous to the proof of the first trace inequality and therefore omitted here. ∎

2.3.2 Inverse inequalities for finite element functions

In several proofs the following inverse inequalities for a finite element function $v_h \in V_h(T)$ are required. Let $\mathbb{P}^m(\omega)$ be the space of polynomials of degree (at most) m over some domain ω.

Lemma 2.5 *Let T be an arbitrary tetrahedron, E a face of it, and $v_h \in V_h(T) = \mathbb{P}^1(T)$ a finite element function over T. Then*

$$\|C_T^T \nabla v_h\|_T \lesssim \|v_h\|_T \tag{2.14}$$
$$\|v_h\|_{\infty,T} \lesssim |T|^{-1/2} \cdot \|v_h\|_T \tag{2.15}$$
$$\|v_h\|_{\infty,E} \lesssim |E|^{-1/2} \cdot \|v_h\|_E \tag{2.16}$$

hold. The component-wise form of (2.14) is

$$\|\tilde{D}_i v_h\|_T \lesssim h_{i,T}^{-1} \cdot \|v_h\|_T \qquad i = 1 \ldots d.$$

Proof: The proofs are again based on the transformation technique. The norm inequality

$$\|\bar{\nabla}\bar{v}_h\|_{\bar{T}} \lesssim \|\bar{v}_h\|_{\bar{T}} \qquad \forall\, \bar{v}_h \in V_h(\bar{T}) = \mathbb{P}^1(\bar{T})$$

holds on the standard tetrahedron $\bar{T}$ since both norms act on the finite dimensional space $\mathbb{P}^1(\bar{T})$. The transformation via F_A gives for an arbitrary $v_h \in V_h$

$$\|A_T^T \nabla v_h\|_T \lesssim \|v_h\|_T \quad .$$

From (2.4) we obtain

$$\|C_T^T \nabla v_h\|_T = \|C_T^T A_T^{-T} \cdot A_T^T \nabla v_h\|_T \leq \|C_T^T A_T^{-T}\|_{\mathbb{R}^{3\times 3}} \cdot \|A_T^T \nabla v_h\|_T \lesssim \|A_T^T \nabla v_h\|_T$$

and thus the first inequality.

The next two inequalities are derived analogously. From

$$\|v_h\|_{\infty,T}^2 \;=\; \|\bar{v}_h\|_{\infty,\bar{T}}^2 \lesssim \|\bar{v}_h\|_{\bar{T}}^2 = |\det A_T|^{-1} \cdot \|v_h\|_T^2$$

one readily obtains (2.15). For a face E one similarly derives

$$\|v_h\|_{\infty,E} \lesssim |E|^{-1/2} \cdot \|v_h\|_E \qquad .$$

Note that, strictly speaking, inequalities (2.15) and (2.16) constitute norm equivalences (over finite dimensional spaces) whereas (2.14) does not. ■

2.3.3 Bubble functions and their inverse inequalities

Bubble functions and the so-called inverse inequalities related to it play a vital role in our finite element error analysis. Of course different bubble functions can (and have to be) employed for different classes of problems and norms involved. Nevertheless we define here the probably most versatile and commonly used bubble functions. The corresponding inverse inequalities are given and proved.

Other bubble functions that are utilized for an L_2 estimate alone are introduced in the appropriate section 3.4

Let $T \in \mathcal{T}_h$ be an arbitrary tetrahedron, and denote by $\lambda_{T,1}, \cdots, \lambda_{T,4}$ its barycentric coordinates. The *element bubble function* $b_T \in \mathbb{P}^4(T)$ is defined by

$$b_T := 256\, \lambda_{T,1} \cdot \lambda_{T,2} \cdot \lambda_{T,3} \cdot \lambda_{T,4} \qquad \text{on } T \qquad . \tag{2.17}$$

Let E be an inner face (triangle) of $\mathcal{T}_h$, and let T_1 and T_2 be the two tetrahedra that contain E. Enumerate the vertices of T_1 and T_2 such that the vertices of E are numbered first. The *face bubble function* b_E is then defined by

$$b_E := 27\, \lambda_{T_i,1} \cdot \lambda_{T_i,2} \cdot \lambda_{T_i,3} \qquad \text{on } T_i, i = 1,2 \qquad . \tag{2.18}$$

This definition is extended in the obvious way for boundary faces $E \subset \partial\Omega$, i.e. b_E is then defined only on one tetrahedron. For simplicity assume that b_T and b_E are extended by zero outside their original domain of definition. Note that b_E is piecewise cubic on ω_E. Both bubble functions satisfy

$$0 \le b_T(\mathbf{x}), b_E(\mathbf{x}) \le 1 \qquad , \qquad \max b_T = \max b_E = 1 \qquad .$$

The following examples of the corresponding two-dimensional bubble functions give some impression of their shape.

Let E be a face (triangle) of a tetrahedron T. An extension operator $F_{ext} : \mathbb{P}^0(E) \to \mathbb{P}^0(T)$ is given by

$$F_{ext}(\varphi)(\lambda_{T,1}, \lambda_{T,2}, \lambda_{T,3}, \lambda_{T,4}) \;:=\; \varphi(x) = const \qquad \text{for any } x \in E \qquad . \tag{2.19}$$

Then the following inverse inequalities hold.

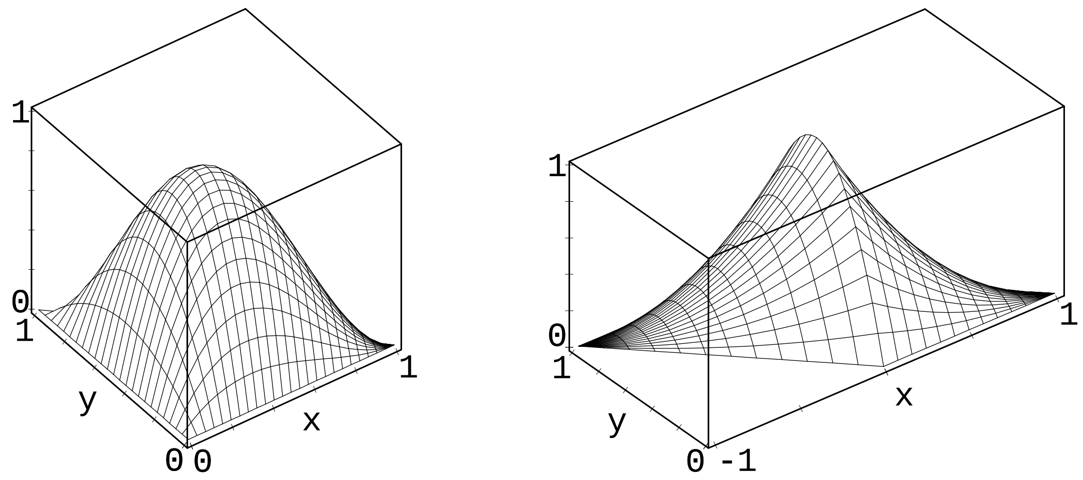


Figure 2.4: Element bubble function b_T and face bubble function b_E (in $\mathbb{R}^2$)

Lemma 2.6 (Inverse inequalities for bubble functions)
Assume that $\varphi_T \in \mathbb{P}^0(T)$ *and* $\varphi_E \in \mathbb{P}^0(E)$. *Then*

$$\|\varphi_T\|_T \sim \|b_T^{1/2} \cdot \varphi_T\|_T \tag{2.20}$$
$$\|\nabla(b_T \cdot \varphi_T)\|_T \lesssim h_{min,T}^{-1} \cdot \|\varphi_T\|_T \tag{2.21}$$
$$\|\varphi_E\|_E \sim \|b_E^{1/2} \cdot \varphi_E\|_E \tag{2.22}$$
$$\|F_{ext}(\varphi_E) \cdot b_E\|_T \lesssim h_E^{1/2} \cdot \|\varphi_E\|_E \qquad \text{for } E \in T \tag{2.23}$$
$$\|\nabla(F_{ext}(\varphi_E) \cdot b_E)\|_T \lesssim h_E^{1/2} \cdot h_{min,T}^{-1} \cdot \|\varphi_E\|_E \qquad \text{for } E \in T \tag{2.24}$$

Proof: For all inequalities the transformation technique is applied.

Obviously $\|b_{\bar{T}}^{1/2} \cdot \|_{\bar{T}}$ and $\| \cdot \|_{\bar{T}}$ are equivalent norms on the finite dimensional space $\mathbb{P}^0(\bar{T})$. The transformation from the unitary tetrahedron $\bar{T}$ to the actual tetrahedron T leads directly to (2.20). Inequality (2.22) is proven in exactly the same way.

Similar to (2.14) one obtains the inequality $\|C_T^T \nabla\psi\|_T \lesssim \|\psi\|_T$ for all $\psi \in \mathbb{P}^4(T)$. Together with (2.9) we derive

$$\|\nabla(b_T \cdot \varphi_T)\|_T = \|H_T^{-1} C_T^T \nabla(b_T \cdot \varphi_T)\|_T \lesssim h_{min,T}^{-1} \cdot \|C_T^T \nabla(b_T \cdot \varphi_T)\|_T \lesssim h_{min,T}^{-1} \cdot \|b_T \cdot \varphi_T\|_T.$$

This proves (2.21) since $0 \leq b_T \leq 1$. The inequality

$$\|F_{ext}(\varphi_E) \cdot b_E\|_T^2 = 6|T| \cdot \|F_{ext}(\bar{\varphi}_E) \cdot b_{\bar{E}}\|_{\bar{T}}^2 \lesssim 6|T| \cdot \|\bar{\varphi}_E\|_{\bar{E}}^2 = h_E \cdot \|\varphi_E\|_E^2$$

holds for all $\varphi_E \in \mathbb{P}^0(E)$ since $\| \cdot \|_{\bar{E}}$ and $\|b_{\bar{E}} \cdot F_{ext}(\cdot)\|_{\bar{T}}$ are equivalent norms over a finite dimensional space of polynomials. Thus (2.23) is proven. Finally (2.24) is obtained using the same techniques as for (2.21) and (2.23). ∎

An even stronger result which will be useful occasionally is given in lemma 2.7.

Lemma 2.7 (Equivalence relations for bubble functions)

$$\|\nabla b_T\|_T \sim h_{min,T}^{-1} \cdot |T|^{1/2} \qquad , \tag{2.25}$$
$$\|\nabla b_E\|_T \sim h_{min,T}^{-1} \cdot |T|^{1/2} \qquad \forall\, E \in T \qquad . \tag{2.26}$$

Proof: Standard scaling arguments and $C_T^T \cdot C_T = H_T^2$ imply

$$\begin{aligned}
\|\nabla b_T\|_T^2 &= 6|T| \cdot \|C_T^{-T} \hat{\nabla} \hat{b}_T\|_{\hat{T}}^2 \\
&= 6|T| \cdot \int_{\hat{T}} \hat{\nabla}^T \hat{b}_T \cdot C_T^{-1} C_T^{-T} \cdot \hat{\nabla} \hat{b}_T \;=\; 6|T| \cdot \int_{\hat{T}} \hat{\nabla}^T \hat{b}_T \cdot H_T^{-2} \cdot \hat{\nabla} \hat{b}_T \\
&> 6|T| \cdot h_{min,T}^{-2} \cdot \int_{\hat{T}} \left(\frac{\partial \hat{b}_T}{\partial \hat{z}} \right)^2 \qquad .
\end{aligned}$$

The reference tetrahedron $\hat{T}$ is uniquely determined by its vertices $(0,0,0)^T$, $(1,0,0)^T$, $(\hat{x}_2, 1, 0)^T$, and $(\hat{x}_3, \hat{y}_3, 1)^T$, with $0 < \hat{x}_2, \hat{x}_3 < 1$ and $|\hat{y}_3| < 1$ (cf. section 2.1.3). Define the compact set

$$S := \Big\{ (\hat{x}_2, \hat{x}_3, \hat{y}_3) \quad : \quad 0 \leq \hat{x}_2, \hat{x}_3 \leq 1 \,,\; |\hat{y}_3| \leq 1 \Big\}$$

which covers all possible tetrahedra $\hat{T}$ (and some more). Obviously $\|\partial \hat{b}_T / \partial \hat{z}\|_{\hat{T}}$ varies continuously over S and thus attains its minimum. This is positive since $\partial \hat{b}_T / \partial \hat{z}$ cannot vanish everywhere on $\hat{T}$. Therefore

$$\|\nabla b_T\|_T^2 \;\gtrsim\; h_{min,T}^{-2} \cdot |T|$$

which, together with (2.21), implies the assertion. For b_E proceed analogously. ■

Remark 2.2 Bubble functions b_T or b_E which are transformed via F_A^{-1} become the corresponding bubble functions on the standard tetrahedron $\bar{T}$, respectively, i.e.

$$b_{\bar{T}} \;=\; \bar{b}_T \;:=\; b_T \circ F_A \qquad \text{and} \qquad b_{\bar{E}} \;=\; \bar{b}_E \;:=\; b_E \circ F_A \qquad .$$

A similar relation holds for the transformation F_C. □

Chapter 3

The Poisson equation

3.1 Analytical Background

The classical formulation of the Poisson problem reads

$$\text{Find } u \in C^2(\Omega) \cap C(\overline{\Omega}) : \qquad \begin{aligned} -\Delta u &= f && \text{in } \Omega \quad , \\ u &= 0 && \text{on } \Gamma_D = \partial\Omega \quad . \end{aligned}$$

Under suitable smoothness assumptions on the data (i.e. f and Ω) it yields a unique solution. In practice, however, the data are rarely as smooth as required. Then the variational or weak formulation is more appropriate:

$$\begin{aligned} \text{Find } u \in H_o^1(\Omega) : \qquad a(u,v) &= (f,v) \qquad \forall\, v \in H_o^1(\Omega) \\ \text{with} \qquad a(u,v) &:= \int_\Omega \nabla^T u \cdot \nabla v \;=\; (\nabla u, \nabla v) \\ (f,v) &= \int_\Omega f \cdot v \end{aligned} \tag{3.1}$$

and $H_o^1(\Omega)$ being the usual Sobolev space of functions from $H^1(\Omega)$ whose trace on the Dirichlet part Γ_D of the boundary (here the whole of $\partial\Omega$) vanishes.

The Lax-Milgram lemma [13] answers the question of the existence and uniqueness of a solution to the positive provided that

- $f \in [H_o^1(\Omega)]^* = H^{-1}(\Omega)$
- $a(\cdot\,,\,\cdot)$ is elliptic, i.e. $\exists \mu_1 > 0 :\; a(v,v) \geq \mu_1 \cdot \|v\|^2_{H_o^1(\Omega)} \qquad \forall\, v \in H_o^1(\Omega)$
- $a(\cdot\,,\,\cdot)$ is bounded, i.e. $|a(v,w)| \leq \mu_2 \cdot \|v\|_{H_o^1(\Omega)} \cdot \|w\|_{H_o^1(\Omega)} \qquad \forall\, v,w \in H_o^1(\Omega) \qquad .$

For the whole of our investigation we demand a stronger smoothness of the right-hand side, namely

$$f \in L_2(\Omega) \qquad ,$$

thus the first assumption is satisfied. The second and third assumption are automatically valid with a domain dependent constant $\mu_1 = \mu_1(\Omega)$ and $\mu_2 = 1$.

For convex domains Ω the right-hand side $f \in L_2(\Omega)$ implies $u \in H^2(\Omega)$, but otherwise one obtains $u \in H^{1+\alpha}(\Omega)$ with $0 < \alpha < 1$ (in general).

Assume now a triangulation $\mathcal{T}_h$ consisting of tetrahedra (3D) or triangles (2D). Let V_h be the space of piecewise linear functions over $\mathcal{T}_h$. Let $V_{o,h} \subset V_h$ be the subspace of those functions of V_h with homogeneous Dirichlet boundary conditions, i.e. $V_{o,h} := V_h \cap H^1_o(\Omega)$. The approximate or FEM solution u_h is obtained via

$$\text{Find } u_h \in V_{o,h} : \qquad a(u_h, v_h) = (f, v_h) \qquad \forall\, v_h \in V_{o,h} \qquad . \tag{3.2}$$

The Poisson equation is one of the simplest boundary value problems, and error estimators for it are long known and well established. Therefore this problem has been chosen to investigate how error estimators perform (or have to be modified) if one encounters an anisotropic solution or utilizes an anisotropic mesh.

Let us specify the framework of this chapter. We try to bound the error $u - u_h$ in

- the energy norm $|||v|||^2_\Omega := a(v, v) = \|\nabla v\|^2$ which coincides with the H^1 seminorm.
- the L_2 norm.

Furthermore different error estimators are investigated. The residual error estimator (section 3.2) and the local Dirichlet problem error estimator (section 3.3) estimate the error in the energy norm. The L_2 error estimator (section 3.4) is self-explanatory.

A anisotropic Zienkiewicz-Zhu like error estimator which aims for the energy norm has been derived as well. Unfortunately this estimator can only be applied to fairly simple meshes of tensor product type (in the sense that the estimator fails to estimate the error on non-tensor product type meshes). Despite much research we did not find an estimator that meets our expectations. Therefore we discuss our estimator only briefly in section 3.5.

At present the author is investigating a local problem estimator which bounds the error in the L_2 norm.

Last but not least it should be mentioned here that certain *interpolation* error estimates play a vital role in our analysis. The kind of the interpolation estimate is apparently strongly related to the FEM error estimate to be obtained. Yet our interpolation error estimate does not hold for arbitrary functions. Roughly speaking, it holds only when the anisotropic mesh corresponds in some way to the anisotropic solution (or more precisely, to the error $u - u_h$). This condition is supported by heuristic arguments — it seems sensible that the tetrahedra are stretched along that direction where the solution varies little (i.e. when the solution shows an almost lower-dimensional behaviour). Sections 3.2.1, 3.2.2 and 3.4.3, 3.4.4 are devoted to this topic.

Despite this knowledge we are still not able to *guarantee* this condition by some computable values since it involves the (unknown) exact solution u. Nevertheless numerical experiments indicate that a 'sensible' mesh yields useful error estimators.

3.2 Residual error estimator

Residual error estimators have been known for a long time, and they were probably the first estimators ever to be analysed [7]. Since then much work has been devoted to this type of error estimator for various problem classes. Verfürth [29] derived lower bounds of the error.

Residual error estimators suitable for anisotropic meshes have been first investigated into by Siebert [28]. There cuboidal and prismatic grids were considered.

The estimator presented here works on tetrahedral and triangular grids which are more difficult to deal with. Moreover, Siebert's estimator is improved slightly (cf. remark 3.5 on page 36).

Because of some (possibly not only technical) reasons the anisotropic error estimator can not be applied to an arbitrary mesh with an arbitrary solution. The mesh and the solution have to correspond in some way. Section 3.2.1 is devoted to this topic.

In order to derive an upper bound of the error, anisotropic interpolation estimates play a vital part. They are derived in section 3.2.2.

Finally, in section 3.2.3 the anisotropic error estimator is defined, and lower and upper bounds on the error are proven.

3.2.1 The set $H^1_{\mathcal{T}}(\Omega)$ of adapted functions

In order to derive interpolation estimates on anisotropic meshes we have to assume that the function to be interpolated corresponds to that mesh. More precisely, terms of the form $\dfrac{h_{i,T}}{h_{min,T}}\,\|\tilde{D}_i v\|_T$ have to be bounded although the aspect ratio $h_{i,T}/h_{min,T}$ can be arbitrarily large. The set $H^1_{\mathcal{T}}(\Omega)$ is introduced to achieve this.

Definition 3.1 (Adapted function) *Let $c_a > 1$ be a fixed constant. Similarly to [28] the mesh (or the triangulation $\mathcal{T}_h$) is said to be* adapted *to the function $v \in H^1(\Omega)$ iff*

$$\sum_{T\in\mathcal{T}_h} h_{min,T}^{-2} \cdot \|C_T^T\,\nabla v\|_T^2 \;\le\; c_a \cdot \|\nabla v\|^2 \tag{3.3}$$

holds. The component-wise form reads as

$$\sum_{T\in\mathcal{T}_h}\sum_{i=1}^{d} \frac{h_{i,T}^2}{h_{min,T}^2}\cdot\|\tilde{D}_i v\|_T^2 \;\le\; c_a \cdot \sum_{T\in\mathcal{T}_h}\sum_{i=1}^{d}\|\tilde{D}_i v\|_T^2 \qquad .$$

We also say that the function v is adapted to the mesh (or the triangulation). For a family $\mathcal{F}$ of triangulations $\{\mathcal{T}_h\}$ the set of adapted functions is denoted by $H^1_{\mathcal{T}}(\Omega)$.

Let $H^1_{o,\mathcal{T}}(\Omega) := H^1_{\mathcal{T}}(\Omega) \cap H^1_o(\Omega)$ be the corresponding set of functions with homogeneous Dirichlet boundary conditions.

Remark 3.1 Note that although $H^1_{\mathcal{T}}(\Omega)$ is a set of functions it does not constitute a subspace itself since the difference of two adapted functions is not necessarily an adapted function again. Furthermore it seems that (3.3) is not only a technical condition. □

3.2.2 Anisotropic interpolation estimates

Interpolation estimates are a major tool in the error analysis performed here. Since the interpolation has to act on functions $v \in H^1_T(\Omega)$ we cannot use the usual Lagrange interpolation. Therefore the interpolation operator introduced in this section follows the lines of Clément [14] instead. All estimates, however, are derived for the use on anisotropic meshes.

A local L_2 projection, along with approximation estimates, will be presented first. Then the Clément interpolation operator is constructed. Finally it is modified in such a way that homogeneous Dirichlet boundary conditions will be preserved.

The local L_2 projection

Consider a node a_j. A so-called *macro element* M_j of this node a_j is defined by

$$M_j := \bigcup_{T \ni a_j} T \qquad ,$$

i.e. M_j consists of all tetrahedra containing a_j. For simplicity the subscript j will be omitted in the next lemma and proof.

Lemma 3.1 *Let a be a node of $\mathcal{T}_h$ and M the corresponding macro element. Let $V_h(M)$ be the space V_h restricted to M. Let the local L_2 projection $P : H^1(M) \mapsto \mathbb{P}^0(M)$ be defined by*

$$\int_M (v - Pv) \cdot \varphi = 0 \qquad \forall\, \varphi \in \mathbb{P}^0(M) \qquad .$$

Then the relations

$$\begin{aligned} \|v - Pv\|_M &\leq \|v\|_M & (3.4)\\ \|v - Pv\|_M &\lesssim \|C^T(\mathbf{x})\nabla v\|_M & (3.5)\\ \|C^T(\mathbf{x})\nabla(v - Pv)\|_M &= \|C^T(\mathbf{x})\nabla v\|_M & (3.6) \end{aligned}$$

hold. The component-wise form of (3.6) is

$$\|\tilde{D}_i(v - Pv)\|_M \lesssim h_{i,M}^{-1} \cdot \sum_{k=1}^{d} h_{k,M}\|\tilde{D}_k v\|_M \qquad i = 1 \dots d \qquad ,$$

with $h_{i,M}$ explained in remark 2.1 on page 15.

Proof: The first inequality is readily obtained using the projection orthogonality:

$$\|v - Pv\|_M^2 = \int_M (v - Pv)(v - Pv) = \int_M (v - Pv) \cdot v \leq \|v - Pv\|_M \cdot \|v\|_M$$

since $Pv \in \mathbb{P}^0(M)$.

The second inequality requires a closer investigation. A continuous mapping F_B that maps a reference domain $\check{M} \in \mathcal{M}$ onto the macro element M will play an important role in the proof. Furthermore, the set $\mathcal{M}$ of reference domains shall be finite. For a start we will construct the reference domains $\check{M}$.

Assume that the macro element M is the union of K tetrahedra $T_1 \dots T_K$. Let the nodes of M be $a_1 \dots a_L$ (apart from node a), where L is bounded because of the mesh requirements. Two macro elements M and M' are said to belong to the same *class* iff

- they consist of the same number of tetrahedra, i.e. $K = K'$,
- the tetrahedra and the nodes can be numbered such that for all $i = 1 \ldots K$ the following holds: If the tetrahedron T_i has the nodes $a, a_{j_1}, a_{j_2}, a_{j_3}$ then the tetrahedron T_i' has the nodes $a', a'_{j_1}, a'_{j_2}, a'_{j_3}$.

This condition implies that the triangulations of both macro elements are topologically equivalent. The number of such topologies is bounded since K is bounded. Therefore the number of classes of macro elements is bounded as well. For a fixed class an arbitrary macro element will be chosen whose node a coincides with the coordinate origin. This macro element is said to be the reference domain of this class. All reference domains form the (finite) set $\mathcal{M}$. Note that a condition on the size of the reference domains is not necessary.

Let now M be an arbitrary macro element and $\check{M}$ be the corresponding reference domain. Because of the construction of the reference domain there exists a continuous, piecewise linear mapping F_B that satisfies

$$\begin{aligned} F_B &: \quad \check{M} \mapsto M \\ F_B = F_i &: \quad \mathbf{x}(\boldsymbol{\mu}) = B_i \boldsymbol{\mu} + \mathbf{a} \qquad \text{on } \check{T}_i, \quad B_i \in \mathbb{R}^{d\times d}, \mathbf{a} \in \mathbb{R}^d \\ \text{with} \quad F_i &: \quad \check{T}_i \mapsto T_i \quad \text{affine linear, } i = 1 \ldots K \qquad . \end{aligned}$$

Temporarily $\check{T}_i$ and T_i shall denote the i-th tetrahedron of $\check{M}$ and M, respectively, and $\mathbf{a}$ denotes the vector corresponding to node a. Variables that are related to the reference domain will be denoted by a $\check{}$ (small check).

The Poincaré inequality holds for the domain $\check{M}$. Its inequality constant can be chosen independent of $\check{M}$ since the number of reference domains $\check{M} \in \mathcal{M}$ is bounded. Thus for $\check{u} \in H^1(\check{M})$

$$\int_{\check{M}} |\check{u}|^2 \lesssim \left| \int_{\check{M}} \check{u} \right|^2 + \int_{\check{M}} |\check{\nabla} \check{u}|^2 \qquad .$$

For a function $v \in H^1(M)$ define an averaging operator $I : H^1(M) \to \mathbb{P}^0(M)$ by

$$Iv := |\check{M}|^{-1} \cdot \sum_{i=1}^{K} \int_{T_i} v \cdot |\det B_i|^{-1} \;=\; const.$$

Set $\check{v} := v \circ F_B \in H^1(\check{M})$. The definition of I gives

$$\int_{\check{M}} \breve{Iv} = |\check{M}| \cdot Iv = \sum_{i=1}^{K} \int_{T_i} v \cdot |\det B_i|^{-1} = \int_{\check{M}} \check{v}$$

and

$$\check{\nabla}(\breve{Iv}) = 0 \qquad .$$

Inserting now $\check{u} := \check{v} - \breve{Iv}$ in the Poincaré inequality results in

$$\int_{\check{M}} |\check{v} - \breve{Iv}\,|^2 \lesssim \int_{\check{M}} |\check{\nabla} \check{v}|^2 \qquad .$$

Obviously $\underline{c} \le |\check{T}_i| \le \bar{c}$ since the number of reference domains $\check{M}$ is bounded, and each $\check{M}$ consists only of a bounded number of tetrahedra $\check{T}_i$. Hence

$$|\det B_i| = |T_i| / |\check{T}_i| \;\sim h_{1,T_i}\, h_{2,T_i}\, h_{3,T_i} \qquad .$$

Since the $h_{i,T}$ cannot change rapidly one obtains

$$|\det B_i| \sim |\det B_j| \qquad \forall\, T_i, T_j \subset M \qquad .$$

Applying the transformation $F_B : \check{M} \mapsto M$ gives

$$\begin{aligned}
\int_M (v - Iv)^2 &= \sum_{i=1}^{K} \int_{T_i} (v - Iv)^2 = \sum_{i=1}^{K} \int_{\check{T}_i} (\check{v} - \breve{Iv})^2 \cdot |\det B_i| \\
&\lesssim \max_{i=1\ldots K} \{|\det B_i|\} \cdot \int_{\check{M}} (\check{v} - \breve{Iv})^2 \\
&\lesssim \max_{i=1\ldots K} \{|\det B_i|\} \cdot \int_{\check{M}} |\check{\nabla} \check{v}|^2 \\
&\lesssim \sum_{i=1}^{K} \int_{\check{T}_i} |\check{\nabla} \check{v}|^2 \cdot |\det B_i| \\
&= \sum_{i=1}^{K} \int_{T_i} |B_i^T \nabla v|^2 = \sum_{i=1}^{K} \int_{T_i} \left| B_i^T C_{T_i}^{-T} \cdot C_{T_i}^T \nabla v \right|^2 \\
&\le \sum_{i=1}^{K} \left\| B_i^T C_{T_i}^{-T} \right\|_{\mathbb{R}^{3\times 3}}^2 \cdot \int_{T_i} \left| C_{T_i}^T \nabla v \right|^2 \qquad .
\end{aligned}$$

Lemma 2.1 on page 12 is now utilized to bound the norm of $B_i^T C_{T_i}^{-T}$. Let $T_i - \mathbf{a}$ be the tetrahedron T_i shifted by $-\mathbf{a}$. By definition the mappings B_i and $C_{T_i}^{-1}$ act as follows:

$$\begin{array}{ccccc}
\check{T}_i & \xmapsto{B_i} & (T_i - \mathbf{a}) & \xmapsto{C_{T_i}^{-1}} & \hat{T}_i, \\
& \check{T}_i & \xmapsto{C_{T_i}^{-1} B_i} & \hat{T}_i & .
\end{array}$$

The number of tetrahedra $\check{T}_j \subset \check{M}$ is bounded. Hence the diameters of the inscribed spheres of all tetrahedra $\check{T}_i$ can be bounded uniformly from below, i.e. $\varrho(\check{T}_i) \gtrsim 1$. The longest edge of $\hat{T}_i$ is bounded from above by $\sqrt{6}$ (see definition of the mapping C_{T_i}). Lemma 2.1 yields readily

$$\left\| B_i^T C_{T_i}^{-T} \right\|_{\mathbb{R}^{3\times 3}} = \left\| C_{T_i}^{-1} B_i \right\|_{\mathbb{R}^{3\times 3}} \le d(\hat{T}_i) \,/\, \varrho(\check{T}_i) \lesssim 1$$

and further

$$\int_M (v - Iv)^2 \lesssim \sum_{i=1}^{K} \int_{T_i} \left| C_{T_i}^T \nabla v \right|^2 = \| C^T(\mathbf{x}) \nabla v \|_M^2 \qquad .$$

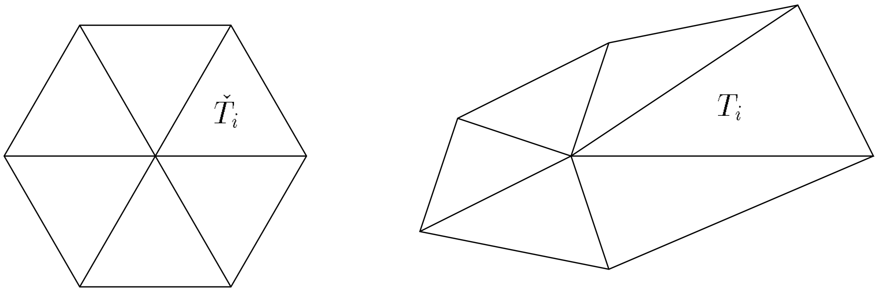
The orthogonality property of the projection and $Pv - Iv \in V_h$ then imply

$$\begin{aligned}
\|v - Pv\|_M^2 &= \int_M (v - Pv)(v - Pv) = \int_M (v - Pv)(v - Iv) \\
&\le \|v - Pv\|_M \cdot \|v - Iv\|_M \\
\text{and} \qquad \|v - Pv\|_M &\le \|v - Iv\|_M \lesssim \|C^T(\mathbf{x}) \nabla v\|_M
\end{aligned}$$

finishing the second part of the proof. Recall that $C(\mathbf{x})$ is the global matrix function defined in (2.2).

The last inequality is obvious since $Pv \in \mathbb{P}^0(M)$. ∎

Remark 3.2 In the case $\Omega \subset \mathbb{R}^2$ the reference domains can be chosen easily. Assume that the macro element consists of K triangles. When a is a inner node then choose $\check{M}$ to be a regular K-polygon with the midpoint in the coordinate origin. Figure 3.1 may serve for visualization.



$$\begin{array}{ccc} \check{M} & \overset{F_B}{\longmapsto} & M \\ \check{T}_i & \overset{F_i}{\longmapsto} & T_i \end{array}$$

Figure 3.1: Continuous, piecewise affine linear mapping F_B for $\Omega \subset \mathbb{R}^2$

If a is a boundary node then let $\check{M}$ be the union of those K (congruent) triangles whose vertices have the polar coordinates $(0,0)$, $(1,(i-1)\pi/2K)$ and $(1, i\pi/2K)$, $i = 1 \ldots K$.

The regular polygon is chosen here only for the convenience of the description, but otherwise completely arbitrary. Any other reference domain could serve the same purpose.

The case $\Omega \subset \mathbb{R}^3$ is more difficult since generally no regular polyhedra exist. Thus we had to utilize the more technical definition of the reference domains here. □

The H^1 interpolation operator

Now the Clément interpolation operator is constructed. Let P_j be the aforementioned local L_2 projection over the macro element M_j of a node a_j. The interpolation operator R is defined by

$$Rv := \sum_{j=1}^{N} (P_j v)(a_j) \cdot \varphi_j$$

with φ_j being the (piecewise affine linear) basis function related to node a_j. Then the following theorem holds.

Lemma 3.2 *The interpolation operator $R : H^1(\Omega) \mapsto V_h$ satisfies*

$$\|v - Rv\| \lesssim \|v\| \qquad \forall\, v \in H^1(\Omega) \qquad .$$

If additionally the mesh is adapted to v, i.e. $v \in H^1_T(\Omega)$ then

$$\begin{aligned} \|h_{min}^{-1}(\mathbf{x}) \cdot (v - Rv)\| &= \left(\sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \|v - Rv\|_T^2 \right)^{1/2} \lesssim \|\nabla v\| \\ \|h_{min}^{-1}(\mathbf{x}) \cdot C^T(\mathbf{x}) \nabla(v - Rv)\| &= \left(\sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \|C_T^T \nabla(v - Rv)\|_T^2 \right)^{1/2} \lesssim \|\nabla v\| \end{aligned}$$

hold. The component-wise form of the last inequality is

$$\left\| \frac{h_i(\mathbf{x})}{h_{min}(\mathbf{x})} \tilde{D}_i(v - Rv) \right\| = \left(\sum_{T \in \mathcal{T}_h} \frac{h_{i,T}^2}{h_{min,T}^2} \|\tilde{D}_i(v - Rv)\|_T^2 \right)^{1/2} \lesssim \|\nabla v\|$$

for all $i = 1 \ldots d$.

Proof: Let T be an arbitrary tetrahedron, and denote the set of its nodes by N_T. Let a_k be an arbitrary but fixed node of N_T. Then R can be represented over T as

$$Rv\Big|_T = \sum_{a_j \in N_T} (P_j v)(a_j) \cdot \varphi_j\Big|_T = P_k v\Big|_T + \sum_{a_j \in N_T} (P_j v - P_k v)(a_j) \cdot \varphi_j\Big|_T,$$

$$\text{since} \qquad P_k v\Big|_T = \sum_{a_j \in N_T} (P_k v)(a_j) \cdot \varphi_j\Big|_T \qquad .$$

The inverse inequality (2.15) and the triangle inequality imply

$$\begin{aligned} |(P_j v - P_k v)(a_j)| &\leq \|P_j v - P_k v\|_{\infty,T} \\ &\lesssim |T|^{-1/2} \cdot \|P_j v - P_k v\|_T \\ &\leq |T|^{-1/2} \cdot \Big(\|v - P_j v\|_T + \|v - P_k v\|_T \Big) \qquad . \end{aligned}$$

The bound $\|\varphi_j\|_T \leq |T|^{1/2}$ gives

$$\begin{aligned} \|(P_j v - P_k v)(a_j) \cdot \varphi_j\|_T &= |(P_j v - P_k v)(a_j)| \cdot \|\varphi_j\|_T \\ &\lesssim \Big(\|v - P_j v\|_T + \|v - P_k v\|_T \Big) \qquad . \end{aligned}$$

Applying this inequality to the representation of R leads to

$$\begin{aligned} \|v - Rv\|_T &\leq \|v - P_k v\|_T + \Big\| \sum_{a_j \in N_T} (P_j v - P_k v)(a_j) \cdot \varphi_j \Big\|_T \\ &\lesssim \|v - P_k v\|_T + \sum_{a_j \in N_T} \Big(\|v - P_j v\|_T + \|v - P_k v\|_T \Big) \\ &\lesssim \sum_{a_j \in N_T} \|v - P_j v\|_T \leq \sum_{a_j \in N_T} \|v - P_j v\|_{M_j} \qquad . \end{aligned}$$

The local approximation inequality (3.4) results in

$$\|v - Rv\|_T \overset{(3.4)}{\lesssim} \sum_{a_j \in N_T} \|v\|_{M_j} \sim \|v\|_{M(T)}$$

with $M(T) := \bigcup_{a_j \in N_T} M_j = \bigcup_{T' \cap T \neq \emptyset} T'$. This holds since every tetrahedron T' is contained in at most four macro elements M_j. Then

$$\|v - Rv\|^2 = \sum_{T \in \mathcal{T}_h} \|v - Rv\|_T^2 \lesssim \sum_{T \in \mathcal{T}_h} \|v\|_{M(T)}^2 \sim \|v\|^2$$

because every tetrahedron T appears only a bounded number of times in the sum. Hence the first inequality is obtained.

For the second inequality we apply (3.5) instead of (3.4) and obtain

$$\|v - Rv\|_T \overset{(3.5)}{\lesssim} \sum_{a_j \in N_T} \|C^T(\mathbf{x})\nabla v\|_{M_j} \sim \|C^T(\mathbf{x})\nabla v\|_{M(T)} \qquad .$$

Similarly this yields

$$\begin{aligned} \|h_{min}^{-1}(\mathbf{x}) \cdot (v - Rv)\|^2 &= \sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \|v - Rv\|_T^2 \lesssim \sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \|C^T(\mathbf{x})\nabla v\|_{M(T)}^2 \\ &\sim \sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \cdot \|C_T^T \nabla v\|_T^2 \end{aligned}$$

since $h_{min,T'}$ does not change rapidly for $T' \subset M(T)$.

In order to bound this last sum we have to assume $v \in H_{\mathcal{T}}^1(\Omega)$. Recalling the definition of this set (cf. (3.3)) yields

$$\|h_{min}^{-1}(\mathbf{x}) \cdot (v - Rv)\|^2 \lesssim \sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \cdot \|C_T^T \nabla v\|_T^2 \lesssim \|\nabla v\|^2 \qquad .$$

Thus the second result is proven.

The last part is derived similarly, and for this reason only major inequalities are given here. As before

$$C_T^T \nabla Rv\Big|_T = C_T^T \nabla P_k v\Big|_T + \sum_{a_j \in N_T} (P_j v - P_k v)(a_j) \cdot C_T^T \nabla \varphi_j\Big|_T \qquad .$$

Recalling the inverse inequality (2.14)

$$\left\|C_T^T \nabla \varphi_j\right\|_T \lesssim \|\varphi_j\|_T \sim |T|^{1/2}$$

leads to

$$\begin{aligned} \left\|(P_j v - P_k v)(a_j) \cdot C_T^T \nabla \varphi_j\right\|_T &= |(P_j v - P_k v)(a_j)| \cdot \left\|C_T^T \nabla \varphi_j\right\|_T \\ &\lesssim \Big(\|v - P_j v\|_T + \|v - P_k v\|_T\Big) \end{aligned}$$

analogously as before. Similarly to the second part one obtains

$$\begin{aligned} \left\|C_T^T \nabla (v - Rv)\right\|_T &\leq \left\|C_T^T \nabla (v - P_k v)\right\|_T + \left\| \sum_{a_j \in N_T} (P_j v - P_k v)(a_j) \cdot C_T^T \nabla \varphi_j \right\|_T \\ &\overset{(3.6)}{\lesssim} \left\|C^T(\mathbf{x}) \nabla v\right\|_{M_k} + \left\|C^T(\mathbf{x}) \nabla v\right\|_{M(T)} \\ &\lesssim \left\|C^T(\mathbf{x}) \nabla v\right\|_{M(T)} \end{aligned}$$

and hence

$$\sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \|C_T^T \nabla (v - Rv)\|_T^2 \lesssim \sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \|C^T(\mathbf{x}) \nabla v\|_{M(T)}^2 \sim \sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \|C_T^T \nabla v\|_T^2.$$

Utilizing $v \in H_{\mathcal{T}}^1(\Omega)$ results immediately in

$$\|h_{min}^{-1}(\mathbf{x}) \cdot C^T(\mathbf{x}) \nabla (v - Rv)\| \lesssim \|\nabla v\| \qquad .$$

$\blacksquare$

The H_o^1 interpolation operator

The interpolation operator R introduced above has the disadvantage that it does not preserve homogeneous Dirichlet boundary conditions. This is remedied in the next theorem.

Definition 3.2 (Clément interpolation operator) *Let N_I be the set of all inner nodes of the triangulation. The* Clément interpolation operator *$R_o : H_o^1(\Omega) \mapsto V_{o,h}$ is defined by*

$$R_o v := \sum_{a_j \in N_I} (P_j v)(a_j) \cdot \varphi_j \qquad . \tag{3.7}$$

The following anisotropic interpolation estimates are valid.

Theorem 3.3 *The interpolation operator $R_o : H_o^1(\Omega) \mapsto V_{o,h}$ satisfies*

$$\|v - R_o v\| \ \lesssim\ \|v\| \qquad \forall\, v \in H_o^1(\Omega) \qquad . \tag{3.8}$$

If additionally the mesh is adapted to $v \in H_o^1(\Omega)$, i.e. $v \in H_{o,\mathcal{T}}^1(\Omega)$ then

$$\|h_{min}^{-1}(\mathbf{x}) \cdot (v - R_o v)\| \ \lesssim\ \|\nabla v\| \tag{3.9}$$

$$\|h_{min}^{-1}(\mathbf{x}) \cdot C^T(\mathbf{x}) \nabla (v - R_o v)\| \ \lesssim\ \|\nabla v\| \tag{3.10}$$

hold. The component-wise form of the last inequality is

$$\left\| \frac{h_i(\mathbf{x})}{h_{min}(\mathbf{x})} \tilde{D}_i (v - R_o v) \right\| \ \lesssim\ \|\nabla v\|$$

for all $i = 1 \ldots d$.

Proof: The definition of the interpolation operator R_o implies

$$R_o v \ =\ Rv \ -\ \sum_{a_j \in \Gamma_D} (P_j v)(a_j) \cdot \varphi_j \qquad .$$

Since we want to utilize the previous lemma it is sufficient to bound terms of the form $\|(P_j v)(a_j) \cdot \varphi_j\|_T$ and $\|(P_j v)(a_j) \cdot C_T^T \nabla \varphi_j\|_T$ for boundary nodes $a_j \in \Gamma_D$

Thus let $a_j \in \Gamma_D$ be fixed. Let $T \subset M_j$ be an arbitrary (but fixed) tetrahedron with a boundary face $E \ni a_j$. The inverse inequality (2.15) yields

$$|(P_j v)(a_j)| \ \le\ \|P_j v\|_{\infty,E} \ \lesssim\ |T|^{-1/2} \cdot \|P_j v\|_T \ \le\ |T|^{-1/2} \cdot (\|v\|_T + \|v - P_j v\|_T) \quad .$$

The estimate of the local L_2 projection, and $\|\varphi_j\|_T \le |T|^{1/2}$ give

$$\|(P_j v)(a_j) \cdot \varphi_j\|_T \ \lesssim\ \|v\|_{M_j} \qquad \forall\, T \subset M_j \qquad .$$

Applying this estimate results in

$$\begin{aligned}
\Big\| \sum_{a_j \in \Gamma_D} (P_j v)(a_j) \cdot \varphi_j \Big\|^2 \ &=\ \sum_{T \cap \Gamma_D \neq \emptyset} \Big\| \sum_{a_j \in T \cap \Gamma_D} (P_j v)(a_j) \cdot \varphi_j \Big\|_T^2 \\
&\le\ 4 \sum_{a_j \in \Gamma_D} \sum_{T \subset M_j} \|(P_j v)(a_j) \cdot \varphi_j\|_T^2 \\
&\lesssim\ \sum_{a_j \in \Gamma_D} \|v\|_{M_j}^2 \ \sim\ \|v\|^2
\end{aligned}$$

following the same arguments as in the previous proof. Finally the first inequality is obtained:

$$\|v - R_o v\| \le \|v - Rv\| + \left\| \sum_{a_j \in \Gamma_D} (P_j v)(a_j) \cdot \varphi_j \right\| \lesssim \|v\| \qquad .$$

The second inequality is derived similarly and thus only major estimates are given. The inverse inequality (2.16) and the trace inequality (2.12) yield

$$\begin{aligned} |(P_j v)(a_j)| &\le \|P_j v\|_{\infty,E} \lesssim |E|^{-1/2} \|P_j v\|_E \\ &= |E|^{-1/2} \|v - P_j v\|_E \qquad \text{since } v = 0 \text{ on } E \\ &\lesssim |T|^{-1/2} \left(\|v - P_j v\|_T + \|C_T^T \nabla (v - P_j v)\|_T \right) \qquad . \end{aligned}$$

Recalling the estimates (3.5) and (3.6) of the local L_2 projection leads to

$$\begin{aligned} |(P_j v)(a_j)| &\lesssim |T|^{-1/2} \cdot \|C^T(\mathbf{x}) \nabla v\|_{M_j} \\ \text{and} \qquad \|(P_j v)(a_j) \cdot \varphi_j\|_T &\lesssim \|C^T(\mathbf{x}) \nabla v\|_{M_j} \qquad \forall T \subset M_j \end{aligned}$$

as in the lines above. Applying this estimate results in

$$\begin{aligned} \left\| h_{min}^{-1}(\mathbf{x}) \sum_{a_j \in \Gamma_D} (P_j v)(a_j) \cdot \varphi_j \right\|^2 &\le 4 \sum_{a_j \in \Gamma_D} \sum_{T \subset M_j} h_{min,T}^{-2} \|(P_j v)(a_j) \cdot \varphi_j\|_T^2 \\ &\lesssim \sum_{a_j \in \Gamma_D} h_{min,M_j}^{-2} \cdot \|C^T(\mathbf{x}) \nabla v\|_{M_j}^2 \\ &\sim \sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \cdot \|C_T^T \nabla v\|_T^2 \lesssim \|\nabla v\|^2 \end{aligned}$$

and and thus the second inequality is proven.

In order to derive the last part of the theorem we proceed analogously. From

$$\left\| C_T^T \nabla \varphi_j \right\|_T \lesssim \|\varphi_j\|_T \le |T|^{1/2}$$

we obtain

$$\left\| (P_j v)(a_j) \cdot C_T^T \nabla \, \varphi_j \right\|_T \lesssim \|C^T(\mathbf{x}) \nabla v\|_{M_j}$$

for a boundary tetrahedron T. The remainder of the proof is similar to the lines above and the previous proof and thus it will be omitted here. ■

Remark 3.3 Note that $v \in H^1_{o,T}(\Omega)$ is required for the last two interpolation estimates (3.9) and (3.10) but not for (3.8). □

3.2.3 Anisotropic residual error estimator

Consider a tetrahedron T. Let P_{L_2} be the L_2 projection from $L_2(\Omega)$ onto the space of piecewise constant functions over the triangulation.

Definition 3.3 (Element and jump residual) *Let $v_h \in V_{o,h}$ be an arbitrary finite element function. The* element residual *over a tetrahedron T is defined by*

$$r_T(v_h) := P_{L_2} f + \Delta v_h \qquad . \tag{3.11}$$

The gradient jump *or* jump residual *of a function across some (interior) face E is defined as*

$$r_E(v_h)(x) := \lim_{t\to+0}\left[\frac{\partial}{\partial n_E}v_h(x+tn_E) \;-\; \frac{\partial}{\partial n_E}v_h(x-tn_E)\right] \quad , \tag{3.12}$$

with $n_E \perp E$ being any of the two unitary normal vectors and $x \in E$.

For the convenience of the notation a discrete, mesh dependent norm is defined by

$$D_{h,m}(v_h) \;:=\; \left(\sum_{E\in\Omega}\frac{h^{2m}_{min,T}}{h_E}\cdot\|r_E(v_h)\|^2_E\right)^{1/2} \qquad m=1,2 \tag{3.13}$$

where $h_{min,T}$ is from one of the two tetrahedra that contain E (cf. remark 2.1 on page 15).

Obviously $r_T(v_h) = P_{L_2}f$ holds for piecewise linear basis functions as considered here. The definition above however allows to extend this theory readily to quadratic basis function. Moreover, this residual of v_h is related to the strong form of the differential operator and as such problem dependent.

Definition 3.4 (Residual error estimator) *The* local residual error estimator $\eta_{R,T}(u_h)$ *for a tetrahedron T is defined by*

$$\eta_{R,T}(u_h) := h_{min,T}\cdot\left(\|r_T(u_h)\|^2_T \;+\; \sum_{E\in\partial T\setminus\Gamma_D} h_E^{-1}\cdot\|r_E(u_h)\|^2_E\right)^{1/2} \quad . \tag{3.14}$$

Theorem 3.4 (Residual error estimator) *Let $u \in H^1_o(\Omega)$ be the exact solution and $u_h \in V_{o,h}$ be the FEM solution.*

Then the error is bounded locally from below by

$$\eta_{R,T}(u_h) \lesssim \|\nabla(u-u_h)\|_{\omega_T} \;+\; h_{min,T}\cdot\|f-P_{L_2}f\|_{\omega_T} \tag{3.15}$$

for all $T \in \mathcal{T}_h$.

Assume further that the mesh is adapted to the error $u-u_h$, i.e. $u-u_h \in H^1_{o,T}(\Omega)$. Then the error is bounded globally from above by

$$\|\nabla(u-u_h)\| \lesssim \left(\sum_{T\in\mathcal{T}_h}\eta^2_{R,T}(u_h) \;+\; \sum_{T\in\mathcal{T}_h} h^2_{min,T}\cdot\|f-P_{L_2}f\|^2_T\right)^{1/2} \tag{3.16}$$

or, alternatively

$$\|\nabla(u-u_h)\| \;\lesssim\; \|h_{min}(\mathbf{x})\,f\| \;+\; D_{h,1}(u_h) \quad .$$

Proof: Firstly, estimate (3.15) will be proven. We start with the norm $\|r_T(u_h)\|_T$ of the element residual $r_T = r_T(u_h) := P_{L_2}f + \Delta u_h$. Since we use linear ansatz functions $r_T \in \mathbb{P}^0(T)$ holds. For $\mathbf{x} \in T$ let

$$w(\mathbf{x}) := r_T(u_h)(\mathbf{x})\cdot b_T(\mathbf{x}) \qquad \in \mathbb{P}^4(T)\cap H^1_o(T) \quad .$$

Integration by parts yields

$$\begin{aligned}
\int_T r_T \cdot w &= \int_T (f + \Delta u_h) \cdot w \; + \; \int_T (P_{L_2} f - f) \cdot w \\
&= \int_T \nabla^T (u - u_h) \cdot \nabla w \; + \; \int_T (P_{L_2} f - f) \cdot w \\
\left| \int_T r_T \cdot w \right| &\le \|\nabla (u - u_h)\|_T \cdot \|\nabla w\|_T \; + \; \|f - P_{L_2} f\|_T \cdot \|w\|_T \qquad .
\end{aligned}$$

Recalling (2.20), (2.21), and $0 \le b_T \le 1$ gives the following bounds

$$\left.\begin{aligned}
\left| \int_T r_T \cdot w \right| &= \|b_T^{1/2} \cdot r_T\|_T^2 \gtrsim \|r_T\|_T^2 \\
\|\nabla w\|_T &= \|\nabla (b_T \cdot r_T)\|_T \lesssim h_{min,T}^{-1} \cdot \|r_T\|_T \\
\|w\|_T &= \|b_T \cdot r_T\|_T \le \|r_T\|_T
\end{aligned}\right\} \tag{3.17}$$

that result in

$$\begin{aligned}
\|r_T\|_T^2 &\lesssim \|\nabla (u - u_h)\|_T \cdot h_{min,T}^{-1} \cdot \|r_T\|_T \; + \; \|f - P_{L_2} f\|_T \cdot \|r_T\|_T \\
\text{and} \qquad h_{min,T}^2 \cdot \|r_T\|_T^2 &\lesssim \|\nabla (u - u_h)\|_T^2 \; + \; h_{min,T}^2 \cdot \|f - P_{L_2} f\|_T^2 \qquad .
\end{aligned}$$

Now we aim for a bound of the norm $\|r_E(u_h)\|_E$ of the gradient jump across some inner face (triangle) E. Since we use linear ansatz functions $r_E \in \mathbb{P}^0(E)$ holds. Let T_1 and T_2 be the two tetrahedra that E belongs to. Assume that the right hand side $f = -\Delta u$ is in $L_2(\Omega)$. Integration by parts yields for any function $w \in H_o^1(\omega_E)$

$$\begin{aligned}
0 &= \int_{\omega_E} \nabla^T w \nabla u \; - \; \int_{\omega_E} w \cdot f \\
\text{and} \qquad - \int_E w \cdot r_E(u_h) &= \sum_{i=1}^{2} \int_{\partial T_i} w \cdot \frac{\partial u_h}{\partial n} = \sum_{i=1}^{2} \left(\int_{T_i} \nabla^T w \nabla u_h \; + \; \int_{T_i} w \cdot \Delta u_h \right) \\
&= \sum_{i=1}^{2} \left(\int_{T_i} \nabla^T w \nabla u_h \; + \; \int_{T_i} w \cdot (r_{T_i} - P_{L_2} f) \right) \\
&= \sum_{i=1}^{2} \left(\int_{T_i} \nabla^T w \nabla (u_h - u) \; + \; \int_{T_i} w \cdot (r_{T_i} + f - P_{L_2} f) \right) \qquad .
\end{aligned}$$

Let now the function $w \in H_o^1(\omega_E)$ be defined by

$$w := F_{ext}(r_E(u_h)) \cdot b_E \qquad ,$$

with F_{ext} being the extension operator of (2.19). Because of $w|_E = r_E \cdot b_E|_E$ we conclude

$$\left| \int_E r_E^2 \cdot b_E \right| \le \sum_{i=1}^{2} \left(\|\nabla (u - u_h)\|_{T_i} \cdot \|\nabla w\|_{T_i} \; + \; (\|r_{T_i}\|_{T_i} + \|f - P_{L_2} f\|_{T_i}) \cdot \|w\|_{T_i} \right) \qquad .$$

The function w is piecewise cubic on ω_E. The inverse inequalities (2.22) – (2.24) imply

$$\begin{aligned}
\left| \int_E r_E^2 \cdot b_E \right| &= \|b_E^{1/2} \cdot r_E\|_E^2 \gtrsim \|r_E\|_E^2 \\
\|\nabla w\|_{T_i} &= \|\nabla (F_{ext}(r_E) \cdot b_E)\|_{T_i} \lesssim h_E^{1/2} \, h_{min,T_i}^{-1} \cdot \|r_E\|_E \\
\text{and} \qquad \|w\|_{T_i} &= \|F_{ext}(r_E) \cdot b_E\|_{T_i} \lesssim h_E^{1/2} \cdot \|r_E\|_E
\end{aligned}$$

and subsequently lead to

$$\|r_E\|_E^2 \lesssim \sum_{i=1}^{2} \Big(\|\nabla(u-u_h)\|_{T_i} \cdot h_E^{1/2}\, h_{min,T_i}^{-1}\, \|r_E\|_E \; + $$
$$ + \left(\|r_{T_i}\|_{T_i} + \|f - P_{L_2} f\|_{T_i}\right) \cdot h_E^{1/2}\, \|r_E\|_E \Big) \qquad .$$

The dimensions $h_E = h_{E,T_i}$ and h_{min,T_i} cannot change rapidly for adjacent tetrahedra. Recalling the bound of $\|r_T\|_T$ from above we conclude

$$\|r_E\|_E \lesssim h_E^{1/2}\, h_{min,T_1}^{-1} \cdot \Big(\|\nabla(u-u_h)\|_{\omega_E} \; + \; h_{min,T_1} \|f - P_{L_2} f\|_{\omega_E} \Big) \qquad .$$

For a fixed tetrahedron $T = T_1$ we sum up over all (inner) faces $E \subset \partial T \setminus \Gamma_D$ and obtain

$$\sum_{E \subset \partial T \setminus \Gamma_D} \frac{h_{min,T}^2}{h_E} \cdot \|r_E(u_h)\|_E^2 \;\lesssim\; \Big(\|\nabla(u-u_h)\|_{\omega_T}^2 \; + \; h_{min,T}^2 \, \|f - P_{L_2} f\|_{\omega_T}^2 \qquad .$$

This accomplishes the proof of (3.15).

Secondly, in order to derive (3.16) we utilize the orthogonality property of the error

$$(\nabla(u-u_h), \nabla v_h) = 0 \qquad \forall\, v_h \in V_{o,h} \qquad .$$

Integration by parts gives for all $v \in H^1_{o,\mathcal{T}}(\Omega)$

$$\begin{aligned}
(\nabla(u-u_h), \nabla v) &= (\nabla(u-u_h), \nabla(v - R_o v)) \\
&= \sum_{T \in \mathcal{T}_h} (f + \Delta u_h, v - R_o v)_T \;-\; \sum_{E \in \Omega} (r_E(u_h), v - R_o v)_E \\
&= \sum_{T \in \mathcal{T}_h} \Big[(f + \Delta u_h, v - R_o v)_T \;-\; \frac{1}{2} \sum_{E \in \partial T \setminus \Gamma_D} (r_E(u_h), v - R_o v)_E \Big] \\
&\le \sum_{T \in \mathcal{T}_h} \Big[\|f + \Delta u_h\|_T \cdot \|v - R_o v\|_T \;+\; \frac{1}{2} \sum_{E \in \partial T \setminus \Gamma_D} \|r_E(u_h)\|_E \cdot \|v - R_o v\|_E \Big] \\
&\le \sum_{T \in \mathcal{T}_h} \Big[h_{min,T} \|f + \Delta u_h\|_T \cdot h_{min,T}^{-1} \|v - R_o v\|_T \\
&\qquad + \frac{1}{2} \sum_{E \in \partial T \setminus \Gamma_D} \frac{h_{min,T}}{h_{E,T}^{1/2}} \|r_E(u_h)\|_E \cdot \frac{h_{E,T}^{1/2}}{h_{min,T}} \|v - R_o v\|_E \Big].
\end{aligned}$$

Applying the Cauchy-Schwarz inequality yields

$$\begin{aligned}
&(\nabla(u-u_h), \nabla v) \le \\
&\quad \le \Big(\sum_{T \in \mathcal{T}_h} h_{min,T}^2 \|f + \Delta u_h\|_T^2 \Big)^{\frac{1}{2}} \cdot \Big(\sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \|v - R_o v\|_T^2 \Big)^{\frac{1}{2}} + \\
&\qquad + \frac{1}{2} \Big(\sum_{T \in \mathcal{T}_h} \sum_{E \in \partial T \setminus \Gamma_D} \frac{h_{min,T}^2}{h_{E,T}} \|r_E(u_h)\|_E^2 \Big)^{\frac{1}{2}} \cdot \Big(\sum_{T \in \mathcal{T}_h} \sum_{E \in \partial T \setminus \Gamma_D} \frac{h_{E,T}}{h_{min,T}^2} \|v - R_o v\|_E^2 \Big)^{\frac{1}{2}}.
\end{aligned}$$

The second root term is readily bounded by means of the H_o^1 interpolation theorem 3.3

$$\|h_{min}^{-1}(\mathbf{x}) \cdot (v - R_o v)\| = \left(\sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \|v - R_o v\|_T^2 \right)^{1/2} \lesssim \|\nabla v\| \qquad .$$

In order to bound the last term the trace inequality (2.12) is applied to $\|v - R_o v\|_E$ giving

$$\sum_{E \in \partial T \setminus \Gamma_D} \frac{h_{E,T}}{h_{min,T}^2} \|v - R_o v\|_E^2 \lesssim h_{min,T}^{-2} \left(\|v - R_o v\|_T^2 \; + \; \|C_T^T \nabla (v - R_o v)\|_T^2 \right) .$$

Recalling again the H_o^1 interpolation theorem results in

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} \sum_{E \in \partial T \setminus \Gamma_D} \frac{h_{E,T}}{h_{min,T}^2} \|v - R_o v\|_E^2 \; &\lesssim \; \|h_{min}^{-1}(\mathbf{x}) \cdot (v - R_o v)\|^2 \\ &\qquad + \|h_{min}^{-1}(\mathbf{x}) \cdot C^T(\mathbf{x}) \nabla (v - R_o v)\|^2 \\ &\lesssim \; \|\nabla v\|^2 \qquad . \end{aligned}$$

Combining all inequalities yields

$$(\nabla(u - u_h), \nabla v) \;\lesssim\; \left(\sum_{T \in \mathcal{T}_h} h_{min,T}^2 \|f + \Delta u_h\|_T^2 \quad + \quad D_{h,1}^2 \right)^{\frac{1}{2}} \cdot \|\nabla v\| \qquad .$$

Substituting $v := u - u_h \in H_{o,T}^1(\Omega)$ gives the second formulation of the upper bound of the error. Note that $\Delta u_h = 0$ on T since we are using linear basis functions, but this notation indicates how to modify the estimator for higher order basis functions.

Finally, utilizing the triangle inequality $\|f + \Delta u_h\|_T \leq \|r_T(u_h)\|_T + \|f - P_{L_2} f\|_T$ results in the first upper bound of the error. ■

Remark 3.4 The term $P_{L_2} f$ appears both in the definition of the element residual $r_T(u_h)$ as well as in inequalities (3.15) and (3.16).

Assume for the moment that this term is replaced by an arbitrary function from $L_2(\Omega)$. Then one would obtain an upper bound of the error similar to (3.16) but (3.15) would no longer hold. Choosing f instead of $P_{L_2} f$ would, for example, result in

$$\|\nabla(u - u_h)\|^2 \lesssim \sum_{T \in \mathcal{T}_h} \eta_{R,T}^2(u_h)$$

with $r_T(u_h) := f\Big|_T$ which is exactly the second formulation of the upper bound.

There are two reasons for using $P_{L_2} f$. Firstly, this term (or a similar one) is required to derive a lower bound of the error.

Secondly, it may be difficult to evaluate exactly the norm $\|r_T\|$ or the integrals over f, respectively. If f is suitably smooth (e.g. $f \in L_2 \cap C^0(T)$) then Pf may represent a quadrature rule. For example, the midpoint quadrature rule is equivalent to $P : L_2(T) \cap C^0(T) \to \mathbb{P}^0(T) \quad , \quad Pf(\mathbf{x}) := f(\mathbf{x}_{midpoint})$ on T. The term $\|f - Pf\|$ then assesses the quality of the quadrature error.

The main feature of all quadrature rules is that Pf always maps into a finite dimensional space. The proofs for different quadrature rules would be similar to the one here. Hence the L_2 projection serving as $P = P_{L_2}$ may suffice. □

Remark 3.5 Siebert [28] proposes a similar error estimator for rectangular or cuboidal finite elements. There the factor of the gradient jump in the definition of the error estimator equals h_E instead of h_{min}^2/h_E as in our work (cf. (3.14)). Thus Siebert has to impose an additional condition on u_h to give a reliable lower bound of the error. This renders our estimator slightly more general. $\square$

Remark 3.6 The condition $u - u_h \in H^1_{o,\mathcal{T}}(\Omega)$ is discussed in more detail in section 5.2.1 on page 70. $\square$

3.3 A local Dirichlet problem error estimator

3.3.1 Introduction and definition

Local problem error estimators have been known for a long time [8, 9, 30, 31]. In this section we demonstrate on the example of a local Dirichlet problem that these error estimators can be applied to anisotropic meshes as well. As far as we know this is the first rigorous analytical investigation into this type of anisotropic error estimator.

The basic idea is to solve a local problem with a higher accuracy. The difference to the finite element solution serves as error estimator. The remainder of this section is devoted to the definition of the local problem and to the estimator. Then theorem 3.6 states the equivalence of this local problem error estimator $\eta_{D,T}$ and the residual error estimator $\eta_{R,T}$. Lower and upper bounds are given as well. The proofs conclude that section. Finally, in section 3.3.3 it is shown that the local Dirichlet problem is well–conditioned.

Recall that $a(\cdot,\cdot) = (\nabla\cdot, \nabla\cdot)$ is the bilinear form associated with the weak formulation of the Poisson problem. Again, u and u_h denote the exact and the FEM solution, respectively. Let T be an arbitrary but fixed tetrahedron. Recall that the domain ω_T is formed by T and all (at most four) adjacent tetrahedra that have a common face with T. The true error $e = u - u_h$ then satisfies

$$a(e, v) = a(u - u_h, v) = \int_{\omega_T} f \cdot v - \int_{\omega_T} \nabla^T u_h \nabla v \qquad \forall\, v \in H^1_o(\omega_T) \quad .$$

A straightforward approximation of the space $H^1_o(\omega_T)$ by some local, finite dimensional space V_T leads to the problem:

$$\begin{array}{lrcll} \text{Find } e_T \in V_T : & a(e_T, v_T) & = & a(u - u_h, v_T) & \forall\, v_T \in V_T \\ \text{or, equivalently,} & \displaystyle\int_{\omega_T} \nabla^T e_T \nabla v_T & = & \displaystyle\int_{\omega_T} f \cdot v_T - \int_{\omega_T} \nabla^T u_h \nabla v_T & \forall\, v_T \in V_T \quad . \end{array}$$

Then $\|\nabla e_T\|_{\omega_T}$ could serve as error estimator.

From the isotropic error analysis (e.g. [31]) it is known that the local space

$$V_T \;:=\; \operatorname{span}\{b_{T'}, b_E \;:\; T' \subset \omega_T\,,\; E \subset \partial T \setminus \Gamma_D\} \quad \subset H^1_o(\omega_T)$$

is well–suited, with b_T and b_E being the element and face bubble function defined in (2.17) and (2.18), respectively.

Finally, for reasons that were explained in remark 3.4 on page 35 we want to use $P_{L_2} f$ instead of f. Thus the local problem and the estimator are defined as follows.

Definition 3.5 (Local problem error estimator) *Find a solution $e_T \in V_T$ of the local problem*

$$\begin{array}{rcll} \displaystyle\int_{\omega_T} \nabla^T e_T \nabla v_T & = & \displaystyle\int_{\omega_T} P_{L_2} f \cdot v_T - \int_{\omega_T} \nabla^T u_h \nabla v_T & \forall\, v_T \in V_T \\ \textit{or, equivalently,} \quad a(e_T, v_T) & = & \displaystyle a(u - u_h, v_T) + \int_{\omega_T} (P_{L_2} f - f) \cdot v_T & \forall\, v_T \in V_T\,. \end{array}$$

Then

$$\eta_{D,T} := \|\nabla e_T\|_{\omega_T} \tag{3.18}$$

is said to be the local Dirichlet problem error estimator.

The weak formulation in the definition above can be seen as the discrete analogue of the local Dirichlet problem

$$\begin{aligned} -\Delta\varphi &= P_{L_2} f && \text{in } \omega_T \\ \varphi &= u_h && \text{on } \partial\omega_T \end{aligned}$$

which is solved on the manifold $u_h + V_T$.

3.3.2 Equivalence and bounds of the local problem error estimator

Consider an arbitrary tetrahedron T. Let the four tetrahedra of $\omega_T \setminus T$ be denoted by $T_1 \ldots T_4$. Denote the faces of T by $E_i := T \cap T_i$. The modifications for T being a boundary tetrahedron are obvious. The next lemma facilitates the proof of the actual theorem.

Lemma 3.5 *The following inequalities hold for all $v_T \in V_T$.*

$$\|v_T\|_{\omega_T} \lesssim h_{min,T} \cdot \|\nabla v_T\|_{\omega_T} \tag{3.19}$$

$$\|v_T\|_{E_i} \lesssim h_{E_i}^{-1/2} \cdot h_{min,T} \cdot \|\nabla v_T\|_{T_i} \quad . \tag{3.20}$$

Proof: Let us start with inequality (3.19). Assume that $v_T \in V_T$ is represented as

$$v_T = \alpha_0 \cdot b_T + \sum_{i=1}^{4} \alpha_i \cdot b_{T_i} + \sum_{i=1}^{4} \beta_i \cdot b_{E_i} \qquad \alpha_i, \beta_i \in \mathbb{R}.$$

We split the domain ω_T in $\omega_T = \bigcup_{i=1}^{4} T_i \cup T$. The following table outlines the two kinds of subdomains of ω_T, the corresponding representation of v_T and the estimates that we want to prove. Note that, on the subdomain T, the norms of the left-hand side and the right-hand side are taken over different domains (cf. also remark 3.7 on page 41).

subdomain of ω_T	representation of v_T	estimates
T_i	$v_T\Big\vert_{T_i} = \alpha_i \cdot b_{T_i} + \beta_i \cdot b_{E_i}\Big\vert_{T_i}$	$\|v_T\|_{T_i} \lesssim h_{min,T} \cdot \|\nabla v_T\|_{T_i}$
T	$v_T\Big\vert_T = \alpha_0 b_T + \sum_{i=1}^{4} \beta_i \cdot b_{E_i}\Big\vert_T$	$\|\beta_i \cdot b_{E_i}\|_T \lesssim h_{min,T} \cdot \|\nabla v_T\|_{T_i}$
		$\|\alpha_0 \cdot b_T\|_T \lesssim h_{min,T} \cdot \|\nabla v_T\|_{\omega_T}$
		$\Rightarrow \|v_T\|_T \lesssim h_{min,T} \cdot \|\nabla v_T\|_{\omega_T}$

Let us first consider a tetrahedron T_i over which v_T is reduced to

$$v_T\Big|_{T_i} = \alpha_i \cdot b_{T_i} + \beta_i \cdot b_{E_i}\Big|_{T_i} \quad .$$

Without loss of generality a scaling $\alpha_i^2 + \beta_i^2 = 1$ is assumed here.

In this proof the transformation from the reference tetrahedron $\hat{T}_i$ onto T_i is utilized. The reference tetrahedron $\hat{T}_i$ is uniquely determined by its vertices $(0,0,0)^T$, $(1,0,0)^T$,

$(\hat{x}_2, 1, 0)^T$, and $(\hat{x}_3, \hat{y}_3, 1)^T$, with $0 < \hat{x}_2, \hat{x}_3 < 1$ and $|\hat{y}_3| < 1$ (cf. section 2.1.3). Define the compact set

$$S := \left\{ (\hat{x}_2, \hat{x}_3, \hat{y}_3, \alpha_i, \beta_i) \quad : \quad 0 \leq \hat{x}_2, \hat{x}_3 \leq 1 \,,\ |\hat{y}_3| \leq 1 \,,\ \alpha_i^2 + \beta_i^2 = 1 \right\} \qquad .$$

Let $\partial/\partial\hat{z}$ be the partial derivative corresponding to the third coordinate axis. Obviously

$$\begin{aligned} \|\hat{v}_T\|_{\hat{T}_i} &= \|\alpha_i \cdot \hat{b}_{T_i} + \beta_i \cdot \hat{b}_{E_i}\|_{\hat{T}_i} \\ \text{and} \qquad \left\|\frac{\partial}{\partial\hat{z}}\hat{v}_T\right\|_{\hat{T}_i} &= \left\|\alpha_i \cdot \frac{\partial}{\partial\hat{z}}\hat{b}_{T_i} + \beta_i \cdot \frac{\partial}{\partial\hat{z}}\hat{b}_{E_i}\right\|_{\hat{T}_i} \end{aligned}$$

vary continuously over S and thus attain their maximum and minimum, respectively. The terms $\partial\hat{b}_{T_i}/\partial\hat{z}$ and $\partial\hat{b}_{E_i}/\partial\hat{z}$ are polynomials of order 3 and 2, respectively. Thus $\partial\hat{v}_T/\partial\hat{z}$ could only vanish everywhere in $\hat{T}_i$ if $\hat{v}_T \equiv 0$. This is impossible since $\alpha_i^2 + \beta_i^2 = 1$, and therefore the minimum of $\|\partial\hat{v}_T/\partial\hat{z}\|_{\hat{T}_i}$ over S is positive. Hence

$$\|\hat{v}_T\|_{\hat{T}_i} \lesssim \left\|\frac{\partial}{\partial\hat{z}}\hat{v}_T\right\|_{\hat{T}_i} \qquad \forall\, v_T \in V_T \qquad .$$

The transformation $F_C : \hat{T}_i \mapsto T_i$ results in

$$\|v_T\|_{T_i} \lesssim \|\mathbf{e}_3^T \cdot C_{T_i}^T \cdot \nabla v_T\|_{T_i} \leq \|C_{T_i} \cdot \mathbf{e}_3\|_{\mathbb{R}^3} \cdot \|\nabla v_T\|_{T_i} \qquad \forall\, v_T \in V_T \qquad .$$

Because of $\|C_{T_i} \cdot \mathbf{e}_3\|_{\mathbb{R}^3} = \|\mathbf{p}_3\|_{\mathbb{R}^3} = h_{min,T_i} \sim h_{min,T}$ we obtain

$$\|v_T\|_{T_i} \lesssim h_{min,T} \cdot \|\nabla v_T\|_{T_i} \qquad \forall\, v_T \in V_T \,,\ \forall\, i = 1 \ldots 4 \qquad . \tag{3.21}$$

This last inequality does not hold for T instead of T_i (cf. remark 3.7 below), and the arguments from above cannot be utilized. This is due to the fact that $\partial\hat{v}_T/\partial\hat{z}$ can vanish (everywhere in $\hat{T}$) over a set S defined similarly. Hence a more sophisticated investigation is necessary.

The representation of v_T over T is

$$v_T\Big|_T = \sum_{i=1}^{4} \beta_i \cdot b_{E_i}\Big|_T + \alpha_0 \cdot b_T \qquad .$$

The terms $\beta_i \cdot b_{E_i}\Big|_T$ and $\alpha_0 \cdot b_T$ are dealt with separately.

In order to bound the first term we apply exactly the same arguments as above for (3.21) and obtain on the tetrahedron T_i

$$\|\beta_i \cdot b_{E_i}\|_{T_i} \lesssim h_{min,T} \cdot \|\nabla(\alpha_i \cdot b_{T_i} + \beta_i \cdot b_{E_i})\|_{T_i} = h_{min,T} \cdot \|\nabla v_T\|_{T_i}$$

for all $v_T \in V_T$. The estimate

$$\|\beta_i \cdot b_{E_i}\|_T \lesssim \|\beta_i \cdot b_{E_i}\|_{T_i} \lesssim h_{min,T} \cdot \|\nabla v_T\|_{T_i} \qquad \forall\, v_T \in V_T \tag{3.22}$$

is readily obtained since $\|b_{E_i}\|_{T_i} = \sqrt{27/280} \cdot |T_i|^{1/2}$ and $|T| \sim |T_i|$. Note that the left norm is over T but the right one over T_i.

Secondly we want to prove that

$$\|\alpha_0 \cdot b_T\|_T \lesssim h_{min,T} \cdot \|\nabla v_T\|_{\omega_T} \qquad \forall\, v_T \in V_T \qquad . \tag{3.23}$$

The case $\alpha_0 = 0$ is trivial so we assume $\alpha_0 = 1$ without loss of generality. We distinguish two cases.

1. $\forall\, i : |\beta_i| \le c_0 = \frac{256}{1485}\sqrt{55}$.
   We consider the compact set
$$S := \Big\{(\hat{x}_2, \hat{x}_3, \hat{y}_3, \beta_1, \beta_2, \beta_3, \beta_4) \quad : \quad 0 \le \hat{x}_2, \hat{x}_3 \le 1\,,\ |\hat{y}_3| \le 1\,,\ |\beta_i| \le c_0\Big\}.$$
   Obviously $\|\alpha_0 \cdot \hat{b}_T\|_{\hat{T}}$ is constant. The term $\|\partial\hat{v}_T/\partial\hat{z}\|_{\hat{T}}$ varies continuously over S and thus attains its minimum. Furthermore $\partial\hat{b}_T/\partial\hat{z}$ and $\partial\hat{b}_{E_i}/\partial\hat{z}$ are polynomials of order 3 and 2, respectively. Thus $\partial\hat{v}_T/\partial\hat{z}$ does not vanish everywhere in $\hat{T}$ if $\alpha_0 \ne 0$. Hence the minimum of the norm $\|\partial\hat{v}_T/\partial\hat{z}\|_{\hat{T}}$ is positive which results in
$$\|\alpha_0 \cdot \hat{b}_T\|_{\hat{T}} \lesssim \left\|\frac{\partial}{\partial\hat{z}}\hat{v}_T\right\|_{\hat{T}} \qquad .$$
   The transformation onto T is as before and yields
$$\|\alpha_0 \cdot b_T\|_T \lesssim h_{min,T} \cdot \|\nabla v_T\|_T \qquad \forall\, v_T \in V_T \qquad .$$
   Thus the desired result is obtained.

2. $\exists i : |\beta_i| > c_0 = \frac{256}{1485}\sqrt{55}$.
   Here the compactness argument cannot be applied. But a straightforward calculation gives
$$\|\hat{b}_T\|_{\hat{T}} = \frac{64}{3465}\sqrt{77} \qquad \text{and} \qquad \|\hat{b}_{E_i}\|_{\hat{T}} = \frac{3}{140}\sqrt{35}$$
   and thus
$$\|\alpha_0 \cdot \hat{b}_T\|_{\hat{T}} \;=\; c_0 \cdot \|\hat{b}_{E_i}\|_{\hat{T}} \;<\; \|\beta_i \cdot \hat{b}_{E_i}\|_{\hat{T}} \qquad .$$
   The transformation and the application of the previous result (3.22) yield
$$\|\alpha_0 \cdot b_T\|_T \;<\; \|\beta_i \cdot b_{E_i}\|_T \lesssim h_{min,T} \cdot \|\nabla v_T\|_{T_i} \qquad \forall\, v_T \in V_T$$
   and hence the desired inequality.

Recalling inequalities (3.21) – (3.23) and the representation of v_T results in
$$\begin{aligned}
\|v_T\|^2_{\omega_T} &= \sum_{i=1}^{4} \|v_T\|^2_{T_i} \,+\, \|\alpha_0 \cdot b_T \,+\, \sum_{i=1}^{4} \beta_i \cdot b_{E_i}\|^2_T \\
&\le \sum_{i=1}^{4} \|v_T\|^2_{T_i} \,+\, 5\,\|\alpha_0 \cdot b_T\|^2_T \,+\, 5\sum_{i=1}^{4} \|\beta_i \cdot b_{E_i}\|^2_T \\
&\lesssim h^2_{min,T} \cdot \|\nabla v_T\|^2_{\omega_T} \qquad .
\end{aligned}$$
Hence inequality (3.19) is proven.

Inequality (3.20) follows from
$$\begin{aligned}
\|v_T\|_{E_i} \;=\; \|\beta_i \cdot b_{E_i}\|_{E_i} &\;=\; 3 \cdot h_{E_i}^{-1/2} \cdot \|\beta_i \cdot b_{E_i}\|_{T_i} \\
\text{and} \qquad \|\beta_i \cdot b_{E_i}\|_{T_i} &\;\lesssim\; \cdot h_{min,T} \cdot \|\nabla v_T\|_{T_i} \qquad \forall\, v_T \in V_T
\end{aligned}$$
which was proven above.

Finally it is easily verified that both inequalities also hold if T is a boundary tetrahedron. ∎

Remark 3.7 Note that in inequality (3.19) the norm over the whole domain ω_T is necessary. Especially the inequality

$$\|v_T\|_T \lesssim h_{min,T} \cdot \|\nabla v_T\|_T \qquad \forall\, v_T \in V_T$$

does *not* hold.

Consider for example a tetrahedron T with vertices $P_0 = \mathbf{0}$, $P_1 = \mathbf{e}_1$, $P_2 = \mathbf{e}_2$, and $P_3 = h \cdot \mathbf{e}_3$ with $h \to 0$. Choose $v_T := b_{E_1} + b_{E_2}$, with $b_{E_1} = 27xy(1 - x - y - z/h)$ and $b_{E_2} = 27xyz/h$ being two face bubble functions. Then

$$\begin{aligned} \|v_T\|_T &= \sqrt{27/560} \cdot h^{1/2} \\ \text{and} \qquad h_{min,T} \cdot \|\nabla v_T\|_T &= \sqrt{81/35} \cdot \frac{h^{3/2}}{\sqrt{1+2h^2}} \qquad . \end{aligned}$$

Thus the abovementioned inequality does not hold with a multiplicative constant independent of h. Note also that the corresponding isotropic estimates are much easier to derive. □

Remark 3.8 If one knows that the domain ω_T is contained in a rectangular prism with minimal side length $l \sim h_{min,T}$ then (3.19) coincides with the Friedrichs inequality. But we have not shown such a geometrical condition and thus had to proceed in the way described above. □

Now the main theorem will be stated and proven.

Theorem 3.6 (Local problem error estimator) *The local problem error estimator is equivalent to the residual error estimator $\eta_{R,T}$ in the following sense:*

$$\eta_{D,T} \lesssim \sum_{T' \subset \omega_T} \eta_{R,T'} \tag{3.24}$$

$$\eta_{R,T} \lesssim \eta_{D,T} \qquad . \tag{3.25}$$

The lower bound of the error is

$$\eta_{D,T}(u_h) \le \|\nabla(u - u_h)\|_{\omega_T} + c \cdot h_{min,T} \cdot \|f - P_{L_2} f\|_{\omega_T} \qquad . \tag{3.26}$$

Finally assume that the mesh is adapted to the error, i.e. $u - u_h \in H^1_{o,\mathcal{T}}(\Omega)$. Then the error is bounded globally from above by

$$\|\nabla(u - u_h)\| \lesssim \left(\sum_{T \in \mathcal{T}_h} \eta^2_{D,T} + \sum_{T \in \mathcal{T}_h} h^2_{min,T} \cdot \|f - P_{L_2} f\|^2_T \right)^{1/2} \qquad . \tag{3.27}$$

Proof: Let T be an arbitrary but fixed tetrahedron throughout the proofs.

For the first inequality recall the definition (3.18) of $\eta_{D,T}$ and $r_T(u_h)$. By integration by parts we obtain

$$\begin{aligned} \eta^2_{D,T} = \|\nabla e_T\|^2_{\omega_T} &\overset{(3.18)}{=} \int_{\omega_T} P_{L_2} f \cdot e_T - \int_{\omega_T} \nabla^T u_h \nabla e_T \\ &= \sum_{T' \subset \omega_T} \int_{T'} (P_{L_2} f + \Delta u_h) \cdot e_T \; - \sum_{E \subset \partial T \setminus \Gamma_D} \int_E r_E(u_h) \cdot e_T \\ &\le \left(\sum_{T' \subset \omega_T} \|r_{T'}(u_h)\|^2_{T'} \right)^{1/2} \cdot \|e_T\|_{\omega_T} + \sum_{E \subset \partial T \setminus \Gamma_D} \|r_E(u_h)\|_E \cdot \|e_T\|_E . \end{aligned}$$

Now $\|e_T\|_{\omega_T}$ and $\|e_T\|_E$ are bounded using lemma 3.5 on page 38 which results in

$$\eta_{D,T}^2 \lesssim h_{min,T} \cdot \left(\sum_{T' \subset \omega_T} \|r_{T'}(u_h)\|_{T'} + \sum_{E \subset \partial T \setminus \Gamma_D} h_E^{-1/2} \|r_E(u_h)\|_E \right) \cdot \|\nabla e_T\|_{\omega_T} \quad .$$

Recalling $\eta_{D,T} = \|\nabla e_T\|_{\omega_T}$ proves the desired inequality (3.24)

$$\eta_{D,T} \lesssim \sum_{T' \subset \omega_T} \eta_{R,T'} \quad .$$

For the proof of the second inequality we require bounds of $\eta_{R,T}$, and thus of $\|r_{T'}\|_{T'}$ and $\|r_E\|_E$. Let $T' \subset \omega_T$ be an arbitrary tetrahedron. Recall definition (2.17) of the bubble function $b_{T'}$ and set $v_{T'} := b_{T'} \cdot r_{T'} \in V_T$. Inverse inequality (2.20) and integration by parts imply

$$\begin{aligned}
\|r_{T'}\|_{T'}^2 &\lesssim \|b_{T'}^{1/2} \cdot r_{T'}\|_{T'}^2 = \int_{T'} r_{T'} \cdot v_{T'} \\
&= \sum_{T'' \subset \omega_T} \int_{T''} (P_{L_2} f + \Delta u_h) \cdot v_{T'} \qquad \text{since } v_{T'} \in H_o^1(T') \\
&= \int_{\omega_T} P_{L_2} f \cdot v_{T'} - \int_{\omega_T} \nabla^T u_h \cdot \nabla v_{T'} \overset{(3.18)}{=} \int_{\omega_T} \nabla^T e_T \cdot \nabla v_{T'} \\
&\leq \|\nabla e_T\|_{\omega_T} \cdot \|\nabla v_{T'}\|_{\omega_T} \quad .
\end{aligned}$$

The inverse inequality (2.21) states

$$\|\nabla v_{T'}\|_{T'} = \|\nabla(b_{T'} \cdot r_{T'})\|_{T'} \lesssim h_{min,T'}^{-1} \cdot \|r_{T'}\|_{T'} \quad .$$

Combining both inequalities yields

$$\|r_{T'}\|_{T'} \lesssim h_{min,T}^{-1} \cdot \eta_{D,T} \qquad \forall\, T' \in \omega_T$$

since $h_{min,T'}$ does not change rapidly across adjacent tetrahedra T'.

The bound of $\|r_E\|_E$ is obtained similarly. Recall the definitions (2.18) and (2.19) of the bubble function b_E and the extension operator F, respectively, and set $v_E := b_E \cdot F(r_E) \in V_T$. Inverse inequality (2.22) and integration by parts imply

$$\begin{aligned}
\|r_E\|_E^2 &\lesssim \|b_E^{1/2} \cdot r_E\|_E^2 = \int_E r_E \cdot v_E = \sum_{T' \subset \omega_E} \int_{\partial T'} \frac{\partial u_h}{\partial n} \cdot v_E \\
&= \sum_{T' \subset \omega_E} \int_{\partial T'} \Delta u_h \cdot v_E + \int_{\omega_E} \nabla^T u_h \cdot \nabla v_E \\
&\overset{(3.18)}{=} \sum_{T' \subset \omega_E} \int_{\partial T'} r_{T'} \cdot v_E - \int_{\omega_E} \nabla^T e_T \cdot \nabla v_E \\
&\leq \sum_{T' \subset \omega_E} \|r_{T'}\|_{T'} \cdot \|v_E\|_{T'} + \|\nabla e_T\|_{\omega_T} \cdot \|\nabla v_E\|_{\omega_E} \quad .
\end{aligned}$$

The inverse inequalities (2.23) and (2.24) imply

$$\begin{aligned}
\|v_E\|_{T'} &= \|b_E \cdot F(r_E)\|_{T'} \lesssim h_E^{1/2} \cdot \|r_E\|_E \\
\|\nabla v_E\|_{T'} &= \|\nabla(b_E \cdot F(r_E))\|_{T'} \lesssim h_E^{1/2} \cdot h_{min,T'}^{-1} \cdot \|r_E\|_E
\end{aligned}$$

The last three estimates and the previous bound of $\|r_{T'}\|_{T'}$ result in

$$\|r_E\|_E \lesssim h_E^{1/2} \cdot h_{min,T}^{-1} \cdot \eta_{D,T}$$

since $h_{min,T'}$ does not change rapidly across adjacent tetrahedra T'. Finally the desired estimate (3.25) is obtained:

$$\eta_{R,T}^2 = h_{min,T}^2 \cdot \left(\|r_T\|_T^2 + \sum_{E \subset \partial T \setminus \Gamma_D} h_E^{-1} \cdot \|r_E\|_E^2 \right) \lesssim \eta_{D,T}^2 \qquad .$$

In order to prove the third estimate, recall again that

$$\begin{aligned} \eta_{D,T}^2 = \|\nabla e_T\|_{\omega_T}^2 &\overset{(3.18)}{=} \int_{\omega_T} P_{L_2} f \cdot e_T - \int_{\omega_T} \nabla^T u_h \nabla e_T \\ &= \int_{\omega_T} \nabla^T (u - u_h) \cdot \nabla e_T + \int_{\omega_T} (P_{L_2} f - f) \cdot e_T \\ &\leq \|\nabla(u - u_h)\|_{\omega_T} \cdot \|\nabla e_T\|_{\omega_T} + \|f - P_{L_2} f\|_{\omega_T} \cdot \|e_T\|_{\omega_T} \qquad . \end{aligned}$$

Applying estimate (3.19) of lemma 3.5 on page 38 to $e_T \in V_T$ results now readily in

$$\eta_{D,T}^2(u_h) \leq \|\nabla(u - u_h)\|_{\omega_T}^2 + c \cdot h_{min,T}^2 \cdot \|f - P_{L_2} f\|_{\omega_T}^2 \qquad .$$

Note that here the only constant appears at approximation term.

Finally inequality (3.27) follows immediately from (3.16) and (3.25). ■

3.3.3 Condition number of the FEM matrix of the local problem

The error estimator $\eta_{D,T}$ requires the solution of a local finite element problem. It can be shown easily that this problem is well-behaved, i.e. the condition number of the corresponding FEM matrix is bounded independently of the aspect ratio of the elements under consideration.

Let T be an interior tetrahedron. An arbitrary function $v_T \in V_T$ can be written as

$$v_T = \alpha_0 \cdot b_T + \sum_{i=1}^{4} \alpha_i \cdot b_{T_i} + \sum_{i=1}^{4} \beta_i \cdot b_{E_i} \qquad \alpha_i, \beta_i \in \mathbb{R} \qquad .$$

For the remainder of this section, define

$$\mathbf{v} := (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta_1, \beta_2, \beta_3, \beta_4)^T \in \mathbb{R}^9 \qquad .$$

By means of the FEM isomorphism

$$v_T \in V_T \qquad \longleftrightarrow \qquad \mathbf{v} \in \mathbb{R}^9$$

one obtains

$$a(v_T, w_T) = (K_T \mathbf{v}, \mathbf{w}) \qquad \forall v_T, w_T \in V_T \qquad .$$

Here $K_T \in \mathbb{R}^{9 \times 9}$ is the usual FEM stiffness matrix which is symmetric and positive definite.

Theorem 3.7 (Condition number) *The condition number $\kappa(K_T)$ of the local problem stiffness matrix K_T is bounded independently of T:*

$$\kappa(K_T) \lesssim 1 \qquad \forall\, T \in \mathcal{T}_h \qquad .$$

Proof: We start with

$$(K_T\mathbf{v}, \mathbf{v}) = a(v_T, v_T) = \|\nabla v_T\|^2_{\omega_T} \qquad .$$

Lemma 2.7 on page 19 states

$$\|\nabla b_{E_i}\|_{\omega_T} \sim \|\nabla b_T\|_{\omega_T} \sim \|\nabla b_{T_i}\|_{\omega_T} \sim h^{-1}_{min,T} \cdot |T|^{1/2}$$

since $|T_i| \sim |T|$ for adjacent tetrahedra. Thus

$$\begin{aligned}
\|\nabla v_T\|^2_{\omega_T} &\le \|\alpha_0 \cdot \nabla b_T\|^2_{\omega_T} + \sum_{i=1}^{4} \|\alpha_i \cdot \nabla b_{T_i}\|^2_{\omega_T} + \sum_{i=1}^{4} \|\beta_i \cdot \nabla b_{E_i}\|^2_{\omega_T} \\
&\lesssim h^{-2}_{min,T} \cdot |T| \cdot \|\mathbf{v}\|^2_{\mathbb{R}^9} \qquad .
\end{aligned}$$

On the other hand one has from (3.19)

$$\|\nabla v_T\|_{\omega_T} \gtrsim h^{-1}_{min,T} \cdot \|v_T\|_{\omega_T} \qquad .$$

Decompose the domain $\omega_T = \bigcup_{i=1}^{4} T_i \cup T$ as above. A straightforward calculation yields

$$\begin{aligned}
\|v_T\|^2_{T_i} &= \|\alpha_i \cdot b_{T_i} + \beta_i \cdot b_{E_i}\|^2_{T_i} = 6|T_i| \cdot \left\|\alpha_i \cdot \bar{b}_{T_i} + \beta_i \cdot \bar{b}_{E_i}\right\|^2_{\bar{T}_i} \\
&= 6|T_i| \cdot \left(\frac{4096}{155925}\alpha_i^2 + \frac{16}{525}\alpha_i\beta_i + \frac{9}{560}\beta_i^2\right) \\
&\gtrsim |T| \cdot (\alpha_i^2 + \beta_i^2) \qquad .
\end{aligned}$$

(Alternatively, a similar compactness argument as before could be employed.) Analogously

$$\begin{aligned}
\|v_T\|^2_T &= \|\alpha_0 \cdot b_T + \sum_{i=1}^{4} \beta_i \cdot b_{E_i}\|^2_T = 6|T| \cdot \|\alpha_0 \cdot \bar{b}_T + \sum_{i=1}^{4} \beta_i \cdot \bar{b}_{E_i}\|^2_{\bar{T}} \\
&\ge 6|T| \cdot \frac{331\,547 - 17\sqrt{33\,671\,493}}{9\,979\,200} \cdot \left(\alpha_0^2 + \sum_{i=1}^{4} \beta_i^2\right) \qquad .
\end{aligned}$$

(The constants have been evaluated using a computer algebra system.) Together with $|T_i| \sim |T|$ one eventually obtains

$$\|\nabla v_T\|^2_{\omega_T} \gtrsim h^{-2}_{min,T} \cdot \left(\|v_T\|^2_T + \sum_{i=1}^{4} \|v_T\|^2_{T_i}\right) \gtrsim h^{-2}_{min,T} \cdot |T| \cdot \|\mathbf{v}\|^2_{\mathbb{R}^9}$$

and thus

$$h^{-2}_{min,T} \cdot |T| \cdot \|\mathbf{v}\|^2_{\mathbb{R}^9} \lesssim \|\nabla v_T\|^2_{\omega_T} = (K_T\mathbf{v}, \mathbf{v}) \lesssim h^{-2}_{min,T} \cdot |T| \cdot \|\mathbf{v}\|^2_{\mathbb{R}^9}$$

which implies the bounded condition number of K_T. The case of T being a boundary tetrahedron is dealt with analogously. ■

3.4 L_2 error estimator

An L_2 error estimator for non-uniform isotropic meshes has been derived by Eriksson and Johnson [15]. Here we will propose an L_2 error estimator that is suitable for anisotropic meshes. Some ideas of the aforementioned work have been utilized and extended to our case. To our knowledge, an anisotropic L_2 error estimator has not been analysed before.

The framework of the proofs is similar to the one of the residual error estimator of section 3.2. Special bubble functions which are required to prove the lower bound of the error are introduced in section 3.4.1. For these bubble functions inverse estimates similar to (2.20) – (2.24) are desired. This leads to an additional assumption on the mesh which will be discussed in section 3.4.2. This mesh requirement might be purely technical and be due to the techniques used here. Nevertheless it seems to be fairly natural from a heuristic point of view. Section 3.4.3 is devoted to the relation between the anisotropic mesh and the anisotropic solution. Interpolation error estimates are derived in section 3.4.4. The L_2 error estimator is given in section 3.4.5. Firstly, however, some useful notation is introduced.

Let $M := (m_{i,j})_{i,j=1}^d$ be a matrix of L_2 functions $m_{i,j} \in L_2(\omega)$. Let $\|M\|_\omega^2 := \sum\limits_{i,j=1}^d \|m_{i,j}\|_\omega^2$ be the usual L_2 norm of M, and define $|M|^2 := \sum\limits_{i,j=1}^d |m_{i,j}|^2$. Let $D^2 v := \left(\dfrac{\partial^2 v}{\partial x_i \partial x_j}\right)_{i,j=1}^d$.

Finally, note that in sections 3.4.1 and 3.4.5 special bubble functions are used that differ from the general bubble functions defined previously in section 2.3.3. For simplicity the same notation b_T and b_E is used here.

3.4.1 Special L_2 bubble functions

For the proof of the lower bound of the error we utilize bubble functions of a higher smoothness, i.e. we now demand $b_T \in H_o^2(T)$ and $b_E \in H_o^2(\omega_E)$.

Let $T \in \mathcal{T}_h$ be an arbitrary tetrahedron, and denote by $\lambda_{T,1}, \cdots, \lambda_{T,4}$ its barycentric coordinates. The *element bubble function* $b_T \in \mathbb{P}^8(T) \cap H_o^2(T)$ is defined by

$$b_T := 4^8 \cdot \lambda_{T,1}^2 \cdot \lambda_{T,2}^2 \cdot \lambda_{T,3}^2 \cdot \lambda_{T,4}^2 \qquad \text{on } T \qquad . \tag{3.28}$$

We also need a bubble function b_E defined on $\omega_E = T_1 \cup T_2$. The technical definition is due to the smoothness requirement $b_E \in H_o^2(\omega_E)$. More precisely, we will construct (a whole class of) bubble functions b_E that depend on a *guide vector* $\mathbf{a}_E$ associated with the face E. The following definition of the (class of) bubble functions is given for general guide vectors $\mathbf{a}_E$. The flexibility in the choice of the guide vector is exploited in the next section when the guide vector will be specified.

Consider an arbitrary inner face (triangle) E of $\mathcal{T}_h$ and the domain $\omega_E = T_1 \cup T_2$. The bubble function is defined separately on each tetrahedron; so let T be any of the two tetrahedra. Via the transformation A_T^{-1} it is mapped onto the unitary tetrahedron $\bar{T}$ with a face $\bar{E}$. Let the barycentric coordinates of $\bar{T}$ be numbered such that the ones associated with the three nodal points of $\bar{E}$ are $\lambda_1 \ldots \lambda_3$.

Firstly we define three cut-off functions $b_i \in H^2(\bar{T})$, $i = 1 \ldots 3$, by

$$b_i = b_i(\lambda_1, \lambda_2, \lambda_3, \lambda_4) := \begin{cases} -128\lambda_i^3 + 48\lambda_i^2 & \text{if } \lambda_i \le 1/4 \\ 1 & \text{if } \lambda_i > 1/4 \end{cases} \qquad .$$

A function $b_{0,\bar{E}} \in H^2(\bar{E})$ is defined by

$$b_{0,\bar{E}} := \begin{cases} 12^6 \cdot (\lambda_1 - \frac{1}{4})^2 \cdot (\lambda_2 - \frac{1}{4})^2 \cdot (\lambda_3 - \frac{1}{4})^2 & \text{if } \lambda_1, \lambda_2, \lambda_3 \geq 1/4 \\ 0 & \text{otherwise} \end{cases} .$$

Let $\mathbf{a}_E \in \mathbb{R}^3$ be a so-called *guide vector* with $\sphericalangle(\mathbf{a}_E, E) \neq 0$. Define the transformed and normalized vector $\bar{\mathbf{a}} := A_T^{-1}(\mathbf{a}_E) / \left| A_T^{-1}(\mathbf{a}_E) \right|_{\mathbb{R}^3}$. (Note that $\bar{\mathbf{a}}$ depends on $\mathbf{a}_E$ and T.) A function $b_0 \in H^2(\bar{T})$ is then defined by

$$b_0(\mathbf{x}_{\bar{E}} + t \cdot \bar{\mathbf{a}}) := b_{0,\bar{E}}(\mathbf{x}_{\bar{E}}) \qquad \mathbf{x}_{\bar{E}} \in \bar{E}\,,\ t \in \mathbb{R} \ \text{ such that } \mathbf{x}_{\bar{E}} + t \cdot \bar{\mathbf{a}} \in \bar{T} \qquad .$$

A bubble function on the unitary tetrahedron $\bar{T}$ is defined by

$$\bar{b}_{E,T} := b_0 \cdot b_1 \cdot b_2 \cdot b_3 \qquad .$$

By means of the coordinate transformation with A_T we obtain a bubble function $b_{E,T} = F_A(\bar{b}_{E,T}$ on the actual tetrahedron T.

Finally consider the two tetrahedra $T_1 \cup T_2 = \omega_E$, and define the *face bubble function* $b_E \in H_o^2(\omega_E)$ by

$$b_E(\mathbf{x}) := \begin{cases} b_{E,T_1}(\mathbf{x}) & \text{if } \mathbf{x} \in T_1 \\ b_{E,T_2}(\mathbf{x}) & \text{if } \mathbf{x} \in T_2 \end{cases} . \tag{3.29}$$

In order to visualize the construction of the face bubble function in the two-dimensional case, figure 3.2 shows a cut-off function b_1 as well as b_0 and b_E. They are depicted on the standard tetrahedron with the choice $\mathbf{a}_E = (2,1)^T$.

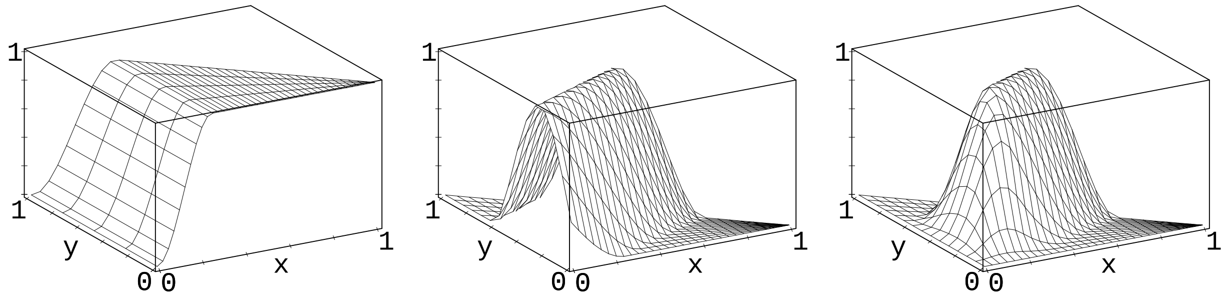


Figure 3.2: Functions b_1, b_0 and b_E (on a single triangle)

Note that (although not reflected by the notation) b_E depends on the guide vector $\mathbf{a}_E$. Strictly speaking, one had to write $b_{E,\mathbf{a}_E}$ instead of b_E. The actual choice of $\mathbf{a}_E$ is discussed in the next section.

To end off, it is easily seen that

$$0 \leq b_T, b_E \leq 1 \qquad \text{and} \qquad \max_{\mathbf{x} \in T} b_T(\mathbf{x}) = \max_{\mathbf{x} \in \omega_E} b_E(\mathbf{x}) = 1$$

holds.

Remark 3.9 To use the same guide vector $\mathbf{a}_E$ on T_1 and T_2 guarantees continuity of the derivative of b_0 and b_E across E which implies $b_E \in C^1(\omega_E) \cap H_o^2(\omega_E)$. The functions $b_1 \ldots b_3$ act as cut-off functions to meet the zero boundary conditions of H_o^2 on $\partial T \setminus E$. $\square$

3.4.2 Additional mesh requirement for the lower error bound and inverse inequalities

The lower bound of the error relies heavily on the use of bubble functions. The smoothness assumption $b_E \in H^2_o(\omega_E)$ lead to the technical definition of the bubble function b_E in the section above. Eventually we want to employ inverse inequalities similar to (2.20) – (2.24). For that reason a further requirement on the mesh is necessary.

Consider an inner face E and $\omega_E = T_1 \cup T_2$. Let the bubble function b_E be defined as above. Both tetrahedra T_1 and T_2 are now transformed, each one via the corresponding matrix A_{T_i}. Thus the guide vector $\mathbf{a}_E$ (which is the same for T_1 and T_2 in the original space) is transformed into two different guide vectors on $\bar{T}_1$ and $\bar{T}_2$ of the transformed spaces. For unambiguous reference, denote the transformed and normalized guide vectors $A_T^{-1}(\mathbf{a}_E)/\left|A_T^{-1}(\mathbf{a}_E)\right|_{\mathbb{R}^3}$ on T_1 and T_2 by $\bar{\mathbf{a}}_{T_1}$ and $\bar{\mathbf{a}}_{T_2}$. Consider the angles $\sphericalangle(\bar{\mathbf{a}}_{T_i}, \bar{E})$ between $\bar{\mathbf{a}}_{T_i}$ and $\bar{E}$, $i = 1, 2$. The additional mesh requirement then reads as follows.

Additional mesh requirement (analytic form)
Consider an inner face E and $\omega_E = T_1 \cup T_2$. There has to be a guide vector $\mathbf{a}_E$ such that $\sphericalangle(\bar{\mathbf{a}}_{T_1}, \bar{E})$ and $\sphericalangle(\bar{\mathbf{a}}_{T_2}, \bar{E})$ are both bounded from below by some angle $\alpha_0 > 0$. This has to be satisfied for all inner faces E of $\mathcal{T}_h$, with the same smallest angle α_0. (3.30)

In the proof of lemma 3.9 it will be shown that this additional mesh requirement ensures that there exists always a bubble function b_E satisfying inverse inequality (3.37). Firstly, however, this mesh requirement will be closer investigated into.

It would be desirable to have not only an analytic condition on $\mathcal{T}_h$ which just demands the existence of some guide vectors. Thus we seek an equivalent condition which can be verified. The following geometric condition turns out to be sufficient for (3.30) but it may not be necessary. Nevertheless it seems to be fairly close to a necessary condition (at least from heuristic arguments). We have to start with some additional (temporary) notation.

Consider an inner face E, and denote its midpoint by M. Denote by A and B the vertices of T_1 and T_2 which are not in E, respectively. The upper part of figure 3.3 depicts this notation.

Consider the plane ε that contains A, B and M (if A, B, and M are on a straight line then any plane ε can be chosen, and (3.30) and (3.31) are naturally satisfied). Its intersection with ω_E is depicted in the lower part of figure 3.3. Let Q_1 and Q_2 be the intersection points of ε and the boundary of the face (triangle) E, respectively. Enumerate Q_1 and Q_2 such that Q_2 is closer to AB. Define $C := AB \cap Q_1Q_2$. Let A_1 and B_1 be those points that satisfy $A_1B_1 \parallel AB, M \in A_1B_1$, and $A_1 \in AQ_1, B_1 \in BQ_1$.

The geometrical mesh requirement then reads as follows.

Additional mesh requirement (geometric form)
With the notation from above we demand

$$\frac{|Q_1M|}{|Q_1C|} \geq c_0 > 0 \qquad , \tag{3.31}$$

with $|Q_1M|$ being the length of the line segment Q_1M. This has to be satisfied for all inner faces E of $\mathcal{T}_h$, with the same c_0.

Lemma 3.8 *The geometric assumption (3.31) implies the analytic assumption (3.30).*

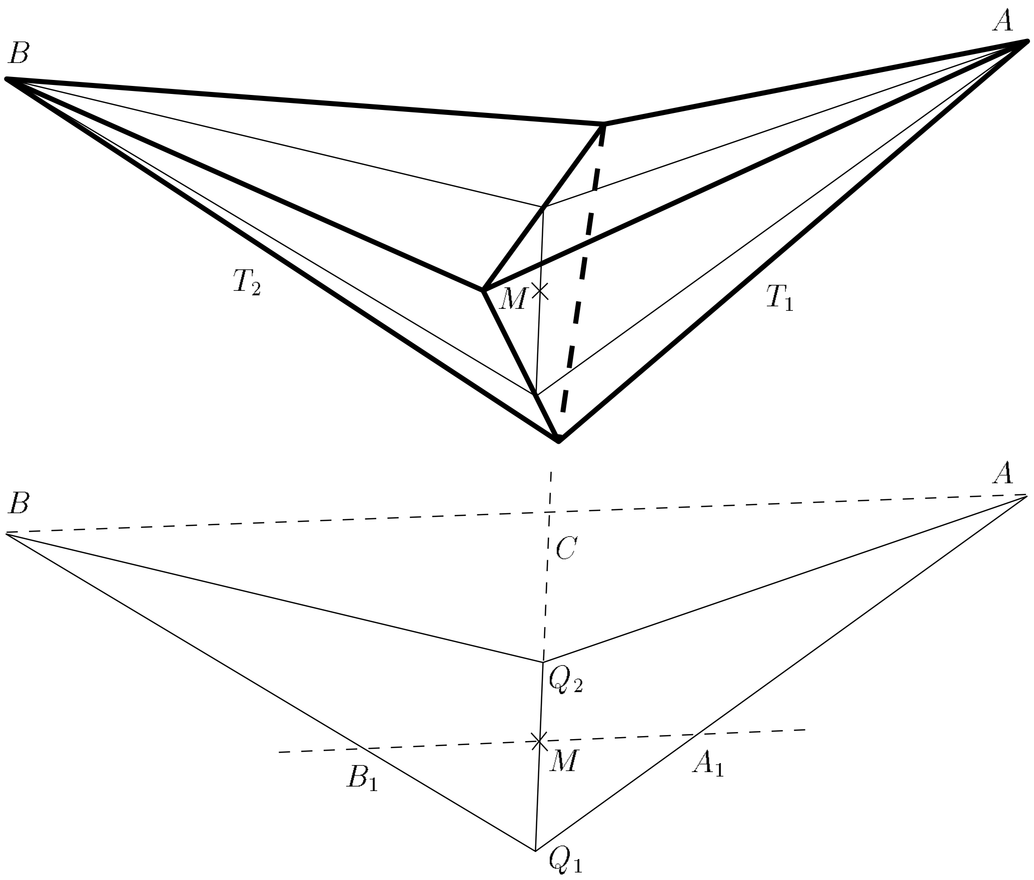


Figure 3.3: Notation of $\omega_E = T_1 \cup T_2$, and intersection of ω_E and ε

Proof: Consider any of the two tetrahedra, say T_1. The transformation via $A_{T_1}^{-1}$ (or more precisely, via F_A^{-1}) maps T_1 onto the standard tetrahedron $\bar{T}_1$. Assume, without loss of generality, that A is mapped onto the vertex $\bar{A} = (1,0,0)^T$, and E is mapped onto the opposite face $\bar{E}$. The transformed point $\bar{M}$ is still the midpoint of $\bar{E}$, the points $\bar{Q}_i$ lie on the boundary of $\bar{E}$, and $\bar{A}_1 \in \bar{A}\bar{Q}_1$ is situated somewhere on the boundary $\partial T_1 \setminus \bar{E}$. Figure 3.4 may visualize this.

Choose now the guide vector $\mathbf{a}_E$ in the direction $\overrightarrow{\bar{M}\bar{A}_1}$. The radius of the inscribed circle of $\bar{E}$ is $\varrho(\bar{E}) = 1/(2+\sqrt{2})$. The distance between $\bar{A}_1$ and $\bar{E}$ equals

$$\operatorname{dist}(\bar{A}_1, \bar{E}) \;=\; \frac{\operatorname{dist}(\bar{A}_1, \bar{E})}{\operatorname{dist}(\bar{A}, \bar{E})} \;=\; \frac{|\bar{Q}_1\bar{A}_1|}{|\bar{Q}_1\bar{A}|} \;=\; \frac{|Q_1A_1|}{|Q_1A|} \;=\; \frac{|Q_1M|}{|Q_1C|} \;\geq\; c_0 > 0$$

since the linear transformation $A_{T_1}^{-1}$ preserves above's ratios of line segments, and because of the similarity theorem (cf. figures 3.3 and 3.4).

Basic geometry then implies

$$\begin{aligned} \tan \sphericalangle(\bar{\mathbf{a}}_{T_1}, \bar{E}) \;&\geq\; \frac{\operatorname{dist}(\bar{A}_1, \bar{E})}{\varrho(\bar{E})} \;\geq\; (2+\sqrt{2}) \cdot c_0 \\ \sphericalangle(\bar{\mathbf{a}}_{T_1}, \bar{E}) \;&\geq\; \alpha_0 \;:=\; \arctan((2+\sqrt{2}) \cdot c_0) \end{aligned}$$

i.e. the angle $\sphericalangle(\bar{\mathbf{a}}_{T_1}, \bar{E})$ is bounded from below by some angle $\alpha_0 > 0$. The other three cases where A is mapped onto an other vertex of $\bar{T}$ are treated completely analogously.

A similar consideration verifies $\sphericalangle(\bar{\mathbf{a}}_{T_2}, \bar{E}) \geq \alpha_0$. Thus $\mathbf{a}_E$ satisfies the analytic mesh requirement (3.30). ■

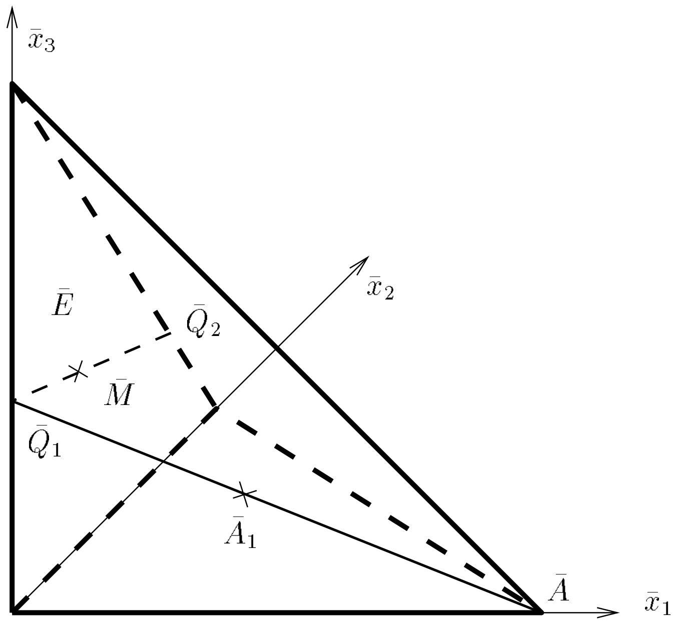


Figure 3.4: Mapped standard tetrahedron $\bar{T}_1$

Remark 3.10 In the two-dimensional case the condition and the proof are similar. Then the lower part of figure 3.3 may be utilized if Q_1 and Q_2 are replaced by both endpoints of E. □

Let now $\mathcal{T}_h$ be a triangulation where either the analytic mesh requirement (3.30) or the geometric mesh requirement (3.31) is satisfied. Then one can fix appropriate guide vectors $\mathbf{a}_E$ for all inner faces E. Each of these guide vectors determines a concrete bubble function from the whole class of functions defined by (3.29).

From now on, we consider these very bubble functions. All inequalities and constants hereafter are such that they do not depend on the actual choice of the bubble functions (or guide vectors, respectively) but only on the (global) minimal angle α_0 of (3.30).

Several inverse inequalities are comprised in the lemma below.

Lemma 3.9 (Inverse inequalities) *Let either the analytic mesh requirement (3.30) or the geometric mesh requirement (3.31) be satisfied. Let F_{ext} be the extension operator of (2.19). The following inverse inequalities hold for all $\varphi_T \in \mathbb{P}^0(T)$ and $\varphi_E \in \mathbb{P}^0(E)$.*

$$\|\varphi_T\|_T \sim \|b_T^{1/2} \cdot \varphi_T\|_T \tag{3.32}$$
$$\|b_T \cdot \varphi_T\|_T \leq \|\varphi_T\|_T \tag{3.33}$$
$$\|\Delta(b_T \cdot \varphi_T)\|_T \lesssim h_{min,T}^{-2} \cdot \|\varphi_T\|_T \tag{3.34}$$
$$\|\varphi_E\|_E \sim \|b_E^{1/2} \cdot \varphi_E\|_E \tag{3.35}$$
$$\|F_{ext}(\varphi_E) \cdot b_E\|_T \lesssim h_E^{1/2} \cdot \|\varphi_E\|_E \qquad \textit{for } E \in T \tag{3.36}$$
$$\|\Delta(F_{ext}(\varphi_E) \cdot b_E)\|_T \lesssim h_E^{1/2} \cdot h_{min,T}^{-2} \cdot \|\varphi_E\|_E \qquad \textit{for } E \in T \tag{3.37}$$

Proof: The inequalities (3.32) and (3.35) are derived analogously to inequalities (2.20) and (2.22) of lemma 2.6. Inequality (3.33) results immediately from $0 \leq b_T \leq 1$.

In order to prove (3.34) we utilize the transformation technique which yields for general $w \in H^2(T)$

$$\begin{aligned}
\|\Delta w\|_T^2 \;\leq\; 3 \cdot \|D^2 w\|^2 &= 3 \int_T |A_T^{-T} \cdot A_T^T \cdot D^2 w \cdot A_T \cdot A_T^{-1}|^2 \\
&\lesssim \|A_T^{-1}\|_{\mathbb{R}^{3\times 3}}^4 \cdot \int_T |A_T^T \cdot D^2 w \cdot A_T|^2 \\
&= \|A_T^{-1}\|_{\mathbb{R}^{3\times 3}}^4 \cdot |\det A_T| \cdot \int_{\bar{T}} |\bar{D}^2 \bar{w}|^2 .
\end{aligned}$$

For $\mathbf{x} \in T$ set now

$$w(\mathbf{x}) := \varphi_T(\mathbf{x}) \cdot b_T(\mathbf{x}) \qquad \in \mathbb{P}^8(T) \cap H_o^2(T) \qquad .$$

The bound $\|A_T^{-1}\|_{\mathbb{R}^{3\times 3}} \lesssim h_{min,T}^{-1}$ of (2.7) and the equivalence of norms over the finite dimensional space $\mathbb{P}^0(\bar{T}) \ni \bar{\varphi}_T$ imply

$$\begin{aligned}
\|\bar{D}^2 \bar{w}\|_{\bar{T}} &= \|\bar{D}^2(\overline{\varphi_T \cdot b_T})\|_{\bar{T}} \lesssim \|\bar{\varphi}_T\|_{\bar{T}} \\
\text{and} \qquad \|\Delta w\|_T &\lesssim h_{min,T}^{-2} \cdot |\det A_T|^{1/2} \cdot \|\bar{\varphi}_T\|_{\bar{T}} \;=\; h_{min,T}^{-2} \cdot \|\varphi_T\|_T \qquad .
\end{aligned}$$

Thus (3.34) is obtained.

Inequality (3.36) utilizes the facts that $0 \leq b_E \leq 1$ and that $\varphi_E \in \mathbb{P}^0(E)$ is a constant function. This yields

$$\|F_{ext}(\varphi_E) \cdot b_E\|_T \;\leq\; |T|^{1/2} \cdot |\varphi_E(\mathbf{x})| \;\lesssim\; h_E^{1/2} \cdot \|\varphi_E\|_E$$

and the desired estimate is obtained.

Inequality (3.37) requires a closer investigation. Let E be a face of T and $\mathbf{a}_E$ be a guide vector. The transformation via A_T^{-1} maps T onto the unitary tetrahedron $\bar{T}$. Again, denote by $\bar{E}$ and $\bar{\mathbf{a}}$ the transformed face and (normalized) guide vector, respectively.

Let $\mathbf{l}_1$ and $\mathbf{l}_2$ be two orthogonal unitary vectors in that plane that contains $\bar{E}$, i.e. $|\mathbf{l}_i| = 1$, $\mathbf{l}_1 \perp \mathbf{l}_2$, $\bar{E} \subset \operatorname{span}(\mathbf{l}_1, \mathbf{l}_2)$. The transformed guide vector has been defined such that $|\bar{\mathbf{a}}|_{\mathbb{R}^3} = 1$. The three vectors $\mathbf{l}_1, \mathbf{l}_2$ and $\bar{\mathbf{a}}$ form the basis of a coordinate system which is denoted by $(\mathbf{l}_1, \mathbf{l}_2, \bar{\mathbf{a}})$.

We now investigate the function b_0. Its definition from above yields

$$\left| \frac{\partial^2 b_0}{\partial \mathbf{l}_i \, \partial \mathbf{l}_j}(\mathbf{x}) \right| \lesssim 1 \qquad i,j = 1,2 \qquad \text{and} \qquad \frac{\partial b_0}{\partial \bar{\mathbf{a}}} = 0$$

where $\partial^2 / \partial \mathbf{l}_i \, \partial \mathbf{l}_j$ denotes the second partial directional derivatives with respect to $\mathbf{l}_i$ and $\mathbf{l}_j$. Thus the matrix $D^2_{(\mathbf{l}_1, \mathbf{l}_2, \bar{\mathbf{a}})} b_0$ of the second derivatives (with respect to the system $(\mathbf{l}_1, \mathbf{l}_2, \bar{\mathbf{a}})$) is bounded for all $\mathbf{x} \in \bar{T}$, i.e. all matrix entries are bounded.

We introduce a second coordinate system with the basis $(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3)$ and $\mathbf{l}_3 := \mathbf{l}_1 \times \mathbf{l}_2$. Note that $|\mathbf{l}_3| = 1$ and $\mathbf{l}_3 \perp \bar{E}$. The bases of both coordinate systems are linked via

$$\begin{aligned}
(\mathbf{l}_1, \mathbf{l}_2, \bar{\mathbf{a}}) &= (\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3) \cdot B \\
\text{with} \qquad B &:= \begin{pmatrix} 1 & 0 & (\bar{\mathbf{a}}, \mathbf{l}_1) \\ 0 & 1 & (\bar{\mathbf{a}}, \mathbf{l}_2) \\ 0 & 0 & (\bar{\mathbf{a}}, \mathbf{l}_3) \end{pmatrix}
\end{aligned}$$

and $(\bar{\mathbf{a}}, \mathbf{l}_i) = \cos \sphericalangle(\bar{\mathbf{a}}, \mathbf{l}_i)$ being the usual $\mathbb{R}^3$ scalar product of two vectors.

The analytic mesh requirement (3.30) (or the geometric mesh requirement (3.31) together with lemma 3.8) guarantees that the angle $\sphericalangle(\bar{\mathbf{a}}, \bar{E})$ between the transformed guide vector $\bar{\mathbf{a}}$ and $\bar{E}$ is bounded from below by some angle $\alpha_0 > 0$. Thus we conclude

$$\begin{array}{rcl} |\sphericalangle(\bar{\mathbf{a}}, \mathbf{l}_3)| & \leq & \pi/2 - \alpha_0 \\ |(\bar{\mathbf{a}}, \mathbf{l}_3)| \;=\; |\cos \sphericalangle(\bar{\mathbf{a}}, \mathbf{l}_3)| & \geq & \sin\alpha_0 > 0 \\ \|B^{-1}\|_{\mathbb{R}^{3\times 3}} \;=\; \left\| \begin{pmatrix} 1 & 0 & -(\bar{\mathbf{a}}, \mathbf{l}_1)/(\bar{\mathbf{a}}, \mathbf{l}_3) \\ 0 & 1 & -(\bar{\mathbf{a}}, \mathbf{l}_2)/(\bar{\mathbf{a}}, \mathbf{l}_3) \\ 0 & 0 & 1/(\bar{\mathbf{a}}, \mathbf{l}_3) \end{pmatrix} \right\|_{\mathbb{R}^{3\times 3}} & \lesssim & \dfrac{1}{|(\bar{\mathbf{a}}, \mathbf{l}_3)|} \lesssim 1 \end{array} \quad .$$

The matrices of the second derivatives with respect to both coordinate systems are related according to

$$D^2_{(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_3)}\, b_0 \;=\; B^{-T} \cdot D^2_{(\mathbf{l}_1,\mathbf{l}_2,\bar{\mathbf{a}})}\, b_0 \cdot B^{-1} \quad .$$

Because of

$$|D^2_{(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_3)}\, b_0| \;\lesssim\; \|B^{-1}\|^2_{\mathbb{R}^{3\times 3}} \cdot |D^2_{(\mathbf{l}_1,\mathbf{l}_2,\bar{\mathbf{a}})}\, b_0| \;\lesssim\; 1$$

the matrix $D^2_{(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_3)}\, b_0$ of the second derivatives (with respect to the system $(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3)$) is also bounded for all $\mathbf{x} \in \bar{T}$. A rotation of the coordinate system does not change the sum of the squared second derivatives. Thus we can switch from $(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3)$ to the standard coordinate system of $\bar{T}$ and obtain for all $\mathbf{x} \in \bar{T}$

$$|D^2_{(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_3)}\, b_0| \;=\; |\bar{D}^2\, b_0| \;\lesssim\; 1 \quad .$$

Analogously $|\bar{D}^1\, b_0| \lesssim 1$ is derived, and $b_0 \leq 1$ is obvious. Additionally the cut-off functions b_1, b_2, b_3 and their first and second derivative are bounded, i.e. for all $\mathbf{x} \in \bar{T}$

$$|b_i| \leq 1 \quad , \quad |\bar{D}^1\, b_i| \lesssim 1 \quad , \quad |\bar{D}^2\, b_i| \lesssim 1 \qquad i = 1, 2, 3 \quad .$$

All this implies

$$|\bar{D}^2\,(b_0 \cdot b_1 \cdot b_2 \cdot b_3)| \;=\; |\bar{D}^2\, \bar{b}_E| \lesssim 1$$

for all $\mathbf{x} \in \bar{T}$.

Let now $w := F_{ext}(\varphi_E) \cdot b_E$. Since $\varphi_E \in \mathbb{P}^0(E)$ is a constant function we conclude

$$\|\bar{D}^2\, \bar{w}\|_{\bar{T}} \;=\; \|\bar{D}^2\, (\overline{F_{ext}(\varphi_E)} \cdot \bar{b}_E)\|_{\bar{T}} \;=\; |\varphi_E(\mathbf{x})| \cdot \|\bar{D}^2\, \bar{b}_E\|_{\bar{T}} \;\lesssim\; \|\bar{\varphi}_E\|_{\bar{E}} \quad .$$

Now we utilize an inequality of the proof of (3.34) from above giving

$$\begin{array}{rcl} \|\Delta w\| & \lesssim & \|A_T^{-1}\|^2_{\mathbb{R}^{3\times 3}} \cdot |\det A_T|^{1/2} \cdot \|\bar{D}^2 \bar{w}\|_{\bar{T}} \\ & \lesssim & h^{-2}_{min,T} \cdot |\det A_T|^{1/2} \cdot \|\bar{\varphi}_E\|_{\bar{E}} \\ & = & h^{-2}_{min,T} \cdot h_E^{1/2} \cdot \|\varphi_E\|_E \end{array} \quad .$$

Thus the desired estimate is proven. ■

3.4.3 The set $H^2_{\mathcal{T}}(\Omega)$ of L_2 adapted functions

For the error estimator we require that the anisotropy of the solution of the dual problem is reflected in the mesh in some way. This leads to the following

Definition 3.6 (L_2 adapted function) *Let $c_b > 1$ be a fixed constant. Then a function $v \in H^2(\Omega)$ is said to be L_2* adapted *to the mesh if*

$$\sum_{T\in\mathcal{T}_h} h_{min,T}^{-4} \cdot \|C_T^T \cdot D^2 v \cdot C_T\|_T^2 \;\le\; c_b \cdot \|D^2 v\|_\Omega^2 \tag{3.38}$$

holds. Denote the set of L_2 adapted functions by $H^2_{\mathcal{T}}(\Omega)$.

3.4.4 Anisotropic interpolation estimates

The imbedding theorem for Sobolev spaces implies $H^2(\Omega) \hookrightarrow C^0(\Omega)$. Hence for a function $v \in H^2(\Omega)$ the Lagrange interpolate $\mathrm{Int}(v) \in C^0(\bar\Omega)$ is well-defined.

First we state well-known interpolation estimates on the unitary tetrahedron $\bar T$.

Lemma 3.10 *Let $\bar v \in H^2(\bar T)$. The following estimates hold:*

$$\begin{aligned}\|\bar v - \mathrm{Int}\,\bar v\|_{\bar T} &\lesssim \|\bar D^2 \bar v\|_{\bar T}\\ \|\bar\nabla(\bar v - \mathrm{Int}\,\bar v)\|_{\bar T} &\lesssim \|\bar D^2 \bar v\|_{\bar T}\end{aligned} \quad .$$

Note that all derivatives are with respect to the reference coordinate system.

Using scaling arguments we now obtain interpolation estimates for an L_2 adapted function on the actual tetrahedron T .

Lemma 3.11 *Let v be an L_2 adapted function, i.e. $v \in H^2_{\mathcal{T}}(\Omega)$. The interpolation estimates*

$$\begin{aligned}\sum_{T\in\mathcal{T}_h} h_{min,T}^{-4} \cdot \|v - \mathrm{Int}\,v\|_T^2 &\lesssim \|D^2 v\|_\Omega^2\\ \sum_{T\in\mathcal{T}_h} h_{min,T}^{-4} \cdot \|C_T^T \nabla\,(v - \mathrm{Int}\,v)\|_T^2 &\lesssim \|D^2 v\|_\Omega^2\end{aligned}$$

hold with a constant c independent of v or $\mathcal{T}_h$.

Proof: The transformation technique yields

$$\|v - \mathrm{Int}\,v\|_T^2 = |\det A_T| \cdot \|\bar v - \mathrm{Int}\,\bar v\|_{\bar T}^2 \lesssim |\det A_T| \cdot \|\bar D^2 \bar v\|_{\bar T}^2 \quad .$$

The second derivative is transformed via

$$\bar D^2 \bar v \;=\; A_T^T \cdot D^2 v \cdot A_T$$

resulting in

$$\begin{aligned}|\det A_T| \cdot \|\bar D^2 \bar v\|_{\bar T}^2 &= |\det A_T| \cdot \int_{\bar T} |\bar D^2 \bar v|^2\\ &= \int_T |A_T^T \cdot D^2 v \cdot A_T|^2 = \int_T |A_T^T C_T^{-T} \cdot C_T^T \cdot D^2 v \cdot C_T \cdot C_T^{-1} A_T|^2\\ &\lesssim \|C_T^{-1} A_T\|_{\mathbb{R}^{3\times3}}^4 \cdot \int_T |C_T^T \cdot D^2 v \cdot C_T|^2 \overset{(2.3)}{\lesssim} \|C_T^T \cdot D^2 v \cdot C_T\|_T^2 \quad .\end{aligned}$$

Recalling $v \in H^2_{\mathcal{T}}(\Omega)$ completes the first part of the proof:

$$\sum_{T\in\mathcal{T}_h} h^{-4}_{min,T} \cdot \|v - \mathrm{Int}\, v\|^2_T \lesssim \sum_{T\in\mathcal{T}_h} h^{-4}_{min,T} \cdot \|C^T_T \cdot D^2 v \cdot C_T\|^2_T \lesssim \|D^2 v\|^2_\Omega \quad .$$

The second part of the proof utilizes (2.4) giving

$$\begin{aligned} \|C^T_T \cdot \nabla(v - \mathrm{Int}\, v)\|^2_T &= \|C^T_T A^{-T}_T \cdot A^T_T \cdot \nabla(v - \mathrm{Int}\, v)\|^2_T \\ &\le \|C^T_T A^{-T}_T\|^2_{\mathbb{R}^{3\times 3}} \cdot \|A^T_T \cdot \nabla(v - \mathrm{Int}\, v)\|^2_T \\ &\overset{(2.4)}{\lesssim} |\det A_T| \cdot \|\bar\nabla(\bar v - \mathrm{Int}\, \bar v)\|^2_{\bar T} \lesssim |\det A_T| \cdot \|\bar D^2 \bar v\|^2_{\bar T} \\ &\lesssim \|C^T_T \cdot D^2 v \cdot C_T\|^2_T \end{aligned}$$

as above. Recalling $v \in H^2_{\mathcal{T}}(\Omega)$ we conclude

$$\sum_{T\in\mathcal{T}_h} h^{-4}_{min,T} \cdot \|C^T_T \nabla\,(v - \mathrm{Int}\, v)\|^2_T \lesssim \sum_{T\in\mathcal{T}_h} h^{-4}_{min,T} \cdot \|C^T_T \cdot D^2 v \cdot C_T\|^2_T \lesssim \|D^2 v\|^2_\Omega$$

analogously to the first part of the proof. ■

Lemma 3.12 *Let v be an L_2 adapted function, i.e. $v \in H^2_{\mathcal{T}}(\Omega) \cap H^1_o(\Omega)$. The estimates*

$$\begin{aligned} |(f, v - \mathrm{Int}\, v)| &\lesssim \left(\sum_{T\in\mathcal{T}_h} h^4_{min,T} \cdot \|f\|^2_T\right)^{1/2} \cdot \|D^2 v\|_\Omega \qquad \forall\, f \in L_2(\Omega) \\ |(\nabla w_h, \nabla(v - \mathrm{Int}\, v))| &\lesssim \left(\sum_{E\in\Omega} \frac{h^4_{min,T}}{h_E} \cdot \|r_E(w_h)\|^2_E\right)^{1/2} \cdot \|D^2 v\|_\Omega \qquad \forall\, w_h \in V_{o,h} \end{aligned}$$

hold.

Proof: The first result is readily obtained by Cauchy's inequality and lemma 3.11.

$$\begin{aligned} |(f, v - \mathrm{Int}\, v)| &= \left|\sum_{T\in\mathcal{T}_h} \int_T f \cdot (v - \mathrm{Int}\, v)\right| \\ &\le \sum_{T\in\mathcal{T}_h} h^2_{min,T} \|f\|_T \cdot h^{-2}_{min,T} \|v - \mathrm{Int}\, v\|_T \\ &\le \left(\sum_{T\in\mathcal{T}_h} h^4_{min,T}\, \|f\|^2_T\right)^{1/2} \cdot \left(\sum_{T\in\mathcal{T}_h} h^{-4}_{min,T}\, \|v - \mathrm{Int}\, v\|^2_T\right)^{1/2} \\ &\lesssim \left(\sum_{T\in\mathcal{T}_h} h^4_{min,T}\, \|f\|^2_T\right)^{1/2} \cdot \|D^2 v\|_\Omega \quad . \end{aligned}$$

To prove the second estimate we integrate by parts and apply Cauchy's inequality to conclude for any $g \in H^1_o(\Omega)$

$$\begin{aligned} (\nabla w_h, \nabla g) &= \sum_{T\in\mathcal{T}_h} \int_T \nabla^T w_h \cdot \nabla g = \sum_{T\in\mathcal{T}_h} \int_{\partial T} \frac{\partial w_h}{\partial n} \cdot g = \sum_{E\in\Omega} \int_E r_E(w_h) \cdot g \\ &\le \sum_{E\in\Omega} \frac{h^2_{min,T}}{h^{1/2}_E} \|r_E(w_h)\|_E \cdot \frac{h^{1/2}_E}{h^2_{min,T}} \|g\|_E \le D_{h,2}(w_h) \cdot \left(\sum_{E\in\Omega} \frac{h_E}{h^4_{min,T}} \|g\|^2_E\right)^{\frac{1}{2}} . \end{aligned}$$

Utilizing the trace inequality (2.12)

$$\|g\|_E^2 \lesssim h_E^{-1} \left(\|g\|_T^2 + \|C_T^T \nabla g\|_T^2\right)$$

and rewriting the sum over all faces E as a sum over all tetrahedra T implies

$$\sum_{E\in\Omega} \frac{h_E}{h_{min,T}^4} \|g\|_E^2 \lesssim \sum_{T\in\mathcal{T}_h} h_{min,T}^{-4} \left(\|g\|_T^2 + \|C_T^T \nabla g\|_T^2\right) \qquad .$$

Substituting $g := v - \operatorname{Int} v \in H_o^1(\Omega)$, recalling $v \in H_{\mathcal{T}}^2$, and applying lemma 3.11 results immediately in

$$\sum_{E\in\Omega} \frac{h_E}{h_{min,T}^4} \|g\|_E^2 \lesssim \|D^2 v\|_\Omega^2 \qquad .$$

Thus the desired estimate is proven. ∎

3.4.5 Anisotropic L_2 error estimator

Definition 3.7 (L_2 error estimator) *For an arbitrary tetrahedron T let the L_2 error estimator $\eta_{R,L_2,T}(u_h)$ be defined by*

$$\eta_{R,L_2,T}(u_h) := \left(h_{min,T}^4 \cdot \|r_T(u_h)\|_T^2 + \sum_{E\in\partial T\setminus\Gamma_D} \frac{h_{min,T}^4}{h_E} \cdot \|r_E(u_h)\|_E^2 \right)^{1/2} \qquad . \tag{3.39}$$

In order to obtain an upper bound of the L_2 error we utilize the Aubin-Nitsche trick [6, 19]. The following theorem is valid.

Theorem 3.13 (L_2 error estimator) *Let $u \in H_o^1(\Omega)$ be the exact solution and $u_h \in V_{o,h}$ be the FEM solution.*

If the additional mesh requirement (3.31) is satisfied then the error (in the L_2 norm) is bounded locally from below by

$$\eta_{R,L_2,T}(u_h) \lesssim \left(\|u - u_h\|_{\omega_T}^2 + h_{min,T}^4 \cdot \|f - P_{L_2} f\|_{\omega_T}^2\right)^{1/2} \tag{3.40}$$

for all $T \in \mathcal{T}_h$.

Assume that Ω is a convex polygonal domain. Let $v_D \in H^2(\Omega)$ be the solution of the dual problem

$$-\Delta v_D = u - u_h \quad \text{in } \Omega, \qquad v_D = 0 \quad \text{on } \partial\Omega.$$

Suppose that v_D is an L_2 adapted function, i.e. $v_D \in H_{\mathcal{T}}^2(\Omega)$. Then the error (in the L_2 norm) is bounded globally from above by

$$\|u - u_h\| \lesssim \left(\sum_{T\in\mathcal{T}_h} \eta_{R,L_2,T}^2(u_h) + \sum_{T\in\mathcal{T}_h} h_{min,T}^4 \|f - P_{L_2} f\|_T^2\right)^{1/2} \tag{3.41}$$

or, alternatively

$$\|u - u_h\| \lesssim \|h_{min}^2(\mathbf{x})\, f\| + D_{h,2}(u_h) \qquad .$$

Proof: Firstly, estimate (3.40) will be proven.

We start with the norm $\|r_T(u_h)\|_T$ of the element residual $r_T = r_T(u_h) := P_{L_2} f + \Delta u_h$. Since we use linear ansatz functions $r_T \in \mathbb{P}^0(T)$ holds. For $\mathbf{x} \in T$ let

$$w(\mathbf{x}) := r_T(u_h)(\mathbf{x}) \cdot b_T(\mathbf{x}) \qquad \in \mathbb{P}^8(T) \cap H_o^2(T) \qquad .$$

Integration by parts and $w \in H_o^2(T)$ then results in

$$\begin{aligned} \int_T r_T \cdot w &= \int_T (f + \Delta u_h) \cdot w + \int_T (P_{L_2} f - f) \cdot w \\ &= \int_T (u_h - u) \cdot \Delta w + \int_T (P_{L_2} f - f) \cdot w \\ \left| \int_T r_T \cdot w \right| &\leq \|u - u_h\|_T \cdot \|\Delta w\|_T + \|f - P_{L_2} f\|_T \cdot \|w\|_T \qquad . \end{aligned}$$

Recalling the inverse estimates (3.32) – (3.34) we conclude

$$\begin{aligned} \|r_T\|_T^2 &\lesssim \|u - u_h\|_T \cdot h_{min,T}^{-2} \cdot \|r_T\|_T + \|f - P_{L_2} f\|_T \cdot \|r_T\|_T \\ \text{and} \qquad h_{min,T}^4 \cdot \|r_T\|_T^2 &\lesssim \|u - u_h\|_T^2 + h_{min,T}^4 \cdot \|f - P_{L_2} f\|_T^2 \qquad . \end{aligned}$$

Now we aim for a bound of the norm $\|r_E(u_h)\|_E$ of the gradient jump across some inner face (triangle) E. Since we use linear ansatz functions $r_E \in \mathbb{P}^0(E)$ holds. Let T_1 and T_2 be the two tetrahedra that E belongs to. The additional mesh requirement (3.31) and lemma 3.9 on page 49 imply that there exists a bubble function b_E which satisfies the inverse inequalities (3.35) – (3.37). Let the function $w \in H_o^2(\omega_E)$ be defined by

$$w := F_{ext}(r_E(u_h)) \cdot b_E \qquad ,$$

with F_{ext} being the extension operator of (2.19). Assume that the right hand side $f = -\Delta u$ is in $L_2(\Omega)$. Integration by parts yields

$$\begin{aligned} -\int_E w \cdot r_E(u_h) &= \int_{\omega_E} \nabla^T w \nabla (u_h - u) + \sum_{i=1}^{2} \int_{T_i} w \cdot (r_{T_i} + f - P_{L_2} f) \\ &= \int_{\omega_E} \Delta w \cdot (u - u_h) + \sum_{i=1}^{2} \int_{T_i} w \cdot (r_{T_i} + f - P_{L_2} f) \qquad . \end{aligned}$$

Because of $w|_E = r_E \cdot b_E|_E$ we conclude

$$\left| \int_E r_E^2 \cdot b_E \right| \leq \sum_{i=1}^{2} \Big(\|u - u_h\|_{T_i} \cdot \|\Delta w\|_{T_i} + (\|r_{T_i}\|_{T_i} + \|f - P_{L_2} f\|_{T_i}) \cdot \|w\|_{T_i} \Big) \qquad .$$

Utilizing the inverse inequalities (3.35) – (3.37) results in

$$\begin{aligned} \|r_E\|_E^2 \lesssim \sum_{i=1}^{2} \Big(&\|u - u_h\|_{T_i} \cdot h_E^{1/2}\, h_{min,T_i}^{-2}\, \|r_E\|_E + \\ &+ (\|r_{T_i}\|_{T_i} + \|f - P_{L_2} f\|_{T_i}) \cdot h_E^{1/2}\, \|r_E\|_E \Big) \qquad . \end{aligned}$$

The dimensions $h_E = h_{E,T_i}$ and h_{min,T_i} cannot change rapidly for adjacent tetrahedra. Recalling the bound of $\|r_T\|_T$ from above we conclude

$$\|r_E\|_E \lesssim h_E^{1/2}\, h_{min,T_1}^{-2} \cdot \Big(\|u - u_h\|_{\omega_E} \;+\; h_{min,T_1}^2 \|f - P_{L_2} f\|_{\omega_E} \Big) \qquad .$$

For a fixed tetrahedron $T = T_1$ we sum up over all (inner) faces $E \subset \partial T \setminus \Gamma_D$ and obtain

$$\sum_{E \subset \partial T \setminus \Gamma_D} \frac{h_{min,T}^4}{h_E} \cdot \|r_E(u_h)\|_E^2 \;\lesssim\; \|u - u_h\|_{\omega_T}^2 \;+\; h_{min,T}^4 \, \|f - P_{L_2} f\|_{\omega_T}^2 \qquad .$$

This accomplishes the proof of (3.40).

Secondly, in order to derive (3.41) we integrate by parts, utilize the dual solution $v_D \in H_{\mathcal{T}}^2(\Omega)$, and apply lemma 3.12 yielding

$$\begin{aligned}
\|u - u_h\|^2 \;&=\; (u - u_h, -\Delta v_D) \;=\; (\nabla(u - u_h), \nabla v_D) \;=\; (\nabla(u - u_h), \nabla(v_D - \operatorname{Int} v_D)) \\
&=\; (f, v_D - \operatorname{Int} v_D) \;-\; (\nabla u_h, \nabla(v_D - \operatorname{Int} v_D)) \\
&\lesssim\; \left(\left(\sum_{T \in \mathcal{T}_h} h_{min,T}^4 \cdot \|f\|_T^2 \right)^{1/2} \;+\; D_{h,2} \right) \cdot \|D^2 v_D\| \\
&\lesssim\; \left(\left(\sum_{T \in \mathcal{T}_h} h_{min,T}^4 \cdot \|f\|_T^2 \right)^{1/2} \;+\; D_{h,2} \right) \cdot \|u - u_h\|
\end{aligned}$$

since $\|D^2 v_D\| \;\leq\; c_\Omega \cdot \|\Delta v_D\| \;=\; c_\Omega \cdot \|u - u_h\|$ holds. This corresponds to the second formulation of the upper bound of the error. Utilizing the triangle inequality $\|f\|_T \leq \|P_{L_2} f\|_T + \|f - P_{L_2} f\|_T$ results in the first upper bound of the error. ∎

Remark 3.11 The problem in applying this error estimation lies clearly in verifying the assumption $v_D \in H_{\mathcal{T}}^2(\Omega)$ for the solution of the dual problem.

Additionally one may argue that the dual solution procedure is inappropriate for an anisotropic solution where probably even singularities occur. □

3.5 An anisotropic Zienkiewicz-Zhu like error estimator

An (isotropic) error estimator that is based on an averaged gradient has been proposed first by Zienkiewicz and Zhu [36]. The analysis shows this estimator to be equivalent to a *modified* residual error estimator. (Strictly speaking, one has to show first that the modified residual 'error estimator' is an estimator indeed.) Later the estimator has been improved by the 'superconvergent patch recovery' [37]. We tried to derive an anisotropic version of the earlier estimator but failed for the general case. Because of that we present only the results of the anisotropic version of the original estimator and do not consider the latter one.

3.5.1 Cuboidal or rectangular mesh

Although cuboidal meshes do not fit into the framework of this paper we will present the results for two reasons. Firstly, no anisotropic Zienkiewicz-Zhu like error estimator has been derived so far, and secondly, the structure of this estimator might give some clue for tetrahedral meshes.

An analysis of this estimator for an isotropic mesh which consists of rectangles (2D) and bilinear basis functions is done by Rank and Zienkiewicz [23]. The extension to rectangular prisms (3D) and trilinear basis functions is obvious.

The modification of the estimator for anisotropic rectangular or cuboidal meshes is almost straightforward. In order to get some impression of the kind of these modifications our result is stated here.

Assume a rectangular prism T whose edges are aligned with the coordinate axes. Denote the edge lengths by $h_{1,T}, h_{2,T}, h_{3,T}$, and define $H_T := \operatorname{diag}(h_{1,T}, h_{2,T}, h_{3,T})$. Let $h_{min} := \min\{h_{1,T}, h_{2,T}, h_{3,T}\}$.

Define the *recovered gradient* $\nabla^R u_h$ as the trilinear Lagrange interpolate at the nodes of the mesh. The nodal value at $\mathbf{x}$ is given by

$$\nabla^R u_h(\mathbf{x}) := \frac{1}{8} \sum_{T \ni \mathbf{x}} \nabla u_h \Big|_T \qquad .$$

(For a boundary node set the recovered derivative which is normal to the boundary equal to the true derivative). Now define the *anisotropic Zienkiewicz-Zhu* error estimator by

$$\eta_{Z,T}(u_h) := h_{min,T} \cdot \| H_T^{-1} \left(\nabla^R u_h - \nabla u_h \right) \|_T$$

and the *modified residual error estimator* by

$$\eta_{\tilde{R},T}(u_h) := h_{min,T} \cdot \left(\sum_{E \in \partial T \setminus \Gamma_D} h_E^{-1} \cdot \| r_E(u_h) \|_E^2 \right)^{1/2} \qquad ,$$

i.e. only the jump residuals are utilized here, and the element residual is omitted. Simple algebra shows the anisotropic Zienkiewicz-Zhu like estimator to be equivalent to the modified residual error estimator, i.e.

$$\eta_{Z,T} \sim \eta_{\tilde{R},T}$$

which is also true when suitable different weights (non-negative and bounded away from 0) of the recovered gradient are used.

3.5.2 Tetrahedral mesh of tensor product type

For isotropic triangular grids (in two dimensions) a proof completely different from above is given by Rodriguez [25]. We have extended his ideas to the isotropic three-dimensional case and arbitrary non-negative weights of the recovered gradient, and corrected a minor mistake.

Unfortunately we failed to obtain an error estimator for general tetrahedral, anisotropic meshes. Only *tensor product type meshes* can be considered, i.e. where six tetrahedra can be found that form a rectangular prism. Hence we do not achieve the geometrical flexibility that we were aiming for by using tetrahedral elements. Because of this unsatisfactory result we omit the derivation and only present the estimator.

For the remainder of this section assume a tetrahedral mesh of tensor product type. Assume further that the tensor product mesh is aligned with the coordinate axes (i.e. each circumscribing rectangular prism). Denote by $\tilde{h}_{1,T}, \tilde{h}_{2,T}, \tilde{h}_{3,T}$ the side lengths of the circumscribing rectangular prisms of a tetrahedron T. Set $\tilde{H}_T := \operatorname{diag}\{\tilde{h}_{1,T}, \tilde{h}_{2,T}, \tilde{h}_{3,T}\}$.

Define the modified residual error estimator and the estimator based on a recovered gradient by

$$\begin{aligned}
\eta^2_{\tilde{R},T}(u_h) &:= \sum_{E\in\partial T\setminus\Gamma_D} \frac{h^2_{min,T}}{h_E}\cdot\|r_E(u_h)\|^2_E \\
\eta_{Z,T}(u_h) &:= h_{min,T}\cdot\left\|\tilde{H}_T^{-T}\left(\nabla^R u_h - \nabla u_h\right)\right\|_T \qquad ,
\end{aligned}$$

respectively. For the proof we require node-related quantities which can be derived easily. With N_T denoting the set of nodes of T, one has

$$\begin{aligned}
\eta^2_{Z,T}(u_h) &= h^2_{min,T}\cdot\left\|\tilde{H}_T^{-T}\left(\nabla^R u_h - \nabla u_h\right)\right\|^2_T \\
&\sim h^2_{min,T}\cdot|T|\cdot\sum_{x\in N_T}\left\|\tilde{H}_T^{-T}\left(\nabla^R u_h - \nabla u_h\right)(\mathbf{x})\right\|^2_{\mathbb{R}^3} \qquad ,
\end{aligned}$$

implying the definition

$$\eta^2_{Z,x} := \sum_{T:x\in N_T}\left\|\tilde{H}_T^{-T}\left(\nabla^R u_h - \nabla u_h\Big|_T\right)(\mathbf{x})\right\|^2_{\mathbb{R}^3} \qquad .$$

From

$$\eta^2_{\tilde{R},T}(u_h) = \sum_{E\in\partial T\setminus\Gamma_D}\frac{h^2_{min,T}}{h_E}\cdot\|r_E(u_h)\|^2_E = 3|T|\sum_{E\in\partial T\setminus\Gamma_D}\frac{h^2_{min,T}}{h^2_E}\cdot r^2_E(u_h)$$

one easily identifies the node related quantity as

$$\eta^2_{\tilde{R},x}(u_h) := \sum_{\substack{E\ni x\\ E\not\subset\Gamma_D}} h_E^{-2}\cdot r^2_E(u_h) \qquad .$$

Theorem 3.14 *Assume a tetrahedral, tensor product type mesh which is aligned with the coordinate axes. Then the following relations hold.*

$$\begin{aligned}
\eta_{Z,x} &\sim \eta_{\tilde{R},x} \qquad , \\
\sum_{T\in\mathcal{T}_h}\eta^2_{Z,T} &\sim \sum_{T\in\mathcal{T}_h}\eta^2_{\tilde{R},T} \qquad ,
\end{aligned}$$

$$\eta_{Z,T} \lesssim \sum_{T'\cap T\neq\emptyset}\eta_{\tilde{R},T'} \qquad \textit{and} \qquad \eta_{\tilde{R},T} \lesssim \sum_{T'\cap T\neq\emptyset}\eta_{Z,T'} \qquad .$$

Chapter 4

A singularly perturbed reaction–diffusion equation

4.1 Analytical Background

Important real life problems where anisotropic solutions can occur include diffusion-convection-reaction problems, for example convection dominated problems or singularly perturbed problems. There so-called interior layers or boundary layers (of different kind) with strong anisotropic behaviour can evolve. In order to decide if error estimators can be applied in conjunction with anisotropic meshes, we have chosen the following model problem.

Let us consider the singularly perturbed reaction diffusion equation whose classical formulation reads

$$\text{Find } u \in C^2(\Omega) \cap C(\overline{\Omega}): \qquad \begin{array}{rcll} -\varepsilon\Delta u + u &=& f & \text{in } \Omega, \\ u &=& 0 & \text{on } \Gamma_D = \partial\Omega \quad . \end{array} \tag{4.1}$$

The positive parameter ε is supposed to be very small, $\varepsilon \ll 1$, and has much influence on the solution. Under suitable smoothness assumptions on the data (i.e. f and Ω) the differential equation (4.1) yields a unique solution. The corresponding variational or weak formulation is

$$\begin{array}{rrcl} \text{Find } u \in H_o^1(\Omega): & a(u,v) &=& (f,v) \qquad \forall\, v \in H_o^1(\Omega) \\ \text{with} & a(u,v) &:=& \displaystyle\int_\Omega \varepsilon \cdot \nabla^T u \nabla v \,+\, u\,v \\ & (f,v) &=& \displaystyle\int_\Omega f \cdot v \qquad . \end{array} \tag{4.2}$$

The Lax-Milgram lemma ensures that there exists a unique solution of (4.2) provided that

- $f \in [H_o^1(\Omega)]^* = H^{-1}(\Omega)$
- $a(\cdot\,,\,\cdot)$ is elliptic, i.e. $a(v,v) \geq \mu_1 \cdot \|v\|^2_{H_o^1(\Omega)} \qquad \forall\, v \in H_o^1(\Omega)$
- $a(\cdot\,,\,\cdot)$ is bounded, i.e. $|a(v,w)| \leq \mu_2 \cdot \|v\|_{H_o^1(\Omega)} \cdot \|w\|_{H_o^1(\Omega)} \qquad \forall\, v,w \in H_o^1(\Omega) \qquad .$

For the whole of our investigation we demand a stronger smoothness of the right-hand side, namely

$$f \in L_2(\Omega) \qquad ,$$

thus the first assumption is satisfied. The second and third assumption are automatically valid with constants $\mu_1 = \varepsilon$ and $\mu_2 = 1$.

The finite element method is exactly the same as in section 3.1, i.e.

$$\text{Find } u_h \in V_{o,h}: \qquad a(u_h, v_h) = (f, v_h) \qquad \forall\, v_h \in V_{o,h} \qquad . \tag{4.3}$$

Note that the energy norm is defined by the bilinear form and depends on ε:

$$|||v|||^2 := a(v,v) = \varepsilon\|\nabla v\|^2 + \|v\|^2 \qquad .$$

The model problem (4.1) is of interest since one can usually expect boundary layers when a non-vanishing right-hand side f meets homogeneous Dirichlet boundary conditions. Inside Ω and sufficiently far away from boundary the solution in usually smooth provided f is smooth enough too. Thus the boundary layers mark the domain of interest, and their resolution requires increased numerical effort. Note however that (4.1) is only a model problem insofar as

- the differential operator is still symmetric and elliptic.
- it can be solved using standard FEM, i.e. no modifications like the Galerkin least squares method or the streamline diffusion method are necessary.

For a more detailed introduction to the analysis and numerical treatment of singularly perturbed differential equations (convection-diffusion and flow problems) see Roos, Stynes and Tobiska [27], and the literature cited therein. Miller, O'Riordan and Shishkin [18] investigate singularly perturbed problems with emphasis on numerical methods and *a priori* estimates.

We are interested in error estimators in particular. *Isotropic* estimators for diffusion-convection-reaction problems can be roughly divided into two major classes. *A priori* error estimators (in conjunction with adapted numerical methods) are known for some time.

For *a posteriori* error estimators, however, the knowledge has been unsatisfactory for a long time. Most estimators yield upper and lower bounds on the error that are not asymptotically equivalent. By this we mean that the upper and lower bound differ by a factor that increases, for example, as the discretization parameter $h \to 0$, or as $\varepsilon \to 0$ in the case of a singularly perturbed problem. The first *a posteriori* error estimate with asymptotically equivalent upper and lower bound on the error is, to our knowledge, due to Angermann [2]. He measures the error in the somewhat strange norm

$$\|v\|_{V_0} := \sup_{v \in V_0} \frac{a(v,v)}{\|v\|_{H^1}}$$

which is weaker than the energy norm, i.e. $\sqrt{\varepsilon}\,|||v||| \le \|v\|_{V_0} \lesssim |||v|||$. Angermann himself stated that estimates in this norm are mainly of theoretical interest.

Only recently Verfürth [32] derived the first *a posteriori* error estimator in the energy norm for the model problem (4.1) where upper and lower bounds are asymptotically equivalent.

In the remainder of this chapter an *a posteriori* error estimator for model problem (4.1) is derived that can applied to *anisotropic* meshes. The upper and lower error bounds involve the same terms and are asymptotically equivalent. Our estimator is partially influenced by Verfürth's isotropic version.

4.2 Residual error estimator

4.2.1 Special face bubble functions

In this section special face bubble functions are defined, and the corresponding inverse inequalities will be derived. The definition and the proof are given first for the standard tetrahedron $\bar{T}$ and then for the actual tetrahedron T.

Consider the standard tetrahedron $\bar{T}$ and the face $\bar{E}_1$ thereof. For a real number $\delta \in (0,1]$ define a linear mapping $F_\delta : \mathbb{R}^d \mapsto \mathbb{R}^d$ by

$$\begin{array}{rcll} F_\delta(x_1,\ldots,x_n) & := & (\delta \cdot x_1, x_2, \ldots, x_n)^T & \\ \text{or} \qquad F_\delta(\mathbf{x}) & = & B_\delta \cdot \mathbf{x} & \text{with } \; B_\delta = \operatorname{diag}\{\delta, 1, \ldots, 1) \in \mathbb{R}^{d\times d} \qquad . \end{array}$$

Obviously this yields

$$|\det B_\delta| = \delta \qquad \text{and} \qquad \|B_\delta^{-1}\|_{\mathbb{R}^{d\times d}} = \delta^{-1} \qquad .$$

Set $\bar{T}_\delta := F_\delta(\bar{T})$, i.e. $\bar{T}_\delta$ is the tetrahedron with the face $\bar{E}_1$ and a vertex at $\delta \cdot \mathbf{e}_1$.

Let $\bar{b}_{E_1}$ be the usual face bubble function of $\bar{E}_1$ on $\bar{T}$ (cf. (2.18)). Define the special face bubble function $\bar{b}_\delta$ by

$$\bar{b}_\delta \;=\; \bar{b}_{E_1,\delta} \;:=\; \bar{b}_{E_1} \circ F_\delta^{-1}$$

i.e. $\bar{b}_\delta$ is the usual face bubble function of the face $\bar{E}_1$ on the tetrahedron $\bar{T}_\delta$. For clarity we recall $\bar{b}_\delta = 0$ on $\bar{T} \setminus \bar{T}_\delta$.

Then the following inverse inequalities hold.

Lemma 4.1 (Inverse inequalities on the standard tetrahedron) *Assume $\bar{\varphi} \in \mathbb{P}^0(\bar{E}_1)$, and let F_{ext} be the extension operator of (2.19). The following inverse inequalities hold.*

$$\begin{array}{rcl} \|\bar{b}_\delta \cdot F_{ext}(\bar{\varphi})\|_{\bar{T}} & \lesssim & \delta^{1/2} \cdot \|\bar{\varphi}\|_{\bar{E}_1} \\ \|\nabla(\bar{b}_\delta \cdot F_{ext}(\bar{\varphi}))\|_{\bar{T}} & \lesssim & \delta^{-1/2} \cdot \|\bar{\varphi}\|_{\bar{E}_1} \end{array}$$

Proof: We employ standard scaling techniques via F_δ and utilize the inverse inequalities (2.23) and (2.24) on $\bar{T}$. Hence the desired estimates

$$\begin{array}{rrcl} & \|\bar{b}_\delta \cdot F_{ext}(\bar{\varphi})\|_{\bar{T}} & = & \|\bar{b}_\delta \cdot F_{ext}(\bar{\varphi})\|_{\bar{T}_\delta} \;=\; |\det B_\delta|^{1/2} \cdot \|\bar{b}_{E_1} \cdot F_{ext}(\bar{\varphi})\|_{\bar{T}} \\ & & \overset{(2.23)}{\lesssim} & |\det B_\delta|^{1/2} \cdot h_{\bar{E}_1,\bar{T}}^{1/2} \cdot \|\bar{\varphi}\|_{\bar{E}_1} \;=\; \delta^{1/2} \cdot \|\bar{\varphi}\|_{\bar{E}_1} \\ \text{and} & \|\nabla(\bar{b}_\delta \cdot F_{ext}(\bar{\varphi}))\|_{\bar{T}} & = & \|\nabla(\bar{b}_\delta \cdot F_{ext}(\bar{\varphi}))\|_{\bar{T}_\delta} \\ & & = & |\det B_\delta|^{1/2} \cdot \|B_\delta^{-T} \cdot \nabla(\bar{b}_{E_1} \cdot F_{ext}(\bar{\varphi}))\|_{\bar{T}} \\ & & \lesssim & \delta^{1/2} \cdot \|B_\delta^{-1}\|_{\mathbb{R}^{d\times d}} \cdot \|\nabla(\bar{b}_{E_1} \cdot F_{ext}(\bar{\varphi}))\|_{\bar{T}} \\ & & \overset{(2.24)}{\lesssim} & \delta^{-1/2} \cdot h_{\bar{E}_1,\bar{T}}^{1/2} \cdot h_{min,\bar{T}}^{-1} \cdot \|\bar{\varphi}\|_{\bar{E}_1} \;\sim\; \delta^{-1/2} \cdot \|\bar{\varphi}\|_{\bar{E}_1} \end{array}$$

are obtained. ■

Remark 4.1 All inverse inequalities of this previous lemma are valid for any face $\bar{E}$ of $\bar{T}$ (i.e. not only for $\bar{E}_1$) if the face bubble function $\bar{b}_\delta$ is defined correspondingly. □

Consider now an actual tetrahedron T. The special face bubble function $b_\delta = b_{E,\delta} \in H^1(T)$ of a face E of T is defined by

$$b_\delta \;=\; b_{E,\delta} \;:=\; \bar{b}_{E,\delta} \circ F_A^{-1} \qquad . \tag{4.4}$$

Lemma 4.2 (Inverse inequalities on the actual tetrahedron)
Let E be an arbitrary face of T. Assume $\varphi_E \in \mathbb{P}^0(E)$. The following inverse inequalities hold.

$$\|b_\delta \cdot F_{ext}(\varphi_E)\|_T \;\lesssim\; \delta^{1/2} \cdot h_E^{1/2} \cdot \|\varphi_E\|_E \tag{4.5}$$

$$\|\nabla(b_\delta \cdot F_{ext}(\varphi_E))\|_T \;\lesssim\; \delta^{-1/2} \cdot h_E^{1/2} \cdot h_{min,T}^{-1} \cdot \|\varphi_E\|_E \tag{4.6}$$

Proof: Standard scaling arguments and the previous lemma readily imply

$$\|b_\delta \cdot F_{ext}(\varphi_E)\|_T^2 = 6|T| \cdot \|\bar{b}_\delta \cdot F_{ext}(\bar{\varphi}_E)\|_{\bar{T}}^2 \lesssim 6|T| \cdot \delta \cdot \|\bar{\varphi}_E\|_{\bar{E}}^2 = \delta \cdot h_E \cdot \|\varphi_E\|_E^2 \qquad .$$

The other inequality is derived completely analogously and thus left to the reader. ■

4.2.2 Anisotropic interpolation estimates

The interpolation estimates sought contain the energy norm $|||\cdot|||$ on the right-hand side. For this reason the term ε (which is related to the differential operator and not to the interpolation operator) enters the left-hand side. More precisely, define the auxiliary term

$$\alpha_T := \min\{1, \varepsilon^{-1/2} \cdot h_{min,T}\} \qquad . \tag{4.7}$$

The following lemma is valid.

Lemma 4.3 *Let R_o be the Clément interpolation operator defined in (3.7). Assume that the mesh is adapted to v, i.e. $v \in H^1_{o,\mathcal{T}}(\Omega)$. Then the interpolation estimates*

$$\sum_{T \in \mathcal{T}_h} \alpha_T^{-2} \cdot \|v - R_o v\|_T^2 \;\lesssim\; |||v|||^2 \tag{4.8}$$

$$\varepsilon^{1/2} \sum_{T \in \mathcal{T}_h} \sum_{E \in \partial T \setminus \Gamma_D} \alpha_T^{-1} \cdot \frac{h_{E,T}}{h_{min,T}} \|v - R_o v\|_E^2 \;\lesssim\; |||v|||^2 \tag{4.9}$$

hold.

Proof: The definition of α_T implies

$$\alpha_T^{-1} \;=\; \max\left\{1, \varepsilon^{1/2} \cdot h_{min,T}^{-1}\right\} \qquad .$$

The anisotropic interpolation estimates of theorem 3.3 on page 30 result in

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} \alpha_T^{-2} \cdot \|v - R_o v\|_T^2 \;&=\; \sum_{\substack{T \in \mathcal{T}_h \\ 1 \geq \varepsilon \cdot h_{min,T}^{-2}}} \|v - R_o v\|_T^2 \;+\; \sum_{\substack{T \in \mathcal{T}_h \\ 1 < \varepsilon \cdot h_{min,T}^{-2}}} \varepsilon h_{min,T}^{-2} \cdot \|v - R_o v\|_T^2 \\ &\leq\; \|v - R_o v\|^2 \;+\; \varepsilon \cdot \|h_{min}^{-1}(\mathbf{x}) \cdot (v - R_o v)\|^2 \\ &\lesssim\; \|v\|^2 \;+\; \varepsilon \cdot \|\nabla v\|^2 \;=\; |||v|||^2 \end{aligned}$$

which proves the first inequality.

For the second estimate the trace inequality (2.13) is invoked giving

$$h_{E,T} \cdot \|v - R_o v\|_E^2 \lesssim \|v - R_o v\|_T \cdot \left(\|v - R_o v\|_T + \|C_T^T \nabla(v - R_o v)\|_T\right).$$

Utilizing the first result (4.8), the Cauchy–Schwarz inequality, and theorem 3.3 on page 30 results in

$$\begin{aligned}
\varepsilon^{1/2} \sum_{T\in\mathcal{T}_h} \sum_{E\in\partial T\setminus\Gamma_D} & \alpha_T^{-1} \cdot \frac{h_{E,T}}{h_{min,T}} \|v - R_o v\|_E^2 \lesssim \\
&\lesssim \varepsilon^{1/2} \sum_{T\in\mathcal{T}_h} \left[\alpha_T^{-1} \cdot \|v - R_o v\|_T \cdot h_{min,T}^{-1} \cdot \left(\|v - R_o v\|_T + \|C_T^T \nabla(v - R_o v)\|_T\right)\right] \\
&\lesssim \varepsilon^{1/2} \cdot \left(\sum_{T\in\mathcal{T}_h} \alpha_T^{-2} \cdot \|v - R_o v\|_T^2\right)^{1/2} \cdot \\
&\quad \cdot \left(\|h_{min}^{-1}(\mathbf{x}) \cdot (v - R_o v)\|^2 + \|h_{min}^{-1}(\mathbf{x}) \cdot C^T(\mathbf{x}) \nabla(v - R_o v)\|^2\right)^{1/2} \\
&\lesssim \varepsilon^{1/2} \cdot |||v||| \cdot \|\nabla v\| \leq |||v|||^2 \qquad .
\end{aligned}$$

Hence the second estimate is proven. ∎

4.2.3 Anisotropic residual error estimator

Let the element residual over a tetrahedron T be defined by

$$r_T(v_h) := P_{L_2} f - (-\varepsilon \cdot \Delta v_h + v_h) \qquad . \tag{4.10}$$

Obviously this residual of v_h is related to the strong form of the differential operator. Therefore the definition of r_T is problem dependent and in particular different to the definition for the Poisson equation.

Definition 4.1 (Residual error estimator) *The* local residual error estimator $\eta_{R,T}(u_h)$ *for a tetrahedron T is defined by*

$$\eta_{R,T}(u_h) := \left(\alpha_T^2 \cdot \|r_T(u_h)\|_T^2 + \varepsilon^{3/2} \cdot \alpha_T \cdot \sum_{E\in\partial T\setminus\Gamma_D} \frac{h_{min,T}}{h_E} \cdot \|r_E(u_h)\|_E^2\right)^{1/2} \qquad . \tag{4.11}$$

Theorem 4.4 (Residual error estimator) *Let $u \in H_o^1(\Omega)$ be the exact solution and $u_h \in V_{o,h}$ be the FEM solution. Then the error is bounded locally from below by*

$$\eta_{R,T}(u_h) \lesssim \left(|||u - u_h|||_{\omega_T}^2 + \alpha_T^2 \cdot \|f - P_{L_2} f\|_{\omega_T}^2\right)^{1/2} \tag{4.12}$$

for all $T \in \mathcal{T}_h$.

Assume further that the mesh is adapted to the error $u - u_h$, i.e. $u - u_h \in H_{o,\mathcal{T}}^1(\Omega)$. Then the error is bounded globally from above by

$$|||u - u_h||| \lesssim \left(\sum_{T\in\mathcal{T}_h} \eta_{R,T}^2(u_h) + \sum_{T\in\mathcal{T}_h} \alpha_T^2 \cdot \|f - P_{L_2} f\|_T^2\right)^{1/2} \qquad . \tag{4.13}$$

Proof: The proof of the first estimate (4.12) employs some standard techniques already utilized for the Poisson equation. A more detailed investigation can be found there.

We start with the norm $\|r_T(u_h)\|_T$ of the element residual $r_T = P_{L_2} f + \varepsilon \cdot \Delta u_h - u_h)$. Since we use linear ansatz functions $r_T \in \mathbb{P}^0(T)$ holds. For $\mathbf{x} \in T$ let

$$w(\mathbf{x}) := r_T(u_h)(\mathbf{x}) \cdot b_T(\mathbf{x}) \qquad \in \mathbb{P}^4(T) \cap H_o^1(T) \qquad ,$$

with b_T being the usual bubble functions introduced in section 2.3.3. Integration by parts yields

$$\begin{aligned}
\int_T r_T \cdot w &= \int_T (f + \varepsilon \cdot \Delta u_h - u_h) \cdot w \;+\; \int_T (P_{L_2} f - f) \cdot w \\
&= \int_T \varepsilon \cdot \nabla^T (u - u_h) \cdot \nabla w + (u - u_h) \cdot w \;+\; \int_T (P_{L_2} f - f) \cdot w \\
\left| \int_T r_T \cdot w \right| &\le \varepsilon \cdot \|\nabla(u - u_h)\|_T \cdot \|\nabla w\|_T \;+\; \|u - u_h\| \cdot \|w\|_T \;+\; \|f - P_{L_2} f\|_T \cdot \|w\|_T \quad .
\end{aligned}$$

Bounds of $\left|\int_T r_T \cdot w\right|$, $\|\nabla w\|_T$ and $\|w\|_T$ have already been derived in (3.17). Hence one readily obtains

$$\begin{aligned}
\|r_T\|_T^2 &\lesssim \varepsilon^2 \cdot h_{min,T}^{-2} \cdot \|\nabla(u - u_h)\|_T^2 \;+\; \|u - u_h\|_T^2 \;+\; \|f - P_{L_2} f\|_T^2 \\
\alpha_T^2 \cdot \|r_T\|_T^2 &\lesssim \min\{\varepsilon \cdot h_{min,T}^{-2}, 1\} \cdot \varepsilon \cdot \|\nabla(u - u_h)\|_T^2 \;+ \\
&\quad + \alpha_T^2 \cdot \|u - u_h\|_T^2 \;+\; \alpha_T^2 \cdot \|f - P_{L_2} f\|_T^2 \\
&\le \varepsilon \cdot \|\nabla(u - u_h)\|_T^2 \;+\; \|u - u_h\|_T^2 \;+\; \alpha_T^2 \cdot \|f - P_{L_2} f\|_T^2 \\
&= |||u - u_h|||_T^2 \;+\; \alpha_T^2 \cdot \|f - P_{L_2} f\|_T^2 \quad . \qquad (4.14)
\end{aligned}$$

Now we aim for a bound of the norm $\|r_E(u_h)\|_E$ of the gradient jump across some inner face (triangle) E. Since we use linear ansatz functions $r_E \in \mathbb{P}^0(E)$ holds. Let T_1 and T_2 be the two tetrahedra that E belongs to. Assume that the right hand side $f = -\varepsilon \Delta u + u$ is in $L_2(\Omega)$. Integration by parts yields for any function $w \in H_o^1(\omega_E)$

$$\begin{aligned}
0 &= \int_{\omega_E} \varepsilon \nabla^T u \nabla w \;+\; u \cdot w \;-\; f \cdot w \\
-\varepsilon \int_E r_E(u_h) \cdot w &= \varepsilon \sum_{i=1}^{2} \int_{\partial T_i} w \cdot \frac{\partial u_h}{\partial n} = \varepsilon \sum_{i=1}^{2} \int_{T_i} \left(\nabla^T u_h \nabla w \;+\; \Delta u_h \cdot w \right) \\
&= \sum_{i=1}^{2} \int_{T_i} \left(\varepsilon \nabla^T u_h \nabla w \;+\; (r_{T_i} - P_{L_2} f + u_h) \cdot w \right) \\
&= \sum_{i=1}^{2} \int_{T_i} \left(\varepsilon \nabla^T (u_h - u) \nabla w \;+\; (u_h - u) \cdot w \;+\; (r_{T_i} + f - P_{L_2} f) \cdot w \right)
\end{aligned}$$

since $\varepsilon \Delta u_h = r_{T_i} - P_{L_2} f + u_h$ on T_i. Let now the function w be defined by

$$w := \begin{cases} b_{E,\delta_1} \cdot F_{ext}(r_E(u_h)) & \text{on } T_1 \\ b_{E,\delta_2} \cdot F_{ext}(r_E(u_h)) & \text{on } T_2 \end{cases} \qquad ,$$

with F_{ext} being the extension operator of (2.19) and b_{E,δ_i} being the special face bubble functions defined above. The real numbers δ_i will be chosen later.

Note that $w \in H_o^1(\omega_E)$ since $b_{E,\delta_1}\Big|_E = b_{E,\delta_2}\Big|_E = b_E\Big|_E$. Hence we conclude

$$\varepsilon\ \|b_E^{1/2}\cdot r_E\|_E^2 \le \sum_{i=1}^{2}\Big(\varepsilon\|\nabla(u-u_h)\|_{T_i}\cdot\|\nabla w\|_{T_i} + (\|u-u_h\|_{T_i}+\|r_{T_i}\|_{T_i}+\|f-P_{L_2}f\|_{T_i})\cdot\|w\|_{T_i}\Big).$$

The inverse inequalities (4.5) and (4.6) are used to bound $\|w\|_{T_i}$ and $\|\nabla w\|_{T_i}$, respectively, and subsequently imply

$$\begin{aligned}\|r_E\|_E \ &\lesssim\ \sum_{i=1}^{2} h_E^{1/2}\cdot\Big(h_{min,T_i}^{-1}\cdot\delta_i^{-1/2}\cdot\|\nabla(u-u_h)\|_{T_i}\ + \\ &\qquad + \ \varepsilon^{-1}\cdot\delta_i^{1/2}\cdot(\|u-u_h\|_{T_i}+\|r_{T_i}\|_{T_i}+\|f-P_{L_2}f\|_{T_i})\Big)\qquad .\end{aligned}$$

Now we choose $\delta_i := \varepsilon^{1/2}\cdot h_{min,T_i}^{-1}\cdot\alpha_{T_i}\le 1$ and insert estimate (4.14) which provides a bound of $\|r_{T_i}\|_{T_i}$. One obtains

$$\begin{aligned}&\varepsilon^{3/2}\cdot\alpha_T\cdot\frac{h_{min,T}}{h_E}\cdot\|r_E(u_h)\|_E^2\ \le\\ &\quad\lesssim\ \sum_{i=1}^{2}\varepsilon\cdot\|\nabla(u-u_h)\|_{T_i}^2+\alpha_{T_i}^2\cdot\|u-u_h\|_{T_i}^2+|||u-u_h|||_{T_i}^2+\alpha_{T_i}^2\cdot\|f-P_{L_2}f\|_{T_i}^2\\ &\quad\lesssim\ |||u-u_h|||_{\omega_E}^2\ +\ \alpha_T^2\cdot\|f-P_{L_2}f\|_{\omega_E}^2\end{aligned}$$

since h_{min,T_i} and α_{T_i} do not change rapidly across adjacent tetrahedra, and since $\alpha_{T_i}\le 1$. Summing up over all faces E of T, recalling the definition of $\eta_{R,T}(u_h)$ and applying (4.14) finishes the proof of the lower error bound (4.12).

Secondly, in order to derive (4.13) we utilize the orthogonality property of the error

$$a(u-u_h,v_h)\ =\ 0\qquad\forall\, v_h\in V_{o,h}\qquad .$$

Integration by parts gives for all $v\in H_{o,\mathcal{T}}^1(\Omega)$

$$\begin{aligned}a(u-u_h,v) &= a(u-u_h,v-R_ov)\\ &=\ \varepsilon(\nabla(u-u_h),\nabla(v-R_ov))\ +\ (u-u_h,v-R_ov)\\ &=\ \sum_{T\in\mathcal{T}_h}(f+\varepsilon\Delta u_h-u_h,v-R_ov)_T\ -\ \varepsilon\sum_{E\in\Omega}(r_E(u_h),v-R_ov)_E\\ &=\ \sum_{T\in\mathcal{T}_h}\Big[(r_T(u_h)+f-P_{L_2}f,v-R_ov)_T\ -\ \frac{1}{2}\cdot\varepsilon\sum_{E\in\partial T\setminus\Gamma_D}(r_E(u_h),v-R_ov)_E\Big]\\ &\le\ \sum_{T\in\mathcal{T}_h}\Big[\alpha_T(\|r_T(u_h)\|_T+\|f-P_{L_2}f\|_T)\cdot\alpha_T^{-1}\|v-R_ov\|_T\ +\\ &\qquad +\ \frac{1}{2}\sum_{E\in\partial T\setminus\Gamma_D}\varepsilon^{3/4}\alpha_T^{1/2}\frac{h_{min,T}^{1/2}}{h_{E,T}^{1/2}}\|r_E(u_h)\|_E\ \cdot\ \varepsilon^{1/4}\alpha_T^{-1/2}\frac{h_{E,T}^{1/2}}{h_{min,T}^{1/2}}\|v-R_ov\|_E\Big].\end{aligned}$$

The Cauchy-Schwarz inequality and the interpolation estimate (4.8) yield

$$\sum_{T\in\mathcal{T}_h} \alpha_T(\|r_T(u_h)\|_T + \|f - P_{L_2}f\|_T)\cdot \alpha_T^{-1}\|v - R_o v\|_T \;\le$$

$$\le\; \Big(2\sum_{T\in\mathcal{T}_h}\alpha_T^2\Big(\|r_T(u_h)\|_T^2 + \|f - P_{L_2}f\|_T^2\Big)\Big)^{1/2}\cdot\Big(\sum_{T\in\mathcal{T}_h}\alpha_T^{-2}\|v - R_o v\|_T^2\Big)^{1/2}$$

$$\lesssim\; \Big(\sum_{T\in\mathcal{T}_h}\alpha_T^2\Big(\|r_T(u_h)\|_T^2 + \|f - P_{L_2}f\|_T^2\Big)\Big)^{1/2}\cdot |||v|||$$

since $v \in H^1_{o,\mathcal{T}}(\Omega)$. Analogously

$$\sum_{T\in\mathcal{T}_h}\sum_{E\in\partial T\setminus\Gamma_D} \varepsilon^{3/4}\alpha_T^{1/2}\frac{h_{min,T}^{1/2}}{h_{E,T}^{1/2}}\|r_E(u_h)\|_E \;\cdot\; \varepsilon^{1/4}\alpha_T^{-1/2}\frac{h_{E,T}^{1/2}}{h_{min,T}^{1/2}}\|v - R_o v\|_E \;\le$$

$$\le\; \Big(\varepsilon^{3/2}\sum_{T\in\mathcal{T}_h}\sum_{E\in\partial T\setminus\Gamma_D}\alpha_T\frac{h_{min,T}}{h_{E,T}}\|r_E(u_h)\|_E^2\Big)^{1/2}\cdot$$

$$\cdot\Big(\varepsilon^{1/2}\sum_{T\in\mathcal{T}_h}\sum_{E\in\partial T\setminus\Gamma_D}\alpha_T^{-1}\frac{h_{E,T}}{h_{min,T}}\|v - R_o v\|_E^2\Big)^{1/2}$$

$$\lesssim\; \Big(\varepsilon^{3/2}\sum_{T\in\mathcal{T}_h}\sum_{E\in\partial T\setminus\Gamma_D}\alpha_T\frac{h_{min,T}}{h_{E,T}}\|r_E(u_h)\|_E^2\Big)^{1/2}\cdot |||v|||$$

is derived with the help of interpolation estimate (4.9). Combining these estimates results in

$$a(u - u_h, v) \;\lesssim\; \Big(\sum_{T\in\mathcal{T}_h}\Big[\alpha_T^2\Big(\|r_T(u_h)\|_T^2 + \|f - P_{L_2}f\|_T^2\Big) \;+$$

$$+\;\varepsilon^{3/2}\alpha_T\sum_{E\in\partial T\setminus\Gamma_D}\frac{h_{min,T}}{h_{E,T}}\|r_E(u_h)\|_E^2\Big]\Big)^{1/2}\cdot |||v||| \qquad .$$

Substituting $v := u - u_h \in H^1_{o,\mathcal{T}}(\Omega)$ finishes the proof. ∎

Chapter 5

Numerical examples

5.1 Scope of and introduction to the numerical experiments

After the introductory words of chapter 1, the reader certainly expects an example where an adaptive anisotropic strategy and its superiority is demonstrated. Hopefully the reader is not too disappointed that we will not proceed that way. In order to justify this, recall the steps of an adaptive anisotropic strategy, namely

1. Estimate the error for a solution on a given mesh.

2. Obtain information for a new, better mesh. This includes:

   - Detect regions of anisotropic behaviour of the solution.
   - Determine a (quasi) optimal aspect ratio and stretching direction of the finite elements.
   - Determine the element size.

3. Based on this information, construct a new mesh or perform a mesh refinement.

4. Solve the arising finite element system.

We have investigated task 1 and will test our error estimators numerically.

The second step, the extraction of mesh information, is much less clear. As far as we know only heuristic considerations are invoked. For example, the analysis of the partial second derivatives is frequently employed (cf. [21, 24, 35]). In practice however, this partially yields unrealistic and unpractically high optimal aspect ratios, and therefore an artificial maximum aspect ratio has to be specified. Other approaches suffer similar drawbacks, e.g. it is not clear how to extend the use of Lagrange multipliers (see e.g. Rachowicz [22]) in order to obtain the stretching direction.

The third task is to construct or refine an anisotropic mesh. It requires much programming effort and a sophisticated and efficient data structure. We could not afford to, and did not want to dedicate the whole energy to this remeshing. Also, we did not have access to an anisotropic mesh construction tool (which might be a topic of future work or collaboration). Mainly because of these reasons we do not present an adaptive strategy but focus mainly on error estimators instead.

Finally, the FEM system is solved using a standard FEM package, and not much thought is devoted to this step yet.

Based on the scope of and our ability for numerical validation we have chosen two examples. They include the following tasks and tests.

1. Utilize two anisotropic meshes which were constructed on *a priori* knowledge.
2. Test if the condition $u - u_h \in H^1_{o,\mathcal{T}}(\Omega)$ is satisfied.
3. Test the lower and upper bound on the error for
   - the residual error estimator $\eta_{R,T}$,
   - the local Dirichlet problem error estimator $\eta_{D,T}$,
   - the L_2 residual error estimator $\eta_{R,L_2,T}$.

The following two-dimensional Poisson problem is chosen as test problem.

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega = [0,1] \times [0,1] \qquad , \\ u &= 0 \qquad \text{on } \Gamma_D = \partial\Omega \qquad . \end{aligned}$$

The exact solution u is prescribed to be

$$u(x,y) := \left(1 - e^{-\alpha x} - (1 - e^{-\alpha})\, x\right) \cdot 4\, y\, (1-y)$$

with a parameter $\alpha = 1000$. The right-hand side f is chosen accordingly. The exact solution exhibits an exponential layer with an initial steepness of $\alpha = 1000$ along the boundary at $x = 0$. Figure 5.1 shows a rough image of u.

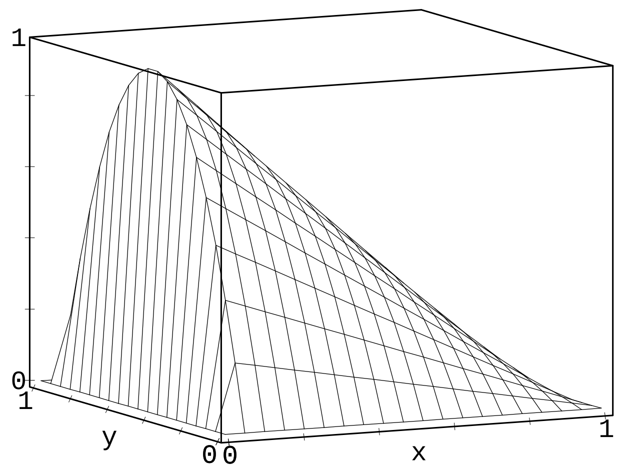


Figure 5.1: Exact solution u

We also have to mention the present shortcomings of the tests in order to distinguish between what the numerical experiments can tell us and what still remains hidden. The shortcomings are:

- Some three-dimensional experiments were carried out but they are not shown here.
- We do not consider a family $\mathcal{F}$ of triangulations $\mathcal{T}_h$ but only one given mesh at a time.
- The error estimator for the singularly perturbed problem is not investigated.

- Since the exact solution u is known *a priori*, the condition $u - u_h \in H^1_{o,\mathcal{T}}(\Omega)$ can be tested numerically. Usually this is not possible. Furthermore, the assumption $v_D \in H^2_{\mathcal{T}}(\Omega)$ for the L_2 error estimator can not be tested since the dual solution is not known.

We are aware that our two very limited examples give no numerical validation of our estimator but we did not aim at that at present. More precisely, we want to show that error estimators can be applied to anisotropic meshes, despite the theoretical shortcoming $u - u_h \in H^1_{o,\mathcal{T}}(\Omega)$ or $v_D \in H^2_{\mathcal{T}}(\Omega)$, respectively. Thus further research in this topic is justified.

5.2 Two numerical examples

Let us consider the two meshes depicted in figure 5.2. The adaption to the boundary layer can be seen clearly.

Mesh 1 (the left one) is an unstructured mesh and has been designed to give a small error in the H^1 seminorm. The largest interior angle of its triangles is about 179.77 degree. *Mesh 2* is a structured, tensor product type mesh and shall result in a small error in the L_2 norm.

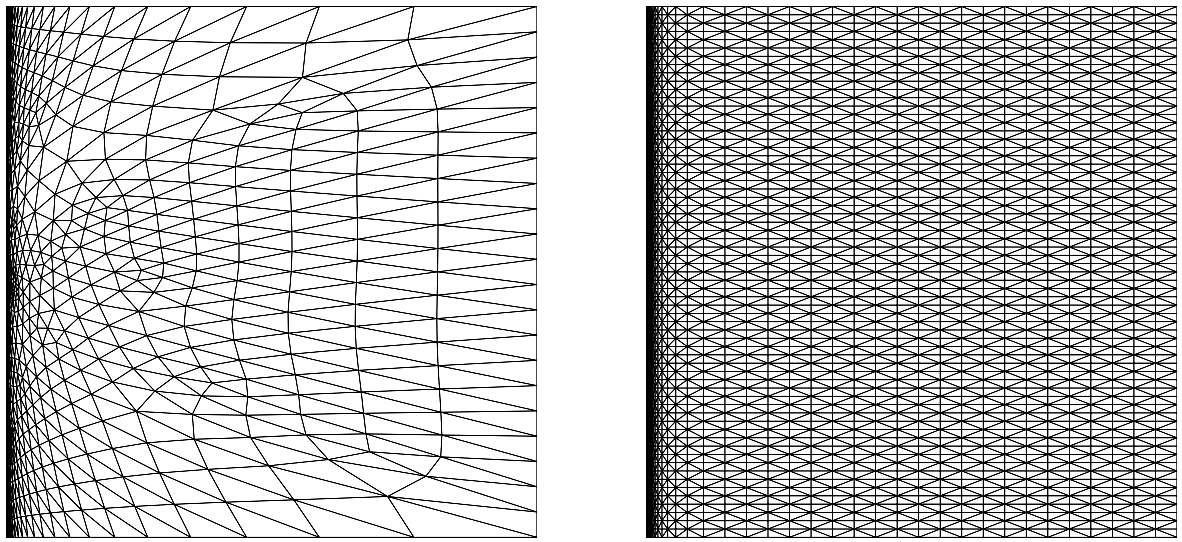


Figure 5.2: Mesh 1 and mesh 2

The following table displays some additional details of both meshes.

	Mesh 1	Mesh 2
Number of elements	7796	8192
Number of nodes	4020	4225
Maximum aspect ratio	1091.9	210.3
Nodal L_∞ error	7.767E-3	1.614E-3

5.2.1 The condition on $u - u_h$

If the residual error estimator (for both the Poisson problem and the singularly perturbed reaction diffusion problem) or the local Dirichlet problem error estimator is to be applied in order to obtain reliable upper bounds on the error, one has to assume $u - u_h \in H^1_{o,\mathcal{T}}(\Omega)$. In explicit writing, the inequality

$$\sum_{T\in\mathcal{T}_h} h_{min,T}^{-2} \cdot \|C_T^T \nabla(u-u_h)\|_T^2 \;\leq\; c_a \cdot \|\nabla(u-u_h)\|^2$$

has to hold with a bounded constant c_a. (Recall that, strictly speaking, this has to hold for a family $\mathcal{F}$ of triangulations).

For real life problems this condition can not be tested since the exact solution u is not known. Therefore we have spent much time to find a *computable* assumption that implies the condition above. Unfortunately, however, we failed despite much research.

Our investigations and several examples strengthened our impression that the condition reflects indeed how good the anisotropic mesh is adapted to (or matches) an anisotropic solution. In particular, heuristic considerations (for the stretching direction, the aspect ratio etc.) usually lead to anisotropic meshes where error estimators provide acceptable and useful error bounds. In this sense we are aware that the condition $u - u_h \in H^1_{o,\mathcal{T}}(\Omega)$ cannot be tested numerically but we equally think that nevertheless error estimators should be applied.

For our examples, the smallest possible constant c_a has been evaluated by dividing the left-hand side of above's inequality by $\|\nabla(u-u_h)\|^2$. The constants equal

$$\begin{array}{lll} \text{Mesh 1} : & c_a = 3.351 & , \\ \text{Mesh 2} : & c_a = 2.702 & . \end{array}$$

Since c_a is always larger than 1, one can certainly state that c_a is not too large and thus $u - u_h \in H^1_{o,\mathcal{T}}(\Omega)$ is satisfied.

5.2.2 Error bounds in the energy norm

The condition $u - u_h \in H^1_{o,\mathcal{T}}(\Omega)$ is satisfied for our meshes, as we have seen. Then the upper bound of the error for the residual error estimator or the local Dirichlet problem error estimator are both of the form

$$\|\nabla(u-u_h)\| \lesssim \left(\sum_{T\in\mathcal{T}_h} \eta_{*,T}^2(u_h) \;+\; \sum_{T\in\mathcal{T}_h} h_{min,T}^2 \cdot \|f - P_{L_2} f\|_T^2 \right)^{1/2}$$

where $\eta_{*,T}$ is either $\eta_{R,T}$ or $\eta_{D,T}$, respectively. For the validation of the error bound we computed the following quantities.

Term	Formula	Mesh 1	Mesh 2
Exact error	$\|\nabla(u-u_h)\|$	0.3863	0.4923
Residual error estimate	$\left(\sum_{T\in\mathcal{T}_h} \eta_{R,T}^2(u_h)\right)^{1/2}$	1.8774	2.5020
Local problem error estimate	$\left(\sum_{T\in\mathcal{T}_h} \eta_{D,T}^2(u_h)\right)^{1/2}$	0.6130	0.7947
Approximation error	$\left(\sum_{T\in\mathcal{T}_h} h_{min,T}^2 \cdot \|f - P_{L_2} f\|_T^2\right)^{1/2}$	0.1798	0.0487

These numbers show that, for both examples, the error is bounded from above indeed. More precisely, the error is overestimated. The approximation term plays a minor role although the right-hand side f is very large in the boundary layer.

The lower bounds on the error hold unconditionally, and they are of the form

$$\begin{aligned} \eta_{R,T}(u_h) &\lesssim \|\nabla(u-u_h)\|_{\omega_T} + h_{min,T} \cdot \|f - P_{L_2} f\|_{\omega_T} \\ \text{and} \qquad \eta_{D,T}(u_h) &\leq \|\nabla(u-u_h)\|_{\omega_T} + c \cdot h_{min,T} \cdot \|f - P_{L_2} f\|_{\omega_T} \qquad . \end{aligned}$$

Therefore we consider the following ratios which have to be bounded from above, and also give the range of its value over all elements T.

Ratio	Mesh 1	Mesh 2
$\dfrac{\eta_{R,T}}{\|\nabla(u-u_h)\|_{\omega_T} + h_{min,T} \cdot \|f - P_{L_2} f\|_{\omega_T}}$	$0.34 \ldots 3.6$	$0.24 \ldots 3.2$
$\dfrac{\eta_{D,T}}{\|\nabla(u-u_h)\|_{\omega_T} + h_{min,T} \cdot \|f - P_{L_2} f\|_{\omega_T}}$	$0.15 \ldots 1.0$	$0.20 \ldots 1.0$

These values clearly validate the lower bounds of the error.

5.2.3 Error bounds in the L_2 norm

In order to obtain an upper bound on the error, the condition $v_D \in H^2_{\mathcal{T}}(\Omega)$ on the dual solution v_D has to be satisfied. This condition can not be tested but it appears that both meshes reflect the anisotropic solution well enough. Thus we hope that the error estimator yields realistic bounds. The upper bound is of the form

$$\|u - u_h\| \lesssim \left(\sum_{T \in \mathcal{T}_h} \eta^2_{R,L_2,T}(u_h) + \sum_{T \in \mathcal{T}_h} h^4_{min,T} \cdot \|f - P_{L_2} f\|^2_T \right)^{1/2} \qquad .$$

For the validation of the error bound we computed the following quantities.

Term	formula	Mesh 1	Mesh 2
Exact error	$\|u - u_h\|$	3.768E-03	3.285E-04
L_2 error estimate	$\left(\sum_{T \in \mathcal{T}_h} \eta^2_{R,L_2,T}(u_h) \right)^{1/2}$	4.234E-02	2.567E-03
Approximation error	$\left(\sum_{T \in \mathcal{T}_h} h^4_{min,T} \cdot \|f - P_{L_2} f\|^2_T \right)^{1/2}$	7.279E-04	2.292E-05

These numbers show that the error is again overestimated. The approximation error is negligible compared with the exact error.

The lower bound on the error holds if the additional mesh requirement (3.30) is satisfied, which is the case for both meshes. The error bound then reads

$$\eta_{R,L_2,T}(u_h) \lesssim \|u - u_h\|_{\omega_T} + h^2_{min,T} \cdot \|f - P_{L_2} f\|_{\omega_T}$$

Therefore we consider the following ratio which has to be bounded from above, and also give the range of its value over all elements T.

Ratio	Mesh 1	Mesh 2
$\dfrac{\eta_{R,L_2,T}}{\Vert u-u_h\Vert_{\omega_T} \;+\; h^2_{min,T}\cdot\Vert f-P_{L_2}f\Vert_{\omega_T}}$	$0.16\ldots 24$	$0.05\ldots 16$

Although the lower bound is validated, the range of the corresponding ratios is much larger now.

5.3 Conclusions

The first test has been aiming at the condition $u-u_h\in H^1_{o,\mathcal{T}}(\Omega)$ which is satisfied for both meshes. This means that this condition *can* be satisfied. Moreover, what is done heuristically by anisotropic mesh generation seems to correspond to (or be reflected by) this condition.

Secondly, the energy error estimators show the anticipated behaviour. The local Dirichlet problem error estimator performs better than the residual error estimator. This coincides with the different quality of the lower bounds on the error, cf. (3.15) and (3.26). Note however, that the local problem estimator requires more computational effort.

Lastly, the L_2 residual error estimator yields less sharp bounds. This might be partially due to the fact that the energy norm is naturally related to the differential equation but the choice of the L_2 norm is somewhat arbitrary.

Chapter 6

Summary

This work has been aiming at error estimators suitable for anisotropic tetrahedral or triangular grids, respectively.

Several estimators known from the isotropic case have been investigated as to whether they can be modified for and applied to simplicial anisotropic meshes. The anisotropic *residual error estimator* has been derived and proven to be equivalent to the error. The condition $u - u_h \in H^1_{o,\mathcal{T}}(\Omega)$ has been discussed in section 5.2.1.

Several more estimators have been devised for the Poisson problem. The *local Dirichlet problem error estimator* has been shown to be equivalent to the residual error estimator and thus, to the error too. A general *Zienkiewicz-Zhu like error estimator* could not be derived; only special cases have been considered and proven. An L_2 *error estimator* is given as well. The conditions to guarantee the L_2 error bounds are, however, rather difficult.

As promised in the introduction, a singularly perturbed reaction diffusion equation has been investigated into too, and an anisotropic *residual error estimator* has been found. It is apparent that the error estimator (and the proofs of the error bounds, of course) depend rather strongly on the underlying differential equations.

Two simple numerical examples have shown that error estimators *can* be applied successfully to anisotropic meshes. Hence further research is, in our opinion, justified.

6.1 Open points and future work

The present state of research suggests that the following topics could be investigated in the future. Partially they are already under consideration.

1. It would be very desirable to find means to guarantee the condition $u - u_h \in H^1_{o,\mathcal{T}}(\Omega)$. If not possible, this condition should be tested thoroughly numerically.

   It could be sensible to approximate the term $\nabla(u - u_h)$ appearing in the condition by some heuristically chosen value, for example by means of a recovered gradient $\nabla^R u_h - \nabla u_h$.

2. Other boundary conditions (i.e. Neumann and Robin boundary conditions) have already been considered but not been applied yet.

3. At present a local problem error estimator for the L_2 norm is investigated into.

4. All estimators will be tested extensively (including 3D).
5. An adaptive strategy would be desirable. For this, mesh information have to be extracted, and a remeshing is necessary.
6. Investigations on error estimators based on the complementary energy principle could possibly yield upper bounds on the error with a constant 1.

Bibliography

[1] M. Ainsworth and J. T. Oden. A unified approach to a posteriori error estimation using element residual methods. *Numer. Math.*, 65:23–50, 1993.

[2] L. Angermann. Balanced a-posteriori error estimates for finite volume type discretizations of convection-dominated elliptic problems. *Computing*, 55(4):305–323, 1995.

[3] T. Apel and M. Dobrowolski. Anisotropic interpolation with applications to the finite element method. *Computing*, 47:277–293, 1992.

[4] T. Apel and G. Lube. Anisotropic mesh refinement in stabilized Galerkin methods. *Numer. Math.*, 74(3):261–282, 1996.

[5] T. Apel and S. Nicaise. Elliptic problems in domains with edges: anisotropic regularity and anisotropic finite element meshes. In J. Cea, D. Chenais, G. Geymonat, and J. L. Lions, editors, *Partial Differential Equations and Functional Analysis (In Memory of Pierre Grisvard)*, pages 18–34. Birkhäuser, Boston, 1996. Shortened version of Preprint SPC94_16, TU Chemnitz-Zwickau, 1994.

[6] J. P. Aubin. Behaviour of the error of the approximate solution of boundary value problems for linear elliptic operators by Galerkin's and finite difference methods. *Ann. Scuola Norm. Sup. Pisa*, 21:599–637, 1967.

[7] I. Babuška and W. C. Rheinboldt. A-posteriori error estimates for the finite element method. *Int. J. Num. Meth. Eng.*, 12:1597–1615, 1978.

[8] I. Babuška and W. C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15(4):736–754, 1978.

[9] R. E. Bank and A. Weiser. Some a posteriori error estimators for elliptic partial differential equations. *Math. Comput.*, 44(170):283–301, 1985.

[10] R. Beinert and D. Kröner. Finite volume methods with local mesh alignment in 2–D. In *Adaptive Methods – Algorithms, Theory and Applications*, volume 46 of *Notes on Numerical Fluid Mechanics*, pages 38–53, Braunschweig, 1994. Vieweg.

[11] J. U. Brackbill. An adaptive grid with directional control. *J. Comput. Physics*, 108:38–50, 1993.

[12] M. J. Castro Díaz and F. Hecht. Anisotropic surface mesh generation. Report 2672, INRIA, 1995.

[13] P. Ciarlet. *The finite element method for elliptic problems.* North-Holland Publishing Company, Amsterdam, 1978.

[14] P. Clément. Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.*, 2:77–84, 1975.

[15] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems I: A linear model problem. *SIAM J. Num. Anal.*, 28:43–77, 1991.

[16] J. Fröhlich, J. Lang, and R. Roitzsch. Selfadaptive finite element computations with smooth time controller and anisotropic refinement. report 96–16, ZIB, 1996.

[17] R. Kornhuber and R. Roitzsch. On adaptive grid refinement in the presence of internal and boundary layers. *IMPACT of Computing in Sci. and Engrg.*, 2:40–72, 1990.

[18] J. J. H. Miller, E. O'Riordan, and G. Shishkin. *Fitted numerical methods for singularly perturbed problems. Error Estimates in the maximum norm for linear problems in one and two dimensions.* World Scientific Publications, Singapore, 1996.

[19] J. A. Nitsche. Ein Kriterium für die Quasi-Optimalität des Ritzschen Verfahrens. *Numer. Math.*, 21:138–160, 1968.

[20] R. H. Nochetto. Pointwise a posteriori error estimators for elliptic problems on highly graded meshes. *Math. Comp.*, 64:1–22, 1995.

[21] J. Peraire, M. Vahdati, K. Morgan, and O. C. Zienkiewicz. Adaptive remeshing for compressible flow computation. *J. Comp. Phys.*, 72:449–466, 1987.

[22] W. Rachowicz. An anisotropic h-type mesh refinement strategy. *Comput. Methods Appl. Mech. Engrg.*, 109:169–181, 1993.

[23] E. Rank and O. C. Zienkiewicz. A simple error estimator in the finite element method. *Comm. Appl. Num. Meth.*, 3:243–249, 1987.

[24] W. Rick, H. Greza, and W. Koschel. FCT-solution on adapted unstructured meshes for compressible high speed flow computations. In E. H. Hirschel, editor, *Flow simulation with high-performance computers I*, volume 38 of *Notes on Num. Fluid Mechanics*, pages 334–438. Vieweg, 1993.

[25] R. Rodriguez. Some remarks on the Zienkiewicz-Zhu estimator. *Int. J. Numer. Meth. in PDE*, 10:625–635, 1994.

[26] H.-G. Roos. Layer-adapted grids for singular perturbation problems. report MATH-NM-03-1997, TU Dresden, 1997.

[27] H.-G. Roos, M. Stynes, and L. Tobiska. *Numerical methods for singularly perturbed differential equations. Convection-diffusion and flow problems.* Springer, Berlin, 1996.

[28] K. G. Siebert. An a posteriori error estimator for anisotropic refinement. *Numer. Math.*, 73(3):373–398, 1996.

[29] R. Verfürth. A posteriori error estimators for the Stokes equation. *Numer. Math.*, 55:309–325, 1989.

[30] R. Verfürth. A posteriori error estimation and adaptive mesh–refinement techniques. *Journal of Computational and Applied Mathematics*, 50:67–83, 1994.

[31] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh–refinement techniques*. Wiley-Teubner, Chichester; Stuttgart, 1996.

[32] R. Verfürth. Robust a posteriori error estimators for the singularly perturbed Helmholtz equation. *Numer. Math.*, to appear.

[33] R. Vilsmeier and D. Hänel. Computational aspects of flow simulation in three dimensional, unstructured, adaptive grids. In E. H. Hirschel, editor, *Flow simulation with high-performance computers II*, volume 52 of *Notes on Num. Fluid Mechanics*, pages 431–446. Vieweg, 1996.

[34] G. Zhou and R. Rannacher. Mesh orientation and anisotropic refinement in the streamline diffusion method. Preprint 93-57, Universität Heidelberg, IWR, SFB 359, 1993.

[35] O. C. Zienkiewicz and J. Wu. Automatic directional refinement in adaptive analysis of compressible flows. *Int. J. Num. Meth. Eng.*, 37:2189–2210, 1994.

[36] O. C. Zienkiewicz and J. Z. Zhu. A simple error estimator and adaptive procedure for practical engineering analysis. *Int. J. Num. Meth. Eng.*, 24:337–357, 1987.

[37] O. C. Zienkiewicz and J. Z. Zhu. The superconvergent patch recovery (SPR) and adaptive finite element refinement. *Comput. Methods Appl. Mech. Eng.*, 101(1-3):207–224, 1992.

Other titles in the SFB393 series:

96-01 V. Mehrmann, H. Xu. Chosing poles so that the single-input pole placement problem is well-conditioned. Januar 1996.

96-02 T. Penzl. Numerical solution of generalized Lyapunov equations. January 1996.

96-03 M. Scherzer, A. Meyer. Zur Berechnung von Spannungs- und Deformationsfeldern an Interface-Ecken im nichtlinearen Deformationsbereich auf Parallelrechnern. March 1996.

96-04 Th. Frank, E. Wassen. Parallel solution algorithms for Lagrangian simulation of disperse multiphase flows. Proc. of 2nd Int. Symposium on Numerical Methods for Multiphase Flows, ASME Fluids Engineering Division Summer Meeting, July 7-11, 1996, San Diego, CA, USA. June 1996.

96-05 P. Benner, V. Mehrmann, H. Xu. A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils. April 1996.

96-06 P. Benner, R. Byers, E. Barth. HAMEV and SQRED: Fortran 77 Subroutines for Computing the Eigenvalues of Hamiltonian Matrices Using Van Loans's Square Reduced Method. May 1996.

96-07 W. Rehm (Ed.). Portierbare numerische Simulation auf parallelen Architekturen. April 1996.

96-08 J. Weickert. Navier-Stokes equations as a differential-algebraic system. August 1996.

96-09 R. Byers, C. He, V. Mehrmann. Where is the nearest non-regular pencil? August 1996.

96-10 Th. Apel. A note on anisotropic interpolation error estimates for isoparametric quadrilateral finite elements. November 1996.

96-11 Th. Apel, G. Lube. Anisotropic mesh refinement for singularly perturbed reaction diffusion problems. November 1996.

96-12 B. Heise, M. Jung. Scalability, efficiency, and robustness of parallel multilevel solvers for nonlinear equations. September 1996.

96-13 F. Milde, R. A. Römer, M. Schreiber. Multifractal analysis of the metal-insulator transition in anisotropic systems. October 1996.

96-14 R. Schneider, P. L. Levin, M. Spasojević. Multiscale compression of BEM equations for electrostatic systems. October 1996.

96-15 M. Spasojević, R. Schneider, P. L. Levin. On the creation of sparse Boundary Element matrices for two dimensional electrostatics problems using the orthogonal Haar wavelet. October 1996.

96-16 S. Dahlke, W. Dahmen, R. Hochmuth, R. Schneider. Stable multiscale bases and local error estimation for elliptic problems. October 1996.

96-17 B. H. Kleemann, A. Rathsfeld, R. Schneider. Multiscale methods for Boundary Integral Equations and their application to boundary value problems in scattering theory and geodesy. October 1996.

96-18 U. Reichel. Partitionierung von Finite-Elemente-Netzen. November 1996.

96-19 W. Dahmen, R. Schneider. Composite wavelet bases for operator equations. November 1996.

96-20 R. A. Römer, M. Schreiber. No enhancement of the localization length for two interacting particles in a random potential. December 1996. to appear in: Phys. Rev. Lett., March 1997

96-21 G. Windisch. Two-point boundary value problems with piecewise constant coefficients: weak solution and exact discretization. December 1996.

96-22 M. Jung, S. V. Nepomnyaschikh. Variable preconditioning procedures for elliptic problems. December 1996.

97-01 P. Benner, V. Mehrmann, H. Xu. A new method for computing the stable invariant subspace of a real Hamiltonian matrix or Breaking Van Loan's curse? January 1997.

97-02 B. Benhammouda. Rank-revealing 'top-down' ULV factorizations. January 1997.

97-03 U. Schrader. Convergence of Asynchronous Jacobi-Newton-Iterations. January 1997.

97-04 U.-J. Görke, R. Kreißig. Einflußfaktoren bei der Identifikation von Materialparametern elastisch-plastischer Deformationsgesetze aus inhomogenen Verschiebungsfeldern. March 1997.

97-05 U. Groh. FEM auf irregulären hierarchischen Dreiecksnetzen. March 1997.

97-06 Th. Apel. Interpolation of non-smooth functions on anisotropic finite element meshes. March 1997

97-07 Th. Apel, S. Nicaise. The finite element method with anisotropic mesh grading for elliptic problems in domains with corners and edges.

97-08 L. Grabowsky, Th. Ermer, J. Werner. Nutzung von MPI für parallele FEM-Systeme. March 1997.

97-09 T. Wappler, Th. Vojta, M. Schreiber. Monte-Carlo simulations of the dynamical behavior of the Coulomb glass. March 1997.

97-10 M. Pester. Behandlung gekrümmter Oberflächen in einem 3D-FEM-Programm für Parallelrechner. April 1997.

97-11 G. Globisch, S. V. Nepomnyaschikh. The hierarchical preconditioning having unstructured grids. April 1997.

97-12 R. V. Pai, A. Punnoose, R. A. Römer. The Mott-Anderson transition in the disordered one-dimensional Hubbard model. April 1997.

97-13 M. Thess. Parallel Multilevel Preconditioners for Problems of Thin Smooth Shells. May 1997.

97-14 A. Eilmes, R. A. Römer, M. Schreiber. The two-dimensional Anderson model of localization with random hopping. June 1997.

The complete list of current and former preprints is available via
`http://www.tu-chemnitz.de/sfb393/sfb97pr.html`.