Advances in Numerical Mathematics
Hrsg./Eds.: Hans Georg Bock, Wolfgang Hackbusch,
Mitchell Luskin, Rolf Rannacher

Guido Kanschat

Discontinuous Galerkin Methods for Viscous Incompressible Flow

TEUBNER RESEARCH

Guido Kanschat

Discontinuous Galerkin Methods for Viscous Incompressible Flow

TEUBNER RESEARCH

Advances in Numerical Mathematics

Guido Kanschat

Discontinuous Galerkin Methods for Viscous Incompressible Flow

TEUBNER RESEARCH

Bibliographic information published by Die Deutsche Nationalbibliothek
Die Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.d-nb.de>.

Habilitationsschrift Universität Heidelberg, 2004

1st Edition November 2007

Readers: Ute Wrasmann / Anita Wilke

Deutscher Universitäts-Verlag and Teubner Verlag are companies of
Springer Science+Business Media.
www.duv.de
www.teubner.de

Cover design: Regine Zimmer, Dipl.-Designerin, Frankfurt/Main
Printed on acid-free paper
Printed in Germany

ISBN 978-3-8350-4001-4

Preface

Discontinuous Galerkin (DG) methods are an alternative to stabilized continuous finite element methods for advection-dominated flow problems. They are well suited in particular for the construction of stable discretizations of nonlinear hyperbolic operators, where they are a generalization of finite volume methods to higher order schemes. On the other hand, the discretization of second order elliptic operators by DG methods is not straight-forward, since the trial function spaces are not contained in $W^{1,2}$ anymore. Several options have been developed, which can be split roughly into two groups: first, the schemes augmenting the elliptic operator by penalizing the discontinuities of the shape functions (primal schemes). On the other hand, the second order operator can be converted into a system of first order equations. For these, numerical fluxes can be devised much in the fashion of hyperbolic equations to obtain a stable discretization (mixed methods).

Among the primal schemes, the symmetric interior penalty method by Arnold, based on a method by Nitsche for handling boundary conditions in weak form, is handled in detail in Chapter 2. After quoting the standard energy and L^2 results, where the bounds and influence of the stabilization parameter on the scheme are studied in detail, a section is devoted to the recent point-wise and a posteriori error analysis presented by Kanschat and Rannacher in [KR02]. A short comparison with continuous methods, a question still neglected in the literature, finishes this chapter.

Chapter 3 is devoted to the local discontinuous Galerkin (LDG) method. First, the analysis of the so called *standard* LDG method is quoted from the literature. Then, the main section of this chapter is concerned with the so called *superconvergent* LDG method from [CKPS01]. After presenting the convergence analysis in energy and L^2 norms, we show new numerical evidence for possible generalizations of this method in two and three dimensions.

The extension of the LDG method to Stokes equations is discussed and analyzed in Chapter 4, using results from [CKSS02] and [CKS03a]. The addition of advection fields and reaction terms is the topic of Chapter 5. Convergence analysis for advection-diffusion-reaction (in part from [GK03c]) and Oseen equations is presented, the latter from [CKS03b]. Emphasis is on the robustness of the methods close to the singularly perturbed limit. A reaction-diffusion example shows the superiority of the DG scheme compared to the standard continuous method on coarse meshes. We also show examples for Oseen equations exhibiting the robustness of the discretization with respect to the Reynolds number. Closing this chapter, we show results for incompressible Navier-Stokes equations. Since generalization from Oseen equations to the nonlinear problem is not straight-forward, two techniques for overcoming the difficulties are presented, namely the divergence correction scheme from [GRW02] and a method using $H^{\mathrm{div}}(\Omega)$-conforming elements from [CKS05] and [CKS07]

In the 6th chapter, we show that an efficient solution of all the discretizations presented earlier is possible. Up to the results presented here, no optimal solvers for DG methods for elliptic problems were available. Preconditioners are analyzed and tested with respect to their feasibility in Krylov-space iterations. First, the multilevel scheme for the interior penalty method from [GK03c] and [GK03b] is discussed. New data structures for the efficient implementation on locally refined grids are described in detail. Then, it is shown how these schemes can be used for preconditioning of the LDG method (from [GK03a], [Kan03b] and [Kan03a]). Furthermore, their robustness in the advection dominated case is demonstrated in the case of straight and closed streamlines. The chapter closes with new results for incompressible problems. It is shown that a general block preconditioner concept can be used in conjunction with the preconditioners of the earlier sections to obtain very efficient preconditioners for Stokes discretizations. Investigating dependence on the stabilization parameters of the discretization is crucial there. Finally, it is shown that a modification of these preconditioners can be used efficiently for the discretization of Oseen equations in Section 5.2. Good convergence rates are obtained for a wide range of Reynolds numbers, allowing to address stationary Navier-Stokes equations in a next step. Still, the results are not completely satisfying and ongoing research is devoted to the problem of preconditioning linear systems arising during the solution of stationary, incompressible flow problems.

This volume is an extended version of the habilitation thesis of the author handed in in November 2003. It was augmented by more recent results on Navier-Stokes equations and on solvers on locally refined meshes. The computational examples were implemented using the deal.II finite element library [BHK07] developed by the author, W. Bangerth and R. Hartmann; the underlying techniques have become integrated part of this library. Furthermore, the author wants to thank the following persons for the collaborative research that led to this volume and for their support: B. Cockburn, J. Gopalakrishnan, I. Perugia, R. Rannacher, Ch. Schwab and D. Schötzau.

Dr. Guido Kanschat

Contents

List of Figures

List of Tables

Chapter 1

Basics

This chapter summarizes definitions and results from general finite element theory for further reference. It sets out from definition of meshes and finite element function spaces. Then, it summarizes results on Sobolev functions and their approximation by polynomials on grids. Section 1.3 presents an abstract finite element analysis so that proofs in later chapters can refer to the specific steps without repeating them several times. Finally, the discontinuous Galerkin method for pure advection problems is presented.

1.1 Meshes and shape functions

1.1.1 Let Ω be a domain in $\mathbb{R}^d$, $d = 2, 3$. We denote the boundary of Ω by $\Gamma = \partial\Omega$. In each point $x \bullet\!\!\bullet \partial\Omega$, $\mathbf{n} = \mathbf{n}(x)$ denotes the *outward unit normal vector*. For elliptic problems, we distinguish between the Dirichlet boundary Γ_D and the Neumann boundary Γ_N. The Neumann boundary is an open subset of $\partial\Omega$ in the topology inherited by $\partial\Omega$ from $\mathbb{R}^d$ and $\partial\Omega = \Gamma_D \bullet\!\!\blacksquare \Gamma_N$.

1.1.2 Let the domain Ω be subdivided into a *triangulation*[*] $\mathbb{T}$, consisting of closed triangular or quadrilateral and tetrahedral or hexahedral grid cells[†] T in two and three space dimensions, respectively. Quadrilaterals and hexahedra will be referred to as *tensor product cells* and will always be assumed convex. This assumption is no restriction since it is necessary in order to have a uniformly positive determinant of the mapping from the. Triangles and tetrahedra are refered to as *simplicial cells*.

1.1.3 Definition: The *reference cell* T_0 is either the unit simplex spanned by the origin and the d coordinate vectors (simplicial cells), or the unit hyper-cube $[\bullet\!\!\blacksquare 1, 1]^d$.

Grid cells T of a triangulation $\mathbb{T}$ are images of the reference cell subject to a C^∞ mapping Ψ_T. This mapping will be required to be conforming in the sense that the faces of neighboring cells are on top of ech other and do neither produce overlap nor gaps.

[*] Following common practice, we will commit the etymological crime to use the term triangulation synonymously for subdivision even if the mesh cells are not triangles.

[†] These names are to be understood in a topological sense, since we will allow for curved boundaries of the cells

1.1.4 Assumption: *For the analysis of the methods presented here, we will consider only domains, where* $\bar{\Omega} = \bigcup T$. *Furthermore, we will assume that each edge or face of a mesh cell which is subset of the boundary* $\partial\Omega$ *is either a subset of* Γ_D *or of* Γ_N. *This assumption is purely technical and proofs can be extended to the general case by generic arguments, provided that the mapping* Ψ_T *reproduces the boundary sufficiently accurately.*

1.1.5 We will consider sequences of triangulations generated by successive (possibly local) refinement of the *coarse grid* $\mathbb{T}_0$,

$$\mathbb{T}_0 \bullet\!\!\to \ldots \bullet\!\!\to \mathbb{T}_\ell \bullet\!\!\to \ldots \bullet\!\!\to \mathbb{T}_L,$$

where we denote possibly by the *refinement operator* "$\bullet\!\!\to$"; Global refinement, i.e. refinement of all mesh cells is denoted by "$\sqsubset$".

Here, refinement of a cell means dividing the reference cell into 2^d similar children and mapping those with suitable mappings Ψ_T each.

1.1.6 With each triangulation $\mathbb{T}_\ell$, we associate the mesh size function $h \bullet\!\!\to L^\infty(\Omega)$, such that $h(x) = h_T = \operatorname{diam}(T)$ if $x \bullet\!\!\to T$ with $T \bullet\!\!\to \mathbb{T}_\ell$. We will denote by h_ℓ the maximum of $h(x)$ for a triangulation $\mathbb{T}_\ell$. If we want to stress the mesh size dependence of the triangulation, we write $\mathbb{T}_{h_\ell}$ instead of $\mathbb{T}_\ell$. We write $\mathbb{T}_h$ instead of $\mathbb{T}_{h_\ell}$ if no confusion between grid levels can arise.

A grid cell T in this hierarchy belongs to exactly one level ℓ, which is the number of refinement steps necessary to create it from one of the coarse grid cells.

1.1.7 Definition: By $\mathbb{E}$ ($\mathbb{E}_\ell$, $\mathbb{E}_h$) we denote the set of closed *faces* (edges and surfaces) E of the grid cells in two and three dimensions; a face is the $d \bullet\!\!\to 1$-dimensional intersection of two mesh cells. We do not require that the edges of cells mach exactly, such that a face may be the a subset of such an edge The set $\mathbb{E}$ is subdivided into the subsets of edges on the Dirichlet and Neumann boundaries and the subset of interior edges,

$$\begin{aligned}
\mathbb{E}_h^\Gamma &= \left\{ E \bullet\!\!\to \mathbb{E}_h \middle| E \bullet\!\!\to \Gamma \right\} \\
\mathbb{E}_h^i &= \mathbb{E}_h \bullet\!\!\to (\mathbb{E}_h^D \bullet\!\!\to \mathbb{E}_h^\Gamma) \\
\mathbb{E}_h^D &= \left\{ E \bullet\!\!\to \mathbb{E}_h \middle| E \bullet\!\!\to \Gamma^D \right\} \\
\mathbb{E}_h^N &= \left\{ E \bullet\!\!\to \mathbb{E}_h \middle| E \bullet\!\!\to \Gamma^N \right\}.
\end{aligned}$$

In the case of local refinement, an edge E of a cell T may not be in the set $\mathbb{E}_h$. Then, E is the union of 2^{d-1} elements of $\mathbb{E}_h$, being the edges of the children of the refined neighbor (see Figure 1.1). The use of the word edge here and below will always imply the word surface in three space dimensions.

1.1.8 On the reference cell, we define the polynomial spaces

$$\begin{aligned}
\mathbb{P}_k &= \left\{ p(x) \right\} \\
\mathbb{Q}_k &= \left\{ \prod_{i=1}^d p_i(x_i) \right\},
\end{aligned}$$

where p and p_i are polynomials of degree at most k. We will refer to $\mathbb{P}_k$ as the *complete polynomial space* of degree k and to $\mathbb{Q}_k$ as the *tensor product polynomials* of degree k. If the dimension of the domain is important, we will write $\mathbb{P}_k^d$ and $\mathbb{Q}_k^d$ instead of $\mathbb{P}_k$ and $\mathbb{Q}_k$, respectively.

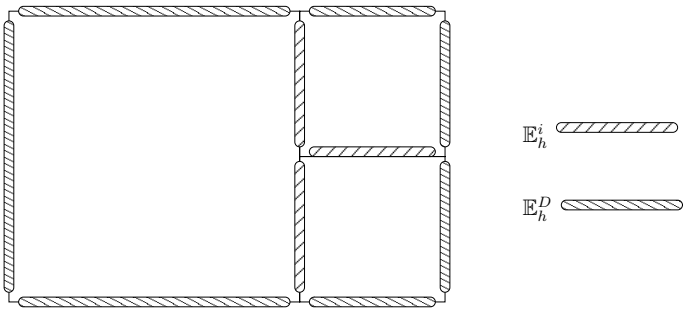

Figure 1.1: The set $\mathbb{E}_h^i$ and $\mathbb{E}_h^D$ (if $\Gamma_D = \Gamma$).

1.1.9 We choose a convenient basis for $\mathbb{P}_k$ and $\mathbb{Q}_k$, respectively. This may either be Legendre polynomials for both spaces or Lagrange polynomials for a grid of support points (usually equidistant) for $\mathbb{Q}_k$. Although Legendre polynomials seem to be more convenient due to their orthogonality properties, this advantage may be lost on general quadrilaterals due to the mapping defined below. The basis functions are called φ_i, where $i = 1, \ldots, n_k$ and n_k is the dimension of $\mathbb{P}_k$ and $\mathbb{Q}_k$, respectively.

1.1.10 The mapping Ψ_T will be either in $(\mathbb{P}_k)^d$ or in $(\mathbb{Q}_k)^d$ for simplicial and tensor product cells, respectively. Remark that the surfaces of hexahedral grid cells are not necessarily plane quadrilaterals; they are bilinear mappings of those.

1.1.11 Definition: The *shape functions* on the grid cell T are the basis functions on the reference cell T_0 defined in Section 1.1.9 subject to the mapping Ψ_T, that is, the ith shape function of mesh cell T is

$$\varphi_{T,i}(x) = \varphi_i(\Psi_T^{-1}(x)). \tag{1.1}$$

The *shape function space* of the cell T is the space spanned by all functions $\bullet\,\varphi_{T;i}\,\bullet_{i=1\ldots n_k}$ and we denote it by $\mathbb{P}_k(T)$ and $\mathbb{Q}_k(T)$, respectively. These are also called *mapped polynomial space*s, but the functions are not polynomials due to the mapping Ψ_T^{-1}. With n_T, we denote the number of shape functions on cell T[‡].

We remark that the spaces $\mathbb{P}_k(T)$ have optimal approximation properties only if Ψ_T is an affine mapping on all cells T (see [ABF02]).

1.1.12 Finite element spaces on $\mathbb{T}_h$ are generated by juxtaposition of the shape functions of all mesh cells:

$$X_{h,k}^P = \bigoplus_{T\in\mathbb{T}_h} \mathbb{P}_k(T)$$
$$X_{h,k}^Q = \bigoplus_{T\in\mathbb{T}_h} \mathbb{Q}_k(T)$$

[‡]By choosing the double subscript $T; i$ for basis functions, the application to hp-methods becomes evident, since numbering on different cells is completely independent

The generic symbol $X_{h,k}$ will be used for both of these spaces. Whenever we need a basis for these spaces, we will use the basis functions in (1.1) with double indices (T, i); this notation is very convenient for discontinuous Galerkin methods.

1.1.13 Assumption: *We will often assume that the triangulation $\mathbb{T}$ is* shape regular*, that is, for each cell T, there is a mapping Ψ_T, such that the minimal and maximal eigenvalues of $\bullet\ \Psi_T$ admit the estimate*

$$\frac{\max_{x\in T_0} \lambda_{\max}(\bullet\ \Psi_T(x))}{\min_{x\in T_0} \lambda_{\min}(\bullet\ \Psi_T(x))} \leqslant C_s, \tag{1.2}$$

with a constant C_s independent of the grid cell T. In the same way, a family of triangulation $\bullet\mathbb{T}_\ell\bullet$ is shape regular if C is independent of ℓ, too.

1.1.14 Assumption: *Usually, we will assume that the mesh size variation of the triangulation $\mathbb{T}_h$ is bounded, that is, there is a constant $c_{\mathbb{T}} > 1$ such that for any two cells T and T' holds:*

$$\frac{h'_T}{h_T} \leqslant c_{\mathbb{T}}\big(1 + \operatorname{dist}(T, T')\big). \tag{1.3}$$

This implies in particular that for two adjacent cells

$$\frac{h'_T}{h_T} \leqslant c_{\mathbb{T}}. \tag{1.4}$$

1.1.15 Remark: A common type of local mesh refinement allows that the refinement level of two neighboring cells may differ by one (*one-irregular mesh*). If such a mesh is generated from a single parallelogram cell, this constant is $c_{\mathbb{T}} = 2^d$, where d is the dimension of the domain. If the coarse mesh consists of more than one cell, $c_{\mathbb{T}} \leqslant 2^j c_{\mathbb{T}_0}$, where $c_{\mathbb{T}_0}$ is the corresponding value for the coarse mesh and j is the maximum difference of refinement levels achievable at a single point of the coarse grid. Similar relations can be derived for cells with general tensor product shape.

1.1.16 Definition: By $\breve{T}$, we denote the union of all mesh cells sharing at least part of an edge, that is

$$\breve{T} = \bigcup_{T\cap S\in\mathbb{E}_h} S. \tag{1.5}$$

This union explicitly includes T itself. In Figure 1.2, all the set $\breve{T}$ of the dark grey cell T is the whole grey area.

1.1.17 Lemma: *Let assumptions 1.1.13 and 1.1.14 hold for a mesh $\mathbb{T}_h$. Then, there is a constant μ independent of the grid cell such that the number of neighbors of a cell T does not exceed μ, that is,*

$$\#\{S\in\mathbb{T}_h;\ S\cap T\neq\emptyset,\ S\subset\breve{T}\} \leqslant \mu \quad \forall T\in\mathbb{T}_h. \tag{1.6}$$

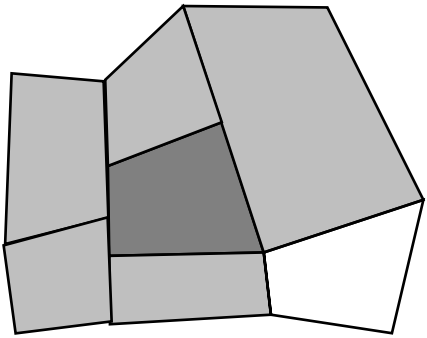

Figure 1.2: The cel T (dark grey) and the neighbor set $\breve{T}$ (grey). Note that a common corner does not result in neighborhood.

1.2 Spaces and approximation

1.2.1 For a set $S \bullet \bullet \mathbb{R}^d$, $L^p(S)$, denotes the space of functions which are p-integrable on the set S. Its norm will be denoted by $\bullet \vdots\vdots \bullet_{L^p(S)}$. If $S = \Omega$, we abbreviate $\bullet \vdots\vdots \bullet_p = \bullet \vdots\vdots \bullet_{L^p(\Omega)}$ and $\bullet \vdots\vdots \bullet = \bullet \vdots\vdots \bullet_2$. Furthermore, the spaces $W^{k,p}(S)$ are the usual Sobolev spaces of functions in $L^p(S)$ with weak derivatives of order up to k in $L^p(S)$ (cf. e. g. [Ada75]). Their norms will be indexed by the space. We sometimes write $H^k(S)$ instead of $W^{k,2}(S)$. The space $W^{-k,q}(S)$ with $1/p + 1/q = 1$ is the dual space of $W^{k,p}(S)$. The norm on $W^{k,p}(S)$ is denoted by $\|.\|_{W^{k,p}(S)} = \|.\|_{k,p;S}$.

Furthermore, we define the space of functions with divergence in $L^2(\Omega)$, namely (cf. for instance [Mon03]),

$$H^{\mathrm{div}}(\Omega) := \bullet \vdots v \bullet \vdots L^2(\Omega; \mathbb{R}^d) : \bullet \ \vdots \nabla \cdot \vdots L^2(\Omega) \bullet \vdots \tag{1.7}$$

1.2.2 Theorem (Sobolev embedding): *A continuous embedding of $W^{k_1,p_1}(\Omega)$ into $W^{k_2,p_2}(\Omega)$ exists for bounded $\Omega \bullet \vdots \mathbb{R}^{d_1}$, if*

$$k_2 \bullet \vdots \frac{d_1}{p_2} \leqslant k_1 \bullet \vdots \frac{d_1}{p_1}. \tag{1.8}$$

In this case, we have for any $u \bullet \vdots W^{k_1,p_1}(\Omega)$:

$$\|u\|_{k_2,p_2;\Omega} \leqslant C \|u\|_{k_1,p_1;\Omega}. \tag{1.9}$$

Let S be a smooth manifold of dimension d_2 in the closure of Ω. Then, a trace of functions $u \bullet \vdots W^{k_1,p_1}(\Omega)$ exists in $W^{k_2,p_2}(S)$ if

$$k_2 \bullet \vdots \frac{d_2}{p_2} \leqslant k_1 \bullet \vdots \frac{d_1}{p_1} \tag{1.10}$$

and (1.9) *holds accordingly with S replacing Ω on the left.*

Proof. This result can be found for instance in [Ada75]. □

1.2.3 Lemma (Friedrichs inequality): *For a function $u \in W^{1,2}(\Omega)$ on Lipschitz bounded domains Ω, the following two estimates hold with constants c depending on Ω:*

$$\|u\|_\Omega^2 \leqslant c^2\Big(\|\nabla u\|_\Omega^2 + \|u\|_{\partial\Omega}^2\Big) \tag{1.11}$$

$$\|u\|_\Omega^2 \leqslant c^2\Big(\|\nabla u\|_\Omega^2 + \Big(\int_{\partial\Omega} u\Big)^2\Big) \tag{1.12}$$

1.2.4 On a domain $S \subset \mathbb{R}^d$, we abbreviate The $L^2(S)$-scalar product

$$\big(u, v\big)_S := \int_S u(x)v(x)\,dx.$$

For a surface F in $\mathbb{R}^d$, we define analogously

$$\big\langle u, v\big\rangle_F := \int_F u(x)v(x)\,ds.$$

On the sets $\mathbb{T}_h$ and E_h (E_h^i and so on analogously), we abbreviate again

$$\big(u, v\big)_{\mathbb{T}_h} := \sum_{T\in\mathbb{T}_h} \big(u, v\big)_T$$
$$\big\langle u, v\big\rangle_{\mathbb{E}_h} := \sum_{E\in\mathbb{E}_h} \big\langle u, v\big\rangle_E.$$

1.2.5 Definition: The *broken Sobolev space* $W^{k,p}(\mathbb{T}_h)$ on the triangulation $\mathbb{T}_h$ and its norm are defined as

$$W^{k,p}(\mathbb{T}_h) = \big\{v \in L^p(\Omega)\big|\, \forall T \in \mathbb{T}_h : v|_T \in W^{k,p}(T)\big\}$$
$$\|v\|_{W^{k,p}(\mathbb{T}_h)} = \left(\sum_{T\in\mathbb{T}_h} \|v\|_{W^{k,p}(T)}^p\right)^{\frac{1}{p}}$$

1.2.6 Definition: Since functions in broken spaces do not have well-defined values on the edges of a triangulation, *numerical flux* operators will be defined there. Their purpose is extracting a unique value from a double valued function. These fluxes will be defined for the particular methods below in terms of mean value and jump operators. On an interior edge between the grid cells K^+ and K^-, let u^+ and u^- be the traces of the function u from the cells K^+ and K^-, respectively. Then, we define for each point x on an interior edge E the *mean value operator*

$$\{u\}(x) := \frac{u^+(x) + u^-(x)}{2}.$$

Let $\mathbf{n}^+$ and $\mathbf{n}^-$ be the outward unit normal vectors of the two cells. Then, for any multiplicative operation $\circ$, we define the *jump operator*

$$[\![\mathbf{n} \circ u]\!] := \mathbf{n}^+ \circ u^+ + \mathbf{n}^- \circ u^-.$$

For instance, for vector valued functions u we have, for the scalar product, the vector product and the Kronecker product, respectively,

$$\begin{aligned}\llbracket \mathbf{n} \bullet u \rrbracket &= \mathbf{n}^+ \bullet u^+ + \mathbf{n}^- \bullet u^-\\ \llbracket \mathbf{n} \bullet u \rrbracket &= \mathbf{n}^+ \bullet u^+ + \mathbf{n}^- \bullet u^-\\ \llbracket \mathbf{n} \bullet u \rrbracket &= \mathbf{n}^+ \bullet u^+ + \mathbf{n}^- \bullet u^-.\end{aligned}$$

Additionally, we define

$$\llbracket u \rrbracket = u^+ \bullet u^-.$$

This jump will only occur squared, such that the product does not depend on the orientation of the face normal.

1.2.7 Definition: Given a vector b in the point $x \in E$ for any edge E not parallel to b, we define the *upwind flux* $u^\uparrow$

$$u^\uparrow(x) = \lim_{\varepsilon \searrow 0} u(x - \varepsilon b). \tag{1.13}$$

The *downwind flux* $u^\downarrow$ is the limit from the opposite side, i.e.,

$$u^\downarrow(x) = \lim_{\varepsilon \searrow 0} u(x + \varepsilon b).$$

If E is parallel to b, we let $u^\uparrow = u^\downarrow = 0$. This definition will be consistent since these fluxes are always multiplied with a term of the form $b \bullet \mathbf{n}_E$.

For a vector field b on Ω, this definition has to be applied pointwise.

1.2.8 Lemma: *The mean value and jump operators defined above admit the following rules*

$$\llbracket uv \rrbracket = \llbracket u \rrbracket \{ v \} + \{ u \} \llbracket v \rrbracket, \tag{1.14}$$

$$\{ uv \} = \{ u \} \{ v \} + \tfrac{1}{4} \llbracket u \rrbracket \llbracket v \rrbracket. \tag{1.15}$$

Furthermore,

$$\{ u \}^2 \leqslant \{ u^2 \}. \tag{1.16}$$

Proof. elementary computation. □

1.2.9 Definition (Broken H^1-norm): on the space $W^1, 2(\mathbb{T}_h)$, we define the norm $\|.\|_{1;h}$ by

$$\|u\|_{1;h}^2 := \|u\|_{W^{1,2}(\mathbb{T}_h)}^2 + \sum_{E \in \mathbb{E}_H^\downarrow} \kappa_E \|\llbracket u \rrbracket\|_E^2 + \sum_{E \in \mathbb{E}_H^D} \kappa_E \|u\|_E^2, \tag{1.17}$$

which is the discontinuous Galerkin equivalent of the $W^{1,2}$-seminorm. In order to obtain optimal estimates, the parameter κ_E is chosen positive and of the order of k^2/h_E. For details on its choice, confer Lemma 2.1.5.

1.2.10 Remark: Friedrichs' inequality holds for this norm as well by applying it cellwise, such that we have

$$\big\|u\big\|_{L^2(\Omega)} \leqslant c\big\|u\big\|_{1;h}. \tag{1.18}$$

Furthermore, we have the Sobolev inequality

$$\big\|u\big\|_{L^4(\Omega)} \leqslant c\big\|u\big\|_{1;h}. \tag{1.19}$$

Both of these inequalities hold because their cellwise versions hold.

1.2.11 The *L^2-projection* $\Pi_h : L^2(\Omega) \bullet \bullet X_{h,k}$ of a function u is defined by the orthogonality relation

$$\big(u \bullet \bullet \Pi_h u, v\big) = 0 \quad \bullet\bullet v \bullet\bullet X_{h,k}.$$

We will denote the *projection error* by $\eta_u = u \bullet \bullet \Pi_h u$.

1.2.12 Lemma: *Let $\Psi : \mathbb{R}^d \bullet \bullet \mathbb{R}^d$ be the scaling of a vector by the factor $h > 0$, that is $\Psi(x) = hx$ and T a bounded domain. Then, for any function $u \bullet \bullet W^{k,p}(T)$*

$$\big|u \bullet\bullet \Psi\big|_{p,k;\Psi T} = h^{d/p-k}\big|u\big|_{p,k;T}. \tag{1.20}$$

Proof. For any multi-index α with $\bullet\bullet\alpha\bullet\bullet = k$ holds

$$\big\|\partial^\alpha u \bullet\bullet \Psi\big\|_{p;\Psi T} = \int_T \big(\partial^\alpha\big(\Psi^{-1}(x)\big)\partial^\alpha u(x)\big)^p \bullet\bullet \det \Psi(x) \bullet\bullet dx = h^{d-kp}\big\|\partial^\alpha u\big\|_{p;T}.$$

□

1.2.13 Lemma: *Assume that $v \bullet\bullet W^{k+1,p}(T)$ and that Assumption 1.1.13 holds. Then, the L^2-projection onto the space $\mathbb{P}_k(T)$ on an arbitrary cell T admits the projection error estimate*

$$\bullet\bullet v \bullet \bullet \Pi_h v \bullet\bullet_{m,q;T} \leqslant c \bullet\bullet T \bullet\bullet^{\frac{1}{q}-\frac{1}{p}} h_T^{k-m+1} \bullet\bullet v \bullet\bullet_{k+1,p;T}. \tag{1.21}$$

Proof. see [Cia78, Theorem 3.1.5] □

1.2.14 Lemma: *For any function $u \bullet \bullet W^{1,p}(T)$ with $1 < p < \bullet$, the trace estimates*

$$\big\|u\big\|_{p;\partial T} \leqslant c\big\|\bullet \ u\big\|_{p;T}^{\frac{1}{p}} \big\|u\big\|_{p;T}^{\frac{1}{q}} \tag{1.22}$$

$$\leqslant c\frac{h_T^{\frac{1}{q}}}{p}\big\|\bullet \ u\big\|_{p;T} + c\frac{h_T^{-\frac{1}{p}}}{q}\big\|u\big\|_{p;T} \tag{1.23}$$

hold with $1/p + 1/q = 1$. In particular, for $u \bullet \bullet W^{1,2}(\mathbb{T})$, we have

$$\big\|u\big\|_{\partial T} \leqslant ch_T^{\frac{1}{2}}\big\|\bullet \ u\big\|_T + h^{-\frac{1}{2}}\big\|u\big\|_T \tag{1.24}$$

Proof. On the reference cell T_0, the estimate

$$\big\|u_0\big\|_{p;\partial T_0} \leqslant c\big\|\nabla u_0\big\|_{p;T_0}^{\frac{1}{p}} \big\|u_0\big\|_{p;T_0}^{\frac{1}{q}},$$

holds (cf. [BS02, Theorem 1.6.6]. Since the terms on both sides scale with the same order of h, estimate (1.22) holds.

The second estimate follows by applying Young inequality $ab \leqslant a^p/p + b^q/q$ and balancing the scaling appropriately. □

1.2.15 Corollary: *Assume that $v \in W^{k+1,p}(\mathbb{T}_h)$ and Assumption 1.1.13 hold. Then, the L^2-projection Π_h admits the estimate*

$$\big|v - \Pi_h v\big|_{m,p;\partial T} \leqslant c h_T^{k-m+\frac{1}{q}} \big|v\big|_{k+1,p;T}. \tag{1.25}$$

Proof. This result is an immediate consequence of estimates (1.21) and (1.22). □

1.2.16 Lemma: *For any function $u_h \in X_{h,k}^P$, the inverse estimate*

$$\big\|\nabla u_h\big\|_T \leqslant c\frac{1}{h_T}\big\|u_h\big\|_T, \tag{1.26}$$

holds.

Proof. The estimate holds on the reference cell by norm equivalence in discrete spaces. The power of h_T follows again by a scaling argument. □

1.2.17 Corollary: *For any function $u_h \in X_{h,k}^P$, the trace estimate*

$$\big\|\partial_n u\big\|_{p;\partial T} \leqslant c h^{-\frac{1}{p}} \big\|\nabla u\big\|_{p;T}. \tag{1.27}$$

holds for $1 < p < \infty$.

Proof. This corollary is an immediate consequence of lemmas 1.2.14 and 1.2.16. □

1.2.18 Lemma: *Let u be a polynomial either in $\mathbb{P}_k$ or in $\mathbb{Q}_k$, such that*

$$\int_{T_0} \nabla u \cdot \tau \, dx = 0 \quad \forall \tau \in (\mathcal{P})^d, \tag{1.28}$$

where $\mathcal{P} \supseteq \mathbb{P}_{k-1}$. If furthermore $u = 0$ on an open subset of at least one of the edges of T_0, then $u = 0$.

Proof. First, if $u \in \mathbb{P}_k$, then for any directional derivative $\partial_i u \in \mathbb{P}_{k-1}$. Therefore, (1.28) implies that u is constant on T_0 and $u \equiv 0$.

If $u \in \mathbb{Q}_k$ and $d \geqslant 2$, the directional derivative $\partial_1 u$ is in $\mathbb{P}^1_{k-1} \times \mathbb{P}^{d-1}_k$. Therefore, using test as functions polynomials in x, y, and z separately in (1.28) implies that $\partial_1 u = \alpha l_k(x_2) \cdots l_k(x_d)$, where l_k is the kth Legendre polynomial on $[-1, 1]$ and α an arbitrary number. Integrating yields

$$u(x_1 \dots x_d) = \alpha x_1 l_k(x_2) \cdots l_k(x_d) + c.$$

Analogous results are obtained for the remaining components, such that u is constant if $k \geqslant 2$ or, if $k = 1$

$$u(x_1 \dots x_d) = \alpha \prod x_i + c.$$

Finally, if this polynomial is zero on an open subset of the boundary of T_0, it is zero everywhere on T_0. □

1.3 Generic finite element analysis

1.3.1 Convergence analysis of finite element discretizations usually follows a well-established path. Since most of the proofs in later chapters follow this path, we present these techniques here in an abstract way. Later we will only refer to the steps of this outline.

1.3.2 Let $a(.,.)$ be a bounded bilinear form on the space V and f a bounded linear functional on V. Then, the generic differential equation subject to discretization will be written in the weak form

$$a(u, v) = f(v) \quad \forall v \in V. \tag{1.29}$$

We assume that this equation is uniquely solvable and we will call u the *continuous solution* or solution to the continuous problem.

1.3.3 Remark: Since we will use this analysis in the context of discontinuous methods, we assume that boundary conditions are imposed weakly in the forms $a(.,.)$ and f; we will not treat them separately in the following paragraphs.

1.3.4 On a sequence of meshes $\{\mathbb{T}_h\}$, we introduce finite dimensional subspaces V_h and the (bi-)linear forms $a_h(.,.)$ and f_h. These forms are assumed to be *consistent* in the following sense: if u is the solution to (1.29), then for any h:

$$a_h(u, v) = f_h(v) \quad \forall v \in V_h \oplus V. \tag{1.30}$$

Furthermore, the *discrete solution* is the function $u_h \in V_h$, such that

$$a_h(u_h, v) = f_h(v) \quad \forall v \in V_h. \tag{1.31}$$

1.3.5 Assumption: *There exists a norm* $\|.\|_A$ *on* V_h *called* energy norm*, such that the* stability estimate

$$\|v\|_A^2 \leqslant c a_h(v,v) \quad \forall v \in V_h, \tag{1.32}$$

holds with a constant c *independent of* h*. If* $a_h(.,.)$ *is a symmetric positive definite bilinear form, then a typical choice is* $\|v\|_A = \sqrt{a_h(v,v)}$*, such that* (1.32) *holds trivially with* $w = v$ *and* $c = 1$*.*

1.3.6 Lemma (Lax-Milgram): *Let* $a(.,.)$ *be a bilinear form on a Hilbert space* V *such that for a constant* c *and* $u, v \in V$ *holds*

$$\begin{aligned} a(u,v) &\leqslant c\|u\|_V\|v\|_V \\ \sup_{v\in V} \frac{a(u,v)}{\|v\|_V} &\geqslant \frac{1}{c}\|u\|_V \\ \sup_{u\in V} \frac{a(u,v)}{\|u\|_V} &\geqslant \frac{1}{c}\|v\|_V, \end{aligned}$$

then, for any linear functional f *on* V *there is a unique solution* u *to*

$$a(u,v) = f(v)$$

and

$$\|u\|_V \leqslant c^2\|f\|_{V'}.$$

This version of the Lax-Milgram lemma can for instance be found in [BF91].

1.3.7 Assumption: *There is a projection* $\Pi_h : V \to V_h$ *admitting the projection error estimate*

$$\|v - \Pi_h v\|_A \leqslant c|h^{s-1}v|_{W^{s,2}(\mathbb{T}_h)}, \tag{1.33}$$

for any $v \in W^{s,2}(\mathbb{T}_h)$*. The projection error of* u *will be abbreviated by* $\eta_u := u - \Pi_h u$*. Typically, with second order elliptic problems, and discretization with polynomials of degree* k *this estimate holds for* $s \leqslant k+1$

1.3.8 Assumption: *If* $u \in W^{s,2}(\mathbb{T}_h)$ *is the solution of* (1.29)*, then the projection error* $\eta_u = u - \Pi_h u$ *and the bilinear form* $a_h(.,.)$ *admit the estimate*

$$a_h(\eta_u, v) \leqslant c|h^{s-1}u|_{W^{s,2}(\mathbb{T}_h)}\|v\|_A. \qquad \forall v \in V_h \tag{1.34}$$

1.3.9 Theorem: *Let* $u \in W^{s,2}(\mathbb{T}_h)$ *be a solution to* (1.30) *and* $u_h \in V_h$ *be a solution to* (1.31) *and* $e_u = u - u_h$*. Furthermore, let assumptions 1.3.5 to 1.3.8 be fulfilled, then*

$$\|e_u\|_A \leqslant c|h^{s-1}u|_{W^{s,2}}. \tag{1.35}$$

Proof. First, we apply triangle inequality,

$$\big\|u-u_h\big\|_A \leqslant \big\|u-\Pi_h u\big\|_A + \big\|\Pi_h u-u_h\big\|_A, \tag{1.36}$$

and estimate the first term by (1.33). For the second, we subtract equations (1.30) and (1.31), yielding the *Galerkin orthogonality* relation

$$a(e_u, v) = 0 \quad \forall v \in V_h. \tag{1.37}$$

Let now $w \in V_h$ with $\big\|w\big\|_A = \big\|\Pi_h u - u_h\big\|_A$ be a function such that the stability estimate (1.32) holds for $v = \Pi_h u - u_h$. Then, Galerkin orthogonality and Assumption 1.3.8 yield

$$\begin{aligned} c\big\|\Pi_h u - u_h\big\|_A^2 &\leqslant a_h(\Pi_h u - u_h, w) \\ &= |a_h(\eta_u, w)| \\ &\leqslant c\big\|\Pi_h u - u_h\big\|_A \big|h^{s-1}u\big|_{W^{s,2}}. \end{aligned} \tag{1.38}$$

Entering this in (1.36) concludes the proof of the theorem. □

1.3.10 Estimates of the L^2-norm and of weaker norms of the error are conducted using the duality argument introduced by Aubin and Nitsche. This argument requires additional assumptions:

1.3.11 Assumption: *Let $t \in \mathbb{N}_0$ and let for arbitrary $f \in W^{t,2}(\Omega)$ the solution z of the dual problem*

$$a(\varphi, z) = \big(f, \varphi\big) \quad \forall \varphi \in V. \tag{1.39}$$

admit the estimate

$$\big|z\big|_{W^{t+2,2};\Omega} \leqslant c\big\|f\big\|_{W^{t,2}(\Omega)}. \tag{1.40}$$

In the case $t = 0$, this assumption is often referred to as elliptic regularity. *In the case of Poisson's equation* (2.1)*, a sufficient condition for* (1.40) *is for instance that $a \in W^{1,\infty}(\Omega)$, $\Gamma_N = \emptyset$ and every corner of the domain is convex. More detailed analysis can be found in [Gri85].*

1.3.12 Assumption: *If $z \in W^{t+2,2}(\mathbb{T}_h)$ is the solution of* (1.29) *and $v \in V_h$, then the bilinear form $a_h(.,.)$ admits the estimate*

$$a_h(v, \eta_z) \leqslant c\big\|v\big\|_A \big|h^{t+1}z\big|_{W^{t+2,2}(\mathbb{T}_h)}. \tag{1.41}$$

1.3.13 Assumption: *In extension of Assumption 1.3.7, the estimate*

$$a_h(u - \Pi_h u, z - \Pi_h z) \leqslant c\big|h^{s-1}u\big|_{W^{s,2}(\mathbb{T}_h)} \big|h^{t+1}u\big|_{W^{t+2,2}(\mathbb{T}_h)}, \tag{1.42}$$

for functions $u \in W^{s,2}(\mathbb{T}_h)$ and $z \in W^{t+2,2}(\mathbb{T}_h)$.

1.3.14 Theorem: *Let Theorem 1.3.9 hold and $t \in \mathbb{N}_0$. If furthermore Assumptions 1.3.11 to 1.3.13 hold, the $W^{-t,2}(\tilde{\Omega})$-norm of the error, where $\tilde{\Omega} \subseteq \Omega$, admits the estimate*

$$\|e_u\|_{W^{-t,2}(\tilde{\Omega})} \leqslant c|h^{s+t}u|_{W^{s,2}(\mathbb{T}_h)}. \tag{1.43}$$

Proof. Let $\tilde{V}$ be the space of functions $\varphi \in W^{t,2}(\Omega)$ with $\varphi(x) = 0$ almost everywhere on $\Omega \setminus \tilde{\Omega}$. By definition of the negative norms, for $t \neq 0$ we have to estimate

$$\|e_u\|_{W^{-t,2}(\tilde{\Omega})} = \sup_{\varphi \in \tilde{V}} \frac{(e_u, \varphi)}{\|\varphi\|_{W^{t,2}(\tilde{\Omega})}}. \tag{1.44}$$

In the case $t = 0$, we take $\varphi = \chi(\tilde{\Omega})e_u$ instead of the supremum. Here $\chi(\tilde{\Omega})$ is the characteristic function of the set $\tilde{\Omega}$:

$$\|e_u\|_{L^2(\tilde{\Omega})} \leqslant \frac{(e_u, \chi(\tilde{\Omega})e_u)}{\|e_u\|_{L^2(\tilde{\Omega})}}. \tag{1.45}$$

Using the dual problem (1.39) and consistency (1.30), we write

$$\begin{aligned}(e_u, \varphi) &= a_h(e_u, z) = a_h(e_u, \eta_z) \\ &= a_h(\Pi_h u - u_h, \eta_z) + a_h(\eta_u, \eta_z).\end{aligned}$$

Due to estimate (1.41), we have

$$a_h(\Pi_h u - u_h, \eta_z) \leqslant c\|\Pi_h u - u_h\|_A |h^{t+1}z|_{W^{t+2,2}(\mathbb{T}_h)},$$

which, together with (1.38) and (1.42) yields

$$(e_u, \varphi) \leqslant c|h^{s-1}u|_{W^{s,2}(\mathbb{T}_h)}|h^{t+1}z|_{W^{t+2,2}(\mathbb{T}_h)}.$$

In virtue of the dual regularity estimate (1.40) we continue

$$(e_u, \varphi) \leqslant |h^{s+t}u|_{W^{s,2}(\mathbb{T}_h)}|\varphi|_{W^{t,2}(\tilde{\Omega})}.$$

Entering this result into (1.44) and (1.45), respectively, yields the result of the theorem □

1.3.15 Remark: The powers of h involved in this section were adapted to second order problems with full regularity. The same mechanism is applicable with different h-dependence. If the powers of h in estimates (1.33) and (1.41) and the gain of regularity in (1.40) are different, it is obvious how to change the results of the theorems to account for the new situation. An example is Section 1.4 on advection problems below.

1.4 Advection Problems

1.4.1 The origin of discontinuous Galerkin methods lies in the discretization of Neutron transport problems investigated by LeSaint and Raviart [LR74]. We will discuss this scheme in a slightly generalized form in the following paragraphs.

1.4.2 We consider the following stationary *linear advection* problem as a prototype for hyperbolic problems: Let $b \in L^\infty(\Omega;\mathbb{R}^d)$ be a vector field with $\nabla \cdot b \in L^\infty(\Omega)$ and $\nabla \cdot b(x) \leqslant 0$ almost everywhere. Then, find u such that

$$\begin{aligned} b\cdot\nabla u &= f && \text{in } \Omega \\ u &= g && \text{on } \Gamma_-, \end{aligned} \tag{1.46}$$

where the *inflow boundary* Γ_- is given by

$$\Gamma_- = \big\{x \in \partial\Omega \;\big|\; \mathbf{n}(x)\cdot b < 0\big\}. \tag{1.47}$$

1.4.3 The discontinuous Galerkin formulation of (1.46) is constructed by multiplying with a test function v and integrating by parts on each cell T of a triangulation $\mathbb{T}_h$, yielding

$$-\big(u, \nabla\cdot(bv)\big)_T + \big\langle u, (b\cdot\mathbf{n})v\big\rangle_{\partial T} = \big(f, v\big)_T. \tag{1.48}$$

This form is well-defined in T, if u and v belong to the space

$$V = \left\{ v \in L^2(\Omega) \middle| \forall T \in \mathbb{T}_h : b\cdot\nabla v|_T \in L^2(T) \right\},$$

1.4.4 Simplifying the notation, we define the form

$$\big\langle u, v\big\rangle_{b;E} := \big\langle u, v\,|b\cdot\mathbf{n}|\big\rangle_E, \tag{1.49}$$

for any edge E of the triangulation and on sets of edges $\mathbb{E}_h$ and so on, accordingly. Here, $\mathbf{n}$ is any of the two normal vectors, since change of orientation does not change the definition. Considering $\nabla\cdot(bu) = u\nabla\cdot b + b\cdot\nabla u$, equation (1.48) becomes

$$-\big(u, b\cdot\nabla v\big)_T - \big(\nabla\cdot bu, v\big)_T - \big\langle u, v\big\rangle_{b;\partial T_-} + \big\langle u, v\big\rangle_{b;\partial T_+} = \big(f, v\big)_T.$$

1.4.5 The traces of the function u are not well-defined on interior edges, since the values from both adjacent cells are not required to coincide. Therefore, the trace of u is replaced by the *numerical flux* $\hat{u}$. The flux chosen here is the upwind flux $\hat{u} = u^\uparrow$ of Definition 1.2.7.

Thus, the discontinuous formulation is (integrating by parts again): find $u \in V$ such that

$$\begin{aligned} \beta_h\big(u, v\big) \equiv \big(b\cdot\nabla u, v\big)_{\mathbb{T}_h} + \big\langle u^\downarrow - u^\uparrow, v^\downarrow\big\rangle_{b;\mathbb{E}_h^i} + \big\langle u^\downarrow, v^\downarrow\big\rangle_{b;\mathbb{E}_h^-} \\ = \big(f, v\big)_\Omega + \big\langle u^D, v^\downarrow\big\rangle_{b;\mathbb{E}_h^-} \quad \forall v \in V, \end{aligned} \tag{1.50}$$

where $\mathbb{E}_h^- = \{E \in \mathbb{E}_h | E \subset \Gamma_-$ and Γ_- is the *inflow boundary*

$$\Gamma_- = \big\{x \in \partial\Omega \big| b\cdot\mathbf{n}(x) < 0\big\}.$$

1.4.6 Remark: Heuristically, the form $\beta_h\big(.,.\big)$ is constructed in such a way that the value of u at the inflow boundary of each cell is weakly prescribed to be equal to the value in the upstream cell.

1.4.7 Lemma: *The form $\beta_h(.,.)$ is positive semi-definite on $W^{1,2}(\mathbb{T}_h)$.*

Proof. Following [LR74], we integrate by parts to obtain

$$\big(b\cdot\nabla u,u\big)_T = -\big(u,b\cdot\nabla u\big)_T - \big(\nabla\cdot b\,u,u\big)_T - \big\langle u,u\big\rangle_{b;\partial T_-} + \big\langle u,u\big\rangle_{b;\partial T_+},$$

yielding

$$2\big(b\cdot\nabla u,u\big)_T = -\big(\nabla\cdot b\,u,u\big)_T + \big\langle u,u\big\rangle_{b;\partial T_+} - \big\langle u,u\big\rangle_{b;\partial T_-}. \tag{1.51}$$

Summing up over all cells T, we get

$$2\big(b\cdot\nabla u,u\big)_{\mathbb{T}_h} = -\big(\nabla\cdot b\,u,u\big)_{\mathbb{T}_h} + \big\langle u^\uparrow,u^\uparrow\big\rangle_{b;\mathbb{E}_h^i\cup\mathbb{E}_h^+} - \big\langle u^\downarrow,u^\downarrow\big\rangle_{b;\mathbb{E}_h^i\cup\mathbb{E}_h^-}.$$

By entering this result into (1.50), we obtain

$$\begin{aligned}2\beta_h\big(u,u\big) &= -\big(\nabla\cdot b\,u,u\big)_{\mathbb{T}_h} + \big\langle u^\uparrow,u^\uparrow\big\rangle_{b;\mathbb{E}_h^i\cup\mathbb{E}_h^+} - \big\langle u^\downarrow,u^\downarrow\big\rangle_{b;\mathbb{E}_h^i\cup\mathbb{E}_h^-}\\ &\qquad + 2\big\langle u^\downarrow-u^\uparrow,v^\downarrow\big\rangle_{b;\mathbb{E}_h^i} + 2\big\langle u^\downarrow,v^\downarrow\big\rangle_{b;\mathbb{E}_h^-}\\ &= -\big(\nabla\cdot b\,u,u\big)_{\mathbb{T}_h} + \big\langle u^\uparrow,u^\uparrow\big\rangle_{b;\mathbb{E}_h^+} + \big\langle u^\downarrow,u^\downarrow\big\rangle_{b;\mathbb{E}_h^-} + \big\langle[\![u]\!],[\![u]\!]\big\rangle_{b;\mathbb{E}_h^i} \geqslant 0\end{aligned} \tag{1.52}$$

□

1.4.8 Definition: The energy norm for the form $\beta_h(.,.)$ is

$$\big\|u\big\|_\beta := \sqrt{\beta_h\big(u,u\big) + \|h^{1/2}b\cdot\nabla u\|^2_{\mathbb{T}_h}}. \tag{1.53}$$

This term is a seminorm due to the preceding lemma and becomes definite on the space V by the additional term. Due to (1.52), we have

$$\big\|u\big\|^2_\beta = \frac{1}{2}\big\langle u^\uparrow,u^\uparrow\big\rangle_{b;\mathbb{E}_h^+} + \frac{1}{2}\big\langle u^\downarrow,u^\downarrow\big\rangle_{b;\mathbb{E}_h^-} + \frac{1}{2}\big\langle[\![u]\!],[\![u]\!]\big\rangle_{b;\mathbb{E}_h^i} + \|h^{1/2}b\cdot\nabla u\|^2_{\mathbb{T}_h}$$

1.4.9 The discontinuous Galerkin discretization of (1.46) is achieved finally by replacing the spaces V by a finite element subspace $V_h = X_{h,k}$.

1.4.10 Remark: The following stability result is a special case of a result in [GK03c] which will be generalized to advection-diffusion-reaction problems in Lemma 5.1.5 on page 108. A more general result including L^p-norms ($1 \leqslant p < \infty$) was obtained earlier in [JP86] by a technique using a global exponential function decaying in direction b.

1.4.11 Lemma: *If the vector field β is constant on each mesh cell, there exists constants c and γ independent of the mesh size such that for any $u_h \in V_h$ the following stability estimate holds:*

$$\big\|u_h\big\|_\beta^2 \leqslant c\beta_h\big(u_h, u_h + \gamma hb\cdot\nabla u_h\big). \tag{1.54}$$

Proof. We omit the index h of u_h for simplicity and start with

$$\begin{aligned}\beta_h\big(u, u+\gamma hb\cdot\nabla u\big) = \beta_h\big(u,u\big) &+ \gamma\big\|h^{1/2}b\cdot\nabla u\big\|_{\mathbb{T}_h}^2 \\ &+ \gamma\big\langle u^\uparrow - u^\downarrow, h^\uparrow b\cdot\nabla u^\uparrow\big\rangle_{b;\mathbb{E}_h^i} + \gamma\big\langle u^\uparrow, h^\uparrow b\cdot\nabla u^\uparrow\big\rangle_{b;\mathbb{E}_h^-}\end{aligned} \tag{1.55}$$

Applying Cauchy-Schwarz and Young inequalities to the boundary terms, and the trace estimate (1.27) to $\big\|b\cdot\nabla u^\uparrow\big\|_E$, we obtain

$$\begin{aligned}\beta_h\big(u, u+\gamma hb\cdot\nabla u\big) \geqslant \beta_h\big(u,u\big) + \gamma\big\|h^{1/2}b\cdot\nabla u\big\|_{\mathbb{T}_h}^2 &- c\delta\gamma\big\|h^{1/2}b\cdot\nabla u\big\|_{\mathbb{T}_h}^2 \\ &- \frac{\gamma}{4\delta}\Big(\big\langle[\![u]\!],[\![u]\!]\big\rangle_{b;\mathbb{E}_h^i} + \big\langle u,u\big\rangle_{b;\mathbb{E}_h^-}\Big).\end{aligned}$$

We can now set $\delta = 1/2c$ and use (1.52) to choose γ to obtain (1.54) with c independent of h. □

1.4.12 Remark: This lemma can be extended to smooth vector fields β by using a suitable projection in the definition of the test function in (1.55); still, the nature of this extension is not completely clear, since it must break down in the case that β has closed integral curves (see the remark at the end of this chapter).

1.4.13 Theorem: *Let $u \in W^{s,2}(\Omega)$ be the solution to* (1.46) *with $3/2 \leqslant s \leqslant k+1$ and $u_h \in V_h = X_{h,k}$ be the finite element solution to* (1.50)*. Then, the error $e_u = u - u_h$ admits the estimate*

$$\big\|e_u\big\|_\beta \leqslant \big|h^{s-\frac{1}{2}}u\big|_{s,2;\mathbb{T}_h}. \tag{1.56}$$

Proof. The proofs follows the generic Theorem 1.3.9. Stability Assumption 1.3.5 was proven in Lemma 1.4.11. Therefore, we begin by proving approximation Assumption 1.3.7, in this case with the approximation order $h^{s-1/2}$. From the projection estimate (1.21), we infer for the projection error $\eta_u = u - \Pi_h u$ that

$$\big\|\eta_u\big\| + \big\|h\nabla\eta_u\big\| \leqslant c\big|h^s u\big|_{s,2;\mathbb{T}_h}.$$

Applying trace inequality (1.24), we deduce

$$\big\|\eta_u\big\|_\beta \leqslant \big|h^{s-\frac{1}{2}}u\big|_{s,2;\mathbb{T}_h}. \tag{1.57}$$

Reverting to the error e_u, we use triangle inequality to obtain

$$\big\|e_u\big\|_\beta \leqslant \big\|\eta_u\big\|_\beta + \big\|\Pi_h e_u\big\|_\beta.$$

The first term is already covered by the previous approximation result. Since $\Pi_h e_u \in V_h$, we use Lemma 1.4.11 and Galerkin orthogonality, yielding

$$\begin{aligned}\big\|\Pi_h e_u\big\|_\beta^2 &\leqslant \beta_h\big(\Pi_h e_u, \Pi_h e_u + \gamma h b\cdot\nabla(\Pi_h e_u)\big)\\ &= \beta_h\big(\eta_u, \Pi_h e_u + \gamma h b\cdot\nabla(\Pi_h e_u)\big).\end{aligned}$$

We estimate this term using the following argument: using integration by parts yields

$$\begin{aligned}\beta_h\big(v,w\big) = \big(v, -b\cdot\nabla w\big)_{\mathbb{T}_h} - \big(\nabla\cdot b\, v, w\big)_{\mathbb{T}_h}\\ + \big\langle v^\uparrow, w^\uparrow - w^\downarrow\big\rangle_{b;\mathbb{E}_h^i} + \big\langle v^\uparrow, w^\uparrow\big\rangle_{b;\mathbb{E}_h^+}.\end{aligned} \tag{1.58}$$

Therefore, abbreviating $\xi = \Pi_h e_u$,

$$\begin{aligned}\beta_h\big(\eta_u, \xi + \gamma h b\cdot\nabla\xi\big) &\leqslant \big(1+\|b\|_\infty\big)\big\|h^{-\frac12}\eta_u\big\|_{\mathbb{T}_h}\big\|h^{\frac12}\xi\big\|_{\mathbb{T}_h} + \big\|\eta_u\big\|_{\mathbb{E}_h^i}\big\langle[\![\xi]\!],[\![\xi]\!]\big\rangle_{b;\mathbb{E}_h^i}\\ &\quad + \big\|\eta_u\big\|_{\mathbb{E}_h^+}\big\langle\xi^\uparrow,\xi^\uparrow\big\rangle_{b;\mathbb{E}_h^+} + \gamma\big\|h^{\frac12}b\cdot\nabla\eta_u\big\|_{\mathbb{T}_h}\big\|h^{\frac12}b\cdot\nabla\xi\big\|_{\mathbb{T}_h}\\ &\quad + \gamma\big\langle[\![u]\!],[\![u]\!]\big\rangle_{b;\mathbb{E}_h^i}\big\|h^{\downarrow}b\cdot\nabla\xi^{\downarrow}\big\|_{\mathbb{E}_h^i} + \gamma\big\langle[\![u]\!],[\![u]\!]\big\rangle_{b;\mathbb{E}_h^-}\big\|h^{\downarrow}b\cdot\nabla\xi^{\downarrow}\big\|_{\mathbb{E}_h^-}\\ &\leqslant c\Big(\big\|\eta_u\big\|_\beta + \big\|h^{-\frac12}\eta_u\big\|_{\mathbb{T}_h} + \big\|\eta_u\big\|_{\mathbb{E}_h^i\cup\mathbb{E}_h^+}\Big)\big\|\xi\big\|_\beta.\end{aligned}$$

Finally, the result follows by applying trace estimate (1.24), inverse estimate and approximation Lemma 1.2.13 to the first term and dividing by $\big\|\xi\big\|_\beta$. □

1.4.14 Corollary: *Under the assumptions of Theorem 1.4.13, the L^2-norm of the error e_u admits the estimate*

$$\big\|e_u\big\| \leqslant c\big|h^{s-\frac12}u\big|_{s,2;\mathbb{T}_h}. \tag{1.59}$$

1.4.15 Remark: For solutions $u \in W^{k+1,2}(\Omega)$, estimate (1.56) is optimal with respect to the approximation result (1.21), while (1.59) is sub-optimal by $h^{1/2}$. For Cartesian grids we have the following superconvergence result due to LeSaint and Raviart [LR74].

1.4.16 Theorem: *Let the triangulation $\mathbb{T}_h$ consist of rectangles only and let b be constant in Ω. Additionally, $V_h = X_{h,k}^Q$ and the continuous solution $u \in W^{k+2,2}(\Omega)\cap W^{k+1,\infty}(\Omega)$. Then, the error e_u admits the estimate*

$$\big\|e_u\big\| \leqslant Ch^{k+1}\big\|u\big\|_{W^{k+2,2}(\Omega)} \tag{1.60}$$

1.4.17 A solution to (1.46) can be constructed by considering the set of *characteristic curves*, which are the integral curves of the vector field b. Let $X(t)$ be such a curve. Then,

$$b\cdot\nabla u(X(t)) = \frac{d}{dt}u(X(t)).$$

Let t_0 be the lower limit of all t such that $X(t)\in\Omega$ and t_1 the least upper bound of those t. Assume that $X(t)$ is Lipschitz continuous on $[t_0,t_1]$ and $X([t_0,t_1])\cap\Omega$ is simply connected. Then, u can be obtained as the solution of the ordinary differential equation

$$u'(t) = f(X(t)), \tag{1.61}$$

with start value $u(t_0) = u^D(X(t_0))$.

1.4.18 Remark: If the set of integral curves of the vector field b contains closed loops, then (1.46) has a solution, if and only if for any such loop

$$\oint f(x(\tau))\, d\tau = 0.$$

This follows immediately by solving (1.61) on such a loop with $X(t_0) = X(T_1)$. Therefore, a stability result cannot be obtained for this case.

1.4.19 Remark: If the vector field b does not have closed integral curves and the mesh is sufficiently fine, the grid cells can be ordered in such a way, that the discrete linear system can be solved cell by cell from inflow to outflow boundary (cf. [JP86]).

Chapter 2

Linear Diffusion I

This chapter treats Poisson equation in its primal form. It sets out from Nitsche's method for weakly imposed Dirichlet boundary conditions (see Nitsche [Nit71]) to define the interior penalty method (see Arnold [Arn82]). The standard energy norm and L^2 error analysis of these methods is quoted from the cited articles. We add a new study of the behavior of the discretization depending on the stabilization parameter.

Then, the L^∞ analysis by Kanschat and Rannacher (cf. [KR02]) is presented, followed by a posteriori error estimates from the same source. Finally, we compare continuous and discontinuous methods.

2.0.1 In this chapter, we will investigate discontinuous Galerkin discretizations of Poisson's equation

$$-\nabla\cdot\big(\nu\nabla u\big) = f \qquad \text{in } \Omega \tag{2.1}$$

$$u = u^D \qquad \text{on } \Gamma^D \tag{2.2}$$

$$\partial_n u = \sigma^N \qquad \text{on } \Gamma^N \tag{2.3}$$

where $\nu \in L^\infty(\Omega)$ with $\nu(x) \geqslant \nu_0 > 0$ almost everywhere in Ω. Γ^D and Γ^N are the Dirichlet and Neumann parts of $\Gamma = \partial\Omega$, respectively. We have $\Gamma^D \cap \Gamma^N = \emptyset$ and $\Gamma = \bar\Gamma^D \cup \bar\Gamma^N$.

2.0.2 The presentation will focus on the generic equation

$$-\Delta u = f, \tag{2.4}$$

under boundary conditions (2.2) and (2.3). Results for the more general problem (2.1) are recovered by standard analytical techniques, usually requiring that the coefficient fulfills $\nu \geqslant \nu_0 > 0$ and is sufficiently smooth.

2.0.3 Theorem: *Equation* (2.4) *has a unique solution* $u \in W^{1,2}(\Omega)$ *with trace* g *on* Γ_D*, provided* $f \in W^{-1,2}(\Omega)$ *(cf.[GT98]).*

If the boundary of the domain is sufficiently regular and $f \in W^{\alpha-1,2}(\Omega)$ *with* $\alpha > 0$*, then* $u \in W^{1+\alpha,2}(\Omega)$ *and (see [Gri85] for details)*

$$\big\|u\big\|_{W^{1+\alpha,2}(\Omega)} \leqslant \big\|f\big\|_{W^{\alpha-1,2}(\Omega)}. \tag{2.5}$$

In particular, if all corners and edges of the domain are convex and $f \in L^2(\Omega)$*, then* $u \in W^{2,2}(\Omega)$ *and* (2.6) *holds with* $\alpha = 1$*, namely*

$$\|u\|_{W^{2,2}(\Omega)} \leqslant \|f\|_{L^2(\Omega)}. \tag{2.6}$$

We will refer to this last estimate as elliptic regularity.

2.1 Weak boundary conditions

2.1.1 In [Nit71], Nitsche proposed a fully conforming method of treating Dirichlet boundary values in weak form*. The weak formulation for Nitsche's method is

$$\begin{aligned} a_h^N(u,v) \equiv (\nabla u, \nabla v)_\Omega + \langle \kappa_h u, v\rangle_{\partial\Omega} - \langle \partial_n u, v\rangle_{\partial\Omega} - \langle u, \partial_n v\rangle_{\partial\Omega} \\ = f_h^N(v) \equiv (f,v)_\Omega + \langle \kappa_h u^D, v\rangle_{\partial\Omega} - \langle u^D, \partial_n v\rangle_{\partial\Omega}, \end{aligned} \tag{2.7}$$

where $\kappa_h > 0$ is a function constant on each edge $E \in \mathbb{E}_h^D$ achieving stability of the form as well as penalizing violation of the boundary condition $u = u^D$. Suitable values for κ will be determined below. The term $-\langle \partial_n u, v\rangle_{\partial\Omega}$ achieves conformity of the method since it eliminates the natural boundary condition. $-\langle u, \partial_n v\rangle_{\partial\Omega}$ was introduced to symmetrize the operator and has to be matched by $-\langle u^D, \partial_n v\rangle_{\partial\Omega}$ on the right hand side to maintain conformity.

Since this method exhibits the essential properties of the interior penalty method below, we will summarize the most important analytical results.

2.1.2 We define the energy norm for (2.7) as

$$\|u\|_N := \sqrt{|u|^2_{1,2;\Omega} + \langle \kappa_h u, u\rangle_{\partial\Omega}}. \tag{2.8}$$

This form is obviously non-negative. The definiteness is established by the following lemma.

2.1.3 Lemma: *Let* $\kappa_h \in L^\infty(\partial\Omega)$ *and* $\kappa_h \geqslant \kappa_0 > 0$ *almost everywhere on* $\partial\Omega$*. Then, for a function* $u \in W^{1,2}(\Omega)$*,* $\|u\|_N = 0$ *implies* $u = 0$*. Therefore,* $\|.\|_N$ *is in fact a norm on* $W^{1,2}(\Omega)$*.*

Proof. First, $\|\nabla u\|_\Omega = 0$ implies that $u \equiv const$ in Ω (cf. e. g. [GT98]). Since $\kappa_h > 0$, $\langle \kappa_h u, u\rangle_{\partial\Omega} = 0$ is positive definite on the space of functions in $L^2(\partial\Omega)$. Since this space includes the traces of functions in $W^{1,2}(\Omega)$, the trace of u must be zero and consequently u itself. □

2.1.4 Discretization of (2.7) is achieved by choosing a mesh $\mathbb{T}_h$ on Ω and restricting test and trial functions to the space $V_h = X^Q_{h,k} \cap C^0(\overline{\Omega})$.

*For simplicity of the presentation, we will assume that $\Gamma_D = \partial\Omega$

2.1.5 Lemma: *Assume that κ_h in* (2.7) *is chosen such that*

$$\kappa_h = \frac{\kappa_N}{h_F}, \qquad \kappa_N > c_*^2, \tag{2.9}$$

where κ_N is independent of h and h_E is the length of the cell T orthogonal to the face E. c_* is the constant in the trace estimate (cf. corollary 1.2.17)

$$\big\|\partial_n p\big\|_{\partial T\cap\partial\Omega} \leqslant c_* h_T^{-\frac{1}{2}} \big\|\nabla p\big\|_T,$$

for all mesh cells T adjacent to the boundary $\partial\Omega$ and for all polynomials $p \in \mathbb{Q}_k(T)$. Then, the stability estimate

$$\big\|u\big\|_N^2 \leqslant c a_h^N\big(u,u\big), \tag{2.10}$$

holds for all $u \in V_h$ with a positive constant c independent of the mesh size and of the actual value of κ_N.

Proof. Application of Young inequality with $0 < \alpha = c_*^2/\kappa_N < 1$ and the trace estimate (1.23) yields

$$\begin{aligned}
a_h^N\big(u,u\big) &= \big\|\nabla u\big\|_\Omega^2 + \big\langle \kappa_h u, u\big\rangle_{\partial\Omega} - 2\big\langle \partial_n u, u\big\rangle_{\partial\Omega} \\
&\geqslant \big\|u\big\|_N^2 - \sum_{T\in\mathbb{T}_h} \Big\{ \frac{\alpha h_T}{c_*^2}\big\|\partial_n u\big\|_{\partial T\cap\partial\Omega}^2 + \frac{c_*^2}{\alpha h_T}\big\|u\big\|_{\partial T\cap\partial\Omega}^2 \Big\} \\
&\geqslant \big\|u\big\|_N^2 - \alpha\big\|\nabla u\big\|^2 - \alpha\big\|\sqrt{\kappa_h} u\big\|_{\partial\Omega}^2 \\
&= (1-\alpha)\,\big\|u\big\|_N^2.
\end{aligned}$$

□

2.1.6 Theorem: *Assume that the solution u of* (2.1) *is in $W^{s,2}(\Omega)$ with $s \geqslant 2$. Let $V_h = X_{h,k}$ with $k \geqslant s-1$. If κ_h is chosen according to* (2.9)*, then the error $e_u = u - u_h$ admits the estimate*

$$\big\|e_u\big\|_N \leqslant c\big|h^{s-1}u\big|_{W^{s,2}(\Omega)}. \tag{2.11}$$

Proof. In view of Theorem 1.3.9, it is sufficient to estimate the error $I_h u - u_h$. By the stability estimate (2.10), we have

$$\begin{aligned}
c\big\|I_h u - u_h\big\|_N^2 &\leqslant a_h^N\big(\eta_u, I_h u - u_h\big) \\
&\leqslant \big\|\nabla\eta_u\big\|\big\|\nabla(I_h u - u_h)\big\| + \big\|\sqrt{\kappa_h}\eta_u\big\|_{\partial\Omega}\big\|\sqrt{\kappa_h}(I_h u - u_h)\big\|_{\partial\Omega} \\
&\qquad + \big\|\partial_n\eta_u\big\|_{\partial\Omega}\big\|I_h u - u_h\big\|_{\partial\Omega} + \big\|\eta_u\big\|_{\partial\Omega}\big\|\partial_n(I_h u - u_h)\big\|_{\partial\Omega} \\
&\leqslant \frac{c}{2}\big\|I_h u - u_h\big\|_N^2 + c\Big(\big\|\nabla\eta_u\big\|^2 + \big\|\sqrt{\kappa_h}\eta_u\big\|_{\partial\Omega}^2 + \big\|\kappa^{-\frac{1}{2}}\partial_n\eta_u\big\|_{\partial\Omega}^2\Big).
\end{aligned}$$

Applying approximation Lemma 1.2.13 (and $\kappa \sim 1/h$) yields

$$\big\|I_h u - u_h\big\|_N \leqslant c\big|h^{s-1}u\big|_{W^{s,2}(\Omega)}.$$

□

2.1.7 Remark: Obviously, the analysis above extends to the situation where Dirichlet boundary conditions are imposed on an open subset $\Gamma_D \subset \partial\Omega$ only.

2.2 The Interior Penalty Method

2.2.1 Applying Nitsche's method of weakly imposed boundary conditions to each grid cell and and averaging the test functions over neighboring cells yields the *interior penalty* method (IP):

$$\begin{aligned} a_h^{\text{IP}}(u,v) \equiv \big(\nabla u, \nabla v\big)_{\mathbb{T}_h} + \big\langle \kappa_h [\![u]\!], [\![v]\!] \big\rangle_{\mathbb{E}_h^i} + \big\langle 2\kappa_h u, v \big\rangle_{\mathbb{E}_h^D} \\ - \big\langle \{\!\{\nabla u\}\!\}, [\![v\mathbf{n}]\!] \big\rangle_{\mathbb{E}_h^i} - \big\langle [\![u\mathbf{n}]\!], \{\!\{\nabla v\}\!\} \big\rangle_{\mathbb{E}_h^i} - \big\langle \partial_n u, v \big\rangle_{\mathbb{E}_h^D} - \big\langle u, \partial_n v \big\rangle_{\mathbb{E}_h^D} \\ = \big(f, v\big)_\Omega + \big\langle 2\kappa_h u^D, v \big\rangle_{\mathbb{E}_h^D} - \big\langle u^D, \partial_n v \big\rangle_{\mathbb{E}_h^D}. \end{aligned} \tag{2.12}$$

2.2.2 Remark: The form $a_h^{\text{IP}}(.,.)$ defined in (2.12) differs from the definition found in the majority of publications (e.g. [Arn82]) in the stabilization parameter on the boundary. The analysis below shows, that the stabilization in our version is indeed more equilibrated (see also [HL02]). See paragraph 2.2.9 on page 40 for detailed results.

2.2.3 Definition: The energy norm for (2.12) —analogously to (2.8)— is

$$\big\|u\big\|_{1;h} := \sqrt{|u|_{1,2;\mathbb{T}_h}^2 + \big\|\sqrt{\kappa_h}[\![u]\!]\big\|_{\mathbb{E}_h^i}^2 + \big\|\sqrt{2\kappa_h}u\big\|_{\mathbb{E}_h^D}^2}. \tag{2.13}$$

The definiteness of this norm follows with Friedrichs inequality in Lemma 1.2.3 by the same arguments as in the proof of Lemma 2.1.3.

2.2.4 We discretize the differential equation by choosing the space $V_h = X_{h,k}$ with polynomials in $\mathbb{P}_k$ (on affine cells) or $\mathbb{Q}_k$ on each cell and $k \geqslant 1$.

2.2.5 Lemma: *If in* (2.12) *for each edge between cells T^+ and T^-*

$$\kappa_h = \frac{\kappa_E^+ + \kappa_E^-}{2}, \qquad \kappa_E^\pm > \frac{c_\pm^2}{2}, \tag{2.14}$$

with positive $\alpha < 1$ and $c_\pm$ is the constant from the trace estimate (1.27) *in corollary 1.2.17 on page 23 for the adjacent cells T^+ and T^-, respectively. Then, the stability estimate*

$$\big\|v\big\|_{1;h}^2 \leqslant c a_h^{IP}(v,v), \tag{2.15}$$

holds for all $v \in V_h$ with a constant c independent of the mesh size and of κ.

Proof. We start with

$$a_h^{\text{IP}}(v,v) = \big\|v\big\|_{1;h}^2 - 2\big\langle \{\!\{\nabla v\}\!\}, [\![v\mathbf{n}]\!] \big\rangle_{\mathbb{E}_h^i} - 2\big\langle \partial_n v, v \big\rangle_{\mathbb{E}_h^D}.$$

The indefinite term is estimated by considering it cell by cell:

$$
\begin{aligned}
&2\big\langle \{\!\{\nabla v\}\!\}, [\![v\mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i\cup\mathbb{E}_h^D} + 2\big\langle \partial_n v, v\big\rangle_{\mathbb{E}_h^D} \\
&\qquad \leqslant \sum_{T\in\mathbb{T}_h} \Big\{ \frac{\alpha h_T}{c_*^2}\big\|\partial_n v\big\|_{\partial T_i}^2 + \frac{c_*^2}{4\alpha h_T}\big\|[\![v]\!]\big\|_{\partial T_i}^2 + \frac{\alpha h_T}{c_*^2}\big\|\partial_n v\big\|_{\partial T_D}^2 + \frac{c_*^2}{\alpha h_T}\big\|v\big\|_{\partial T_D}^2 \Big\} \\
&\qquad\qquad \leqslant \alpha \sum_{T\in\mathbb{T}_h} \Big\{ \big\|\nabla v\big\|_T^2 + \frac{1}{2}\big\|\sqrt{\kappa}[\![v]\!]\big\|_{\partial T_i}^2 + 2\big\|\sqrt{\kappa} v\big\|_{\partial T_D}^2 \Big\} \leqslant \alpha \big\|v\big\|_{1;h}^2.
\end{aligned}
$$

Therefore, the lemma is proven with $c = 1 - \alpha$. □

2.2.6 Additionally, we introduce the following "extended" energy norm:

$$
\big\|v\big\|_{1;h}^{+} := \left(\sum_{T\in\mathbb{T}_h} \Big\{ \big\|\nabla v\big\|_T^2 + \frac{\kappa}{h_T}\big\|[\![v]\!]\big\|_{\partial T}^2 + \frac{h_T}{\kappa}\big\|\{\!\{\partial_n v\}\!\}\big\|_{\partial T}^2 \Big\} \right)^{1/2}, \tag{2.16}
$$

which is equivalent to the energy norm on V_h by (1.26) and the previous lemma.

2.2.7 Lemma: *For functions $u \in W^{s,2}(\mathbb{T}_h)$ with $s \leqslant k+1$, the interior penalty norm admits the projection error estimate*

$$
\big\|\eta_u\big\|_{1;h} \leqslant c\big|h^{s-1}u\big|_{s,2;\mathbb{T}_h}. \tag{2.17}
$$

Proof. This lemma is an immediate consequence of approximation result 1.2.13 and trace estimate 1.2.14. □

2.2.8 If κ is chosen too small, the definiteness of the form $a_h^{\text{IP}}\big(.,.\big)$ is lost. This is reflected in the fact that the smallest solution $\lambda_{\min}$ of the eigenvalue problem

$$
a_h^{\text{IP}}\big(u_h, v\big) = \lambda\big(u, v\big) \quad \forall v \in V_h, \tag{2.18}
$$

becomes dependent of κ and negative. In Figure 2.1, we show the dependency of $\lambda_{\min}$ on κ for the Cartesian meshes of example A.1.1 on page 167. The values are the minimal eigenvalues of the tridiagonal Lanczos matrix generated by the conjugate gradient method. The mesh width for each polynomial degree was chosen such that the Krylov space had a dimension of at least 80 to ensure sufficient approximation of the discrete eigenvalues. The graphs show that whenever κ becomes lower than a certain threshold, depending on the polynomial degree, stability is lost almost instantly (the step size in κ is 10^{-3}). Whenever κ is larger than this threshold, the smallest eigenvalue of $a_h(.,.)$ is constant. Here, the polynomial space $\mathbb{P}_1$ is an exception: the smallest eigenvalue increases slowly, since the degrees of freedom do not allow that the jumps of the corresponding eigenfunction tend to zero.

Figure 2.2 shows the corresponding eigenfunction for κ just below the stability requirement on the left and for the stable form on the right. While the stable version shows approximation to the continuous eigenfunction $\cos(\pi x)\cos(\pi y)$, the unstable one shows oscillations with the highest frequency, usually observed for the largest eigenvalue. The indefinite jump terms serve here to outweigh the large gradients inside the cells.

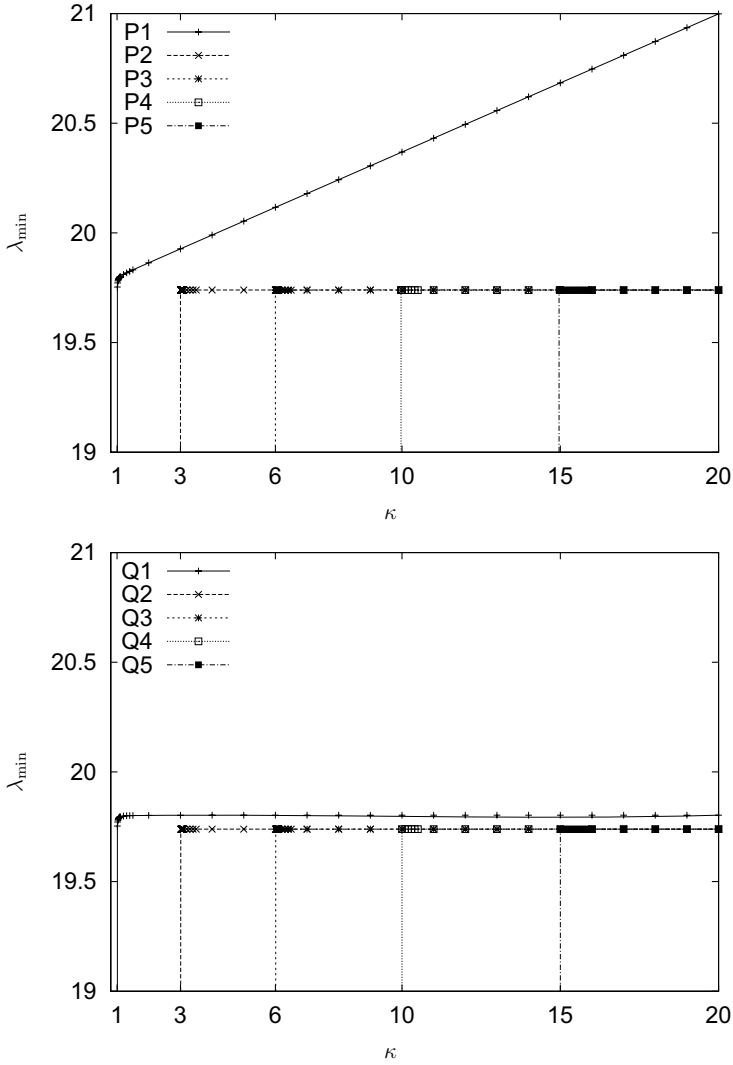

Figure 2.1: Smallest eigenvalue of $a_h(.,.)$ on a mesh of squares

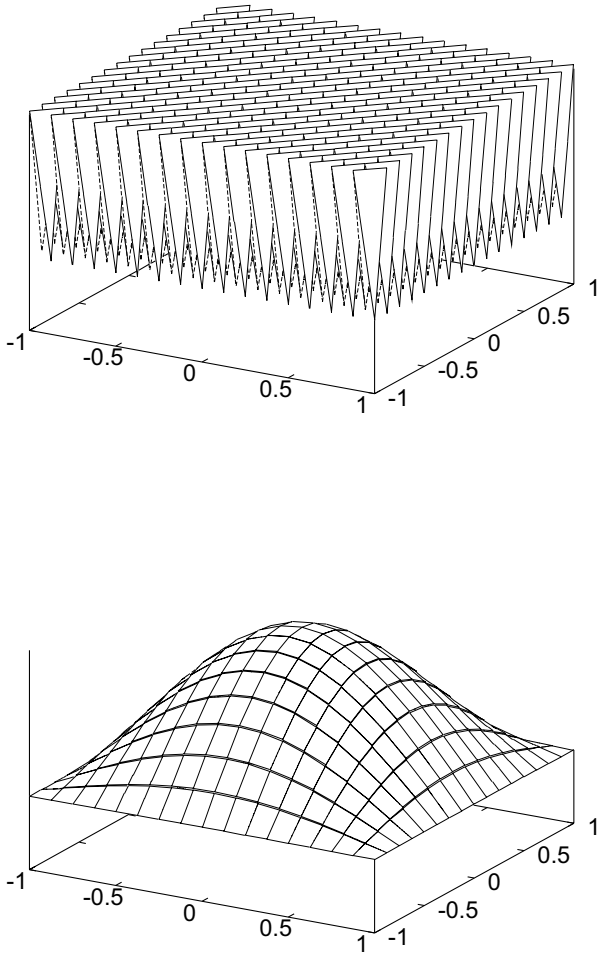

Figure 2.2: Eigenfunction to the smallest eigenvalue of $a_h(.,.)$ for $\kappa = 1.0$ (left) and $\kappa = 1.1$ (right) on a mesh of squares

ϑ	$\mathbb{P}_1$	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_4$	$\mathbb{P}_5$	$\mathbb{Q}_1$	$\mathbb{Q}_2$	$\mathbb{Q}_3$	$\mathbb{Q}_4$	$\mathbb{Q}_5$
1.0	1.34	4.43	9.46	16.41	25.39	1.35	4.44	9.47	16.49	25.50
1.5	1.08	3.30	6.76	11.47	17.50	1.08	3.30	6.77	11.51	17.55
2.0	1.01	3.00	6.00	9.97	14.96	1.01	3.01	6.01	10.00	15.00
4.0	1.00	2.98	5.97	9.85	14.78	1.00	2.99	5.98	9.86	14.83

Table 2.1: The lowest stable κ depending on the boundary stabilization on an equidistant square grid

2.2.9 We already remarked in 2.2.2 that there is some inconsistency of handling stabilization at the boundary in the literature. We investigate this fact by introducing the additional parameter ϑ and the generalized form of equation (2.12)

$$\begin{aligned}\big(\nabla u,\nabla v\big)_{\mathbb{T}_h} + \big\langle \kappa_h[\![u]\!],[\![v]\!]\big\rangle_{\mathbb{E}_h^i} + \big\langle 2\kappa_h[\![u]\!],[\![v]\!]\big\rangle_{\mathbb{E}_h^D}\\ - \big\langle \{\!\{\nabla u\}\!\},[\![v\mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i} - \big\langle [\![u\mathbf{n}]\!],\{\!\{\nabla v\}\!\}\big\rangle_{\mathbb{E}_h^i} - \big\langle \partial_n u, v\big\rangle_{\mathbb{E}_h^D} - \big\langle u,\partial_n v\big\rangle_{\mathbb{E}_h^D}\\ = \big(f,v\big)_\Omega + \big\langle \vartheta\kappa_h u^D, v\big\rangle_{\mathbb{E}_h^D} - \big\langle u^D,\partial_{\mathbf{n}} v\big\rangle_{\mathbb{E}_h^D}.\end{aligned} \tag{2.19}$$

Table 2.1 shows, that the minimal value of κ yielding a stable bilinear form depends on ϑ, as long as $\vartheta < 2$. For $\vartheta \geqslant 2$ this threshold value remains nearly constant. We conclude, that for $\vartheta < 2$, the grid cells causing the instability are located at the boundary. Therefore, all remaining examples in this chapter will be computed with the form (2.12), where $\vartheta = 2$.

2.2.10 Theorem: *Let κ_h be chosen according to* (2.14)*. If the solution u of* (2.4) *is in $W^{s,2}(\mathbb{T}_h)$ with $2 < s \leqslant k+1$, then the error $e_u = u - u_h$ between u and the solution u_h of* (2.12) *admits the estimate*

$$\big\|e_u\big\|_{1;h} \leqslant c\big|h^{s-1}u\big|_{s,2;\mathbb{T}_h}. \tag{2.20}$$

Proof. Again, we follow the outline of Theorem 1.3.9 and estimate the error remaining after projection. The arguments are similar to the proof of Theorem 2.1.6, but we can use the cell-wise L^2-projection Π_h. We abbreviate $\xi_u = \Pi_h u - e_u$ and estimate

$$\begin{aligned}c\big\|\xi_u\big\|_{1;h}^2 \leqslant a_h^{\mathrm{IP}}\big(\xi_u,\xi_u\big) = a_h^{\mathrm{IP}}\big(\eta_u,\xi_u\big)\\ \big|\big(\nabla\eta_u,\nabla\xi_u\big)_{\mathbb{T}_h}\big| + \big|\big\langle \kappa_h[\![\eta_u]\!],[\![\xi_u]\!]\big\rangle_{\mathbb{E}_h^i}\big| + \big|\big\langle 2\kappa_h\eta_u,\xi_u\big\rangle_{\mathbb{E}_h^D}\big|\\ + \big|\big\langle \{\!\{\nabla\eta_u\}\!\},[\![\xi_u]\!]\big\rangle_{\mathbb{E}_h^i}\big| + \big|\big\langle \nabla\eta_u,\xi_u\big\rangle_{\mathbb{E}_h^D}\big|\\ + \big|\big\langle [\![\eta_u]\!],\{\!\{\nabla\xi_u\}\!\}\big\rangle_{\mathbb{E}_h^i}\big| + \big|\big\langle \eta_u,\nabla\xi_u\big\rangle_{\mathbb{E}_h^D}\big|.\end{aligned}$$

By Hölder inequality and approximation results (1.21) and (1.25), we obtain for the first terms

$$\begin{aligned}\big|\big(\nabla\eta_u,\nabla\xi_u\big)_{\mathbb{T}_h}\big| &\leqslant \big\|\nabla\eta_u\big\|_{\mathbb{T}_h}\big\|\xi_u\big\|_{1;h} &&\leqslant c\big\|h^{s-1}u\big\|_{s,2;\mathbb{T}_h}\big\|\xi_u\big\|_{1;h}\\ \big|\big\langle \kappa_h[\![\eta_u]\!],[\![\xi_u]\!]\big\rangle_{\mathbb{E}_h^i}\big| &\leqslant \big\|\sqrt{\kappa_h}\eta_u\big\|_{\mathbb{E}_h^i}\big\|\xi_u\big\|_{1;h} &&\leqslant c\big\|h^{s-1}u\big\|_{s,2;\mathbb{T}_h}\big\|\xi_u\big\|_{1;h}\end{aligned}$$

Here and in the remaining estimates, the boundary terms are treated like the interior terms. We have

$$\big|\big\langle \{\!\{\nabla\eta_u\}\!\},[\![\xi_u]\!]\big\rangle_{\mathbb{E}_h^i}\big| \leqslant \big\|\kappa_h^{-\frac{1}{2}}\{\!\{\nabla\eta_u\}\!\}\big\|_{\mathbb{E}_h^i}\big\langle \kappa_h[\![\xi_u]\!],[\![\xi_u]\!]\big\rangle_{\mathbb{E}_h^i} \leqslant c\big\|h^{s-1}u\big\|_{s,2;\mathbb{T}_h}\big\|\xi_u\big\|_{1;h},$$

again by approximation result (1.25) and since $\kappa_h \approx h^{-1}$. Finally,

$$|\langle [\![\eta_u]\!], \{\!\{\nabla \xi_u\}\!\} \rangle_{\mathbb{E}_h^i}| \leqslant \left\| \sqrt{\kappa_h} [\![\eta_u]\!] \right\|_{\mathbb{E}_h^i} \langle \frac{1}{\kappa_h} \partial_n \xi_u, \partial_n \xi_u \rangle_{\mathbb{E}_h^i} \leqslant c \left\| h^{s-1} u \right\|_{s,2;\mathbb{T}_h} \left\| \xi_u \right\|_{1;h},$$

by approximation estimate (1.25) and inverse estimate (1.26). Summing up and dividing by $\left\| \xi_u \right\|_{1;h}$ yields the result of the theorem. □

2.2.11 Remark: In the last theorem we assumed $u \in W^{2,2}(\mathbb{T}_h)$ to keep the presentation simple. In fact, in [HN01] a technique using weighted norms is presented which allows for solutions with less regularity. This technique applies to the energy estimate above as well as to weaker and pointwise estimates below. Especially in the last case, a combination of weights balancing the singularities of the solution and of the Green function would render the presentation completely unreadable.

2.2.12 Theorem: *Assume that $u \in W^{2,2}(\mathbb{T}_h)$ and* (2.6) *holds for any solution z of* (2.4) *with arbitrary right hand side $f \in L^2(\Omega)$. Then, we have the L^2-error estimate*

$$\left\| e_u \right\| \leqslant c \left| h^2 u \right|_{W^{2,2}(\mathbb{T}_h)}. \tag{2.21}$$

Proof. The proof follows the lines of 1.3.14 with the concrete elliptic regularity estimate (2.6) entering in place of (1.40). □

2.2.13 We present experimental confirmation of the results above for $\mathbb{Q}_k$ shape functions in Table 2.2. We solve example A.2.2 on page 170 on the Cartesian (see example A.1.1 on page 167) and distorted meshes (see example A.1.3), respectively. The table shows the L^2-norm and the $W^{1,2}(\mathbb{T}_h)$-norm of the error $e_u = u - u_h$, as well as the experimental order of convergence determined by the formula

$$\mathrm{ord}_L = -\log_2 \frac{\mathrm{err}_L}{\mathrm{err}_{L-1}}.$$

On both types of meshes, the results are in good correspondence with the theory presented above.

2.2.14 We perform the same experiments with shape function spaces $\mathbb{P}_k$ in Table 2.3. On Cartesian grids, we obtain the same orders of convergence as in Table 2.2, even if the L^2-errors of the $\mathbb{P}_2$-element are considerably larger than those of the $\mathbb{Q}_2$-element.

Since the distorted meshes consist of non-affine grid cells, the $\mathbb{P}_k$-elements suffer from a lack of approximation (cf. [ABF02]). This is clearly visible in the reduced approximation orders. Therefore, $\mathbb{P}_k$-elements should be avoided on grids not consisting of parallelogram cells.

2.2.15 The situation of the distorted grid in the previous paragraph occurs if the computational grid is the result of a grid generator. If the grid is generated by consecutive refinement of a coarse mesh, grid cells approximate parallelograms on coarser meshes. We investigate behavior in this case discretizing on the circle $\Omega = B_1(0)$ (for meshes see example A.1.6 on page 170). The solution is again the exponential function. Table 2.4 shows that on such a sequence of grids the $\mathbb{P}_1$-element converges again with the same order as the $\mathbb{Q}_1$-element.

L	$\|e_u\|$		$\|e_u\|_{1,2;\mathbb{T}_h}$		$\|e_u\|$		$\|e_u\|_{1,2;\mathbb{T}_h}$	
	err	ord	err	ord	err	ord	err	ord
	$\mathbb{Q}_1$, $\kappa = 2$				$\mathbb{Q}_2$, $\kappa = 6$			
	Cartesian mesh							
2	1.64e-2	1.85	3.56e-1	0.99	1.98e-05	2.99	1.15e-2	1.99
3	4.38e-3	1.90	1.73e-1	1.04	1.39e-06	3.00	2.88e-3	2.00
4	1.15e-3	1.93	8.45e-2	1.04	9.16e-08	3.00	7.20e-4	2.00
5	2.96e-4	1.96	4.16e-2	1.02	5.88e-09	3.00	1.80e-4	2.00
6	7.53e-5	1.97	2.06e-2	1.01	3.86e-10	3.00	4.50e-5	2.00
7	1.90e-5	1.99	1.02e-2	1.01	8.17e-11	3.00	1.13e-5	2.00
	distorted mesh							
	$\mathbb{Q}_1$, $\kappa = 6$				$\mathbb{Q}_2$, $\kappa = 18$			
2	1.87e-02	1.74	3.23e-01	0.97	4.31e-04	2.74	1.22e-02	1.87
3	5.26e-03	1.83	1.66e-01	0.97	6.12e-05	2.82	3.30e-03	1.89
4	1.43e-03	1.88	8.46e-02	0.97	8.45e-06	2.86	8.75e-04	1.91
5	3.79e-04	1.91	4.30e-02	0.98	1.14e-06	2.89	2.29e-04	1.94
6	9.89e-05	1.94	2.17e-02	0.98	1.51e-07	2.92	5.90e-05	1.95
7	2.55e-05	1.96	1.10e-02	0.99	1.96e-08	2.94	1.51e-05	1.97

Table 2.2: Convergence of the interior penalty method with tensor product polynomials on Cartesian and distorted meshes

L	$\|e_u\|$		$\|e_u\|_{1,2;\mathbb{T}_h}$		$\|e_u\|$		$\|e_u\|_{1,2;\mathbb{T}_h}$	
	err	ord	err	ord	err	ord	err	ord
	$\mathbb{P}_1$, $\kappa = 2$				$\mathbb{P}_2$, $\kappa = 6$			
	Cartesian mesh							
2	2.24e-2	1.94	4.77e-1	0.98	8.08e-4	3.00	2.80e-2	1.97
3	5.83e-3	1.94	2.37e-1	1.01	9.99e-5	3.02	7.04e-3	1.99
4	1.50e-3	1.96	1.17e-1	1.01	1.24e-5	3.01	1.76e-3	2.00
5	3.83e-4	1.97	5.82e-2	1.01	1.55e-6	3.00	4.41e-4	2.00
6	9.70e-5	1.98	2.90e-2	1.01	1.93e-7	3.00	1.10e-4	2.00
7	2.44e-5	1.99	1.44e-2	1.00	2.42e-8	3.00	2.76e-5	2.00
	distorted mesh							
	$\mathbb{P}_1$, $\kappa = 6$				$\mathbb{P}_2$, $\kappa = 18$			
2	4.71e-2	1.08	6.66e-1	0.76	3.18e-3	2.86	7.20e-2	1.77
3	1.73e-2	1.44	3.54e-1	0.91	3.85e-4	3.04	1.95e-2	1.88
4	5.27e-3	1.72	1.91e-1	0.89	5.15e-5	2.90	5.54e-3	1.82
5	1.53e-3	1.78	1.12e-1	0.77	8.90e-6	2.53	1.79e-3	1.63
6	4.81e-4	1.67	7.93e-2	0.49	1.91e-6	2.22	7.04e-4	1.35
7	1.74e-4	1.46	6.74e-2	0.24	4.52e-7	2.08	3.17e-4	1.15

Table 2.3: Convergence of the interior penalty method with polynomials $\mathbb{P}_k$ on Cartesian and distorted meshes

L	$\|e_u\|$		$\|e_u\|_{1,2;\mathbb{T}_h}$		$\|e_u\|$		$\|e_u\|_{1,2;\mathbb{T}_h}$	
	err	ord	err	ord	err	ord	err	ord
	$\mathbb{Q}_1$, $\kappa = 4$				$\mathbb{P}_1$, $\kappa = 6$			
2	1.67e-2	1.80	3.20e-1	0.89	4.24e-2	1.36	6.32e-1	0.88
3	4.46e-3	1.90	1.62e-1	0.98	1.33e-2	1.67	3.27e-1	0.95
4	1.16e-3	1.94	8.08e-2	1.00	3.82e-3	1.80	1.66e-1	0.98
5	2.97e-4	1.96	4.02e-2	1.01	1.05e-3	1.87	8.31e-2	1.00
6	7.54e-5	1.98	2.00e-2	1.01	2.76e-4	1.92	4.14e-2	1.01
7	1.90e-5	1.99	9.99e-3	1.00	7.12e-5	1.95	2.06e-2	1.01

Table 2.4: Convergence of the interior penalty method with $\mathbb{Q}_1$- and $\mathbb{P}_1$-elements on a circle with bilinear mapping

2.2.16 Lemma: *Let the space $\breve{V}_h$ be the largest continuous subspace of V_h and $\breve{V}_h^D$ the subspace such that for all $\breve{v} \in \breve{V}_h^D$ holds*

$$\left\langle u^D - \breve{v}, w \right\rangle_{\Gamma_D} = 0 \quad \forall w \in \breve{V}_h. \tag{2.22}$$

Then, if $\kappa \to \infty$, the solution $u_{h;\kappa}$ of (2.12) *converges to the solution $\breve{u} \in \breve{V}_h^D$ of the variational problem*

$$\big(\nabla u, \nabla v\big) = \big(f, v\big) \quad \forall v \in \breve{V}_h^0,$$

where $\breve{V}_h^0$ is the subspace of $\breve{V}_h$ with zero traces on Γ_D.

Proof. By Theorem 2.2.10, we have that

$$\big\|\Pi_h u - u_{h;\kappa}\big\|_{IP;\kappa} \leqslant c \inf_{v \in V_h} \big\|u - v\big\|_{IP;\kappa},$$

with c independent of κ. Now, we first assume $u = 0$ on Γ_D and choose $v \in \breve{V}_h$; then, the right hand side does not depend on κ, yielding

$$\big\|[\![u - u_{h;\kappa}]\!]\big\|^2 \leqslant \frac{c}{\kappa}.$$

Inhomogeneous boundary conditions are included by projecting them into the finite element space, first. □

2.2.17 The result of this lemma is verified in Figure 2.3. It shows that a very small error is achieved with κ only slightly (by $\frac{3}{2}$) larger than the stability threshold. If $\mathbb{Q}_1$-elements are used, the error converges to a fixed value very fast. With $\mathbb{P}_1$-elements, the error starts to increase slowly with growing κ. In fact, if $\kappa \to \infty$, the $\mathbb{P}_1$ solution must converge to a globally linear function, therefore not approximating the solution of the differential equation at all.

2.3 Local error estimates

2.3.1 The following paragraphs establish estimates of the error weighted with smooth approximations to Dirac functionals. In the case of quasi-uniform grids, we obtain optimal L^∞-error

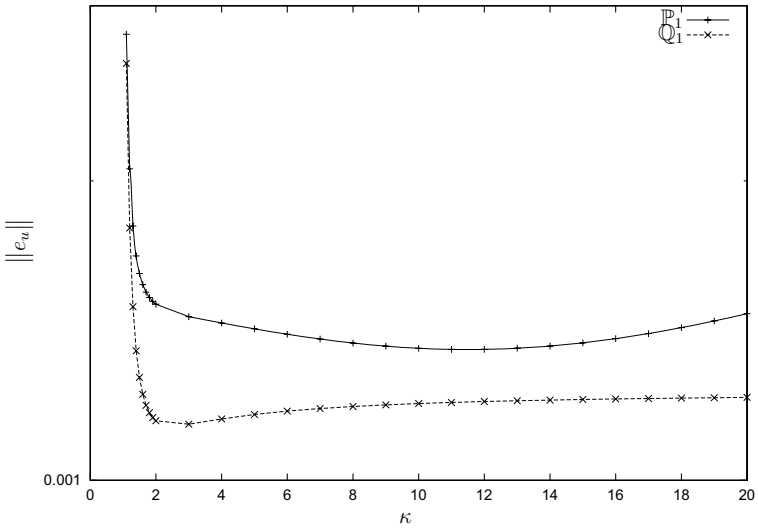

Figure 2.3: Accuracy of the interior penalty method depending on the parameter κ.

estimates. The results are taken from [KR02], following the outlines in [FR76]. Therefore, in this whole section, we restrict ourselves to the two-dimensional case $d = 2$ and to bilinear shape functions $\mathbb{Q}_1$. In order to keep the presentation of the rather lengthy computations simpler, we will use the convention

$$[\![v]\!]|_{\Gamma_D} = \sqrt{2}v$$
$$\{\!\{v\}\!\}|_{\Gamma_D} = \sqrt{\frac{1}{2}}v,$$

throughout this whole section. This way, the boundary terms in (2.12) assume the same form as the interior terms. Furthermore, we assume for simplicity that $\Gamma_N = \emptyset$.

The analysis follows essentially the framework for weak norms in Section 1.3. In order to obtain optimal estimates, this has to be performed in weighted norms. Therefore, we will begin this section by collecting estimates for weight functions and Green functions. The main result of this section can be found in Theorem 2.3.7 on page 48.

2.3.2 Definition: For some point $a \in \overline{\Omega}$ and a parameter $\varepsilon > 0$, let $\mathcal{U}_a^\varepsilon \subset \overline{\Omega}$ be a set, such that

$$a \in \mathcal{U}_a^\varepsilon, \quad \operatorname{diam}(\mathcal{U}_a^\varepsilon) = \varepsilon, \quad c\varepsilon^2 \leqslant |\mathcal{U}_a^\varepsilon| \leqslant \varepsilon^2. \tag{2.23}$$

Typically, we consider a ball of radius ε around a or a patch of mesh cells containing a. Further, let $\delta_a^\varepsilon \in L^\infty(U_a^\varepsilon)$ be a weighting function (*regularized Dirac function*) with support in $\mathcal{U}_a^\varepsilon$ and

$$\int_{\mathcal{U}_a^\varepsilon} \delta_a^\varepsilon(x)dx = 1, \quad \left\|\delta_a^\varepsilon\right\|_{L^\infty(\mathcal{U}_a^\varepsilon)} \leqslant c\varepsilon^{-2}. \tag{2.24}$$

We remark, that the function being constant $1/|\mathcal{U}_a^\varepsilon|$ inside $\mathcal{U}_a^\varepsilon$ and zero otherwise fulfills these conditions. Still, the function δ_a^ε is not required to be constant on the set $\mathcal{U}_a^\varepsilon$. In particular, projections of the Dirac functional into the discrete space V_h will be allowed.

Finally, we define the evaluation functional

$$J_a^\varepsilon(u) = \int_{\mathcal{U}_a^\varepsilon} u(x)\delta_a^\varepsilon(x)dx, \tag{2.25}$$

which approximates the point evaluation $u(a)$ if u is sufficiently regular and $\varepsilon \to 0$.

2.3.3 Definition: We introduce

$$\varrho = \varrho_a^\varepsilon := \sqrt{|x-a|^2+\varepsilon^2}, \tag{2.26}$$

the *regularized distance function* with respect to a point a. With T_a, we denote a mesh cell containing a and $h_a = h_{T_a}$. Furthermore, we define ϱ_h as the L^2-projection of ϱ into the space $X_{h,0}$ of cell-wise constant functions and denote by $\varrho_{h,T}$ its value on the cell T.

2.3.4 Lemma: *Let the shape regularity Assumption 1.1.13 and the assumption on mesh size variation 1.1.14 hold. Then, there exist constants c independent of ε, h and T, such that as soon as $h_a < \varepsilon$, the following estimates hold with $\ell(\varepsilon) = 1 + |\log(\varepsilon)|$.*

$$h_T \leqslant c\varrho_{h,T} \qquad \forall T \in \mathbb{T}_h, \tag{2.27}$$

$$\max_T \varrho \leqslant c \min_T \varrho \qquad \forall T \in \mathbb{T}_h, \tag{2.28}$$

$$\left\|\varrho^{-1}\right\| \leqslant c\sqrt{\ell(\varepsilon)}. \tag{2.29}$$

Proof. We begin proving (2.28). Let T be a cell, such that $\operatorname{dist}(a,T) \leqslant 2\varepsilon$. Then, $\min_T \varrho \geqslant \varepsilon$ and by (1.4)

$$(\max_T \varrho)^2 \leqslant (2\varepsilon + h_T)^2 + \varepsilon^2 \leqslant (5 + 2c_{\mathbb{T}} + c_{\mathbb{T}}^2)\varepsilon^2.$$

Now, if $\operatorname{dist}(a,T) \geqslant 2\varepsilon$ holds, then $h_T \leqslant c(h_a + \operatorname{dist}(T_a,T))$ by Assumption 1.1.14. Since $\operatorname{dist}(T_a,T) \leqslant \operatorname{dist}(a,T)$, we have for $x \in T$

$$\varrho(x)^2 \leqslant (\operatorname{dist}(a,T) + h_T)^2 + \varepsilon^2 \leqslant c\operatorname{dist}(a,T)^2,$$
$$\varrho(x)^2 \geqslant \operatorname{dist}(a,T)^2 - \varepsilon^2 \geqslant \frac{1}{2}\operatorname{dist}(a,T)^2.$$

Therefore, (2.28) is proven and (2.27) follows since the mean values of ϱ on a cell is between the minimum and the maximum.

In order to prove (2.29), we extend the norm to a ball $B_R(a)$ in $\mathbb{R}^2$ around a and containing Ω, yielding

$$\begin{aligned}\left\|\varrho^{-1}\right\|^2 &\leqslant \left\|\varrho^{-1}\right\|_{\mathbb{R}^2}^2 \\ &= 2\pi \int_0^R \frac{r}{r^2+\varepsilon^2}\,dr \\ &= \pi(\log(R^2+\varepsilon^2) - \log\varepsilon^2).\end{aligned}$$

Under the reasonable assumptions that $\varepsilon \ll \operatorname{diam}(\Omega)$ and $\varepsilon < 1$, estimate (2.29) follows. □

2.3.5 Definition: The *regularized Green function* $g = g_a^\varepsilon \in W^{2,2}(\Omega)$ associated with the regularized Dirac function δ_a^ε is the solution of the problem

$$\begin{aligned} -\Delta g_a^\varepsilon &= \delta_a^\varepsilon && \text{in } \Omega, \\ g_a^\varepsilon &= 0 && \text{on } \Gamma_D. \end{aligned} \tag{2.30}$$

Its Ritz projection $g_h \in V_h$ is the solution to the discrete problem

$$a_h^N\big(v, g_h\big) = (\delta_a^\varepsilon, v) \quad \forall v \in V_h. \tag{2.31}$$

2.3.6 Lemma: *Let $d = 2$ and $V_h = X_{h,1}^Q$. Let furthermore the elliptic regularity estimate* (2.6) *hold. Then, the regularized and discrete Green functions defined above admit the estimates*

$$\big\|\nabla g\big\| + \big\|\varrho_h \nabla^2 g\big\| \leqslant c\sqrt{\ell(\varepsilon)}, \tag{2.32}$$

$$\big\|g - g_h\big\| + h\big\|\nabla(g - g_h)\big\|_{\mathbb{T}_h} \leqslant ch\,\ell(\varepsilon), \tag{2.33}$$

$$\big\|\varrho_h(g - \Pi_h g)\big\| + h\big\|\varrho_h \nabla(g - \Pi_h g)\big\|_{\mathbb{T}_h} \leqslant ch\,\ell(\varepsilon). \tag{2.34}$$

Proof. We begin with proving (2.32). By definition, there holds

$$\big\|\nabla g\big\|^2 = (g, \delta_a^\varepsilon) \leqslant c \max_{\mathcal{U}_a^\varepsilon} |g|. \tag{2.35}$$

Now, let $G(x, y)$ be the true Green function of the Laplacian on the domain Ω, for which we have the following bound (see [FR76]):

$$|G(x, y)| \leqslant c\big|\log|x - y|\big| + c.$$

Therefore,

$$|g(x)| = |(G(x, \cdot), \delta)| \leqslant \frac{1}{|\mathcal{U}_a^\varepsilon|} \int_{\mathcal{U}_a^\varepsilon} |G(x, y)| dy \leqslant c\,\ell(\varepsilon).$$

This implies $\big\|\nabla g\big\| \leqslant c\,\ell(\varepsilon)^{1/2}$.

Next, we estimate using the elliptic regularity estimate (2.6),

$$\begin{aligned} \big\|\varrho_h \nabla^2 g\big\|^2 &\leqslant c\big\|\varrho \nabla^2 g\big\|^2 = c \sum_{i=1}^d \big\|x_i \nabla^2 g\big\|^2 + \vartheta^2 \varepsilon^2 \big\|\nabla^2 g\big\|^2 \\ &\leqslant c \sum_{i=1}^d \big\|\nabla^2 (x_i g)\big\|^2 + c\big\|\nabla g\big\|^2 + \vartheta^2 \varepsilon^2 \big\|\nabla^2 g\big\|^2 \\ &\leqslant c \sum_{i=1}^d \big\|\Delta (x_i g)\big\|^2 + \big\|\nabla g\big\|^2 + \vartheta^2 \varepsilon^2 (\big\|\Delta g\big\|^2 + \big\|\nabla g\big\|^2) \\ &\leqslant c \sum_{i=1}^d \big\|x_i \Delta g\big\|^2 + c\big\|\nabla g\big\|^2 + \vartheta^2 \varepsilon^2 \big\|\Delta g\big\|^2 \\ &\leqslant c\big\|\varrho \Delta g\big\|^2 + c\big\|\nabla g\big\|^2. \end{aligned}$$

Consequently, by the definition of g,

$$\big\|\varrho_h\nabla^2 g\big\|^2 \leqslant c\big\|\varrho\delta_a^\varepsilon\big\|^2 + c\big\|\nabla g\big\|^2 \leqslant \ell(\varepsilon).$$

which implies the asserted bound (2.32). Using the estimates (1.21) and (1.25) and the Sobolev inequality

$$\big\|\nabla v\big\| \leqslant c\big\|\nabla^2 v\big\|_{L^1(\Omega)}, \quad v \in V \cap W^{2,1}(\Omega), \tag{2.36}$$

we conclude for $v \in V \cap W^{2,2}(\Omega)$ and $v \in V \cap W^{2,1}(\Omega)$, respectively, the global approximation estimate

$$\big\|v - \Pi_h v\big\|_{1;h}^+ \leqslant c \begin{cases} h\big\|\nabla^2 v\big\|_{L^2(\Omega)} \\ \big\|\nabla^2 v\big\|_{L^1(\Omega)} \end{cases}. \tag{2.37}$$

For $v, w \in V \oplus V_h$, there holds

$$|a_h(v,w)| \leqslant \sum_{T\in\mathbb{T}_h} \Big\{ \big\|\nabla v\big\|_T \big\|\nabla w\big\|_T + \kappa h_T \big\|[\![v]\!]\big\|_{\partial T} \big\|[\![w]\!]\big\|_{\partial T} \\ + \big\|\{\!\{\partial_n v\}\!\}\big\|_{\partial T} \big\|[\![w]\!]\big\|_{\partial T} + \big\|[\![v]\!]\big\|_{\partial T} \big\|\{\!\{\partial_n w\}\!\}\big\|_{\partial T} \Big\},$$

and, consequently,

$$|a_h(v,w)| \leqslant c\big\|v\big\|_{1;h}^+ \big\|w\big\|_{1;h}^+. \tag{2.38}$$

By the coercivity estimate (2.15) and Galerkin orthogonality, we conclude for the dual error $e_g := g - g_h$ that

$$\begin{aligned} \big\|\Pi_h e_g\big\|_{1;h}^2 &\leqslant a_h(\Pi_h e_g, \Pi_h e_g) = a_h(e_g, \Pi_h e_g) + a_h(g - \Pi_h g, \Pi_h e_g) \\ &= a_h(g - \Pi_h g, \Pi_h e_g). \end{aligned}$$

Consequently, by (2.38) and the equivalence of the energy norms,

$$\big\|e_g\big\|_{1;h} \leqslant \big\|\Pi_h e_g\big\|_{1;h} + \big\|g - \Pi_h g\big\|_{1;h} \leqslant c\big\|g - \Pi_h g\big\|_{1;h}^+. \tag{2.39}$$

Hence, approximation estimate (2.37) yields

$$\big\|e_g\big\|_{1;h} \leqslant c\big\|\nabla^2 g\big\|_{L^1(\Omega)}. \tag{2.40}$$

Next, we employ a duality argument. Let $z \in V$ be the solution of

$$-\Delta z = e_g\big\|e_g\big\|^{-1} \text{in} \Omega, \tag{2.41}$$

satisfying the a priori estimate $\big\|z\big\|_{H^2} \leqslant c$. By (2.37) and elliptic regularity, we have

$$\big\|z - \Pi_h z\big\|_{1;h}^+ \leqslant ch\big\|\nabla^2 z\big\| \leqslant ch. \tag{2.42}$$

Using Galerkin orthogonality, there holds

$$\big\|e_g\big\| = a_h(e_g, z) = a_h(e_g, z - \Pi_h z).$$

Consequently,

$$\|e_g\| \leqslant c\|e_g\|_{1;h}^{+}\|z-\Pi_h z\|_{1;h}^{+} \leqslant ch\|e_g\|_{1;h}^{+} \leqslant ch\|\nabla^2 g\|_{L^1(\Omega)}. \tag{2.43}$$

It remains to bound the L^1 norm on the right. There holds

$$\|\nabla^2 g\|_{L^1(\Omega)} \leqslant \|\varrho^{-1}\|\|\varrho\nabla^2 g\| \tag{2.44}$$

which, in view of (2.29) and (2.32) yields the asserted L^2 error estimate (2.33).

Finally, we derive the interpolation estimate (2.34). In virtue of the local interpolation estimates (1.21), we have

$$\begin{aligned}\|\varrho_h(g-\Pi_h g)\|^2 &\leqslant \sum_{T\in\mathbb{T}_h} \varrho_{h,T}^2\|g-\Pi_h g\|_T^2\\ &\leqslant ch^4\sum_{T\in\mathbb{T}_h}\varrho_{h,T}^2\|\nabla^2 g\|_T^2 = ch^4\|\varrho_h\nabla^2 g\|^2,\end{aligned}$$

and analogously,

$$\|\varrho_h\nabla(g-\Pi_h g)\|_{\mathbb{T}_h}^2 \leqslant ch^2\|\varrho_h\nabla^2 g\|^2.$$

Combining this with the a priori bound (2.32) completes the proof. □

2.3.7 Theorem: *Let $\varepsilon>0$ and for some point $a\in\overline{\Omega}$, let $B=B_r(a)$ be a ball around a with radius $r>\varepsilon$. Then, if $h<\frac{1}{2}r$ and $h_a<\varepsilon$, there holds*

$$|J_a^\varepsilon(u)-J_a^\varepsilon(u_h)| \leqslant c\,\ell(\varepsilon)\|h^2u\|_{L^\infty(B)} + c|h^2u|_{2,2;\mathbb{T}_h}. \tag{2.45}$$

Proof. Using Galerkin orthogonality, for e_u and e_g, we conclude

$$\begin{aligned}(e_u,\delta_a^\varepsilon) &= a_h(e_u,g) = a_h(e_u,e_g) = a_h(\eta_u,e_g)\\ &= \sum_{T\in\mathbb{T}_h}\Big\{(\nabla\eta_u,\nabla e_g)_T + \frac{\kappa}{h_T}([\![\eta_u]\!],[\![e_g]\!])_{\partial T} - \frac{1}{2}(\{\!\{\partial_n\eta_u\}\!\},[\![e_g]\!])_{\partial T}\\ &\quad - \frac{1}{2}([\![\eta_u]\!],\{\!\{\partial_n e_g\}\!\})_{\partial T}\Big\},\end{aligned}$$

and abbreviate

$$|(e_u,\delta_a^\varepsilon)| \leqslant \Theta_1^{1/2}\Theta_2^{1/2},$$

where, with some parameter $\gamma\in(0,1]$,

$$\begin{aligned}\Theta_1 &= \sum_{T\in\mathbb{T}_h}\Big\{\|\varrho_h^{-1}\nabla\eta_u\|_T^2 + \frac{c\kappa}{\gamma h_T}(\{\!\{\varrho_h^2\}\!\}^{-1}[\![\eta_u]\!],[\![\eta_u]\!])_{\partial T}\\ &\qquad + \frac{h_T}{\kappa}(\{\!\{\varrho_h^2\}\!\}^{-1}\{\!\{\partial_n\eta_u\}\!\},\{\!\{\partial_n\eta_u\}\!\})_{\partial T}\Big\},\\ \Theta_2 &= \sum_{T\in\mathbb{T}_h}\Big\{\|\varrho_h\nabla e_g\|_T^2 + \frac{\kappa}{h_T}(\{\!\{\varrho_h^2\}\!\}[\![e_g]\!],[\![e_g]\!])_{\partial T}\\ &\qquad + \gamma h_T(\{\!\{\varrho_h^2\}\!\}\{\!\{\partial_n e_g\}\!\},\{\!\{\partial_n e_g\}\!\})_{\partial T}\Big\}.\end{aligned}$$

The two terms Θ_1 and Θ_2 will be estimated separately. First, we estimate Θ_1. For $\breve{T}$ defined in (1.5), we conclude from the interpolation estimates (1.21) and (1.25), observing assumptions 1.1.13 and 1.1.14, that

$$\big\|\nabla\eta_u\big\|_T + h_T^{-1/2}\big\|[\![\eta_u]\!]\big\|_{\partial T} + h_T^{1/2}\big\|\{\!\{\partial_n\eta_u\}\!\}\big\|_{\partial T} \leqslant ch_T\big\|\nabla^2 u\big\|_{\breve{T}},$$

Now, we split the summation over $T \in \mathbb{T}_h$ as follows:

$$\Theta_1 = \sum_{T\in\mathbb{T}_h,\breve{T}\subset B_a} \big\{\dots\big\} + \sum_{T\in\mathbb{T}_h,\breve{T}\not\subset B_a} \big\{\dots\big\}.$$

On cells $T \not\subset B_a$, we have $\varrho_h^{-1} \leqslant c$ and therefore estimate

$$\begin{aligned}\Theta_1 &\leqslant \frac{c}{\gamma}\sum_{T\in\mathbb{T}_h,\breve{T}\subset B_a}\frac{h_T^2}{\varrho_h^2}\big\|\nabla^2 u\big\|_T^2 + \frac{c}{\gamma}\sum_{T\in\mathbb{T}_h,\breve{T}\not\subset B_a} h_T^2\big\|\nabla^2 u\big\|_T^2\\ &\leqslant \frac{c}{\gamma}h^2\big\|\nabla^2 u\big\|_{L^\infty(B_a)}^2\big\|\varrho_h^{-1}\big\|^2 + \frac{c}{\gamma}h^2\big\|\nabla^2 u\big\|^2.\end{aligned}$$

From this, we obtain

$$\Theta_1 \leqslant \frac{c}{\gamma}h^2\,\ell(\varepsilon)\big\|\nabla^2 u\big\|_{L^\infty(B_a)}^2 + \frac{c}{\gamma}h^2\big\|\nabla^2 u\big\|^2. \tag{2.46}$$

Next, we estimate Θ_2. Observing that ϱ_h is constant on each cell T, we have

$$\begin{aligned}a_h(\varrho_h^2 e_g, e_g) &= \sum_{T\in\mathbb{T}_h}\Big\{(\nabla(\varrho_h^2 e_g), \nabla e_g)_T + \frac{\kappa}{h_T}([\![\varrho_h^2 e_g]\!], [\![e_g]\!])_{\partial T}\\ &\qquad - \frac{1}{2}([\![\varrho_h^2 e_g]\!], \{\!\{\partial_n e_g\}\!\})_{\partial T} - \frac{1}{2}(\{\!\{\partial_n(\varrho_h^2 e_g)\}\!\}, [\![e_g]\!])_{\partial T}\Big\}\\ &= \sum_{T\in\mathbb{T}_h}\Big\{(\varrho_h^2\nabla e_g, \nabla e_g)_T + \frac{\kappa}{h_T}(\{\!\{\varrho_h^2\}\!\}[\![e_g]\!], [\![e_g]\!])_{\partial T}\\ &\qquad + \frac{\kappa}{h_T}([\![\varrho_h^2]\!]\{\!\{e_g\}\!\}, [\![e_g]\!])_{\partial T} - \frac{1}{2}([\![\varrho_h^2]\!]\{\!\{e_g\}\!\}, \{\!\{\partial_n e_g\}\!\})_{\partial T}\\ &\qquad - (\{\!\{\varrho_h^2\}\!\}[\![e_g]\!], \{\!\{\partial_n e_g\}\!\})_{\partial T} - \frac{1}{8}([\![\varrho_h^2]\!][\![\partial_n e_g]\!], [\![e_g]\!])_{\partial T}\Big\}.\end{aligned}$$

This leads us to

$$\Theta_2 = a_h(\varrho_h^2 e_g, e_g) + \Theta_3, \tag{2.47}$$

where

$$\begin{aligned}\Theta_3 = \sum_{T\in\mathbb{T}_h}\Big\{&\gamma h_T(\{\!\{\varrho_h^2\}\!\}\{\!\{\partial_n e_g\}\!\}, \{\!\{\partial_n e_g\}\!\})_{\partial T} + \frac{1}{2}([\![\varrho_h^2]\!]\{\!\{e_g\}\!\}, \{\!\{\partial_n e_g\}\!\})_{\partial T}\\ &+ (\{\!\{\varrho_h^2\}\!\}[\![e_g]\!], \{\!\{\partial_n e_g\}\!\})_{\partial T} + \frac{1}{8}([\![\varrho_h^2]\!][\![\partial_n e_g]\!], [\![e_g]\!])_{\partial T}\\ &+ \frac{\kappa}{h_T}([\![\varrho_h^2]\!]\{\!\{e_g\}\!\}, [\![e_g]\!])_{\partial T}\Big\}.\end{aligned}$$

We proceed with the first term on the right in (2.47). Using Galerkin orthogonality of e_g, we obtain

$$a_h(\varrho_h^2 e_g, e_g) = a_h(\varrho_h^2 e_g - \Pi_h(\varrho_h^2 e_g), e_g). \tag{2.48}$$

Since ϱ_h is constant on T and the projection Π_h is local, we have $\Pi_h(\varrho_h^2 e_g) = \varrho_h^2 \Pi_h e_g$. Hence, it follows that

$$a_h(\varrho_h^2 e_g, e_g) = \sum_{T\in\mathbb{T}_h} \Big\{ (\nabla(\varrho_h^2\eta_g), \nabla e_g)_T + \frac{\kappa}{h_T}([\![\varrho_h^2\eta_g]\!], [\![e_g]\!])_{\partial T} \\ - \frac{1}{2}([\![\varrho_h^2\eta_g]\!], \{\!\{\partial_n e_g\}\!\})_{\partial T} - \frac{1}{2}(\{\!\{\partial_n(\varrho_h^2\eta_g)\}\!\}, [\![e_g]\!])_{\partial T} \Big\}.$$

The four terms on the right hand side are now estimated separately using the trace inequalities (1.24), and the interpolation estimates (1.21), (1.25). Furthermore, we employ the a priori bound (2.33) and the error estimate (2.33) for the Green function g and the dual error η_g. Note that in order to estimate the average values $\{\!\{\partial_n\eta_u\}\!\}|_{\partial T}$, it suffices to estimate each of the terms $\partial_n\eta_u|_{\partial T}$ separately. In virtue of Assumption 1.1.14, the weights ϱ_h can be estimated by

$$|[\![\varrho_h^2]\!]|_{\partial T} \leqslant ch_T\varrho_{h,T} \leqslant ch_T|\{\!\{\varrho_h\}\!\}|_{\partial T}, \tag{2.49}$$

and we use that $\nabla^2 g_{h|T} \equiv 0$. By (2.28), $\varrho_h \leqslant c\varrho$. We will use a free parameter $\gamma \in (0,1]$.

For the first term, we find

$$\Big|\sum_{T\in\mathbb{T}_h} (\nabla(\varrho_h^2\eta_g), \nabla e_g)_T\Big| \leqslant c\sum_{T\in\mathbb{T}_h} \Big\{\frac{1}{\gamma}h_T^2\big\|\varrho\nabla^2 g\big\|_T^2 + \gamma\big\|\varrho_h\nabla e_g\big\|_T^2\Big\} \\ \leqslant \frac{c}{\gamma}\,\ell(\varepsilon)h^2 + c\gamma\Theta_2,$$

and, analogously, for the second term,

$$\Big|\sum_{T\in\mathbb{T}_h} \frac{\kappa}{h_T}([\![\varrho_h^2\eta_g]\!], [\![e_g]\!])_{\partial T}\Big| \\ \leqslant c\sum_{T\in\mathbb{T}_h} \Big\{\frac{\kappa}{\gamma h_T}\big\|\{\!\{\varrho_h\}\!\}^{-1}[\![\varrho_h^2\eta_g]\!]\big\|_{\partial T}^2 + \frac{\gamma\kappa}{h_T}\big\|\{\!\{\varrho_h\}\!\}[\![e_g]\!]\big\|_{\partial T}^2\Big\} \\ \leqslant c\sum_{T\in\mathbb{T}_h} \Big\{\frac{1}{\gamma}h_T^2\big\|\varrho\nabla^2 g\big\|_T^2 + \frac{\gamma\kappa}{h_T}\big\|\{\!\{\varrho_h\}\!\}[\![e_g]\!]\big\|_{\partial T}^2\Big\} \\ \leqslant \frac{c}{\gamma}\,\ell(\varepsilon)h^2 + c\gamma\Theta_2.$$

The third term is estimated by

$$\Big|\sum_{T\in\mathbb{T}_h} ([\![\varrho_h^2\eta_g]\!], \{\!\{\partial_n e_g\}\!\})_{\partial T}\Big| \\ = \Big|\sum_{T\in\mathbb{T}_h} \Big\{\big(2[\![\varrho_h]\!]\{\!\{\varrho_h\}\!\}\{\!\{\eta_g\}\!\} + \frac{1}{4}[\![\varrho_h]\!]^2\{\!\{\eta_g\}\!\} + \{\!\{\varrho_h\}\!\}^2[\![\eta_g]\!], \{\!\{\partial_n e_g\}\!\}\big)_{\partial T}\Big\}\Big| \\ \leqslant c\sum_{T\in\mathbb{T}_h} \big\|\varrho_h\eta_g\big\|_{\partial T}\big\|\varrho_h\partial_n e_g\big\|_{\partial T} \\ \leqslant c\sum_{T\in\mathbb{T}_h} h_T^{3/2}\big\|\varrho\nabla^2 g\big\|_T\big(h_T^{-1/2}\big\|\varrho_h\nabla e_g\big\|_T + h_T^{1/2}\big\|\varrho\nabla^2 g\big\|_T\big) \\ \leqslant c\sum_{T\in\mathbb{T}_h} \Big\{\gamma\big\|\varrho_h\nabla e_g\big\|_T^2 + \frac{1}{\gamma}h_T^2\big\|\varrho\nabla^2 g\big\|_T^2\Big\} \\ \leqslant c\gamma\Theta_2 + \frac{c}{\gamma}\,\ell(\varepsilon)h^2.$$

Finally, for the fourth term, we find

$$\begin{aligned}\Big|\sum_{T\in\mathbb{T}_h}(\{\!\{\partial_n(\varrho_h^2\eta_g)\}\!\},[\![e_g]\!])_{\partial T}\Big| &\leqslant \sum_{T\in\mathbb{T}_h}\big\|\{\!\{\varrho_h\}\!\}^{-1}\{\!\{\partial_n(\varrho_h^2\eta_g)\}\!\}\big\|_{\partial T}\big\|\{\!\{\varrho_h\}\!\}[\![e_g]\!]\big\|_{\partial T}\\ &\leqslant c\sum_{T\in\mathbb{T}_h}\Big\{\frac{1}{\gamma}\big(\|\varrho_h\nabla\eta_g\|_T^2+h_T^2\|\varrho_h\nabla^2\eta_g\|_T^2\big)+\gamma\frac{\kappa}{h_T}\big\|\{\!\{\varrho_h\}\!\}[\![e_g]\!]\big\|_{\partial T}^2\Big\}\\ &\leqslant c\sum_{T\in\mathbb{T}_h}\Big\{\frac{1}{\gamma}h_T^2\|\varrho\nabla^2 g\|_T^2+\frac{\gamma\kappa}{h_T}\big\|\{\!\{\varrho_h\}\!\}[\![e_g]\!]\big\|_{\partial T}^2\Big\}\\ &\leqslant \frac{c}{\gamma}\,\ell(\varepsilon)h^2+c\gamma\Theta_2.\end{aligned}$$

Combining these estimates yields

$$|a_h(\varrho_h^2 e_g,e_g)|\leqslant c\gamma\Theta_2+\frac{c}{\gamma}\,\ell(\varepsilon)h^2. \tag{2.50}$$

Next, we estimate the five terms in Θ_3 separately again by using the local trace estimates (1.24), and the a priori bound and error estimate (2.32) and (2.33). For the first term in Θ_3, we obtain

$$\begin{aligned}\Big|\sum_{T\in\mathbb{T}_h}\gamma h_T(\{\!\{\varrho_h^2\}\!\}\{\!\{\partial_n e_g\}\!\},\{\!\{\partial_n e_g\}\!\})_{\partial T}\Big| &\leqslant c\gamma\sum_{T\in\mathbb{T}_h}\Big\{\|\varrho_h\nabla e_g\|_T^2+h_T^2\|\varrho_h\nabla^2 g\|_T^2\Big\}\\ &\leqslant c\gamma\Theta_2+\frac{c}{\gamma}\,\ell(\varepsilon)h^2,\end{aligned}$$

and analogously for the second term,

$$\begin{aligned}&\Big|\sum_{T\in\mathbb{T}_h}\frac{1}{2}([\![\varrho_h^2]\!]\{\!\{e_g\}\!\},\{\!\{\partial_n e_g\}\!\})_{\partial T}\Big|\\ &\leqslant\sum_{T\in\mathbb{T}_h}\Big\{\frac{c}{\gamma}h_T\big\|\{\!\{e_g\}\!\}\big\|_{\partial T}^2+c\gamma h_T\big\|\{\!\{\varrho_h\}\!\}\{\!\{\partial_n e_g\}\!\}\big\|_{\partial T}^2\Big\}\\ &\leqslant\sum_{T\in\mathbb{T}_h}\Big\{\frac{c}{\gamma}\big(\|e_g\|_T^2+h_T^2\|\nabla e_g\|_T^2\big)+c\gamma\big(\|\varrho_h\nabla e_g\|_T^2+h_T^2\|\varrho_h\nabla^2 g\|_T^2\big)\Big\}\\ &\leqslant c\gamma\Theta_2+\frac{c}{\gamma}\,\ell(\varepsilon)h^2.\end{aligned}$$

The third term in Θ_3 is the most critical one, since it does non contain a factor h_T. Therefore, we have to absorb this term into the other definite terms in Θ_2 using the stabilization parameter κ. We recall that for $v_h\in V_h$,

$$|(\{\!\{\partial_n v_h\}\!\},[\![v_h]\!])_{\partial T}|\leqslant\frac{c_0^2}{2\kappa}\big\|\nabla v_h\big\|_{\breve{T}}^2+\frac{\kappa}{2h_T}\big\|[\![v_h]\!]\big\|_{\partial T}^2, \tag{2.51}$$

with $\breve{T}$ defined in (1.5). Splitting the dual error like $e_g=\eta_g+\Pi_h e_g$ with $\eta_g=g-\Pi_h g$, we have

$$(\{\!\{\varrho_h^2\}\!\}[\![e_g]\!],\{\!\{\partial_n e_g\}\!\})_{\partial T}=(\{\!\{\varrho_h^2\}\!\}[\![e_g]\!],\{\!\{\partial_n\eta_g\}\!\})_{\partial T}+(\{\!\{\varrho_h^2\}\!\}[\![e_g]\!],\{\!\{\partial_n\Pi_h e_g\}\!\})_{\partial T}$$

The first term on the right is treated analogously as the other terms before leading to the estimate

$$\Big|\sum_{T\in\_H}(\{\!\{\varrho_h^2\}\!\}[\![e_g]\!],\{\!\{\partial_n\eta_g\}\!\})_{\partial T}\Big|\leqslant\gamma\Theta_2+\frac{c}{\gamma}h^2\,\ell(\varepsilon). \tag{2.52}$$

The second term is treated as follows:

$$\Big|\sum_{T\in\mathbb{T}_h}(\{\!\{\varrho_h^2\}\!\}[\![e_g]\!],\{\!\{\partial_n\Pi_h e_g\}\!\})_{\partial T}\Big| \leqslant \sum_{T\in\mathbb{T}_h}\Big\{\frac{\kappa}{2h_T}(\{\!\{\varrho_h^2\}\!\}[\![e_g]\!],[\![e_g]\!])_{\partial T}+\frac{h_T}{2\kappa}(\{\!\{\varrho_h^2\}\!\}\{\!\{\partial_n\Pi_h e_g\}\!\},\{\!\{\partial_n\Pi_h e_g\}\!\})_{\partial T}\Big\}.$$

Using relation (2.51), we conclude by a lengthy but standard calculation:

$$\begin{aligned}|(\{\!\{\varrho_h^2\}\!\}\{\!\{\partial_n\Pi_h e_g\}\!\},\{\!\{\partial_n\Pi_h e_g\}\!\})_{\partial T}| &\leqslant (\varrho_{h,T}^2+ch_T\varrho_{h,T})\big\|\{\!\{\partial_n\Pi_h e_g\}\!\}\big\|_{\partial T}^2\\ &\leqslant \frac{c_0^2}{h_T}(\varrho_{h,T}^2+ch_T\varrho_{h,T})\big\|\nabla\Pi_h e_g\big\|_T^2\\ &\leqslant \frac{c_0^2}{h_T}\Big\{\Big(1+\frac{\vartheta}{4}\Big)\big\|\varrho_h\nabla\Pi_h e_g\big\|_T^2+\frac{c}{\vartheta}h_T^2\big\|\nabla\Pi_h e_g\big\|_T^2\Big\},\end{aligned}$$

with an arbitrary constant $\vartheta\in(0,1]$. Consequently,

$$\begin{aligned}&\sum_{T\in\mathbb{T}_h}\frac{h_T}{2\kappa}(\{\!\{\varrho_h^2\}\!\}\{\!\{\partial_n\Pi_h e_g\}\!\},\{\!\{\partial_n\Pi_h e_g\}\!\})_{\partial T}\\ &\qquad\leqslant\sum_{T\in\mathbb{T}_h}\frac{c_0^2\mu}{2\kappa}\Big\{\Big(1+\frac{\vartheta}{4}\Big)\big\|\varrho_h\nabla\Pi_h e_g\big\|_T^2+\frac{c}{\vartheta}h_T^2\big\|\nabla\Pi_h e_g\big\|_T^2\Big\}.\end{aligned}$$

From

$$\big\|\nabla\Pi_h e_g\big\|_T^2\leqslant\Big(1+\frac{\vartheta}{4}\Big)\big\|\nabla e_g\big\|_T^2+\frac{c}{\vartheta}\big\|\nabla\eta_g\big\|_T^2,$$

follows

$$\begin{aligned}&\sum_{T\in\mathbb{T}_h}\frac{h_T}{2\kappa}(\{\!\{\varrho_h^2\}\!\}\{\!\{\partial_n\Pi_h e_g\}\!\},\{\!\{\partial_n\Pi_h e_g\}\!\})_{\partial T}\\ &\qquad\leqslant\sum_{T\in\mathbb{T}_h}\frac{c_0^2\mu}{2\kappa}(1+\vartheta)\big\|\varrho_h\nabla e_g\big\|_T^2\\ &\qquad\quad+c\sum_{T\in\mathbb{T}_h}\Big\{h_T^2\big\|\nabla e_g\big\|_T^2+\big\|\varrho_h\nabla\eta_g\big\|_T^2+h_T^2\big\|\nabla\eta_g\big\|_T^2\Big\}.\end{aligned}$$

Hence, observing the results of Lemma 2.3.6, we obtain

$$\sum_{T\in\mathbb{T}_h}\frac{h_T}{2\kappa}(\{\!\{\varrho_h^2\}\!\}\{\!\{\partial_n\Pi_h e_g\}\!\},\{\!\{\partial_n\Pi_h e_g\}\!\})_{\partial T}\leqslant\frac{c_0^2\mu}{2\kappa}(1+\vartheta)\Theta_2+\frac{c}{\vartheta}\,\ell(\varepsilon)h^2. \tag{2.53}$$

Finally, the fourth and the fifth term are estimated in a similar way as before by

$$\begin{aligned}&\Big|\sum_{T\in\mathbb{T}_h}([\![\varrho_h^2]\!][\![\partial_n e_g]\!],[\![e_g]\!])_{\partial T}\Big|\\ &\quad\leqslant\sum_{T\in\mathbb{T}_h}\Big\{\frac{c}{\gamma}h_T\big\|\{\!\{e_g\}\!\}\big\|_{\partial T}^2+c\gamma h_T\big\|\{\!\{\varrho_h\}\!\}\{\!\{\partial_n e_g\}\!\}\big\|_{\partial T}^2\Big\}\\ &\quad\leqslant\sum_{T\in\mathbb{T}_h}\Big\{\frac{c}{\gamma}\big(\big\|e_g\big\|_T^2+h_T^2\big\|\nabla e_g\big\|_T^2\big)+c\gamma\big(\big\|\varrho_h\nabla e_g\big\|_T^2+h_T^2\big\|\varrho_h\nabla^2 g\big\|_T^2\big)\Big\}\\ &\quad\leqslant c\gamma\Theta_2+\frac{c}{\gamma}\,\ell(\varepsilon)h^2,\end{aligned}$$

and by

$$\begin{aligned}
&\Big|\sum_{T\in\mathbb{T}_h}\frac{\kappa}{h_T}([\![\varrho_h^2]\!]\{\!\{e_g\}\!\},[\![e_g]\!])_{\partial T}\Big|\\
&\quad\leqslant\sum_{T\in\mathbb{T}_h}\frac{\kappa}{h_T}\Big\{\big\|[\![\varrho_h]\!]\big\|\big(\frac{\gamma}{h_T}\big\|\{\!\{\varrho\}\!\}[\![e_g]\!]\big\|_{\partial T}^2+\frac{h_T}{\gamma}\big\|\{\!\{e_g\}\!\}\big\|_{\partial T}^2\big)\Big\}\\
&\quad\leqslant\sum_{T\in\mathbb{T}_h}\Big\{c\gamma\frac{\kappa}{h_T}(\{\!\{\varrho_h^2\}\!\}[\![e_g]\!],[\![e_g]\!])_{\partial T}+c\big\|e_g\big\|_T^2+ch_T^2\big\|\nabla e_g\big\|_T^2\Big\}\\
&\quad\leqslant c\gamma\Theta_2+\frac{c}{\gamma}\,\ell(\varepsilon)h^2.
\end{aligned}$$

Collecting these results, we obtain

$$\Theta_3\leqslant\Big(c\gamma+\frac{c_0^2\mu}{2\kappa}(1+\vartheta)\Big)\Theta_2+\Big(\frac{c}{\gamma}+\frac{c}{\vartheta}\Big)\,\ell(\varepsilon)h^2, \tag{2.54}$$

and consequently, in virtue of (2.47) and (2.50),

$$\Theta_2\leqslant\Big(c\gamma+\frac{c_0^2\mu}{2\kappa}(1+\vartheta)\Big)\Theta_2+\Big(\frac{c}{\gamma}+\frac{c}{\vartheta}\Big)\,\ell(\varepsilon)h^2. \tag{2.55}$$

Now, we fix κ according to condition (2.14) and then choose ϑ and γ sufficiently small, such that

$$\Big(c\gamma+\frac{c_*^2}{2\alpha^2\kappa}(1+\vartheta)\Big)\leqslant\mu<1. \tag{2.56}$$

With this choice of the parameters, we obtain that

$$\Theta_2\leqslant c(\vartheta,\gamma)\,\ell(\varepsilon)h^2. \tag{2.57}$$

which together with (2.46) completes the proof. □

2.3.8 Corollary: *Assume that there is no local refinement around the point a, i. e., that there is a mesh-independent constant c such that*

$$ch\leqslant h_a\leqslant h. \tag{2.58}$$

Then, Theorem 2.3.7 implies the point-wise error estimate

$$|e_u(a)|\leqslant c\,\ell(h)\big\|h^2u\big\|_{L^\infty(B)}+c\big|h^2u\big|_{2,2;\mathbb{T}_h}. \tag{2.59}$$

Proof. Let $\varepsilon=h_a$ and $\mathcal{U}_a^\varepsilon=T_a$ in (2.25). There exists a $\delta_a^\varepsilon\in P(T_a)$, such that for any $q\in P(T_a)$ there holds

$$q(a)=\int_{T_a}q(x)\delta_a^\varepsilon(x)dx. \tag{2.60}$$

	$\mathbb{Q}_1$	
	$\|e_h\|_\infty$	$\tilde{e}_{h_L} - \tilde{e}_{h_{L-1}}$
0	5.0e-01	–
1	2.0e-01	-2.83e-01
2	6.3e-02	-2.27e-01
3	2.0e-02	-2.50e-01
4	5.7e-03	-2.04e-01
5	1.6e-03	-2.11e-01
6	4.6e-04	-2.15e-01
7	1.3e-04	-2.20e-01

Table 2.5: L^∞-errors and their scaled differences for discontinuous $\mathbb{Q}_1$ elements.

Choosing $q \equiv 1$, we see that

$$\int_{T_a} \delta_a^\varepsilon(x)dx = 1, \quad \max_{T_a} |\delta_a^\varepsilon| \leqslant c|T_a|^{-1}, \tag{2.61}$$

as is required for the application of Theorem 2.3.7. Using this construction, we derive the estimate

$$|e(a)| \leqslant 2 \max_{T_a} |u - \Pi_h u| + \Big| \int_{T_a} e_u \delta_a^\varepsilon \, dx \Big|. \tag{2.62}$$

Hence, by the approximation properties of $\mathbb{Q}_1$ elements and the result of Theorem 2.3.7, we obtain

$$|e(a)| \leqslant ch^2 \big\|\nabla^2 u\big\|_{L^\infty(T_a)} + ch^2 \, \ell(h) \big\|\nabla^2 u\big\|_{L^\infty(B)} + ch^2 \big\|\nabla^2 u\big\|. \tag{2.63}$$

which completes the proof. □

2.3.9 Remark: It is well known that the logarithmic term in corollary 2.3.8 is observed only on irregular grids in two dimensions with $\mathbb{Q}_1$ or $\mathbb{P}_1$ finite elements. We check this result by solving example A.2.1 on page 170 on the irregular grid in example A.1.4 on page 167.

As can be seen in Figure 2.4, the maxima of the error are located at the two irregular points of the mesh. Since the right graph is scaled up by a factor of 16 (i. e. h^{-2}), the peaks are indeed growing by the logarithmic factor.

2.3.10 The scaled error $\overline{e_u} := h^{-2}\|u - u_h\|_\infty$ admits the asymptotic estimate

$$h_1^{-2}|u(a) - u_{h_1}(a)| - h_2^{-2}|u(a) - u_{h_2}(a)|\tilde{e}_{h_1} - \tilde{e}_{h_2} \approx c \log \frac{h_2}{h_1}.$$

Therefore, this difference should remain constant under refinement, if (2.59) is sharp. In Table 2.5, we display these values together with the L^∞-norm of the errors. This table clearly supports our theoretical result.

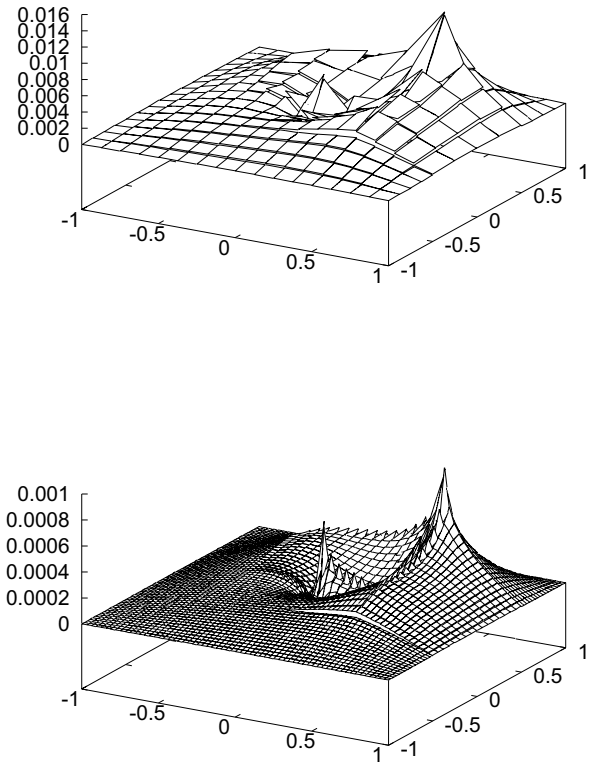

Figure 2.4: Pointwise errors on irregular meshes

2.4 A posteriori error analysis

2.4.1 We will now derive *a posteriori* estimates for the interior penalty scheme (2.12). A similar analysis based on duality arguments has been developed by Becker et al. in [BHL03] and [BHS03], but there the emphasis is on L^2-norm error bounds and on non-matching meshes in the context of domain decomposition. Let $J(\cdot)$ be an arbitrary *linear* functional on V with respect to which the error $e = u - u_h$ is to be estimated. Examples are local averages as considered above, contour integrals or integrals over subdomains:

$$J(u) = |B_\varepsilon|^{-1} \int_{B_\varepsilon} u(x)dx, \quad J(u) = \int_\Gamma u(x)ds, \quad J(u) = \int_{\Omega_0} u(x)dx;$$

for more examples see [BR01]. With the functional $J(\cdot)$, we associate a *dual solution* $z \in V$ as the solution of the auxiliary problem

$$(\nabla\varphi, \nabla z) = J(\varphi) \quad \forall\varphi \in V. \tag{2.64}$$

If the functional $J(\cdot)$ has an L^2 representation j, then the dual problem can be written in strong form as

$$-\Delta z = j \quad \text{in } \Omega, \quad z_{|\partial\Omega_D} = 0, \;\; \partial_n z_{|\partial\Omega_N} = 0. \tag{2.65}$$

Recalling that

$$J(\varphi) = a_h(\varphi, z), \quad \varphi \in V \oplus V_h,$$

we have by Galerkin orthogonality

$$J(e_u) = a_h(e_u, z) = a_h(e_u, z - \psi_h),$$

for arbitrary $\psi_h \in V_h$. Integrating cell-wise by parts and reordering terms, we conclude by elementary but tedious calculation that, for $\eta_z := z - \psi_h$,

$$\begin{aligned}
J(e_u) &= a_h(e_u, z - \zeta_h) \\
&= \sum_{T\in\mathbb{T}_h} \Big\{ (-\Delta e_u, \eta_z)_T + (\partial_n e_u, \eta_z)_{\partial T} - \frac{1}{2}(\{\!\{\partial_n e_u\}\!\}, [\![\eta_z]\!])_{\partial T} \\
&\qquad - \frac{1}{2}([\![e_u]\!], \{\!\{\partial_n \eta_z\}\!\})_{\partial T} + \frac{\kappa}{h_T}([\![e_u]\!], [\![\eta_z]\!])_{\partial T} \Big\} \\
&= \sum_T \Big\{ (f + \Delta u_h, \eta_z)_T - \frac{1}{2}([\![\partial_n u_h]\!], \{\!\{\eta_z\}\!\})_{\partial T} \\
&\qquad - \frac{1}{2}([\![u_h]\!], \{\!\{\partial_n \eta_z\}\!\})_{\partial T} + \frac{\kappa}{h_T}([\![u_h]\!], [\![\eta_z]\!])_{\partial T} \Big\}.
\end{aligned}$$

This is rewritten in the form

$$\begin{aligned}
J(e_u) = \sum_T \eta_T^{(0)} = \sum_T \Big\{ &(R(u_h), \eta_z)_T + \frac{1}{2}(\partial r(u_h), \{\!\{\eta_z\}\!\}) \\
&- \frac{1}{2}(r(u_h), \{\!\{\partial_n \eta_z\}\!\} + \frac{\kappa}{h_T}[\![\eta_z]\!])_{\partial T} \Big\},
\end{aligned} \tag{2.66}$$

with the following notation of cell and edge residuals:

$$\begin{aligned} R(u_h)_{|T} &:= f + \Delta u_h, \\ r(u_h)_{|\Gamma} &:= \left\{ \begin{array}{ll} \frac{1}{2}[\![u_h]\!], & \text{if } \Gamma \subset \partial T \setminus \partial\Omega, \\ u^D - u_h, & \text{if } \Gamma \subset \partial\Omega_D, \\ 0, & \text{if } \Gamma \subset \partial\Omega_N, \end{array} \right\}, \\ \partial r(u_h)_{|\Gamma} &:= \left\{ \begin{array}{ll} \frac{1}{2}[\![\partial_n u_h]\!], & \text{if } \Gamma \subset \partial T \setminus \partial\Omega, \\ 0, & \text{if } \Gamma \subset \partial\Omega_D, \\ u^N - \partial_n u_h, & \text{if } \Gamma \subset \partial\Omega_N, \end{array} \right\}. \end{aligned}$$

From the error representation (2.66), we obtain the following error estimate:

2.4.2 Theorem: *For the error $e = u - u_h$ in the interior penalty scheme, there holds the a posteriori error estimate*

$$J(e_u) = \sum_{T\in\mathbb{T}} \left\{ \varrho_T^{(1)}\omega_T^{(1)} + \varrho_T^{(2)}\omega_T^{(2)} + \varrho_T^{(3)}\omega_T^{(3)} \right\}, \tag{2.67}$$

with the cell residuals $\varrho_T^{(i)}$ and weight factors $\omega_T^{(i)}$ being defined by

$$\begin{aligned} \varrho_T^{(1)} &= \|R(u_h)\|_T, & \omega_T^{(1)} &= \|z - \psi_h\|_T, \\ \varrho_T^{(2)} &= h_T^{-1/2}\|\partial r(u_h)\|_{\partial T}, & \omega_T^{(2)} &= h_T^{1/2}\|z - \psi_h\|_{\partial T}, \\ \varrho_T^{(3)} &= h_T^{-3/2}\|r(u_h)\|_{\partial T}, & \omega_T^{(3)} &= h_T^{3/2}\|\{\!\{\partial_n(z - \psi_h)\}\!\} - \kappa h_T^{-1}[\![z - \psi_h]\!]\|_{\partial T}, \end{aligned}$$

for arbitrary $\psi_h \in V_h$.

2.4.3 Corollary: *The mean quadratic error admits the asymptotically optimal a posteriori estimate*

$$\|e_u\| \leqslant c_I c_S \sum_{T\in\mathbb{T}} \left\{ h_T^2 \varrho_T^{(1)2} + h_T^2 \varrho_T^{(2)2} + h_T \varrho_T^{(3)2} \right\}^{1/2}. \tag{2.68}$$

Proof. Theorem 2.4.2 provides a posteriori estimates for arbitrary functionals of the error. This also includes the L^2-error estimates. To see this, we take the special functional

$$J(\varphi) := \frac{(e_u, \varphi)}{\|e_u\|}.$$

The corresponding dual solution $z \in V$ satisfies $u \in H^2(\Omega)$ and the a priori bound

$$\|\nabla^2 z\| \leqslant c_S, \tag{2.69}$$

where the *stability constant* c_S only depends on the domain Ω. From (2.67) and the interpolation estimates (1.21), (1.25), we infer that

$$\|e_u\| \leqslant c_I c_S \sum_{T\in\mathbb{T}} \left\{ h_T^2 \varrho_T^{(1)2} + h_T^2 \varrho_T^{(2)2} + h_T \varrho_T^{(3)2} \right\}^{1/2}. \tag{2.70}$$

□

2.4.4 Next, we state an a posteriori error estimate for the locally averaged error as considered in our a priori error analysis. In this case the dual solution is just the regularized Green function, $z = g_a^\varepsilon$ introduced in the proof of Theorem 2.3.7. We have the estimate

$$|J_\varepsilon^a(e_u)| \leqslant \eta_1(u_h) := \sum_{T\in\mathbb{T}} \left\{ \varrho_T^{(1)}\omega_T^{(1)} + \varrho_T^{(2)}\omega_T^{(2)} + \varrho_T^{(3)}\omega_T^{(3)} \right\}, \tag{2.71}$$

where the residual terms $\varrho_T^{(i)}$ are as defined above and the weights $\omega_T^{(i)}$ can be estimated as follows:

$$\begin{aligned}
\omega_T^{(1)} &= \|g_a^\varepsilon - \psi_h\|_T, \\
\omega_T^{(2)} &= h_T^{1/2}\|g_a^\varepsilon - \psi_h\|_{\partial T}, \\
\omega_T^{(3)} &= h_T^{3/2}\|\{\!\{\partial_n(g_a^\varepsilon - \psi_h)\}\!\} - \frac{\kappa}{h_T}[\![g_a^\varepsilon - \psi_h]\!]\|_{\partial T}.
\end{aligned}$$

2.4.5 We test our error estimate with the same configuration as in the previous section, that is the solution of example A.2.1 on page 170 on the irregular grid of example A.1.4. The error is evaluated at the point $a = (0.5, 0.5)$, which is the point with maximum error in Figure 2.4 (the value is taken from the solution in the upper right cell adjacent to this point) . We compare three types of "error estimators". The first one, $\eta_0(u_h)$, is obtained directly from the error representation (2.66), avoiding the use of triangle and Hölder inequalities:

$$J(e) = \eta_0(u_h) := \sum_T \eta_T^{(0)}. \tag{2.72}$$

The second estimator uses the local refinement indicators $|\eta_T^{(0)}|$:

$$\eta_1(u_h) := \sum_T \eta_T^{(1)} := \sum_T |\eta_T^{(0)}|.$$

The third one, $\eta_2(u_h)$, is given by Theorem 2.4.2 as described above:

$$\eta_2(u_h) := \sum_T \eta_T^{(2)} := \sum_T \{ \varrho_T^{(1)}\omega_T^{(1)} + \varrho_T^{(2)}\omega_T^{(2)} + \varrho_T^{(3)}\omega_T^{(3)} \}.$$

For practical evaluation of the error estimators, we solve the dual problem (2.64) on the current mesh with bi-quadratic polynomials obtaining $\tilde{z} \in \tilde{V}_h$. We decided for exact computation in the higher order space to avoid additional error contributions. For a more efficient computation, $\tilde{z}$ can be obtained by post-processing; see [BR01] for such strategies and their influence on the estimator.

The quality of the resulting approximate error estimators $\tilde{\eta}_i(u_h)$, $i = 0, 1, 2$, is measured by the "effectivity index":

$$I_{\text{eff}} := \frac{|\tilde{\eta}_i(u_h)|}{|J_a^\varepsilon(e)|}.$$

Mesh adaptation is based on "error indicators" $\eta_T^{(1)}$ and $\eta_T^{(2)}$, respectively. For mesh refinement, a fixed fraction (here 20%) of the grid cells with largest indicator are refined.

L	$e(a)$	$\tilde{\eta}_0(u_h)$	I_{eff}	$\tilde{e}(a)$	$\eta_1(u_h)$	I_{eff}
3	4.306e-3	4.288e-3	0.996	1.82e-5	6.127e-3	1.42
4	1.655e-3	1.652e-3	0.998	2.75e-6	2.296e-3	1.39
5	6.540e-4	6.528e-4	0.998	1.27e-6	9.431e-4	1.44
6	2.920e-4	2.917e-4	0.999	2.66e-7	4.182e-4	1.43
7	1.402e-4	1.401e-4	0.999	8.60e-8	2.000e-4	1.43
8	6.756e-5	6.751e-5	0.999	5.47e-8	9.678e-5	1.43
9	3.376e-5	3.373e-5	0.999	2.47e-8	4.875e-5	1.44
10	1.743e-5	1.742e-5	1.000	3.10e-9	2.522e-5	1.45
11	9.069e-6	9.068e-6	1.000	1.74e-9	1.324e-5	1.46

Table 2.6: Point error $e(a)$ and $\tilde{e}(a) = e(a) + \tilde{\eta}_0(u_h)$ and the corresponding a posteriori estimators $\tilde{\eta}_0(u_h)$ and $\tilde{\eta}_1(u_h)$ (L is maximum refinement level).

L	$e(a)$	$\tilde{\eta}_0(u_h)$	I_{eff}
3	3.877e-03	2.021e-02	5.211
4	1.459e-03	7.310e-03	5.009
5	6.508e-04	3.050e-03	4.686
6	3.146e-04	1.430e-03	4.544
7	1.608e-04	6.985e-04	4.344
8	8.205e-05	3.574e-04	4.355
9	4.203e-05	1.834e-04	4.365
10	8.205e-05	3.574e-04	4.355
11	4.203e-05	1.834e-04	4.365

Table 2.7: Point error $e(a)$ and a posteriori estimator $\tilde{\eta}_2(u_h)$

2.4.6 Table 2.6 presents results for estimators η_0 and η_1 obtained by adaptive refinement based on indicators $\eta_T^{(1)}$. Estimator $\eta_0(u_h)$ is asymptotically optimal while estimator $\eta_1(u_h)$ appears to be off by a factor of about $3/2$. The error representation (2.72) suggests to consider

$$\tilde{J}(u_h) := J(u_h) + \tilde{\eta}_0(u_h)$$

as new approximation. This post-processing step can improve accuracy dramatically, as shown in Table 2.6. The mesh obtained in the eighth step of this iteration is shown in Figure 2.5.

2.4.7 In Table 2.7, we show results for estimator $\eta_2(u_h)$. Since this estimator involves Hölder and triangle inequalities on each cell, there is no cancellation between the different terms of the estimator. Therefore, the error is over-estimated by a factor between 4 and 5. The errors resulting from adaptive refinement —based on the estimator $\eta_2(u_h)$ here— are comparable to those of Table 2.6. Therefore, we conclude that both estimators are suited as refinement criteria.

2.4.8 Finally, in Table 2.8, we compare the "error/mesh-ratio" $e(a) \cdot N$ for uniform and adaptive refinement. On uniformly refined meshes, due to the asymptotic behavior $e(a) \approx h^2\, \ell(h)$ as demonstrated in Table 2.5, we expect $e(a) \cdot N$ to grow like $\ell(1/N)$ while the mesh refinement at the point a should suppress this defect. This is demonstrated by Figure 2.8, where the ratio even improves under adaptive refinement.

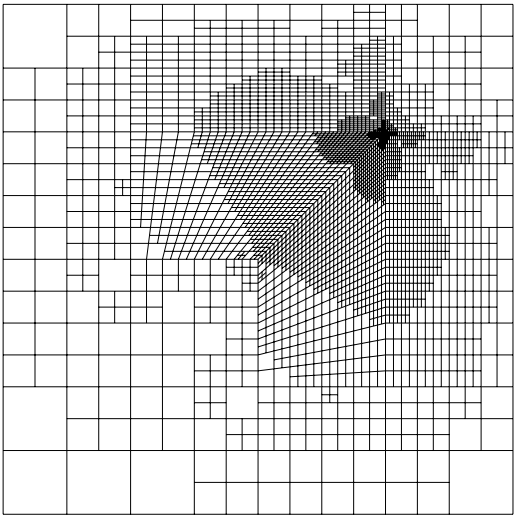

Figure 2.5: Adapted mesh (level $L = 7$) for the point value computation

	uniform refinement		adaptive refinement	
L	$e(a)$	$e(a) \cdot N$	$e(a)$	$e(a) \cdot N$
1	3.811e-02	1.985e+00	3.811e-02	1.985e+00
2	1.126e-02	2.348e+00	1.448e-02	1.433e+00
3	3.212e-03	2.677e+00	4.306e-03	8.404e-01
4	8.956e-04	2.983e+00	1.655e-03	6.261e-01
5	2.464e-04	3.281e+00	6.540e-04	4.804e-01
6	6.716e-05	3.577e+00	2.920e-04	4.238e-01
7	1.818e-05	3.872e+00	1.402e-04	3.945e-01
8	–	–	6.756e-05	3.664e-01
9	–	–	3.376e-05	3.507e-01

Table 2.8: Efficiency of computation of point value $u(a)$ on uniformly and adaptively refined meshes (L refinement level, N number of cells).

2.4.9 Remark: The error representation (2.66) can also be applied to continuous finite element spaces. In that case, $r(u_h) = 0$ on interior edges and we recover the representation in [BR01]. Therefore, (2.66) can be viewed as a straightforward extension of the continuous representation to discontinuous elements.

Here, an advantage of weak boundary conditions becomes clear, since the approximation error at the boundary naturally enters the estimate. Using continuous finite elements with strong boundary conditions, this error must be assessed in a different way.

2.5 Comparison with continuous finite Elements

2.5.1 In this section we do a short experimental comparison with a continuous finite element method (CG). Since we do not have an optimized solver for this method at hand, we only compare discretization errors with respect to degrees of freedom.

2.5.2 Figure 2.6 shows L^2-errors for example A.2.2 on Cartesian grids depending on the number of degrees of freedom for bilinear continuous and discontinuous as well as linear discontinuous polynomial spaces. Four was chosen as the logarithmic base, because this is the approximate increment factor after a single global refinement step. First, we observe that the discontinuous $\mathbb{Q}_1$-element needs about four times as many degrees of freedom than its continuous counterpart. This is exactly the ration of degrees of freedom per cell. Therefore, we conclude that the approximation of continuous and discontinuous elements per grid cell is about the same. For biquadratics, this ratio is slightly better than the expected 9/4.

The spaces $\mathbb{P}_1$ and $\mathbb{P}_2$ have less degrees of freedom per cell than the discontinuous tensor product spaces, while admitting the same asymptotic convergence order. Still, their approximation in relation to the number of degrees of freedom is worse, due to larger interpolation constants. Therefore, the additional numerical effort due to more degrees of freedom due to $\mathbb{Q}_k$-elements pays off.

2.5.3 Considering the number of matrix entries as basis of our comparison, we expect the continuous method to be even more superior, since its stencil is smaller. This can be observed in Figure 2.7 at least compared to the discontinuous tensor product elements. Here, the smaller number of degrees of freedom per cell of the $\mathbb{P}_k$ spaces (which enters quadratically into the number of matrix entries) renders them more efficient than discontinuous $\mathbb{Q}_k$, but still less efficient than continuous elements.

2.5.4 The advantage in degrees of freedom per cell of the continuous element reduces with increasing polynomial degree and the $\mathbb{P}_k$ element eventually has less degrees of freedom per cell. For polynomials of degree 4 in two dimensions, these numbers are 16, 25 and 15 for CG-$\mathbb{Q}_4$, DG-$\mathbb{Q}_4$ and DG-$\mathbb{P}_4$. Nevertheless, Figure 2.8 shows that the efficiencies get closer, but that the missing basis functions of $\mathbb{P}_4$ reduce approximation too much to make it more efficient than the other elements.

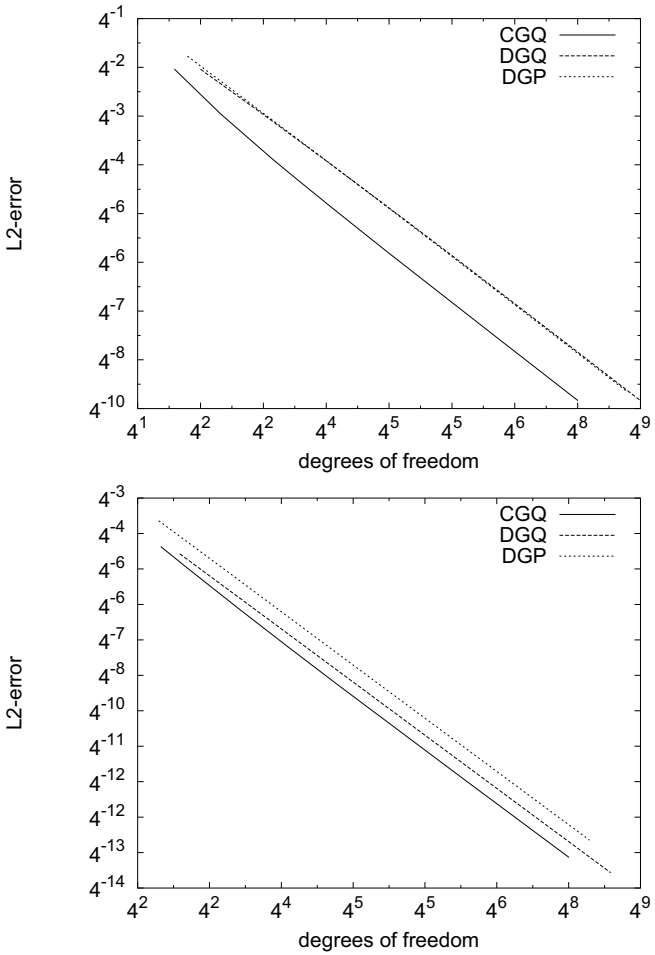

Figure 2.6: L^2-error over number of degrees of freedom for CG-$\mathbb{Q}_1$, DG-$\mathbb{Q}_1$ and DG-$\mathbb{P}_1$ elements (left) and (bi-)quadratic elements (right)

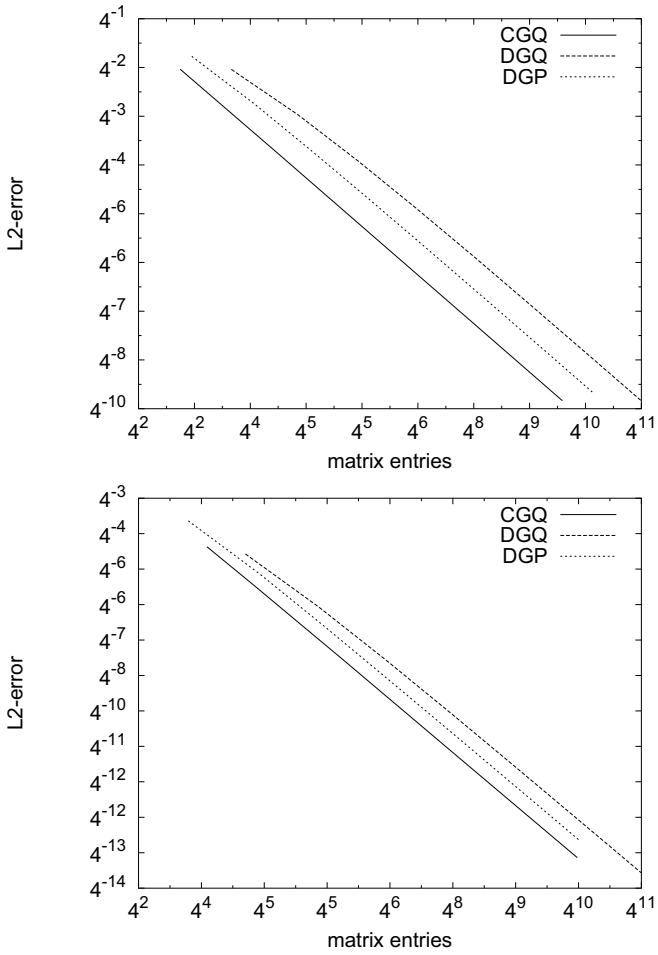

Figure 2.7: L^2-error over number of matrix entries for CG-$\mathbb{Q}_1$, DG-$\mathbb{Q}_1$ and DG-$\mathbb{P}_1$ elements (left) and (bi-)quadratic elements (right)

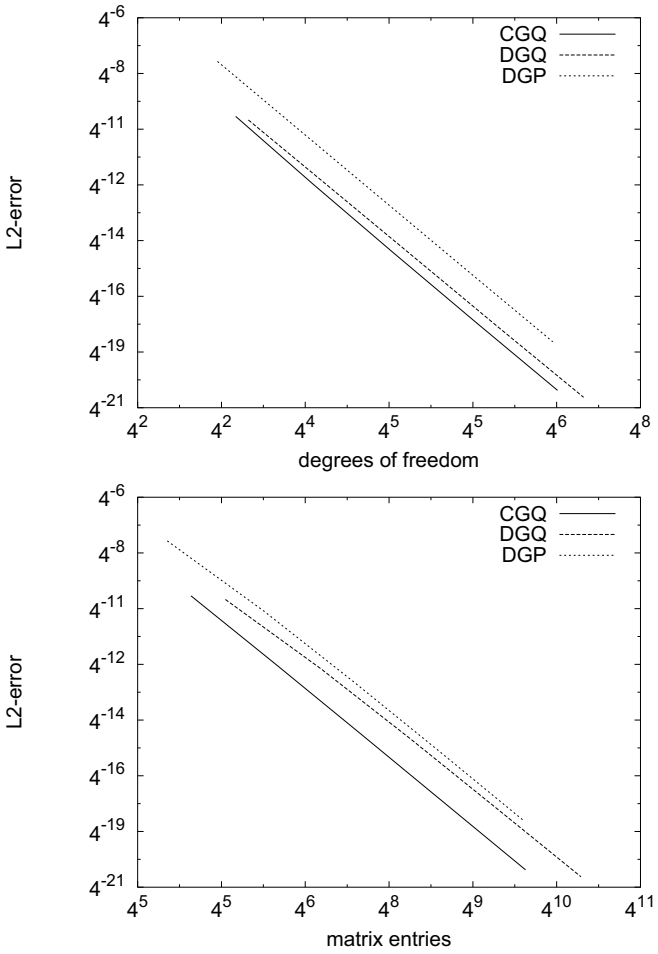

Figure 2.8: L^2-error versus degrees of freedom (left) and matrix entries (right) for polynomials of degree 4

2.5.5 Concluding this comparison, we may say that for Poisson's equation we have to expect a penalty for using discontinuous Galerkin. On the other hand, this penalty is a factor of four in the worst case and decreasing for higher order polynomials. Therefore, we may say that the interior penalty method is not optimal for Poisson equation, but within a reasonable range. Still, we expect it to develop its full potential in singularly perturbed problems, especially with a dominant nonlinear hyperbolic part.

Chapter 3

Linear Diffusion II

This chapter treats discretization of Poisson problems in mixed form by the *local discontinuous Galerkin method* introduced by Cockburn and Shu in [CS98]. We present the analytical results on the so called "standard" method by Castillo et al. from [CCPS00] together with own experimental results assessing the influence of the stabilization parameter. Then, we focus on the "superconvergent" version of this method (see Cockburn, Kanschat et al. [CKPS01]). Here, we show new numerical evidence for applications of the scheme beyond the limits of the presented analysis. Since this method approximates the "stresses" of the mixed formulation better than the standard scheme, it is of special use where accurate values of the derivatives are needed. Finally, a comparison of the methods with respect to computation time (using the efficient solvers from Chapter 6) is conducted.

3.0.1 In this chapter, we consider the mixed formulation of Poisson equation,

$$\begin{aligned} \sigma \quad -\nabla u &= 0 && \text{in} \quad \Omega \\ -\nabla\cdot\sigma \quad\quad &= f \\ u \quad &= u^D && \text{on} \quad \Gamma^D \\ \sigma\cdot\mathbf{n} \quad\quad &= \sigma^N && \text{on} \quad \Gamma^N. \end{aligned} \tag{3.1}$$

3.0.2 Definition: The *dual mixed formulation* of (3.1) is the following variational problem: given the spaces $V^D = L^2(\Omega)$ and

$$\begin{aligned} \Sigma^D = H^{\text{div}}_N(\Omega) &:= \left\{ v \in H^{\text{div}}(\Omega) \middle| v\cdot\mathbf{n} = 0 \text{ on } \Gamma^N \right\}, \\ \left|\tau\right|_{H^{\text{div}}(\Omega)} &:= \left\|\nabla\cdot\tau\right\|_{L^2(\Omega)}. \end{aligned}$$

find $(\sigma, u) \in \Sigma^D \times V^D$ such that

$$\begin{aligned} \left(\sigma,\tau\right) \quad &+ \left(u, \nabla\cdot\tau\right) = 0 \\ -\left(\nabla\cdot\sigma, v\right) \quad &= f, \end{aligned} \tag{3.2}$$

holds for any $\tau \in \Sigma^D$ and $v \in V^D$ In this formulation, $u = 0$ on Γ^D is the natural boundary condition, while $\sigma\cdot\mathbf{n} = 0$ is explicitly built into the space Σ^D (for simplicity, we only consider homogeneous boundary conditions).

3.0.3 Definition: The *primal mixed formulation* of (3.1) is the following variational problem: given the spaces $V^P = W_0^{1,2}(\Omega)$ and $\Sigma^P = L^2(\Omega;\mathbb{R}^d)$ find $(\sigma,u) \in \Sigma^P \times V^P$ such that

$$\begin{array}{ll} (\sigma,\tau) & -(\nabla u,\tau) = 0 \\ (\sigma,\nabla v) & \qquad\quad = f, \end{array} \tag{3.3}$$

holds for any $\tau \in \Sigma^P$ and $v \in V^P$. Here, the natural boundary condition is $\sigma\cdot\mathbf{n} = 0$ on Γ^N and the Dirichlet boundary condition is obeyed by the space V^P.

3.0.4 Remark: Both variational formulations of the mixed Laplacian have a unique solution (u,σ). The primal mixed formulation has properties similar to the purely primal formulation (2.4). The dual mixed formulation allows for accurate solutions if there is a rough diffusion coefficient ν. In this case, $\sigma = \nu\nabla u$ may be a smooth function, even if u is not in $W^{2,2}(\Omega)$.

3.0.5 Discretization of equations (3.2) and (3.3) by finite elements requires an appropriate choice of subspaces $\Sigma_h \subset \Sigma^{D/P}$ and $V_h \subset V^{D/P}$, respectively, such that the *Babuška-Brezzi condition*

$$\inf_{v\in V_h}\sup_{\tau\in\Sigma_h} \frac{(\nabla\cdot\sigma, v)}{|\tau|_{\Sigma^{D/P}}\|v\|_{V^{D/P}}} \geqslant \gamma > 0, \tag{3.4}$$

holds independent of the discretization parameter h (cf. [BF91]). For the dual formulation, the spaces involved are quite intricate (see [BDFM87, BDM85, RT77]), consisting of vector valued shape functions with special continuity requirements at the boundary.

3.0.6 The discretization we introduce now will not set out from a well-posed variational problem like (3.2) or (3.3), but uses stabilization of the bilinear form instead. This is achieved by adding a positive semi-definite bilinear form $c_h(u,v)$ to the second equation in (3.2). This stabilization allows us to conduct a optimal convergence analysis in the pair of Sobolev spaces $\Sigma = W^{1,2}(\mathbb{T}_h;\mathbb{R}^d)$ and $V = W^{1,2}(\mathbb{T}_h)$.

The necessary condition (3.4) for unique solvability can be relaxed to the condition (see[FS91]) that there are two positive numbers γ_1 and γ_2 such that

$$\sup_{\tau\in\Sigma_h} \frac{(\nabla\cdot\tau, v)}{|\tau|_\Sigma} \geqslant \gamma_1\|v\|_V - \gamma_2 c_h(v,v), \tag{3.5}$$

holds for any $v \in V$.

3.1 The LDG Method

3.1.1 Testing (3.1) on a single grid cell $T \in \mathbb{T}_h$ with functions τ and v and integrating by parts yields

$$\begin{array}{lll} (\sigma,\tau)_T & +(u,\nabla\cdot\tau)_T - \langle u, \tau\cdot\mathbf{n}\rangle_{\partial T} = & 0 \\ (\sigma,\nabla v)_T & -\langle\sigma\cdot\mathbf{n}, v\rangle_{\partial T} = & (f,v)_T. \end{array} \tag{3.6}$$

These equations are well-defined for functions $\sigma,\tau \in \Sigma = W^{1,2}(\mathbb{T}_h;\mathbb{R}^d)$ and $u,v \in V = W^{1,2}(\mathbb{T}_h)$.

3.1.2 Assumption: *The discrete subspaces of Σ and V are chosen as $\left(X^Q_{h,k_\sigma}\right)^d$ and X^Q_{h,k_u}, respectively. Here, $k_u \geqslant 1$ and $k_\sigma \in \{k_u - 1, k_u\}$. In the case of parallelogram cells, we also allow the spaces $X^P_{h,k}$.*

3.1.3 Using these discrete spaces and replacing the traces on the cell boundaries in (3.6) by numerical fluxes $\hat{\sigma}_h$ and $\hat{u}_h$ to be specified below, the generic LDG method reads as follows: find $(\sigma_h, u_h) \in \Sigma_h \times V_h$ such that for all $(\tau_h, v_h) \in \Sigma_h \times V_h$

$$\begin{array}{ll} \big(\sigma_h, \tau_h\big)_T & + \big(\hat{u}_h, \nabla\cdot\tau_h\big)_T - \big\langle \hat{u}_h, \tau_h\cdot\mathbf{n}\big\rangle_{\partial T} = 0 \\ \big(\sigma_h, \nabla v_h\big)_T & - \big\langle \hat{\sigma}_h\cdot\mathbf{n}, v_h\big\rangle_{\partial T} = \big(f, v_h\big)_T. \end{array} \tag{3.7}$$

3.1.4 With the flux operators of Definition 1.2.6, we define the numerical fluxes generally $\hat{\sigma}$ and $\hat{u}$ as

$$\hat{\sigma} = \begin{cases} \{\!\{\sigma\}\!\} - \gamma_{uu}[\![u\mathbf{n}]\!] - \gamma_{u\sigma}[\![\sigma\cdot\mathbf{n}]\!] & \text{for } E\in\mathbb{E}_h^i, \\ \sigma - \gamma_{uu}\big(u\mathbf{n} - u^D\mathbf{n}\big) & \text{for } E\in\mathbb{E}_h^D, \\ \sigma^N\cdot\mathbf{n} & \text{for } E\in\mathbb{E}_h^N, \end{cases} \tag{3.8}$$

$$\hat{u} = \begin{cases} \{\!\{u\}\!\} + \gamma_{\sigma u}\cdot[\![u\mathbf{n}]\!] - \gamma_{\sigma\sigma}[\![\sigma\cdot n]\!] & \text{for } E\in\mathbb{E}_h^i, \\ u^D & \text{for } E\in\mathbb{E}_h^D, \\ u - \gamma_{\sigma\sigma}\big(\sigma\cdot\mathbf{n} - \sigma^N\big) & \text{for } E\in\mathbb{E}_h^N. \end{cases} \tag{3.9}$$

The scalar parameters γ_{uu} and $\gamma_{\sigma\sigma}$, as well as the vector valued parameters $\gamma_{u\sigma}$ and $\gamma_{\sigma u}$ will be fixed below. Right away, we choose $\gamma_{\sigma u} = -\gamma_{u\sigma}^T$ to maintain symmetry of the system.

Numerical fluxes defined this way are consistent in the sense that a "continuous" solution $(\sigma, u) \in W^{1,2}(\Omega;\mathbb{R}^{d+1})$ fulfills $\hat{\sigma}\cdot\mathbf{n} = \sigma\cdot\mathbf{n}$ and $\hat{u} = u$.

3.1.5 For the analysis of the scheme (3.7), we rearrange the terms, so that we obtain

$$\begin{array}{ll} a_h(\sigma_h,\tau) & + b_h(u_h,\tau) = F_h(\tau) \qquad \forall \tau\in\Sigma_h \\ -b_h(v,\sigma_h) & + c_h(u_h,v) = G_h(v) \qquad \forall v\in V_h. \end{array} \tag{3.10}$$

The bilinear and linear forms are

$$a_h(\sigma,\tau) = \big(\sigma,\tau\big)_{\mathbb{T}_h} + \big\langle \gamma_{\sigma\sigma}[\![\sigma\cdot\mathbf{n}]\!], [\![\tau\cdot\mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i} + \big\langle \gamma_{\sigma\sigma}\sigma\cdot\mathbf{n}, \tau\cdot\mathbf{n}\big\rangle_{\mathbb{E}_h^N}, \tag{3.11}$$

$$b_h(u,\tau) = \big(u, \nabla\cdot\tau\big)_{\mathbb{T}_h} - \big\langle \{\!\{u\}\!\} + \gamma_{\sigma u}\cdot[\![u\mathbf{n}]\!], [\![\tau\cdot\mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i} - \big\langle u, \tau\cdot\mathbf{n}\big\rangle_{\mathbb{E}_h^N}, \tag{3.12}$$

$$c_h(u,v) = \big\langle \gamma_{uu}[\![u\mathbf{n}]\!], [\![v\mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i} + \big\langle \gamma_{uu}u, v\big\rangle_{\mathbb{E}_h^D}, \tag{3.13}$$

and

$$F_h(\tau) = \big\langle u^D, \tau\cdot\mathbf{n}\big\rangle_{\mathbb{E}_h^D} + \big\langle \gamma_{\sigma\sigma}\sigma^N, \tau\cdot\mathbf{n}\big\rangle_{\mathbb{E}_h^N} \tag{3.14}$$

$$G_h(v) = \big(f, v\big)_{\mathbb{T}_h} + \big\langle \gamma_{uu}u^D, v\big\rangle_{\mathbb{E}_h^D} + \big\langle \sigma^N, v\big\rangle_{\mathbb{E}_h^N}. \tag{3.15}$$

Defining

$$\begin{aligned} \mathcal{L}_h\big(\sigma,u;\tau,v\big) &:= a_h(\sigma,\tau) + b_h(u,\tau) - b_h(v,\sigma) + c_h(u,v) \\ \mathcal{F}_h(\tau;v) &:= F_h(\tau) + G_h(v), \end{aligned}$$

we arrive at the weak formulation: find $(\sigma_h, u_h) \in \Sigma_h \times V_h$ such that

$$\mathcal{L}_h\big(\sigma_h, u_h; \tau, v\big) = \mathcal{F}_h(\tau; v) \quad \forall (\tau, v) \in \Sigma_h \times V_h. \tag{3.16}$$

3.1.6 Lemma: *Let $\Gamma^D \neq \emptyset$ and let the spaces Σ_h and V_h be chosen according to Assumption 3.1.2. Furthermore, let $\gamma_{uu} > 0$ and $\gamma_{\sigma\sigma} \geqslant 0$ on $\mathbb{E}_h$. Then, the discrete system* (3.10) *with fluxes chosen according to* (3.8) *and* (3.9) *is uniquely solvable in $\Sigma_h \times V_h$.*

Proof. (from [CCPS00]) We show that the only solution for homogeneous data is zero. Testing (3.16) with σ_h and u_h yields

$$a_h(\sigma_h, \sigma_h) + c_h(u_h, u_h) = 0.$$

Under the assumptions on the parameters, we conclude

$$\|\sigma_h\|^2 \leqslant a_h(\sigma_h, \sigma_h),$$

and

$$c_h(u_h, u_h) \geqslant 0.$$

Therefore, $\sigma_h = 0$ in Ω and $[\![u_h]\!] = 0$ on $\mathbb{E}_h^i$. Consequently, u_h is continuous in $\bar{\Omega}$. From the first equation of (3.10), we deduce

$$b(u_h, \tau) = 0 \quad \forall \tau \in \Sigma_h.$$

Integrating by parts in (3.12) yields

$$b_h(u_h, \tau) = -\big(\nabla u_h, \tau\big)_{\mathbb{T}_h} + \big\langle \{\!\{\tau\}\!\} - \gamma_{\sigma u}[\![\tau \cdot \mathbf{n}]\!], [\![u_h \mathbf{n}]\!] \big\rangle_{\mathbb{E}_h^i} + \big\langle u_h, \tau \cdot \mathbf{n} \big\rangle_{\mathbb{E}_h^D}, \tag{3.17}$$

and therefore we obtain

$$b_h(u_h, \tau) = \big(\nabla u_h, \tau\big)_{\mathbb{T}_h} = 0 \quad \forall \tau \in \Sigma_h. \tag{3.18}$$

Now, applying Lemma 1.2.18 to all grid cells starting at the Dirichlet boundary yields $u_h \equiv 0$. □

3.1.7 Remark: By choosing $\gamma_{\sigma\sigma} = 0$, the form $a_h(.,.)$ does not contain any boundary terms between elements and is just the L^2-scalar product on Σ_h. Therefore, the scheme is called *local discontinuous Galerkin* (*LDG*) method. Consequently, the resulting matrix is a block diagonal matrix where the size of each block is the number of degrees of freedom on a single cell. Independent of the choice of the polynomial basis, this matrix can be inverted efficiently in a preprocess before starting an iterative linear solver.

3.1.8 The remaining parameters are used for tuning the scheme. In Table 3.1 on the next page, we give an overview over the convergence orders achieved with different parameter values, summarizing the results of the following sections.

3.2 The standard LDG method

3.2.1 Choice of the parameter $\gamma_{u\sigma}$ leads to different flavors of the LDG method. First, in the *standard LDG* method, we set $\gamma_{u\sigma} = 0$ and choose γ_{uu} as

$$\gamma_{uu} = \begin{cases} \widehat{\gamma_{uu}} \max\left\{\dfrac{1}{h_{T^+}}, \dfrac{1}{h_{T^-}}\right\} & \text{for } x \in T^+ \cap T^- \\ 2\dfrac{\widehat{\gamma_{uu}}}{h_T} & \text{for } x \in T \cap \partial\Omega, \end{cases} \tag{3.19}$$

where $\widehat{\gamma_{uu}} > 0$ is independent of the mesh size, but may depend on the shape of the mesh cells. In our experiments, it will always be constant.

γ_{uu}	$\gamma_{u\sigma}$	$\|\sigma - \sigma_h\|$	$\|u - u_h\|$
$\mathcal{O}(1/h)$	0	k	$k+1$
$\mathcal{O}(1/h)$	upwind	k	$k+1$
$\mathcal{O}(1)$	0	k	$k+\frac{1}{2}$
$\mathcal{O}(1)$	upwind	$k+\frac{1}{2}$	$k+1$

Table 3.1: Orders of convergence of the LDG method for different choices of the flux parameters

3.2.2 With the form $\mathcal{L}_h\big(.,.;.,.\big)$, we associate the *energy seminorm* or *LDG seminorm*

$$\big|(\sigma,u)\big|_{\text{LDG}} := \sqrt{\mathcal{L}_h\big(\sigma,u;\sigma,u\big)}. \tag{3.20}$$

In case of the standard LDG method, this seminorm is

$$\big|(\sigma,u)\big|_{\text{LDG}} = \Big(\big\|\sigma\big\|^2 + \big\|\gamma_{uu}^{-\frac{1}{2}}[\![u]\!]\big\|^2_{\mathbb{E}_h^i} + \big\|\gamma_{uu}^{-\frac{1}{2}}u\big\|^2_{\mathbb{E}_h^D}\Big)^{\frac{1}{2}}$$

3.2.3 Theorem: *Let $s \leqslant k_u + 1$ where k_u is the polynomial degree of the space V_h according to Assumption 3.1.2 and let $(\sigma,u) \in W^{s,2}(\mathbb{T}_h;\mathbb{R}^d) \times W^{s,2}(\mathbb{T}_h)$ be a solution to the mixed Poisson problem* (3.1)*. Let $(\sigma_h,u_h) \in \Sigma_h \times V_h$ be a solution to the LDG formulation* (3.10) *with parameters $\gamma_{\sigma\sigma},\gamma_{u\sigma} = 0$ and γ_{uu} according to* (3.19)*. Furthermore, let assumptions 1.1.13 and 1.1.14 on the triangulation hold. Then, the error admits the estimate*

$$\big|(e_\sigma,e_u)\big|_{LDG} \leqslant c\big(\big|h^{s-1}\sigma\big|_{s-1,2;\mathbb{T}_h} + \big|h^{s-1}u\big|_{s,2;\mathbb{T}_h}\big). \tag{3.21}$$

Proof. The proof of Theorem 3.2.3 will follow the generic proof of Theorem 1.3.9. Therefore, we must prove concretizations of assumptions 1.3.5 to 1.3.8. We observe that the stability estimate (1.32) in Assumption 1.3.5 holds with a constant $c = 1$ and $w = v = (\tau_h, v_h)$ due to the definition of the LDG-seminorm in (3.20). The following lemmas provide Assumption 1.3.7 (in its more general form Assumption 1.3.12) and Assumption 1.3.8. □

3.2.4 Lemma: *Let $\sigma \in W^{s-1,2}(\Omega;\mathbb{R}^d)$ and $\tau \in W^{t+1,2}(\Omega;\mathbb{R}^d)$, furthermore $u \in W^{s,2}(\Omega)$ and $v \in W^{t+2,2}(\Omega)$ with $s \leqslant k_u + 1$ and $t \leqslant k_u - 1$, where k_u is the polynomial degree of the space V_h and Σ_h is chosen according to Assumption 3.1.2. Then, the L^2-projection errors η_σ, η_τ, η_u and η_v admit the estimates*

$$\begin{aligned}\big|\mathcal{L}_h\big(\eta_\sigma,\eta_u;\eta_\tau,\eta_v\big)\big| \leqslant c\big(\big|h^{s-1}\sigma\big|_{s-1,2;\mathbb{T}_h} + \big|h^{s-1}u\big|_{s,2;\mathbb{T}_h}\big)\times \\ \times\big(\big|h^{t+1}\tau\big|_{t+1,2;\mathbb{T}_h} + \big|h^{t+1}v\big|_{t+2,2;\mathbb{T}_h}\big),\end{aligned}$$

and

$$0 \leqslant \mathcal{L}_h\big(\eta_\sigma,\eta_u;\eta_\sigma,\eta_u\big) \leqslant c\big(\big|h^{s-1}\sigma\big|^2_{s-1,2;\mathbb{T}_h} + \big|h^{s-1}u\big|^2_{s,2;\mathbb{T}_h}\big)$$

Furthermore, if for the diffusion coefficient holds $\nu \in W^{s-1,\infty}(\mathbb{T}_h)$, then

$$\big|\mathcal{L}_h\big(\eta_\sigma,\eta_u;\eta_\sigma,\eta_u\big)\big| \leqslant c\big|h^{s-1}u\big|^2_{s,2;\mathbb{T}_h}$$

Proof. For functions $\sigma, \tau \in \Sigma \oplus \Sigma_h$ and $u, v \in V \oplus V_h$, we write

$$\mathcal{L}_h\big(\sigma, u; \tau, v\big) = a_h(\sigma, \tau) + b_h(u, \tau) - b_h(v, \sigma) + c_h(u, v),$$

and estimate each term separately. First,

$$a_h(\sigma, \tau) \leqslant \big\|\tfrac{1}{\nu}\big\|_\infty \|\sigma\| \|\tau\|. \tag{3.22}$$

Using form (3.17) of $b_h(.,.)$, we obtain by trace estimate (1.23)

$$\begin{aligned}
b_h(u, \tau) &= -\big(\nabla u, \tau\big)_{\mathbb{T}_h} + \big\langle \{\!\{\tau\}\!\} - \gamma_{\sigma u}[\![\tau \cdot \mathbf{n}]\!], [\![u_h \mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i} + \big\langle u, \tau \cdot \mathbf{n}\big\rangle_{\mathbb{E}_h^D} \\
&\leqslant c \sum_{T \in \mathbb{T}_h} \Big(|u|_{1,2;T} \|\tau\|_T + \big\|h_T^{-\frac{1}{2}} u\big\|_{\partial T} \big\|h_T^{\frac{1}{2}} \tau\big\|_{\partial T} \Big) \\
&\leqslant c \sum_{T \in \mathbb{T}_h} \Big(|u|_{1,2;T} \|\tau\|_T + \big(|u|_{1,2;T} + h_T^{-1} \|u\|_T\big) \big(|h_T \tau|_{1,2;T} + \|\tau\|_T\big) \Big) \\
&\leqslant c\big(|u|_{1,2;\mathbb{T}_h} + \big\|h^{-1} u\big\|\big) \big(|h\tau|_{1,2;\mathbb{T}_h} + \|\tau\|\big).
\end{aligned}$$

Due to symmetry, we obtain the same for the third term. Finally,

$$\begin{aligned}
c_h(u, v) &\leqslant c \sum_{T \in \mathbb{T}_h} \big\|h_T^{-\frac{1}{2}} u\big\|_{\partial T} \big\|h_T^{-\frac{1}{2}} v\big\|_{\partial T} \\
&\leqslant c\big(|u|_{1,2;\mathbb{T}_h} + \big\|h^{-1} u\big\|\big) \big(|v|_{1,2;\mathbb{T}_h} + \big\|h^{-1} v\big\|\big)
\end{aligned}$$

Entering η_σ for σ, η_u for u and so on and approximation result 1.2.13 thus yield the first estimate of the lemma. The second follows by entering η_σ for τ and η_u for v. The third estimate follows by observing that $\sigma = \nu \nabla u \in W^{s-1,2}(\Omega; \mathbb{R}^d)$. □

3.2.5 Lemma: *Assume that $|\gamma_{u\sigma}|$ is independent of h and γ_{uu} chosen according to* (3.19)*. Let $\tau \in \Sigma_h$, $v \in V_h$. Furthermore, let $\sigma \in W^{s-1,2}(\mathbb{T}_h)$ and $u \in W^{s,2}(\mathbb{T}_h)$ with $s \leqslant k_u + 1$ and (η_σ, η_u) the L^2-projection error. Then, the following estimate holds for $h \to 0$:*

$$\big|\mathcal{L}_h\big(\tau, v; \eta_\sigma, \eta_u\big)\big| \leqslant c\big|(\tau, v)\big|_{LDG} \Big(\big|h^{s-1}\sigma\big|_{s-1,2;\mathbb{T}_h} + \big|h^{s-1} u\big|_{s,2;\mathbb{T}_h} \Big). \tag{3.23}$$

Proof. We set out with

$$\big|\mathcal{L}_h\big(\tau, v; \sigma, u\big)\big| \leqslant |a_h(\tau, \eta_\sigma)| + |b_h(v, \eta_\sigma)| + |b_h(\eta_u, \tau)| + |c_h(\eta_u, v)|,$$

and estimate each term separately.

$$|a_h(\tau, \eta_\sigma)| \leqslant \big\|\nu^{\frac{1}{2}}\big\|_\infty \|a_h(\tau, \tau)\|^{\frac{1}{2}} \big\|\eta_\sigma\big\|, \tag{3.24}$$

$$|c_h(v, \eta_u)| \leqslant c\Big(\big\|\gamma_{uu}^{\frac{1}{2}} [\![v]\!]\big\|_{\mathbb{E}_h^i}^2 + \big\|\gamma_{uu}^{\frac{1}{2}} v\big\|_{\mathbb{E}_h^D}^2 \Big)^{\frac{1}{2}} \big\|\gamma_{uu}^{\frac{1}{2}} \eta_u\big\|_{\mathbb{E}_h^i \cup \mathbb{E}_h^D}. \tag{3.25}$$

Since we have to avoid derivatives on u and σ, we use $b(.,.)$ in the forms (3.12) and (3.17), respectively. Using the orthogonality of the L^2-projection, the contributions of the cells vanish and we obtain

$$\begin{aligned}
|b_h(v, \eta_\sigma)| &\leqslant c\big(\big\|\gamma_{uu}^{\frac{1}{2}} [\![v]\!]\big\|_{\mathbb{E}_h^i} + \big\|\gamma_{uu}^{\frac{1}{2}} v\big\|_{\mathbb{E}_h^D} \big) \big\|\gamma_{uu}^{-\frac{1}{2}} \eta_\sigma\big\|_{\mathbb{E}_h^i \cup \mathbb{E}_h^D} \\
&\leqslant c\big|(\tau, v)\big|_{\mathrm{LDG}} \big\|\gamma_{uu}^{-\frac{1}{2}} \eta_\sigma\big\|_{\mathbb{E}_h^i \cup \mathbb{E}_h^D},
\end{aligned} \tag{3.26}$$

and, using inverse estimate,

$$
\begin{aligned}
&|b_h(\eta_u,\tau)| \\
&\qquad\leqslant c\big(\big\|h^{-\frac12}[\![\eta_u]\!]\big\|^2_{\mathbb{E}_h^i}+\big\|h^{-\frac12}\eta_u\big\|^2_{\mathbb{E}_h^N}\big)^{\frac12}\big(\big\|h^{\frac12}[\![\tau\cdot\mathbf{n}]\!]\big\|^2_{\mathbb{E}_h^i}+\big\|h^{\frac12}\tau\cdot\mathbf{n}\big\|^2_{\mathbb{E}_h^N}\big)^{\frac12} \\
&\qquad\leqslant c\big|(\tau,v)\big|_{\mathrm{LDG}}\big\|h^{-\frac12}\eta_u\big\|_{\mathbb{E}_h^i\cup\mathbb{E}_h^N}. \quad (3.27)
\end{aligned}
$$

Summing up yields

$$
\begin{aligned}
\big|\mathcal{L}_h(\tau,v;\sigma,u)\big| &\leqslant c\big|(\tau,v)\big|_{\mathrm{LDG}}\left(\|\eta_\sigma\|+\big\|\gamma_{uu}^{-\frac12}\eta_\sigma\big\|_{\mathbb{E}_h^i\cup\mathbb{E}_h^D}+\big\|\big(h^{-\frac12}+\gamma_{\widehat{u}u}^{\frac12}\big)\eta_u\big\|_{\mathbb{E}_h}\right), \\
&\leqslant c\big|(\tau,v)\big|_{\mathrm{LDG}}\left(\big|h^{s-1}\sigma\big|_{s-1,2;\mathbb{T}_h}+\big|h^{s-1}u\big|_{s,2;\mathbb{T}_h}\right)
\end{aligned}
$$

□

3.2.6 Remark: Reviewing the proof of Theorem 3.2.3, we can extend the result to a more general choice of the parameters. First, if $\gamma_{u\sigma}$ is any quantity obeying the estimate $\gamma_{u\sigma}\leqslant c$ independent of the mesh size h, all estimates remain valid and so does the theorem.

$\gamma_{uu}=\mathcal{O}(h^{-1})$ is the optimal value in view of estimate (3.27). Using $\gamma_{uu}=1$ instead will cause the the energy estimate to be sub-optimal by $\sqrt{h}$ with respect to the approximation of σ due to estimate (3.26)*.

3.2.7 Theorem: *Let the assumptions of Theorem 3.2.3 hold and additionally estimate* (2.5) *with* $\alpha=t+1$, *where* $0\leqslant t\leqslant k_u-1$. *Then, the error* e_u *is bounded by the negative norm estimate*

$$
\big\|e_u\big\|_{-t,2;\Omega}\leqslant\big|h^{s+t}u\big|_{s,2;\Omega}. \quad (3.28)
$$

Proof. In view of the generic Theorem 1.3.14, it will be sufficient to verify Assumptions 1.3.12 and 1.3.13. Let $z_u\in W^{t+2,2}(\Omega)$ be the solution of the dual problem (2.4) with right hand side $\varphi\in W^{t,2}(\Omega)$ and $z_\sigma=\nu\nabla z_u\in W^{t+1,2}(\Omega)$. Then,

$$
\big|z_\sigma\big|_{W^{t+1,2}(\Omega)}\leqslant c\big\|\nu\big\|_{W^{t+1,\infty}(\Omega)}\big|u\big|_{W^{t+2,2}(\Omega)}\leqslant c\big\|\varphi\big\|_{W^{t,2}(\Omega)}.
$$

Therefore, Lemma 3.2.4 and Lemma 3.2.5 (with $s=t+2$) yield the required estimates. □

3.2.8 Corollary: *If the domain is convex in the vicinity of every corner of its boundary and the solution* u *in* $W^{k_u+1,2}(\mathbb{T}_h)$, *then*

$$
\big\|e_u\big\|\leqslant c\big|h^{k_u+1}u\big|_{k_u+1,2;\Omega}
$$

3.2.9 We give numerical evidence of the sharpness of these estimates by solving example A.2.2 on page 170 on Cartesian (see example A.1.1 on page 167) and distorted meshes (see example A.1.3), respectively. Table 3.2 shows indeed the predicted convergence orders for σ and u with $\mathbb{Q}_k$ and $\mathbb{P}_k$ elements†.

*We have not considered $k_\sigma<k_u$ in detail since experiments showed that the gain in accuracy with equal order elements always outweighed the additional degrees of freedom.

†The loss of order in the last row for (bi-)quadratic elements is due to the solver accuracy of 10^{-12}

L	$\|e_\sigma\|$		$\|e_u\|$		$\|e_\sigma\|$		$\|e_u\|$	
	err	ord	err	ord	err	ord	err	ord
	Cartesian mesh							
	$\mathbb{Q}_1$				$\mathbb{Q}_2$			
1	2.01e-1	—	5.13e-2	—	5.04e-3	—	1.82e-3	—
2	1.50e-1	0.42	1.55e-2	1.72	1.22e-3	2.04	2.29e-4	2.99
3	9.39e-2	0.68	4.30e-3	1.85	3.06e-4	2.00	2.86e-5	3.00
4	5.21e-2	0.85	1.14e-3	1.92	7.70e-5	1.99	3.58e-6	3.00
5	2.74e-2	0.93	2.95e-4	1.95	1.93e-5	2.00	4.48e-7	3.00
6	1.41e-2	0.96	7.52e-5	1.97	4.82e-6	2.00	5.65e-8	2.99
7	7.12e-3	0.98	1.90e-5	1.98	1.20e-6	2.00	1.02e-8	2.47
	$\mathbb{P}_1$				$\mathbb{P}_2$			
1	2.01e-1	—	8.31e-2	—	5.09e-3	—	1.82e-3	—
2	1.49e-1	0.43	2.21e-2	1.91	1.22e-3	2.06	2.29e-4	2.99
3	9.39e-2	0.67	5.80e-3	1.93	3.06e-4	2.00	2.86e-5	3.00
4	5.21e-2	0.85	1.50e-3	1.95	7.70e-5	1.99	3.58e-6	3.00
5	2.74e-2	0.93	3.83e-4	1.97	1.93e-5	2.00	4.48e-7	3.00
6	1.41e-2	0.96	9.69e-5	1.98	4.83e-6	2.00	5.65e-8	2.99
7	7.12e-3	0.98	2.44e-5	1.99	1.20e-6	2.00	1.02e-8	2.47
	distorted mesh							
	$\mathbb{Q}_1$				$\mathbb{Q}_2$			
1	2.00e-1	—	5.31e-2	—	7.28e-3	—	2.14e-3	—
2	1.55e-1	0.37	1.69e-2	1.65	1.81e-3	2.00	3.07e-4	2.81
3	9.80e-2	0.66	4.92e-3	1.78	5.11e-4	1.83	4.28e-5	2.84
4	5.45e-2	0.84	1.36e-3	1.86	1.32e-4	1.95	5.82e-6	2.88
5	2.88e-2	0.92	3.65e-4	1.90	3.38e-5	1.97	7.80e-7	2.90
6	1.48e-2	0.95	9.62e-5	1.92	8.76e-6	1.95	1.03e-7	2.93
7	7.57e-3	0.97	2.50e-5	1.95	2.26e-6	1.95	1.34e-8	2.94

Table 3.2: Convergence of the standard LDG method on Cartesian and distorted meshes

3.2.10 Since the size of $\widehat{\gamma_{uu}}$ is not relevant for stability, as long as it is positive, it is interesting to examine its influence on accuracy. Figure 3.1 shows that the accuracy of the scheme for different polynomial degrees (values have been scaled by suitable powers of 2 to fit into the same figure; furthermore, the mesh sizes have been chosen for each polynomial degree such that the error was in the asymptotic regime and well above machine accuracy).

A qualitative difference between odd and even order polynomials can be observed: the norm of the error remains constant at its minimum if $\widehat{\gamma_{uu}}$ tends to infinity and the polynomial degree is odd. in this case, the error grows if $\widehat{\gamma_{uu}}$ becomes small, an effect reducing fast for higher order polynomials.

The figure on the right shows that the error grows with $\widehat{\gamma_{uu}}$, while it remains constant if $\widehat{\gamma_{uu}}$ is small (at least in the range investigated here). This different behavior off odd and even polynomials should be explained by improved analysis, later.

We conclude that the LDG method is much more robust with respect to the stabilization parameter than the interior penalty method. Figure 3.2 shows, that this is not restricted to Cartesian grids but holds on distorted grids, too.

3.2.11 Since the bilinear form $a_h(.,.)$ of the LDG method does not introduce coupling between mesh cells (see Remark 3.1.7), the variable σ can be eliminated from the system on a cell by cell basis, very much like static condensation for standard higher order elements. As result of this condensation, we obtain a bilinear form $s_h(.,.)$ on $V_h \times V_h$, the Schur complement form of the LDG method.

This bilinear form among others is analyzed in [ABCM01]. Here, we will quote the most important results for our work from this article.

3.2.12 Lemma: *For any positive parameter $\widehat{\gamma_{uu}}$, the form $s_h(.,.)$ is a bounded, coercive bilinear form on V_h equipped with the norm $\|.\|_{1;h}$. Namely, there are constants c and C independent of h such that for $u, v \in V_h$*

$$s_h(u,v) \leqslant C\|u\|_{1;h}\|v\|_{1;h} \tag{3.29}$$

$$s_h(u,u) \geqslant c\|(\|_{1;h}u)^2. \tag{3.30}$$

3.2.13 Corollary: *The bilinear forms $a_h^{IP}(.,.)$ and $s_h(.,.)$ are spectrally equivalent in the sense that there are constants c and C independent of h such that for $u \in V_h$*

$$ca_h^{IP}(u,u) \leqslant s_h(u,u) \leqslant Ca_h^{IP}(u,u). \tag{3.31}$$

3.2.14 Remark: Due to the condensation process, the Schur complement form $s_h(.,.)$ contains couplings not only between degrees of freedom of a cell and its neighbours, but also to their neighbors. Therefore, in practical computations, the full system will usually be preferred.

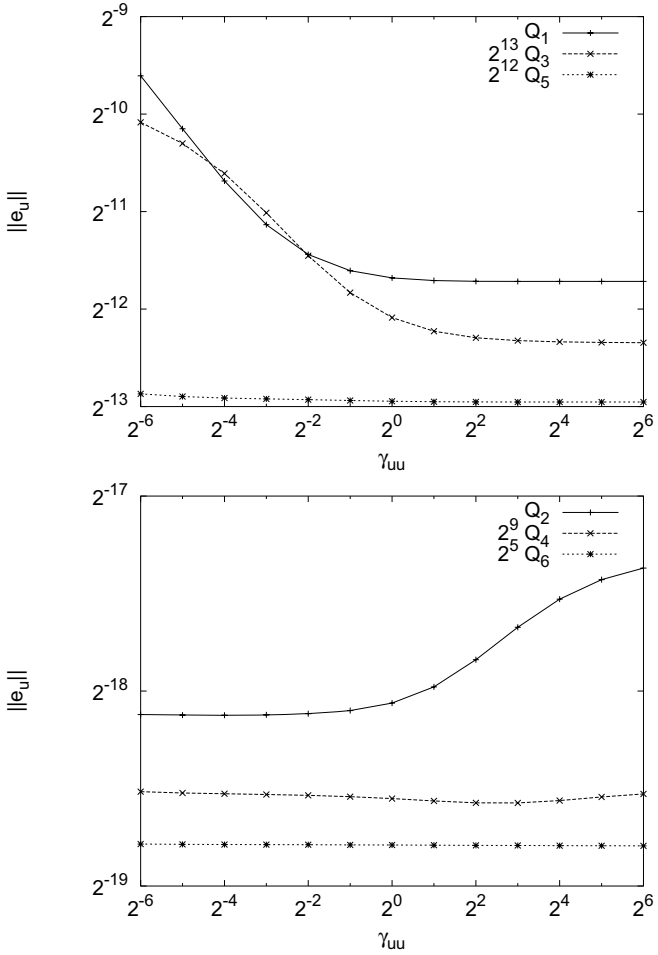

Figure 3.1: Accuracy of the standard LDG method on Cartesian grids depending on $\widehat{\gamma_{uu}}$.

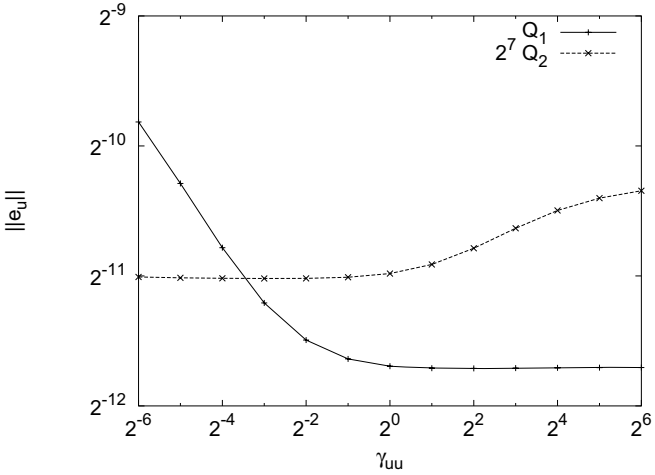

Figure 3.2: Accuracy of the standard LDG method on distorted grids depending on $\widehat{\gamma_{uu}}$.

3.3 The "superconvergent" LDG method

3.3.1 The parameter $\gamma_{u\sigma}$ can improve stability of the LDG method if it is chosen as an upwind flux for the gradient operator. In this case, γ_{uu} should be positive, independent of h.

3.3.2 Remark: The method presented here was introduced in [CKPS01] on Cartesian grids. A modification, allowing to use the method on general grids and improving the condition number of the matrix will be suggested later without mathematical proof of convergence orders.

3.3.3 Definition (Superconvergent fluxes): Let $\mathbf{e} = (1, \ldots, 1)^T \in \mathbb{R}^d$. Then, we choose $\gamma_{uu} = \widehat{\gamma_{uu}}$ on $\mathbb{E}_h$ and $\gamma_{u\sigma}$ on $\mathbb{E}_h^i$ such that

$$\gamma_{u\sigma} \cdot \mathbf{n} = \frac{\operatorname{sign}(\mathbf{e} \cdot \mathbf{n})}{2}. \tag{3.32}$$

Obviously, this definition does not depend on the choice of $\mathbf{n}$ and both values are independent of h.

On each edge E, let $u^\uparrow$ and $u^\downarrow$ be the upwind and downwind values of u with respect to the vector field $\mathbf{e}$ and let $\mathbf{n}^\uparrow$ ($\mathbf{n}^\downarrow$) be the vector normal to E with $\mathbf{n}^\uparrow \cdot \mathbf{e} \geqslant 0$ ($\mathbf{n}^\uparrow \cdot \mathbf{e} < 0$). Then, the bilinear form $b_h(.,.)$ of the superconvergent LDG method assumes the form

$$\begin{aligned} b_h(u,\tau) = &\quad \big(u, \nabla\cdot\tau\big)_{\mathbb{T}_h} + \big\langle u^\uparrow \mathbf{n}^\uparrow, \tau_j^\downarrow - \tau_j^\uparrow \big\rangle_{\mathbb{E}_h^i} - \big\langle u\mathbf{n}, \tau \big\rangle_{\mathbb{E}_h^N}, && (3.33)\\ = &-\big(\nabla u, \tau_j\big)_{\mathbb{T}_h} + \big\langle u^\uparrow - u^\downarrow, \tau_j^\downarrow \cdot \mathbf{n}^\downarrow \big\rangle_{\mathbb{E}_h^i} + \big\langle u, \tau\cdot\mathbf{n}\big\rangle_{\mathbb{E}_h^D}. && (3.34) \end{aligned}$$

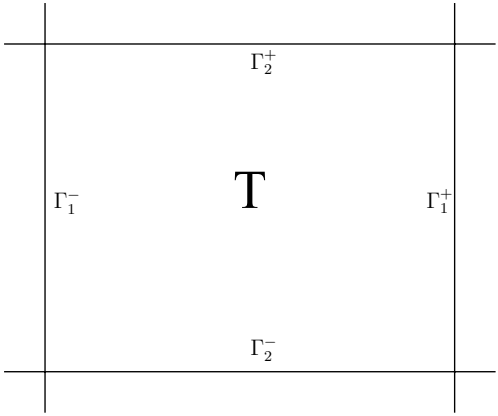

Figure 3.3: Naming of the boundary components of a rectangular grid cell T

3.3.4 Remark: Unique solvability of equation (3.10) with the form $b_h(.,.)$ in (3.33) follows already from Lemma 3.1.6. Furthermore, the proof of Theorem 3.2.3 applies to this method as well, yielding non-optimal convergence estimates in the energy norm. It remains to show improved error estimates with an extra half power of h.

3.3.5 Definition: In order to facilitate the analysis below, we split the Dirichlet and Neumann boundary into the components $\mathbb{E}_h^{D+}$ and $\mathbb{E}_h^{N+}$ with $\mathbf{n} \cdot \mathbf{e} > 0$ and the remaining parts $\mathbb{E}_h^{D-}$ and $\mathbb{E}_h^{N-}$.

The boundary of a grid cell $T = \prod[a_i^-, a_i^+]$ is split into the components $\Gamma_i^\pm = \{x | x_i = a_i^\pm\}$ according to Figure 3.3.

3.3.6 Theorem: *Let the mesh consist of rectangles and let the finite element spaces be* $\Sigma_h = (X_{h,k}^Q)^d$ *and* $V_h = X_{h,k}^Q$. *Let* (σ_h, u_h) *be the solution to* (3.10) *with the form* $b_h(.,.)$ *and* γ_{uu} *chosen according to Definition 3.3.3 and the solution* (σ, u) *of* (3.1) *be in* $W^{k+1,2}(\mathbb{T}_h; \mathbb{R}^d) \times W^{k+2,2}(\mathbb{T}_h)$. *Let furthermore* $u \in W^{k+1,\infty}(\mathbb{E}_h^{N-})$.

Then, the error (e_σ, e_u) *admits the estimates*

$$\big|(e_\sigma, e_u)\big|_{LDG} \leqslant Ch^{k+\frac{1}{2}} \tag{3.35}$$

$$\big\|e_u\big\| \leqslant Ch^{k+1}. \tag{3.36}$$

In particular, this implies

$$\big\|e_\sigma\big\| \leqslant Ch^{k+\frac{1}{2}}. \tag{3.37}$$

Proof. Again, the proof follows the general framework of Theorem 1.3.9. It relies on the introduction of special projection operators and intermediary results in lemmas 3.3.9 and 3.3.10. □

3.3.7 Definition (Upwind projections): For an interval $I = [a^-, a^+] \subset \mathbb{R}$, we define the one-dimensional upwind/downwind projection $\pi^\pm : C^0(I) \to \mathbb{P}_k(I)$ by the conditions

$$\begin{aligned} \big(v - \pi^\pm v, w\big)_I &= 0 \qquad \forall w \in \mathbb{P}_{k-1}(I) \\ v(a^\pm) - \pi^\pm v(a^\pm) &= 0. \end{aligned} \tag{3.38}$$

On a rectangular axiparallel cell $T = \prod[a_i^-, a_i^+]$, we define an *upwinding projection* from the space of continuous functions $C^0(T)$ onto $\mathbb{Q}_k(T)$ by

$$\Pi_u^\pm := \bigotimes_{i=1}^d \pi_i^\pm. \tag{3.39}$$

Here, the subscripts indicate the application of the one dimensional operator $\pi^\pm$ in the corresponding coordinate direction. We will apply this operator to functions in $W^{2,2}(T)$ and $W^{1,\infty}(T)$. These spaces are continuously embedded into $C^0(T)$ for $d = 2, 3$, such that the projection operator is well-defined.

Furthermore, we define projections $\Pi_\sigma^\pm : W^{1,2(T;\mathbb{R}^2)} \to \mathbb{Q}_k^2(T)$ by the conditions

$$\big(\tau - \Pi_\sigma^\pm \tau, \nabla v\big)_T = 0 \qquad \forall\, v \in \mathbb{Q}_k(T) \tag{3.40}$$

$$\big\langle \tau - \Pi_\sigma^\pm \tau, v\mathbf{n} \big\rangle_{\Gamma_T^{i\pm}} = 0 \qquad \forall\, v \in \mathbb{Q}_k(\Gamma_T^{i\pm}). \tag{3.41}$$

On a face $\Gamma_i^\pm$ perpendicular to the x_i-axis, we define the face projector

$$\Pi_{E;i}^\pm = \bigotimes_{j \neq i} \pi_j^\pm. \tag{3.42}$$

3.3.8 Lemma: *Let T be a rectangular mesh cell with edges parallel to the coordinate axes and diameter h_T. Then, the projection errors for functions $v \in W^{s,2}(T)$ and $\tau \in W^{s,2}(T;\mathbb{R}^d)$ with $s \leqslant k+1$ and $k = k_u = k_\sigma$ admit the estimates*

$$\big|v - \Pi_u^\pm v\big|_{m,2;T} \leqslant C h_T^{s-m} \big|v\big|_{s,2;T} \tag{3.43}$$

$$\big|v - \Pi_u^\pm v\big|_{m,2;E} \leqslant C h_T^{s-\frac{1}{2}-m} \big|v\big|_{s,2;T} \tag{3.44}$$

$$\big|\tau - \Pi_\sigma^\pm \tau\big|_{m,2;T} \leqslant C h_T^{s-m} \big|\tau\big|_{s-1,2;T}, \tag{3.45}$$

$$\big|\tau - \Pi_\sigma^\pm \tau\big|_{m,2;E} \leqslant C h_T^{s-\frac{3}{2}-m} \big|\tau\big|_{s-1,2;E}, \tag{3.46}$$

where E is any of the edges(faces) of T. Furthermore, if $v \in W^{s,\infty}(E)$, then

$$\big\|v - \Pi_u^\pm v\big\|_{\infty;E} \leqslant C h_T^s \big|v\big|_{s,\infty;E}. \tag{3.47}$$

Finally, for the edges $\Gamma_i^\pm$ holds

$$\big\|w - \Pi_{E;i}^\pm w\big\|_{\Gamma_i^\pm} \leqslant C h_{\Gamma_i^\pm}^{s-\frac{1}{2}} \big|w\big|_{s,2;T}. \tag{3.48}$$

Proof. Simple calculation shows that the operators $\Pi_u^\pm$ and $\Pi_\sigma^\pm$ preserve polynomials in $\mathbb{Q}_k(T)$. Therefore, the estimates follow by the Bramble-Hilbert lemma and by trace estimates. □

3.3.9 Lemma: *Let $\sigma \in W^{s,2}(\Omega;\mathbb{R}^d)$ and $u \in W^{s,2}(\Omega)$ with $s \leqslant k+1$, where $k = k_u = k_\sigma$ in Assumption 3.1.2. Then, the projection errors $\eta_\sigma^- = \sigma - \Pi_\sigma^- \sigma$ and $\eta_u^+ = u - \Pi_u^+ u$ admit the estimate*

$$0 \leqslant \mathcal{L}_h\big(\eta_\sigma^-, \eta_u^+; \eta_\sigma^-, \eta_u^+\big) \leqslant c\big(\big|h^{s-\frac{1}{2}}\sigma\big|_{s,2;\mathbb{T}_h}^2 + \big|h^{s-\frac{1}{2}}u\big|_{s,2;\mathbb{T}_h}^2\big)$$

Furthermore, if $u \in W^{s+1,2}(\mathbb{T}_h)$ and $\nu \in W^{s,\infty}(\mathbb{T}_h)$, then

$$\big|\mathcal{L}_h\big(\eta_\sigma^-, \eta_u^+; \eta_\sigma^-, \eta_u^+\big)\big| \leqslant c\big|h^{s-\frac{1}{2}}u\big|_{s,2;\mathbb{T}_h}^2$$

Proof. The proof is analogous to Lemma 3.2.4, with the difference that $\gamma_{uu} = \mathcal{O}(1)$ and that powers of h in the form $b_h(.,.)$ are equilibrated in a different way. □

3.3.10 Lemma: *Let $\tau \in \Sigma_h$ and $v \in V_h$, furthermore, $\sigma \in W^{s,2}(\mathbb{T}_h;\mathbb{R}^d)$ and $u \in W^{s,2}(\mathbb{T}_h)$ with $s \leqslant k+1$. Then,*

$$\begin{aligned}|\mathcal{L}_h\big(\tau, v; \eta_\sigma^-, \eta_u^+\big)| \\ \leqslant c\big(\big|h^{s-\frac{1}{2}}\sigma\big|_{s,2;\mathbb{T}_h} + \big|h^{s-\frac{1}{2}}u\big|_{s,2;\mathbb{T}_h} + \big|h^{s-\frac{1}{2}}u\big|_{s,\infty;\mathbb{E}_h^{N-}}\big)\big|(\tau,v)\big|_{LDG},\end{aligned} \quad (3.49)$$

with a constant c independent of the mesh size.

Proof. (from [CKPS01]) For simplicity, we omit the superscripts on $\eta^\pm$ and writing

$$|\mathcal{L}_h\big(\tau, v; \eta_\sigma, \eta_u\big)| \leqslant |a_h(\tau,\eta_\sigma)| + |b_h(\eta_u,\tau)| + |b_h(v,\eta_\sigma)| + |c_h(v,\eta_u)|,$$

we estimate the terms separately. For the first and the last term, we refer to equations (3.24) and (3.25), observing $\gamma_{uu} = \mathcal{O}(1)$ now. The second is used in the form

$$b_h(v,\eta_\sigma) = -(\nabla v, \eta_\sigma)_{\mathbb{T}_h} + \big\langle [\![v\mathbf{n}]\!], \eta_\sigma\big\rangle_{\mathbb{E}_h^i} + \big\langle v, \eta_\sigma\cdot\mathbf{n}\big\rangle_{\mathbb{E}_h^D}.$$

In virtue of the definition of the projection, the first two terms of this sum and the third restricted to the "inflow" boundary vanish. Consequently,

$$\begin{aligned}|b_h(v,\eta_\sigma)| &= |\big\langle v, \eta_\sigma\cdot\mathbf{n}\big\rangle_{\mathbb{E}_h^{D+}}| \\ &\leqslant \big\|\gamma_{uu}^{\frac{1}{2}}v\big\|_{\mathbb{E}_h^{D+}}\big\|\gamma_{uu}^{-\frac{1}{2}}\eta_\sigma\big\|_{\mathbb{E}_h^{D+}} \\ &\leqslant \big|(\tau,v)\big|_{\mathrm{LDG}}\big\|\gamma_{uu}^{-\frac{1}{2}}\eta_\sigma\big\|_{\mathbb{E}_h^{D+}}\end{aligned}$$

The estimate of the third term is more involved and uses the techniques of the superconvergence estimate for advection operators found in [LR74]. We begin with

$$b_h(\eta_u,\tau) = \big(\eta_u, \nabla\cdot\tau\big)_{\mathbb{T}_h} + \big\langle \eta_{u^\uparrow}\mathbf{n}^\uparrow, \tau_j^\downarrow - \tau_j^\uparrow\big\rangle_{\mathbb{E}_h^i} - \big\langle \eta_u\mathbf{n}, \tau\big\rangle_{\mathbb{E}_h^N}.$$

Therefore, the contribution of an interior cell is

$$z_T(\tau,u) = \big(\eta_u, \nabla\cdot\tau\big)_T - \big\langle \eta_u\mathbf{n}, \tau\big\rangle_{\partial^+T} - \big\langle \eta_u^\uparrow\mathbf{n}, \tau\big\rangle_{\partial^-T},$$

where we used that $u^\uparrow$ refers to u of the cell T on $\partial^- T$. Since $u \in W^{1,2}(\Omega)$ and due to the definition of Π^+ we have $\eta_u^\uparrow|_{E_i^-} = (I - \Pi^+_{E;i})u|_{E_i^-}$ and obtain

$$z_T(\tau, u) = \big(\eta_u, \nabla\cdot\tau\big)_T - \sum_{i=1}^{d}\Big(\big\langle u - \Pi^+_{E;i}u, \tau\cdot\mathbf{n}\big\rangle_{E_i^-} + \big\langle u - \Pi^+_{E;i}u, \tau\cdot\mathbf{n}\big\rangle_{E_i^+}\Big). \tag{3.50}$$

Summing up and adding the missing terms at the boundary yields

$$\begin{aligned}|b_h(\tau,u)| \leqslant \sum_{T\in\mathbb{T}_h}|z_T(\tau,u)| \\ + \sum_{E\in\mathbb{E}_h^D}|\big\langle u - \Pi^+_{E;i}u, \tau\cdot\mathbf{n}\big\rangle_E| + \sum_{E\in\mathbb{E}_h^N}|\big\langle u - \Pi^+_{u}u, \tau\cdot\mathbf{n}\big\rangle_E|.\end{aligned}$$

Here, i denotes the coordinate direction orthogonal to E. Since $\Pi^+_u = \Pi^+_{E;i}$ on an edge $E \in \mathbb{E}_h^{N+}\cup\mathbb{E}_i$, we can estimate with (3.48) for any $E \in \mathbb{E}_h^D \cup \mathbb{E}_h^{N+}$:

$$|\big\langle u - \Pi^+_{E;i}u, \tau\cdot\mathbf{n}\big\rangle_E| \leqslant c h_{T_E}^{s-\frac12}|u|_{s,2;T_E}\big\|\tau\big\|_{T_E},$$

with T_E the cell having E on its "downwind" side.

On $\mathbb{E}_h^{N-}$, we estimate by inverse inequality

$$\begin{aligned}\big\langle u - \Pi^+_u u, \tau\cdot\mathbf{n}\big\rangle_E &\leqslant \big\|u - \Pi^+_u u\big\|_{0,\infty,E}\big\|\tau\cdot\mathbf{n}\big\|_{0,1;E} \\ &\leqslant c|E|^{\frac12}\big\|h_{T_E}^{-\frac12}(u - \Pi^+_u u)\big\|_{0,\infty;E}\big\|\tau\cdot\mathbf{n}\big\|_{T_E} \\ &\leqslant c\big|h_{T_E}^{s-\frac12}u\big|_{s,\infty;E}\big\|\tau\cdot\mathbf{n}\big\|_{T_E}.\end{aligned}$$

Consequently,

$$\begin{aligned}|b_h(\tau,u)| \leqslant \sum_{T\in\mathbb{T}_h}|z_T(\tau,u)| \\ + c\Big(\big|h_{T_E}^{s-\frac12}u\big|_{s,2;\mathbb{T}_h}\big|h_{T_E}^{s-\frac12}u\big|_{s,\infty;\mathbb{E}_h^{N-}}\Big)\big|(\tau,v)\big|_{\mathrm{LDG}},\end{aligned}$$

Now, it remains to estimate the terms $z_T(\tau,u)$ with the same powers of h. First, we split

$$z_T(\tau,u) = \sum_{i=1}^{d} z_{T;i}(\tau_i, u).$$

In two dimensions, we have

$$z_{T;1}(\tau_1,u) = \big(u - \Pi^+_u u, \partial_1\tau_1\big)_T - \big\langle u - \Pi^+_{E;1}u, \tau\big\rangle_{\Gamma_T^1}$$

and $z_{T;i}$ with $i > 1$ analogously due to the orthogonality of the edges to the coordinate planes.

We claim that for any $\tau \in \mathbb{Q}_k(T)$

$$z_{T;i}(\tau_i, u) = 0 \qquad \forall\, u \in \mathbb{P}_{k+1}(T). \tag{3.51}$$

Since the cell T is just a scaling of the reference cell $\hat{T} = [-1,1]^d$ and the orthogonality (3.51) is preserved by scaling, we will assume $T = \hat{T}$. Furthermore, we will show the proof for $z_{T;1}(\tau_1, u)$ only, the other directions following the same pattern.

First, we remark that projections Π_u^+ and $\Pi_{E;1}^+$ preserve polynomials. Thus (3.51) holds for $u \in \mathbb{Q}_k$. It remains to show that (3.51) holds for the two polynomials $u(x) = x_1^{k+1}$ and $u(x) = x_2^{k+1}$.

The function $u(x) = x_1^{k+1}$ is constant on Γ_1^+ and Γ_1^-, thus $u - \Pi_{E;1}^+ u = 0$. In the interior, we conclude from $\partial_1 \tau_1 \in \mathbb{P}_{k-1}$ that $u - \Pi^+ u = 0$, therefore $z_{T;1}(\tau_1, x_1^{k+1}) = 0$.

For $u(x) = x_2^{k+1}$, we integrate by parts and obtain

$$\begin{aligned}\big(u - \Pi_u^+ u, \partial_1 \tau_1\big)_T = -\big(\partial_1 (u - \Pi_u^+ u, \tau_1\big)_T \\ + \big\langle u - \Pi_{E;1}^+ u, \tau_1 \big\rangle_{\Gamma_T^{2+}} - \big\langle u - \Pi^+ u, \tau_1 \big\rangle_{\Gamma_T^{2-}}.\end{aligned}$$

Since u is independent of x_1 and $\Pi^+ u = \Pi_{E;1}^+ u$ on Γ_T^{2-}, this sum is zero. We conclude that the linear functional $z_{T;i}(\tau_i, .)$ vanishes on $\mathbb{P}_{k+1}(T)$ for any $\tau \in \mathbb{Q}_k(T)$ and by the Bramble-Hilbert lemma we obtain

$$z_T(\tau, u) \leqslant C h_T^s \big|u\big|_{s,2;T} \big\|\tau\big\|_T,$$

for $s \leqslant k+1$. In higher dimensions, the additional polynomials to be tested are $u(x) = x_i^{k+1}$, $i > 2$. These can be treated the same way as x_2^{k+1}. □

3.3.11 First, we solve example A.2.2 on the Cartesian grids in example A.1.1. The results in Table 3.3 confirm the sharpness of our estimate[‡].

3.3.12 Figure 3.4 shows that the error of the stress σ is concentrated at the boundary. Therefore, we compute approximations for example A.2.1 with homogeneous Dirichlet boundary values for comparison. Table 3.4 shows that the convergence orders of σ are indeed better by one half and the error distribution in Figure 3.5 is equilibrated over the domain. Similar results are obtained with inhomogeneous boundary values lying in the approximation space. We conclude that the projection error of u into the space V_h at the boundary is a major source of error in σ.

3.3.13 In Table 3.5, we present some results *not* covered by our analysis. These results are for the polynomial spaces $\mathbb{P}_0$ and $\mathbb{P}_1$ and inhomogeneous Dirichlet boundary conditions (example A.2.2). They show, that tensor product polynomials are not required to achieve the convergence rates predicted by the theory of the superconvergent method. In particular, piecewise constant functions for u and σ still yield convergence with $\sqrt{h}$ for σ_h in L^2.

3.3.14 Remark: The linear system of the upwind LDG method is not solvable efficiently, since a spectrally equivalent preconditioner is difficult to obtain. Such a preconditioner can be found, if the flux at the Dirichlet boundary is replaced by the standard flux (3.19). This case is not covered by the analysis either, since the negative power of h in front of the jump term cannot be compensated. Nevertheless, Table 3.6 shows that this modification does not affect the convergence rates —compared to Table 3.4— in two and three space dimensions.

[‡]The last row in the $\mathbb{Q}_2$-column is probably spoiled by the accuracy of the linear solver

L	$\|e_\sigma\|$		$\|e_u\|$		$\|e_\sigma\|$		$\|e_u\|$	
	err	ord	err	ord	err	ord	err	ord
	$\mathbb{Q}_1$				$\mathbb{Q}_2$			
	unit square in two dimensions							
1	2.04e-1	—	4.99e-2	—	9.64e-3	—	1.99e-3	—
2	7.25e-2	1.49	1.32e-2	1.92	1.67e-3	2.53	2.73e-4	2.87
3	2.44e-2	1.57	3.62e-3	1.86	2.83e-4	2.56	3.75e-5	2.87
4	8.05e-3	1.60	9.76e-4	1.89	4.78e-5	2.57	4.99e-6	2.91
5	2.68e-3	1.59	2.55e-4	1.93	8.16e-6	2.55	6.47e-7	2.95
6	9.10e-4	1.56	6.54e-5	1.96	1.42e-6	2.53	8.27e-8	2.97
7	3.14e-4	1.53	1.66e-5	1.98	2.69e-7	2.40	1.28e-8	2.70
	unit cube in three dimensions							
1	2.36e-01	—	9.67e-02	—	5.44e-3	—	9.53e-4	—
2	1.07e-01	1.14	2.88e-02	1.75	1.15e-3	2.24	1.79e-4	2.41
3	4.32e-02	1.31	8.15e-03	1.82	2.25e-4	2.36	2.72e-5	2.72
4	1.63e-02	1.40	2.19e-03	1.90	4.19e-5	2.43	3.74e-6	2.86
5	5.98e-03	1.45	5.67e-04	1.95	7.58e-6	2.46	4.90e-7	2.93
6	2.15e-03	1.48	1.44e-04	1.97	1.36e-6	2.48	6.30e-8	2.96

Table 3.3: Convergence results for the superconvergent LDG method in two and three space dimensions

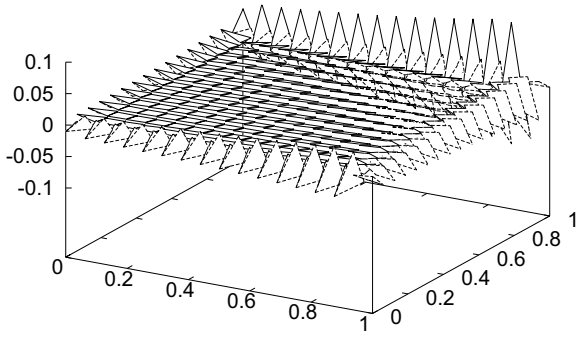

Figure 3.4: Error $\sigma^{(2)} - \sigma_h^{(2)}$ with inhomogeneous Dirichlet boundary values

L	$\|e_\sigma\|$		$\|e_u\|$		$\|e_\sigma\|$		$\|e_u\|$	
	err	ord	err	ord	err	ord	err	ord
	$\mathbb{Q}_1$				$\mathbb{Q}_2$			
	unit square in two dimensions							
1	4.87e-01	—	1.93e-01	—	7.25e-02	—	2.07e-02	—
2	1.62e-01	1.58	4.65e-02	2.06	1.27e-02	2.52	2.67e-03	2.95
3	4.88e-02	1.73	1.18e-02	1.97	1.83e-03	2.79	3.59e-04	2.89
4	1.37e-02	1.83	3.07e-03	1.94	2.40e-04	2.93	4.80e-05	2.90
5	3.66e-03	1.91	7.94e-04	1.95	3.03e-05	2.99	6.25e-06	2.94
6	9.47e-04	1.95	2.03e-04	1.97	3.78e-06	3.00	7.99e-07	2.97
7	2.41e-04	1.97	5.12e-05	1.98	4.73e-07	3.00	1.06e-07	2.92
	unit cube in three dimensions							
1	3.80e-01	—	1.99e-01	—	2.98e-02	—	1.12e-02	—
2	1.00e-01	1.92	5.46e-02	1.87	5.14e-03	2.54	1.64e-03	2.77
3	2.52e-02	1.99	1.36e-02	2.00	6.90e-04	2.90	2.61e-04	2.65
4	6.33e-03	1.99	3.37e-03	2.02	8.65e-05	3.00	3.64e-05	2.84
5	1.59e-03	1.99	8.36e-04	2.01	1.07e-05	3.01	4.78e-06	2.93
6	3.98e-04	2.00	2.08e-04	2.01	1.33e-06	3.01	6.11e-07	2.97

Table 3.4: Convergence results for the superconvergent LDG method with homogeneous Dirichlet boundary values

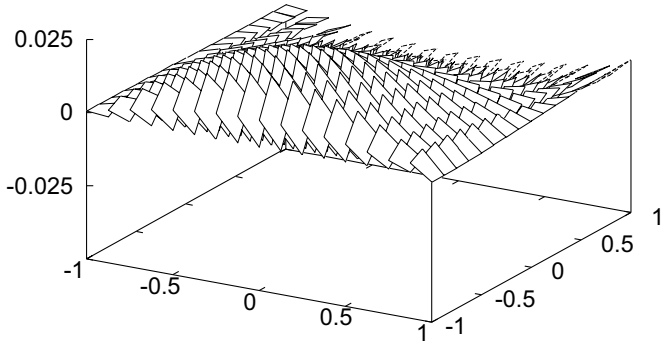

Figure 3.5: Error $\sigma^{(2)} - \sigma_h^{(2)}$ with homogeneous Dirichlet boundary values

L	$\|e_\sigma\|$		$\|e_u\|$		$\|e_\sigma\|$		$\|e_u\|$	
	err	ord	err	ord	err	ord	err	ord
	$\mathbb{P}_0$				$\mathbb{P}_1$			
	unit square in two dimensions							
1	1.01e+0	—	6.84e-1	—	1.99e-1	—	8.14e-2	—
2	6.01e-1	0.75	4.21e-1	0.70	6.73e-2	1.56	2.28e-2	1.83
3	4.21e-1	0.51	2.36e-1	0.84	2.37e-2	1.51	5.91e-3	1.95
4	3.01e-1	0.48	1.24e-1	0.93	8.49e-3	1.48	1.49e-3	1.99
5	2.13e-1	0.50	6.29e-2	0.97	3.05e-3	1.48	3.73e-4	2.00
6	1.50e-1	0.50	3.17e-2	0.99	1.09e-3	1.49	9.34e-5	2.00
7	1.06e-1	0.50	1.59e-2	1.00	3.87e-4	1.49	2.34e-5	2.00
	unit cube in three dimensions							
1	1.92e+0	—	1.44e+0	—	4.43e-1	—	2.27e-1	—
2	1.06e+0	0.86	8.22e-1	0.81	1.51e-1	1.55	6.19e-2	1.87
3	6.60e-1	0.68	4.57e-1	0.85	5.38e-2	1.49	1.59e-2	1.96
4	4.60e-1	0.52	2.42e-1	0.92	1.96e-2	1.46	3.99e-3	1.99
5	3.28e-1	0.49	1.24e-1	0.97	7.08e-3	1.47	9.98e-4	2.00
6	2.33e-1	0.49	6.24e-2	0.99	2.54e-3	1.48	2.49e-4	2.00

Table 3.5: Convergence results for the superconvergent LDG method with $\mathbb{P}_k$ elements

L	$\|e_\sigma\|$		$\|e_u\|$		$\|e_\sigma\|$		$\|e_u\|$	
	err	ord	err	ord	err	ord	err	ord
	$u^D = 0$				$u^D = e^{x+y+z}$			
	unit square in two dimensions							
1	3.74e-1	—	1.97e-1	—	1.44e-1	—	4.70e-2	—
2	1.03e-1	1.86	5.19e-2	1.93	5.80e-2	1.31	1.41e-2	1.74
3	2.58e-2	2.00	1.29e-2	2.01	2.16e-2	1.43	3.93e-3	1.84
4	6.44e-3	2.00	3.24e-3	1.99	7.77e-3	1.47	1.03e-3	1.93
5	1.61e-3	2.00	8.19e-4	1.99	2.76e-3	1.49	2.64e-4	1.97
6	4.01e-4	2.00	2.06e-4	1.99	9.80e-4	1.50	6.66e-5	1.99
7	1.00e-4	2.00	5.17e-5	2.00	3.47e-4	1.50	1.67e-5	1.99
	unit cube in three dimensions							
1	4.06e-1	—	2.23e-1	—	2.56e-1	—	9.68e-2	—
2	1.17e-1	1.79	6.34e-2	1.81	1.19e-1	1.11	2.91e-2	1.73
3	2.95e-2	1.99	1.58e-2	2.00	4.75e-2	1.33	8.33e-3	1.81
4	7.38e-3	2.00	3.98e-3	1.99	1.77e-2	1.42	2.23e-3	1.90
5	1.85e-3	2.00	1.00e-3	1.99	6.39e-3	1.47	5.74e-4	1.96

Table 3.6: Convergence results for the superconvergent LDG method with $1/h$-stabilization at the boundary ($\mathbb{Q}_1 - elements$)

L	$\|e_\sigma\|$		$\|e_u\|$		$\|e_\sigma\|$		$\|e_u\|$	
	err	ord	err	ord	err	ord	err	ord
	irregular				distorted			
1	1.31e-1	—	3.60e-2	—	1.44e-1	—	4.70e-2	—
2	4.53e-2	1.53	9.78e-3	1.88	5.80e-2	1.31	1.41e-2	1.74
3	1.56e-2	1.54	2.56e-3	1.93	2.16e-2	1.43	3.93e-3	1.84
4	5.41e-3	1.53	6.56e-4	1.97	7.77e-3	1.47	1.03e-3	1.93
5	1.89e-3	1.52	1.66e-4	1.98	2.76e-3	1.49	2.64e-4	1.97
6	6.65e-4	1.51	4.18e-5	1.99	9.80e-4	1.50	6.66e-5	1.99
7	2.34e-4	1.50	1.05e-5	2.00	3.47e-4	1.50	1.67e-5	1.99

Table 3.7: Convergence results for the superconvergent LDG method on non-Cartesian grids ($\mathbb{Q}_1 - elements$)

3.3.15 Table 3.7, we show experiments indicating, that the Cartesian meshes required for the analysis are not necessary either in experiments. We show results with the modified scheme (see previous remark) on the unit square (example A.2.2 again), once with the irregular coarse mesh in A.1.4 and once with the distorted meshes in A.1.3. The first mesh contains irregular points spoiling point-wise superconvergence (see 2.3), the other meshes do not converge to parallelogram cells. Even more, the additional half order of convergence observed with homogeneous boundary conditions is still obtained for these examples. A theoretical explanation for these results is still missing.

3.3.16 Remark: LDG methods are based on a mixed system of equations and thus require a larger number of degrees of freedom than the interior penalty method. Furthermore, the solution of the resulting linear systems requires more iteration steps (see Chapter 6). Therefore, a comparison of these methods should consider approximation properties *and* solver efficiency.

The following discussion should give a hint when to choose which of the methods. For the results in Figure 3.6, the linear system was solved down to a residual norm of 10^{-13}. Then, for each method, solution time is plotted over the L^2-norm of the error of the solution u itself (left) and of its derivative σ (right), respectively. Here, the derivatives of the interior penalty solution are simply computed inside the cells. The figure shows that the interior penalty method is faster than the standard LDG by a factor of two, if only the solution u is considered. In this case, the superconvergent LDG is slower by far. On the other hand, if derivatives are important, the picture is just reversed. Superconvergent LDG is fastest due to its half order gain in accuracy, followed by standard LDG. Therefore, a decision between these three methods can only be made according to the quantities that have to be simulated accurately.

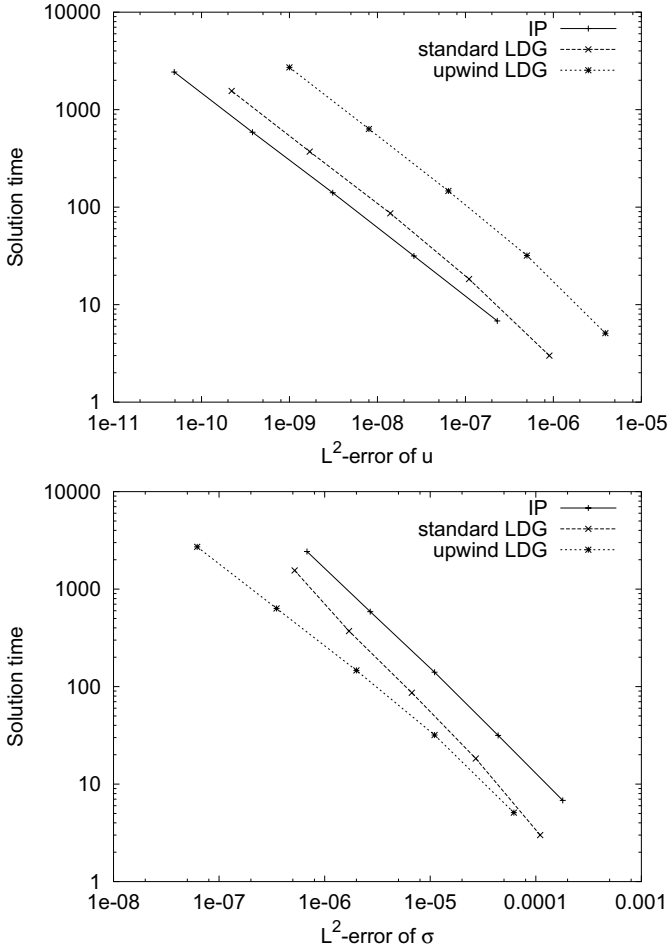

Figure 3.6: Performance of the different DG methods

Chapter 4

Stokes Equations

Several discretizations of Stokes equations by discontinuous functions have been proposed. First, there is the method by Baker et al. (cf. [BJK90, KJ98]). It uses piecewise solenoidal trial functions for the velocity and a "pressure" penalizing the discontinuities of the velocity at boundaries between cells.

A fully discontinuous scheme was devised by Hansbo and Larson in [HL02]. Their method is based on the interior penalty method of Chapter 2 and has the special merit of providing a smooth transition between compressible and incompressible elasticity.

The scheme presented here was introduced and analyzed by Cockburn, Kanschat, Schötzau and Schwab in [CKSS02] and is based on the LDG method in Chapter 3. It is similar to the one by Hansbo and Larson, but is based on the mixed form of the Laplacian, offering improved approximation of the stress tensor. Numerical evidence for the analytical convergence rates is given and the influence of the stabilization parameters is studied. In the end, we discuss stable finite element pairs without stabilization.

4.0.1 We extend the LDG method to Stokes flow problems modeled by *Stokes equations*

$$\begin{aligned} -\nu\Delta u + \nabla p &= f \\ \nabla \cdot u &= 0 \end{aligned} \qquad \text{in } \Omega \subset \mathbb{R}^d, \tag{4.1}$$

with Dirichlet and outflow boundary conditions

$$u = u^D \qquad \text{on } \Gamma_D \tag{4.2}$$

$$\partial_n u + p = 0 \qquad \text{on } \Gamma_N, \tag{4.3}$$

respectively. If Γ_N is empty, it is well known that the pressure is only determined up to a constant. Therefore, we require the additional condition

$$\big(p_h, 1\big)_\Omega = 0. \tag{4.4}$$

and the compatibility condition

$$\int_{\partial\Omega} u \cdot \mathbf{n}\, ds = 0. \tag{4.5}$$

4.0.2 The classical variational problem associated with (4.1) (see [BF91]) is defined as follows: Let

$$\tilde{V} = \big\{v \in W^{1,2}(\Omega;\mathbb{R}^d)\big|v = 0 \text{ on } \Gamma^D\big\} \tag{4.6}$$

$$\tilde{Q} = \begin{cases} L^2(\Omega)/\mathbb{R} = \big\{q \in L^2(\Omega)\big|(q,1)_\Omega = 0\big\} & \text{if } \Gamma_N = \emptyset. \\ L^2(\Omega) & \text{else,} \end{cases} \tag{4.7}$$

where we understand the boundary condition in the sense of traces. Then find $(u,p) \in \tilde{V} \times \tilde{Q}$ such that

$$\begin{array}{rll} \big(\nu\nabla u, \nabla v\big) - \big(\nabla\cdot v, p\big) = f & & \forall v \in \tilde{V}, \\ \big(\nabla\cdot u, q\big) = 0 & & \forall q \in \tilde{Q}. \end{array}$$

4.0.3 Lemma (inf-sup condition): *For any pressure function $q \in \tilde{Q}$ exists a velocity function $v \in \tilde{V}$ with $\|v\|_{1,2} \leqslant \|q\|$, satisfying*

$$\big(\nabla\cdot v, q\big) \geqslant \gamma\|q\|^2, \tag{4.8}$$

where the constant δ depends on Ω only.

4.0.4 In view of the mixed form of the Laplacian (3.1) underlying the LDG method, we transform (4.1) into the system

$$\begin{array}{rcll} \sigma & = & \nu\nabla u & \\ -\nabla\cdot\sigma + \nabla p & = & f & \text{in } \Omega, \\ \nabla\cdot u & = & 0 & \end{array} \tag{4.9}$$

with boundary conditions (4.2) and (4.3). We will assume that $(\sigma,u,p) \in \Sigma \times V \times Q$ and

$$\begin{aligned} \Sigma &= W^{1,2}(\mathbb{T}_h;\mathbb{R}^{d\times d}), \\ V &= W^{1,2}(\mathbb{T}_h;\mathbb{R}^d), \\ Q &= W^{1,2}(\mathbb{T}_h) \cap \tilde{Q} \end{aligned}$$

4.0.5 Assumption: *For estimates of the velocity error in L^2 and weaker norms, we have to assume that the solution (z_u, z_p) of the Stokes problem*

$$\begin{array}{rcrcl} -\nabla\cdot(\nu\nabla z_u) & + & \nabla z_p & = & f_u \\ -\nabla\cdot z_u & & & = & f_p, \end{array} \tag{4.10}$$

admits the stability estimate

$$\big\|\nu^{\frac{1}{2}}u\big\|_{\alpha+1,2;\Omega} + \big\|\nu^{-\frac{1}{2}}p\big\|_{\alpha,2;\Omega} \leqslant c\big(\big\|\nu^{-\frac{1}{2}}f_u\big\|_{\alpha-1,2;\Omega} + \big\|\nu^{\frac{1}{2}}f_p\big\|_{\alpha,2;\Omega}\big). \tag{4.11}$$

4.0.6 Remark: This assumption holds for instance if the domain Ω is bounded and for its boundary holds $\partial\Omega \in C^{\max\{\alpha+1,2\}}$ with Dirichlet boundary values $u^D \in W^{\alpha+1-\frac{1}{2}}(\partial\Omega)$ and $\Gamma_N = \emptyset$ (see [Tém79, Proposition 2.2]).

4.1 LDG discretization

4.1.1 Assumption: *Following Section 3.1, we derive the LDG discretization of 4.9 by choosing finite element subspaces Σ_h, V_h and Q_h of Σ, V and Q, respectively:*

$$\Sigma_h = (X_{h,k_\sigma})^{d\times d}, \qquad V_h = (X_{h,k_u})^d, \qquad Q_h = X_{h,k_p}, \tag{4.12}$$

where $k_u \geqslant 1$ and k_σ, k_p are either equal to k_u or to $k_u - 1$, independent of each other. Again, we limit spaces $X^P_{h,k}$ to the case where the mesh consists of parallelograms only.

4.1.2 Multiplication with a test function on each cell T and integration by parts yield

$$\begin{array}{rrrrrl} \left(\frac{1}{\nu}\sigma,\tau\right)_T & & + \left(u,\nabla\cdot\tau\right)_T & - \left\langle u,\tau\mathbf{n}_T\right\rangle_{\partial T} & = & 0 \\ \left(\sigma,\nabla v\right)_T - \left\langle\sigma, v\otimes\mathbf{n}_T\right\rangle_{\partial T} & & - \left(p,\nabla\cdot v\right)_T & + \left\langle p, v\cdot\mathbf{n}_T\right\rangle_{\partial T} & = & \left(f,v\right)_T \\ & & - \left(u,\nabla q\right)_T & + \left\langle u, q\mathbf{n}_T\right\rangle_{\partial T} & = & 0 \end{array} \tag{4.13}$$

Now, the traces of σ, u and p on ∂T must be replaced by single valued numerical fluxes to render this equation well-defined.

4.1.3 The trace of u appears in the first and third equation of (4.13) and we will denote the corresponding fluxes by $\hat{u}_\sigma$ and $\hat{u}_p$, respectively. Then, we define a general set of consistent fluxes on interior edges by

$$\begin{array}{rrrrl} \hat{\sigma} = \{\!\{\sigma\}\!\} - [\![\sigma\mathbf{n}]\!]\otimes\gamma_{u\sigma} - \gamma_{uu}[\![u\otimes\mathbf{n}]\!], & & & & \\ \hat{u}_\sigma = \{\!\{u\}\!\} & & + \gamma_{u\sigma}^T[\![u\otimes\mathbf{n}]\!], & & \\ \hat{u}_p = \{\!\{u\}\!\} & & + \ \gamma_{up}[\![u\cdot\mathbf{n}]\!] & + \ \gamma_{pp}[\![p\mathbf{n}]\!], & \\ \hat{p} = \{\!\{p\}\!\} & & & - \gamma_{up}\cdot[\![p\mathbf{n}]\!]. & \end{array} \tag{4.14}$$

Remark that $\hat{p}$ does not contain the jump of u, since the corresponding term is already found in the definition of $\hat{\sigma}$. Furthermore, we have chosen the fluxes to render the problem symmetric. The parameters γ_{xy} will be chosen according to the analysis later in this chapter.

4.1.4 On the Dirichlet boundary, these fluxes are taken as

$$\begin{aligned} \hat{\sigma} &= \sigma - \gamma_{uu}(u - u^D)\otimes\mathbf{n}, \\ \hat{u}_\sigma &= u^D, \\ \hat{u}_p &= u^D, \\ \hat{p} &= p, \end{aligned} \tag{4.15}$$

where traces are taken from the interior only.

4.1.5 Definition: The LDG formulation of the mixed Stokes problem (4.9) reads as follows: find functions $(\sigma_h, u_h, p_h) \in \Sigma_h \times V_h \times Q_h$ such that

$$\begin{array}{rrrrll} a_h(\sigma_h,\tau) + b_h(u_h,\tau) & & = f_h^\sigma(\tau) & & \forall\tau\in\Sigma_h \\ - \ b_h(v,\sigma_h) + c_h(u_h,v) + d_h(v,p_h) & & = f_h^u(v) & & \forall v\in V_h \\ - \ d_h(u_h,q) + e_h(p_h,q) & & = f_h^p(q) & & \forall q\in Q_h, \end{array} \tag{4.16}$$

where the bilinear forms are defined as

$$a_h(\sigma,\tau) = \big(\tfrac{1}{\nu}\sigma,\tau\big)_{\mathbb{T}_h}, \tag{4.17}$$
$$b_h(v,\tau) = \big(v,\nabla\cdot\tau\big)_{\mathbb{T}_h} - \big\langle \{\!\{u\}\!\} + [\![u\otimes\mathbf{n}]\!]\gamma_{u\sigma}, [\![\tau\mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i}, \tag{4.18}$$
$$c_h(u,v) = \big\langle \gamma_{uu}[\![u\otimes\mathbf{n}]\!], [\![v\otimes\mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i} + \big\langle \gamma_{uu}u\otimes\mathbf{n}, v\otimes\mathbf{n}\big\rangle_{\mathbb{E}_h^D}, \tag{4.19}$$
$$d_h(v,q) = -\big(q,\nabla\cdot v\big)_{\mathbb{T}_h} + \big\langle \{\!\{q\}\!\} - \gamma_{up}\cdot[\![q\mathbf{n}]\!], [\![v\cdot\mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i} + \big\langle p, v\cdot\mathbf{n}\big\rangle_{\mathbb{E}_h^D}, \tag{4.20}$$
$$e_h(p,q) = \big\langle \gamma_{pp}[\![p\mathbf{n}]\!], [\![q\mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i}. \tag{4.21}$$

The linear forms on the right hand side of (4.16) are

$$f_h^\sigma(\tau) = \big\langle u^D, \tau\mathbf{n}\big\rangle_{\mathbb{E}_h^D}, \tag{4.22}$$
$$f_h^u(v) = \big(f,v\big)_{\mathbb{T}_h} + \big\langle \gamma_{uu}u^D\otimes\mathbf{n}, v\otimes\mathbf{n}\big\rangle_{\mathbb{E}_h^D}, \tag{4.23}$$
$$f_h^p(q) = -\big\langle q, u^D\cdot\mathbf{n}\big\rangle_{\mathbb{E}_h^D}. \tag{4.24}$$

4.1.6 Remark: By integration by parts, the forms $b_h(.,.)$ and $d_h(.,.)$ admit the alternative representation

$$b_h(v,\tau) = -\big(\nabla v,\tau\big)_{\mathbb{T}_h} + \big\langle [\![v\otimes\mathbf{n}]\!], \{\!\{\tau\}\!\} - [\![\tau\cdot\mathbf{n}]\!]\otimes\gamma_{u\sigma}\big\rangle_{\mathbb{E}_h^i} + \big\langle v, \tau\mathbf{n}\big\rangle_{\mathbb{E}_h^D} \tag{4.25}$$
$$d_h(v,q) = \big(v,\nabla q\big)_{\mathbb{T}_h} - \big\langle \{\!\{v\}\!\} + \gamma_{up}[\![v\cdot\mathbf{n}]\!], [\![p\mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i} \tag{4.26}$$

4.1.7 We define the bilinear form associated with the Stokes problem as

$$\mathcal{S}_h\big(\sigma,u,p;\tau,v,q\big) := a_h(\sigma,\tau) + b_h(u,\tau) - b_h(v,\sigma) + c_h(u,v) + d_h(v,p) - d_h(u,q) + e_h(p,q),$$

and the associated energy seminorm

$$\big|(\sigma,u,p)\big|_{\mathrm{S}} := \sqrt{a_h(\sigma,\sigma) + c_h(u,u) + e_h(p,p)}. \tag{4.27}$$

4.1.8 Lemma: *Let γ_{uu} and γ_{pp} be positive and the discrete spaces chosen according to Assumption 4.1.1. Let the Dirichlet boundary Γ_D be nonempty. Then, the discrete equation*

$$\mathcal{S}_h\big(\sigma_h,u_h,p_h;\tau,v,q\big) = f_h^\sigma(\tau) + f_h^u(v) + f_h^p(q) \qquad \forall(\tau,v,q)\in\Sigma_h\times V_h\times Q_h, \tag{4.28}$$

admits a unique solution in $\Sigma_h\times V_h\times Q_h$.

Proof. We show that the homogeneous system admits the only solution 0. Indeed, taking $\tau = \sigma_h$, $v = u_h$ and $q = p_h$ in (4.28) yields

$$\big|(\sigma_h,u_h,p_h)\big|_{\mathrm{S}} = 0.$$

Therefore, $\sigma_h = 0$. Furthermore, $[\![u_h \otimes \mathbf{n}]\!] = 0$ and $[\![p\mathbf{n}]\!] = 0$ on $\mathbb{E}_h^i$ and $u_h = 0$ on $\mathbb{E}_h^D$. Consequently, in view of the first equation in (4.16)

$$\big(\nabla u_h, \tau\big)_{\mathbb{T}_h} = 0 \quad \forall \tau \in \Sigma_h,$$

yielding $u_h = 0$ by the same argument as in the proof of Lemma 3.1.6. Finally, in virtue of the second equation in (4.16),

$$\big(v, \nabla p_h\big)_{\mathbb{T}_h} = 0.$$

Therefore, the pressure p_h is constant on Ω (in the case of tensor product polynomials by Lemma 1.2.18). Now, either $p_h = 0$ on Γ_N or (4.4) yields $p_h = 0$. □

4.1.9 Definition: The standard LDG method for Stokes equations in [CKSS02] uses the following choice of parameters: let E be an edge and T^+, T^- the cells adjacent to E (only T^+ if E is a boundary edge). Then,

$$\tilde{h}_E =: \begin{cases} \min\{h_{T^+}, h_{T^-}\} & \text{if } E \in \mathbb{E}_h^i \\ h_{T^+} & \text{if } E \in \mathbb{E}_h^D, \end{cases}$$

and

$$\gamma_{uu} = \frac{\nu\widehat{\gamma_{uu}}}{\tilde{h}_E}, \tag{4.29}$$

$$\gamma_{pp} = \frac{\widehat{\gamma_{pp}}\tilde{h}_E}{\nu}. \tag{4.30}$$

Here, the parameters $\widehat{\gamma_{uu}}$ and $\widehat{\gamma_{pp}}$ may depend on the shape of the grid cells, but not on h. Furthermore, we choose

$$\gamma_{u\sigma} = 0, \qquad \text{and} \qquad \gamma_{up} = 0.$$

4.1.10 Theorem (energy error): *Assume that the mesh is shape regular and the mesh size variation is bounded according to assumptions 1.1.13 and 1.1.14. Let (σ, u, p) be a solution to equation* (4.9) *and let (σ_h, u_h, p_h) be a solution to* (4.28)*, where the discrete spaces are chosen according to Assumption 4.1.1 and the coefficients according to Definition 4.1.9. If $u \in W^{s,2}(\mathbb{T}_h; \mathbb{R}^d)$ and $p \in W^{s-1,2}(\mathbb{T}_h)$ and $\nu \in W^{s-1,\infty}$ with $s \leqslant k_u + 1$, then the energy seminorm of the error admits the estimate*

$$\big|(e_\sigma, e_u, e_p)\big|_S \leqslant c\big|\nu^{\frac{1}{2}} h^{s-1} u\big|_{s,2;\mathbb{T}_h} + c\big|\nu^{-\frac{1}{2}} h^{s-1} p\big|_{s-1.2;\mathbb{T}_h}. \tag{4.31}$$

Proof. Again, we have stability by definition of the energy seminorm. Therefore, it is sufficient to verify assumptions 1.3.7 and 1.3.8 of the generic theory. This is achieved in lemmas 4.1.11 and 4.1.12 below, where we enter $\sigma = \nu\nabla u \in W^{s-1,2}(\mathbb{T}_h; \mathbb{R}^{d\times d})$. □

4.1.11 Lemma: *Assume that the mesh is shape regular and the mesh size variation is bounded according to assumptions 1.1.13 and 1.1.14. Let Π_σ, Π_u and Π_p be the L^2-projections onto the spaces Σ_h, V_h and Q_h according to Assumption 4.1.1, respectively. Let $\sigma \in W^{s-1,2}(\mathbb{T}_h;\mathbb{R}^{d\times d})$, $u \in W^{s,2}(\mathbb{T}_h;\mathbb{R}^d)$ and $p \in W^{s-1,2}(\mathbb{T}_h)$ and furthermore let $\tau \in W^{t+1,2}(\mathbb{T}_h;\mathbb{R}^{d\times d})$, $v \in W^{t+2,2}(\mathbb{T}_h;\mathbb{R}^d)$ and $q \in W^{t+1,2}(\mathbb{T}_h)$. If γ_{uu} and γ_{pp} are chosen according to Definition 4.1.9, then the L^2-projection errors admit the estimate*

$$\begin{aligned}\mathcal{S}_h\big(\eta_\sigma,\eta_u,\eta_p;\eta_\tau,\eta_v,\eta_q\big)&\\ \leqslant c\big(\big|\nu^{-\frac12}h^{s-1}\sigma\big|_{s-1,2;\mathbb{T}_h}&+\big|\nu^{\frac12}h^{s-1}u\big|_{s,2;\mathbb{T}_h}+\big|\nu^{-\frac12}h^{s-1}p\big|_{s-1,2;\mathbb{T}_h}\big)\\ \big(\big|\nu^{-\frac12}h^{t+1}\tau\big|_{t+1,2;\mathbb{T}_h}&+\big|\nu^{\frac12}h^{t+1}v\big|_{t+2,2;\mathbb{T}_h}+\big|\nu^{-\frac12}h^{t+1}q\big|_{t+1,2;\mathbb{T}_h}\big).\end{aligned}\quad(4.32)$$

In particular,

$$\begin{aligned}\mathcal{S}_h\big(\eta_\sigma,\eta_u,\eta_p;\eta_\sigma,\eta_u,\eta_p\big)&\\ \leqslant c\big(\big|\nu^{-\frac12}h^{s-1}\sigma\big|^2_{s-1,2;\mathbb{T}_h}&+\big|\nu^{\frac12}h^{s-1}u\big|^2_{s,2;\mathbb{T}_h}+\big|\nu^{-\frac12}h^{s-1}p\big|^2_{s-1,2;\mathbb{T}_h}\big).\end{aligned}\quad(4.33)$$

Proof. We split the form $\mathcal{S}_h\big(.,.,.;.,.,.\big)$ into its components and estimate the single forms separately. Since the forms $a_h(.,.)$, $b_h(.,.)$ and $c_h(.,.)$ are the same as in Lemma 3.2.4, only in a vector-valued version, it is sufficient to prove the estimate for $d_h(.,.)$ and $e_h(.,.)$.

For $p, q \in W^{1,2}(\mathbb{T}_h)$, we obtain from trace estimate (1.24)

$$\begin{aligned}e_h(p,q)&\leqslant c\sum_{T\in\mathbb{T}_h}\big\|h_T^{\frac12}\nu^{-\frac12}p\big\|\big\|h_T^{\frac12}\nu^{-\frac12}q\big\|\\ &\leqslant c\big(\big\|\nu^{-\frac12}p\big\|+\big|h\nu^{-\frac12}p\big|_{1,2;\mathbb{T}_h}\big)\big(\big\|\nu^{-\frac12}q\big\|+\big|h\nu^{-\frac12}q\big|_{1,2;\mathbb{T}_h}\big).\end{aligned}$$

Inserting now η_p for p and η_q for q and applying approximation result 1.2.13 yields

$$e_h(\eta_p,\eta_q)\leqslant c\big|\nu^{-\frac12}h^{s-1}p\big|_{s-1,2;\mathbb{T}_h}\big|\nu^{-\frac12}h^{t+1}q\big|_{t+1,2;\mathbb{T}_h}.$$

The estimate of the form $d_h(.,.)$ uses the same trace estimates as the one for $b_h(.,.)$.

Finally, the second estimate follows immediately by letting $\tau=\sigma$, $v=u$ and $q=p$. □

4.1.12 Lemma: *Let the functions σ, u and p as well as the projection operators be as in Lemma 4.1.11. Then, for any $(\tau,v,q)\in\Sigma_h\times V_h\times Q_h$, the estimate*

$$\begin{aligned}\mathcal{S}_h\big(\eta_\sigma,\eta_u,\eta_p;\tau,v,q\big)&\\ \leqslant c\big|(\tau,v,q)\big|_S\big(\big|\nu^{-\frac12}h^{s-1}\sigma\big|_{s-1,2;\mathbb{T}_h}&+\big|\nu^{\frac12}h^{s-1}u\big|_{s,2;\mathbb{T}_h}+\big|\nu^{-\frac12}h^{s-1}p\big|_{s-1,2;\mathbb{T}_h}\big),\end{aligned}\quad(4.34)$$

holds with a constant c independent of the mesh size.

Proof. Again, we split the Stokes bilinear form into its components and estimate each of them separately. For the forms $a_h(.,.)$, $b_h(.,.)$ and $c_h(.,.)$, the estimates proven in Lemma 3.2.5

hold; since the LDG seminorm is bounded by the Stokes seminorm, it is sufficient to prove corresponding estimates for $d_h(.,.)$ and $e_h(.,.)$ here.

By Cauchy-Schwarz inequality and approximation result 1.2.13, we obtain immediately ($\gamma_{pp} = \mathcal{O}(h)$)

$$e_h(\eta_p, q) \leqslant c\big\|\nu^{-\frac{1}{2}} h^{\frac{1}{2}} \eta_p\big\|_{\mathbb{E}_h^i} \sqrt{c_h(q,q)} \leqslant c\big|\nu^{-\frac{1}{2}} h^{s-1} p\big|_{s-1,2;\mathbb{T}_h} \big|(\tau, v, q)\big|_S.$$

Similar to the estimate of $b_h(.,.)$, we use $d_h(.,.)$ in both forms (4.20) and (4.26) to exploit orthogonality of the L^2-projection:

$$\begin{aligned} |d_h(\eta_u, q)| &\leqslant \Big|\big\langle \{\!\{\eta_u\}\!\}, [\![q\mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i}\Big| \leqslant c\big\|h^{-\frac{1}{2}}\eta_u\big\|_{\mathbb{E}_h^i} \big\|\sqrt{h}[\![q]\!]\big\|_{\mathbb{E}_h^i} \\ &\leqslant c\big|\nu^{\frac{1}{2}} h^{s-1} u\big|_{s,2;\mathbb{T}_h} \sqrt{e_h(q,q)}, \end{aligned} \tag{4.35}$$

and

$$\begin{aligned} |d_h(v, \eta_p)| &\leqslant c\Big|\big\langle \{\!\{\eta_p\}\!\}, [\![v\cdot\mathbf{n}]\!]\big\rangle_{\mathbb{E}_h^i} + \big\langle \eta_p, v\cdot\mathbf{n}\big\rangle_{\mathbb{E}_h^D}\Big| \\ &\leqslant c\big|\nu^{-\frac{1}{2}} h^{s-1} p\big|_{s-1,2;\mathbb{T}_h} \sqrt{c_h(v,v)}. \end{aligned} \tag{4.36}$$

□

4.1.13 Lemma (inf-sup condition): *Assume that the parameters of the LDG method are chosen according to Definition 4.1.9 and the discrete spaces according to Assumption 4.1.1. Then, there exist positive constants c_1 and c_2 independent of the mesh size h such that the following* relaxed inf-sup condition *holds: for all $(\tau, v, q) \in \Sigma_h \times V_h \times Q_h$, there is a function $w \in V_h$ with $\big|(0,w,0)\big|_S \leqslant \big\|q\big\|$ such that*

$$\mathcal{S}_h\big(\tau, v, q; 0, w, 0\big) \geqslant c_1\big\|\nu^{-\frac{1}{2}} q\big\|^2 - c_2\big|(\tau, v, q)\big|_S^2. \tag{4.37}$$

Proof. Let (τ, v, q) be given. By the continuous inf-sup condition (4.8), there exists a function $u \in W^{1,2}(\Omega)$ with

$$\big(-\nabla\cdot u, q\big) \geqslant 2c_1\big\|\nu^{-\frac{1}{2}} q\big\|^2 \qquad \big\|\nu^{\frac{1}{2}} u\big\|_{1,2} \leqslant \big\|q\big\|. \tag{4.38}$$

Now, let $w = \Pi_u u$ be the L^2-projection of u into V_h. Then, we have

$$\mathcal{S}_h\big(\tau, v, q; 0, w, 0\big) = -b_h(w, \tau) + c_h(v, w) + d_h(w, q).$$

We continue by estimating these terms separately, using the same estimates as in the proof of Lemma 4.1.12. First,

$$\begin{aligned} |b_h(w,\tau)| &\leqslant |b_h(\eta_u, \tau)| + |b_h(u,\tau)| \\ &\leqslant c\big\|u\big\|_{1,2}\big\|\tau\big\| + \big|\big(\nabla u, \tau\big)_\Omega\big| \leqslant c\big\|\nu^{\frac{1}{2}} u\big\|_{1,2}\big\|\nu^{-\frac{1}{2}}\tau\big\| \leqslant c\big\|q\big\|\big|(\tau, v, q)\big|_S. \end{aligned}$$

Then,

$$|c_h(v,w)| = |c_h(v,\eta_u)| \leqslant c\big|\nu^{\frac{1}{2}}u\big|_{1,2;\mathbb{T}_h}\sqrt{c_h(v,v)} \leqslant c\big\|q\big\|\big|(\tau,v,q)\big|_{\mathrm{S}}.$$

Finally, we split

$$d_h(w,q) = d_h(u,q) - d_h(\eta_u,q).$$

Since u is in $W^{1,2}(\Omega;\mathbb{R}^d)$, the boundary terms in the first part vanish and the term is estimated by (4.38). The remainder is estimated by

$$|d_h(\eta_u,q)| \leqslant c\big\|\nu^{\frac{1}{2}}u\big\|_{1,2;\mathbb{T}_h}\sqrt{e_h(q,q)} \leqslant c\big\|q\big\|\big|(\tau,v,q)\big|_{\mathrm{S}},$$

as soon as $h \leqslant 1$. Combining these estimates, we obtain

$$\begin{aligned}\mathcal{S}_h\big(\tau,v,q;0,w,0\big) &\geqslant d_h(u,q) - |d_h(\eta_u,q)| - |b_h(w,\tau)| - |c_h(v,w)|\\ &\geqslant 2c_1\big\|\nu^{-\frac{1}{2}}q\big\|^2 - c\big\|q\big\|\big|(\tau,v,q)\big|_{\mathrm{S}}\\ &\geqslant c_1\big\|\nu^{-\frac{1}{2}}q\big\|^2 - \frac{c}{4c_1}\big|(\tau,v,q)\big|_{\mathrm{S}}^2,\end{aligned}$$

by Young inequality. □

4.1.14 Theorem: *Under the assumptions of Theorem 4.1.10, the L^2-norm of the error in the pressure admits the estimate*

$$\big\|e_p\big\| \leqslant c\left(\big|\nu h^{s-1}u\big|_{s,2;\mathbb{T}_h} + \big|h^{s-1}p\big|_{s-1,2;\mathbb{T}_h}\right). \tag{4.39}$$

Proof. The key ingredient to the proof of this theorem is the relaxed inf-sup condition of Lemma 4.1.13, which we apply to $q = \Pi_p e_p$. It allows us to estimate

$$\begin{aligned}c_1\big\|\nu^{-\frac{1}{2}}\Pi_p e_p\big\|^2 &\leqslant \mathcal{S}_h\big(\Pi_\sigma e_\sigma, \Pi_u e_u, \Pi_p e_p; 0,w,0\big) + c\big|(\eta_\sigma,\eta_u,\eta_p)\big|_{\mathrm{S}}^2\\ &= \mathcal{S}_h\big(\eta_\sigma,\eta_u,\eta_p;0,w,0\big) + c\big|(\eta_\sigma,\eta_u,\eta_p)\big|_{\mathrm{S}}^2.\end{aligned}$$

Now, we estimate the first term by Lemma 4.1.12 and remark that an estimate for the second is an intermediate result of Theorem 4.1.10. Therefore,

$$\begin{aligned}c_1\big\|\nu^{-\frac{1}{2}}\Pi_p e_p\big\|^2 &\leqslant c\big|(0,w,0)\big|_{\mathrm{S}}\left(\big|\nu^{\frac{1}{2}}h^{s-1}u\big|_{s,2;\mathbb{T}_h} + \big|\nu^{-\frac{1}{2}}h^{s-1}p\big|_{s-1,2;\mathbb{T}_h}\right)\\ &\quad + c\left(\big|\nu^{\frac{1}{2}}h^{s-1}u\big|_{s,2;\mathbb{T}_h} + \big|\nu^{-\frac{1}{2}}h^{s-1}p\big|_{s-1,2;\mathbb{T}_h}\right)^2\\ &\leqslant \frac{c_1}{2}\big\|\Pi_p e_p\big\|^2 + c\left(\big|\nu^{\frac{1}{2}}h^{s-1}u\big|_{s,2;\mathbb{T}_h} + \big|\nu^{-\frac{1}{2}}h^{s-1}p\big|_{s-1,2;\mathbb{T}_h}\right)^2,\end{aligned}$$

by Young inequality. Absorbing the first term to the left yields the estimate for the pressure. □

4.1.15 Theorem (weak estimates): *Let additionally to the assumptions of Theorem 4.1.10 the domain and coefficients of the problem be such that for a non-negative integer $t \leqslant k_u - 1$ the regularity estimate* (2.5) *holds with $\alpha = t+1$. Then, the error admits the estimate*

$$\left\|\nu^{\frac{1}{2}} e_u\right\|_{-t} + \left\|\nu^{-\frac{1}{2}} e_p\right\|_{-t-1} \leqslant C\left(\left|\nu^{\frac{1}{2}} h^{s+t} u\right|_{s,2;\mathbb{T}_h} + \left|\nu^{-\frac{1}{2}} h^{s+t} p\right|_{s-1,2;\mathbb{T}_h}\right). \tag{4.40}$$

In particular, for $t = 0$, we obtain

$$\left\|e_u\right\| \leqslant c\left|h^s u\right|_{s,2;\mathbb{T}_h}. \tag{4.41}$$

Proof. Like in the generic Theorem 1.3.14, let φ_u be either and arbitrary function in $W^{t,2}(\Omega;\mathbb{R}^d)$ if $t > 0$ or the error e_u if $t = 0$. Furthermore, let $\varphi_p \in W^{t+1,2}(\Omega)$ arbitrary. Then, the theorem is proven if we can show that

$$\frac{\big(e_u, \varphi_u\big) + \big(e_p, \varphi_p\big)}{\left\|\varphi_u\right\|_{t,2;\Omega} + \left\|\varphi_p\right\|_{t+1,2;\Omega}} \leqslant c\left(\left|\nu^{\frac{1}{2}} h^{s+t} u\right|_{s,2;\mathbb{T}_h} + \left|\nu^{-\frac{1}{2}} h^{s+t} p\right|_{s-1,2;\mathbb{T}_h}\right).$$

Let (z_u, z_p) be the solution of the dual problem (4.10) with right hand side (φ_u, φ_z) and $z_\sigma = \nu\nabla z_u$. Then, by consistency of the LDG method with the dual problem, we have for any function $(\tau, v, q) \in \Sigma \times V \times Q$:

$$\mathcal{S}_h\big(\tau, v, q; z_\sigma, z_u, z_p\big) = \big(\varphi_u, v\big) + \big(\varphi_p, q\big).$$

Entering $(\tau, v, q) = (e_\sigma, e_u, e_p)$ and applying Galerkin orthogonality yields

$$\begin{aligned}
\big(\varphi_u, e_u\big) &+ \big(\varphi_p, e_p\big) \\
&= \mathcal{S}_h\big(e_\sigma, e_u, e_p; \eta_{z_\sigma}, \eta_{z_u}, \eta_{z_p}\big) \\
&= \mathcal{S}_h\big(\eta_\sigma, \eta_u, \eta_p; \eta_{z_\sigma}, \eta_{z_u}, \eta_{z_p}\big) + \mathcal{S}_h\big(\Pi_\sigma e_\sigma, \Pi_u e_u, \Pi_p e_p; \eta_{z_\sigma}, \eta_{z_u}, \eta_{z_p}\big).
\end{aligned}$$

The first term is estimated by Lemma 4.1.11 and the regularity estimate (4.11):

$$\begin{aligned}
&\mathcal{S}_h\big(\eta_\sigma, \eta_u, \eta_p; \eta_{z_\sigma}, \eta_{z_u}, \eta_{z_p}\big) \\
&\qquad \leqslant c\big(\left|\nu^{\frac{1}{2}} h^{s+t} u\right|_{s,2;\mathbb{T}_h} + \left|\nu^{-\frac{1}{2}} h^{s+t} p\right|_{s-1,2;\mathbb{T}_h}\big) \times \\
&\qquad\qquad \times \big(\left|\nu^{\frac{1}{2}} z_u\right|_{t+2,2;\mathbb{T}_h} + \left|\nu^{-\frac{1}{2}} z_p\right|_{t+1,2;\mathbb{T}_h}\big) \\
&\qquad \leqslant c\big(\left|\nu^{\frac{1}{2}} h^{s+t} u\right|_{s,2;\mathbb{T}_h} + \left|\nu^{-\frac{1}{2}} h^{s+t} p\right|_{s-1,2;\mathbb{T}_h}\big) \times \\
&\qquad\qquad \times \big(\left|\nu^{\frac{1}{2}} \varphi_u\right|_{t,2;\mathbb{T}_h} + \left|\nu^{-\frac{1}{2}} \varphi_p\right|_{t+1,2;\mathbb{T}_h}\big).
\end{aligned}$$

Since the bilinear form $\mathcal{S}_h\big(.,.,.;.,.,.\big)$ is symmetric, we can apply Lemma 4.1.12 to estimate the second term

$$\begin{aligned}
&\mathcal{S}_h\big(\Pi_\sigma e_\sigma, \Pi_u e_u, \Pi_p e_p; \eta_{z_\sigma}, \eta_{z_u}, \eta_{z_p}\big) \\
&\qquad \leqslant \left|(\Pi_\sigma e_\sigma, \Pi_u e_u, \Pi_p e_p)\right|_{\mathrm{S}}\big(\left|\nu^{\frac{1}{2}} h^{t+1} z_u\right|_{t+2,2;\mathbb{T}_h} + \left|\nu^{-\frac{1}{2}} h^{t+1} z_p\right|_{t+1,2;\mathbb{T}_h}\big) \\
&\qquad \leqslant c\big(\left|\nu^{\frac{1}{2}} h^{s+t} u\right|_{s,2;\mathbb{T}_h} + \left|\nu^{-\frac{1}{2}} h^{s+t} p\right|_{s-1,2;\mathbb{T}_h}\big) \times \\
&\qquad\qquad \times \big(\left|\nu^{\frac{1}{2}} z_u\right|_{t+2,2;\mathbb{T}_h} + \left|\nu^{-\frac{1}{2}} z_p\right|_{t+1,2;\mathbb{T}_h}\big) \\
&\qquad \leqslant c\big(\left|\nu^{\frac{1}{2}} h^{s+t} u\right|_{s,2;\mathbb{T}_h} + \left|\nu^{-\frac{1}{2}} h^{s+t} p\right|_{s-1,2;\mathbb{T}_h}\big) \times \\
&\qquad\qquad \times \big(\left|\nu^{\frac{1}{2}} \varphi_u\right|_{t,2;\mathbb{T}_h} + \left|\nu^{-\frac{1}{2}} \varphi_p\right|_{t+1,2;\mathbb{T}_h}\big).
\end{aligned}$$

□

L	$\|e_\sigma\|$	order	$\|e_u\|$	order	$\|e_p\|$	order
	$\mathbb{Q}_1$, $\widehat{\gamma_{uu}} = 1$, $\widehat{\gamma_{pp}} = .06$					
1	8.2e-1	—	3.6e-1	—	5.2e-1	—
2	5.1e-1	0.68	1.0e-1	1.80	5.6e-1	-0.11
3	3.5e-1	0.52	2.5e-2	2.04	4.3e-1	0.39
4	2.2e-1	0.70	5.6e-3	2.16	2.5e-1	0.75
5	1.2e-1	0.86	1.3e-3	2.08	1.4e-1	0.91
6	6.2e-2	0.94	3.3e-4	2.01	7.0e-2	0.96
7	3.2e-2	0.97	8.3e-5	1.99	3.5e-2	0.98
	$\mathbb{Q}_2$, $\widehat{\gamma_{uu}} = 2$, $\widehat{\gamma_{pp}} = .03$					
1	8.5e-2	—	1.4e-2	—	8.0e-2	—
2	2.2e-2	1.96	1.5e-3	3.18	1.8e-2	2.16
3	5.5e-3	1.99	1.8e-4	3.05	4.9e-3	1.87
4	1.4e-3	1.98	2.3e-5	3.01	1.3e-3	1.91
5	3.5e-4	1.99	2.9e-6	3.00	3.3e-4	1.96
6	8.9e-5	1.99	3.6e-7	3.00	8.5e-5	1.98
7	2.2e-5	2.00	4.5e-8	3.00	2.1e-5	1.99
	$\mathbb{Q}_3$, $\widehat{\gamma_{uu}} = 4$, $\widehat{\gamma_{pp}} = .015$					
1	5.4e-3	—	2.7e-3	—	2.9e-3	—
2	1.2e-3	2.14	1.9e-4	3.86	9.7e-4	1.59
3	2.3e-4	2.41	1.3e-5	3.81	2.2e-4	2.13
4	3.4e-5	2.79	9.2e-7	3.85	3.3e-5	2.74
5	4.4e-6	2.94	6.1e-8	3.92	4.4e-6	2.93
6	5.6e-7	2.98	3.9e-9	3.96	5.6e-7	2.98

Table 4.1: Convergence of LDG discretization for Stokes equations

4.1.16 We verify our results approximating the solution in example A.4.2 on page 172 on Cartesian grids (see example A.1.1. We use $\nu = 1$ and refer to Section 5.2 on Oseen equations for different viscosity parameters. Table 4.1 shows that the convergence rates predicted by the analysis are indeed achieved and optimal.

4.1.17 Figure 4.1 shows that the influence of the stabilization parameter $\widehat{\gamma_{uu}}$ on the discretization accuracy is similar to the results for Poisson equation shown earlier in Figure 3.1 on page 76). Moreover, the parameter $\widehat{\gamma_{pp}}$ does not influence the error at all over a wide range. We will exploit this fact later and use this value to adjust the condition number of the Schur complement (Poisseuille flow example A.4.1 on Cartesian 32×32 grid, $\mathbb{Q}_1$ elements).

4.1.18 We demonstrate convergence of the discretization in three dimensions computing Poisseuille flow in a cylindrical pipe (see example A.4.1).The curved boundary is approximated by a bilinear mapping (yielding isoparametric $\mathbb{Q}_1$ elements).) The solution on the grid after one refinement step is shown in Figure 4.2. The convergence results are shown in Table 4.2. They exhibit the same expected convergence rates for σ and u as the results in two dimensions. The pressure exhibits superconvergence here, an effect also observed with stabilized conforming methods (cf. e.g. [Ber01]).

4.1.19 Remark: Like in the case of the LDG method for the Poisson equation, we can eliminate the variable σ by cellwise static condensation. Obviously, our results still hold for the

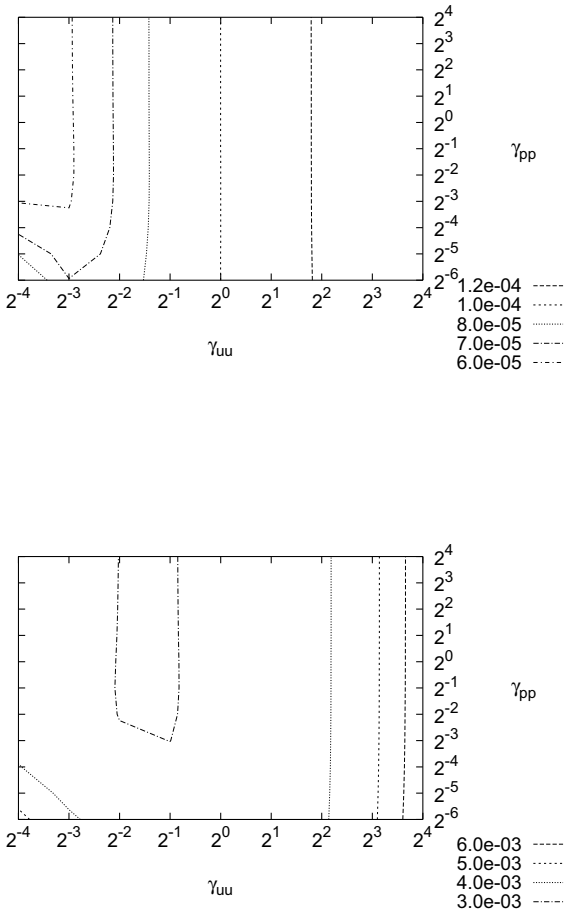

Figure 4.1: Discretization error depending on the two parameters. Errors e_u (top) and e_p (bottom) with homogeneous $\mathbb{Q}_1$ elements

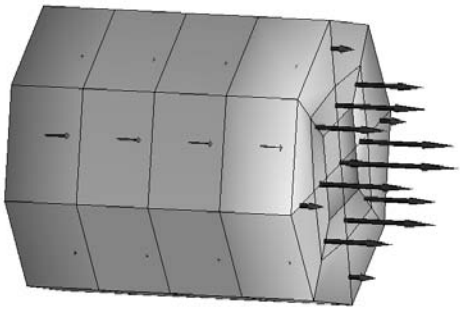

Figure 4.2: Discretization of three dimensional Poisseuille flow, grid after first refinement and solution

L	$\|e_\sigma\|$	order	$\|e_u\|$	order	$\|e_p\|$	order
	$\mathbb{Q}_1$, $\widehat{\gamma_{uu}} = 1$, $\widehat{\gamma_{pp}} = .05$					
1	1.3e+0	—	1.3e-1	—	2.3e+0	—
2	5.3e-1	1.28	3.8e-2	1.79	6.3e-1	1.85
3	2.5e-1	1.08	1.0e-2	1.88	1.7e-1	1.86
4	1.2e-1	1.01	2.8e-3	1.92	4.7e-2	1.88

Table 4.2: Convergence of LDG discretization for Stokes equations in 3D

resulting system, since it is algebraically equivalent. Using the equivalence in Corollary 3.2.13, it is clear that the results of this chapter apply to a Stokes discretization with the interior penalty method as well. In particular, existence and convergence results for u_h and p_h extend to the system

$$\begin{array}{ll} a_h^{\text{IP}}\big(u_h, v_h\big) + d_h(v_h, p_h) = f_h^u(v_h) & \forall\ v_h \in V_h \\ -d_h(u_h, q_h) + e_h(p_h, q_h) = f_h^p(q_h) & \forall\ q_h \in Q_h\,, \end{array} \tag{4.42}$$

where the form $a_h^{\text{IP}}\big(.,.\big)$ is understood to be applied component by component.

4.2 Stable Finite Element Pairs

4.2.1 Up to now we considered only discretizations with a pressure stabilization. Since stable finite element pairs will be of importance in Chapter 5, we will discuss these pairs here. They are analyzed in the context of mixed finite elements (see for instance [BF91]). Since mixed elements are beyond the scope of this book, we will cite the most important results without proof. Stable pairs are characterized by

4.2.2 Assumption (inf-sup condition): *There exists a constant $\gamma > 0$ independent of the mesh-size such that for any $q \in Q_h$ holds*

$$\sup_{0 \neq v \in V_h} \frac{d_h(v, q)}{\|v\|_{1,h}} \geq \gamma \|q\|. \tag{4.43}$$

Equivalently, we may write

$$\forall q \in Q_h \quad \exists v \in V_h \quad \textit{with} \quad \gamma \|v\|_{V_h} \leqslant \|q\|_{Q_h} \quad \textit{and} \quad b(v,q) \geqslant \|q\|_{Q_h}^2. \tag{4.44}$$

4.2.3 Lemma: *If Assumption 4.2.2 holds, the system* (4.16) *admits a unique solution even if* $\gamma_{pp} = 0$*. Theorems 4.1.10 and 4.1.14 hold analogously.*

Proof. This result is an immediate consequence of the well-posedness of the LDG method in the previous chapter and Theorem 1.1 in [BF91, Chapter II]. □

4.2.4 Proofs of Assumption 4.2.2 for special DG schemes can be found for instance in [HL02, Tos02, SST03]. All these proofs rely on proving the assumption first for the space

$$V_h^{\text{div}} = V_h \cap H^{\text{div}}(\Omega). \tag{4.45}$$

Actually, if we can prove (4.44) for V_h^{div} as a subspace of V_h, it is automatically proven for V_h as well.

Therefore, we introduce $H^{\text{div}}(\Omega)$-conforming finite element spaces and investigate their properties.

4.2.5 Definition (Raviart-Thomas elements): The finite element space $RT_k(\mathbb{T}_h)$ of vector valued functions consists of piecewise polynomials with each element p satisfying the following conditions:

1. On the reference cell, we have

   $$\widehat{RT_k} := \mathbb{Q}_k^d + x\mathbb{Q}_k,$$

   that is, it consists of tensor product polynomials of degree k plus in each component v_i the additional set of polynomials of the form $x_i^{k+1} \prod_{j \neq i} v_j(x_j)$ with v_j in $\mathbb{P}_k$.

2. The restriction $v|_T$ to a mesh cell T with transformation Ψ (see Definition 1.1.11) is a polynomial $\hat{v} \in \widehat{RT_k}$ mapped by the Piola transformation

   $$v(\Psi(\hat{x})) = \frac{1}{|\det \nabla \Psi(\hat{x})|} \nabla \Psi(\hat{x}) \hat{v}(\hat{x}) \tag{4.46}$$

3. The normal components of v at interfaces between cells are continuous, that is, on an edge E between cells T_1 and T_2 holds

   $$\mathbf{n} \cdot v|_{T_1} = \mathbf{n} \cdot v|_{T_2}.$$

4.2.6 Remark: From the definition, we see easily $V_h^{\text{div}} = RT_k(\mathbb{T}_h) \subset H^{\text{div}}(\Omega)$ and

$$\nabla \cdot \widehat{RT_K} = \mathbb{Q}_k. \tag{4.47}$$

Furthermore, $RT_k(\mathbb{T}_h)$ is a subset of $H^1(\mathbb{T}_h)$, but not of $H^1(\Omega)$. Therefore, defining jump terms on edges is still reasonable and a discretization of the velocity field in Stokes equations by RT_k is still a DG method.

Since the normal components of functions in V_h^{div} are continuous, their jumps in the form $d_h(.,.)$ vanish. Therefore, the weak incompressibility condition

$$d_h(u,q) = 0 \quad \forall q \in Q_h \tag{4.48}$$

simplifies to

$$\big(\nabla \cdot u, q\big) = 0 \quad \forall q \in Q_h. \tag{4.49}$$

4.2.7 Lemma: *The pair $V_h = RT_0$ and $Q_h = \mathbb{P}_0$ with norms*

and $\|.\|_{L^2(\Omega)}$ and $\|.\|_{1;h}$, respectively fulfills Assumption 4.2.2.

4.2.8 Lemma (Stability of Raviart-Thomas elements): *The pair $V_h = RT_k$ and $Q_h = \mathbb{Q}_k$ with norms $\|.\|_{1,h}$ and $\|.\|_{L^2(\Omega)}$ respectively fulfills Assumption 4.2.2 for any k with a constant γ independent of h.*

4.2.9 An alternative to Raviart-Thomas elements are the Brezzi-Douglas-Marini (BDM) elements in [BDM85]. While the matching pressure space for RT_k is $\mathbb{Q}_k$, the velocity space BDM_k is matched by $\mathbb{P}_{k-1}$.

4.2.10 Definition (BDM element): The space of BDM polynomials of degree k on the reference cell $[-1,1]^2$ is

$$BDM_k = \mathbb{P}_k \oplus \nabla \times (x^{k+1}y) \oplus \nabla \times (xy^{k+1}).$$

These shape functions are mapped to an actual mesh cell by the Piola transformation (4.46).

4.2.11 Remark: This space is chosen such that $\nabla \cdot BDM_k \subset \mathbb{P}_{k-1}$ and for every $u \in BDM_k$, $u \cdot \mathbf{n} \in P_k(F)$ on each face F.

4.2.12 Definition (BDM projection): We define for a sufficiently smooth function u the cell-wise interpolation operator I_T into the space BDM_k by the interpolation conditions

$$\big\langle I_T u \cdot \mathbf{n}, q \big\rangle_F = \big\langle u \cdot \mathbf{n}, q \big\rangle_F \qquad \forall\, q \in \mathbb{P}_k(F) \tag{4.50}$$

$$\big(I_T u, v\big)_T = \big(u, v\big)_T \qquad \forall\, v \in \mathbb{P}_{k-2}^2(T) \tag{4.51}$$

4.2.13 Remark: The interpolation operator in the previous above defines a polynomial $I_T u \in BDM_k$ uniquely. Furthermore, if applied on the whole mesh, its result is in $H^{\text{div}}(\Omega)$ and the commutator property $\nabla \cdot I_{BDM} u = \Pi_h \nabla \cdot u$ holds. For details, we refer the reader to [BF91].

4.2.14 Remark: On triangles, these moments are not sufficient to define the interpolation uniquely. Therefore, the moments

$$\big(I_T u, \varphi\big)_T = \big(u, \varphi\big)_T \qquad\qquad \forall \varphi \in \Phi_k(T),$$

are added, where

$$\Phi_k(T) = \{\Phi \in L^2(T)^2 : D\Psi_T^t \varphi \circ \Psi_T \in \Phi_k(\widehat{T})\}.$$

Here, $\Psi_T : \widehat{T} \to T$ denotes the elemental mapping and $D\Psi_T$ its Jacobian. On the reference triangle $\widehat{T} = \{(\widehat{x}_1, \widehat{x}_2) : \widehat{x}_1 > 0,\ \widehat{x}_1 + \widehat{x}_2 < 1\}$, the space $\Psi_k(\widehat{T})$ is defined by

$$\Psi_k(\widehat{T}) = \{\Psi \in P_k(\widehat{T})^2 :\ \nabla \cdot \Psi = 0 \text{ in } \widehat{T},\ \Psi \cdot n_{\widehat{T}} = 0 \text{ on } \partial\widehat{T}\}.$$

4.2.15 Remark: Both vector spaces BDM_k and RT_k contain the complete polynomial spaces $\mathbb{P}_k^d$. Therefore, the interpolation estimate (1.21) holds for these spaces as well.

Chapter 5

Flow Problems

In the following sections, we study the effect of the addition of lower order terms to the Laplacian or to the Stokes operator. The convergence results of the preceding chapters are used in combination with the stable advection discretization of Section 1.4. This way, results for advection-diffusion-reaction and Oseen equations are easily obtained.

Advection-diffusion problems are a straight-forward extension of the methods developed for Poisson equation (see e. g. Cockburn and Dawson [CD00] and Gopalakrishnan and Kanschat [GK03c]). We present an example with a dominant reaction term, showing that the DG method yields considerably better approximation on coarse grids.

Discontinuous Galerkin methods for Oseen equations were first studied by Cockburn, Kanschat and Schötzau in [CKS03b]. We present an alternative way of analyzing the scheme presented there in Section 5.2.

We close this chapter presenting results on incompressible Navier-Stokes equations. A formulation using discontinuous piecewise solenoidal velocities and continuous pressures was proposed by Karakashian and Jureidini in [KJ98]. The first genuinely discontinuous Galerkin scheme was presented by Girault, Rivière and Wheeler in [GRW02]. It uses a nonsymmetric version of the interior penalty method for the Stokes part of the equation and a stabilized trilinear form for the advection term. We combine their trilinear form with the LDG discretization of Stokes equations and show computational results indicating the feasibility of the method.

Alternatively, the use of divergence conforming finite element spaces for the velocities has been proposed in [CKS05, CKS07]; this approach yields pointwise a. e. divergence free discrete solutions in $H^{\mathrm{div}}(\Omega)$, such that the stability of the linearized problems is guaranteed, but it avoids the construction of a (piecewise) solenoidal basis. For the latter reason, the pressure of this method still approximates the physical quantity. This method is presented in two different flavors at the end of the chapter.

5.0.1 We consider incompressible flow modeled by the stationary, incompressible *Navier-Stokes equations*

$$\begin{aligned} -\nu\Delta u + u\cdot\nabla u + \nabla p &= f \\ \nabla\cdot u &= 0 \end{aligned} \qquad \text{in } \Omega\subset\mathbb{R}^d \tag{5.1}$$

with boundary conditions (4.2) and (4.3) as for the Stokes problem; we restrict the analysis to the pure Dirichlet case, since the question of unique solutions in the presence of natural

boundary conditions is still open. The following analytical results are quoted from [QV94, Chapter 10].

For the question of existence and uniqueness of solutions to this equation as well as for the numerical solution a suitable sequence of linear problems must be solved. This can be achieved by

5.0.2 Definition (Picard iteration): Let $u^0 \in W^{1,2}(\Omega)$ be given, *Picard iteration* defines a sequence $\{u^k\}, k = 1, \ldots$ through the recurrence relation

$$\begin{array}{ll} -\nu\Delta u^k + u^{k-1}\cdot\nabla u^k + \nabla p^k = f \\ \nabla \cdot u^k = 0 \end{array} \tag{5.2}$$

5.0.3 Definition: The resulting linear system is usally refered to as *Oseen equations*. Their weak form is: let $V = W_0^{1,2}(\Omega;\mathbb{R}^d)$ and find $(u,p) \in V \times \tilde{Q}$ (compare (4.7)) such that

$$\begin{array}{lll} -\big(\nu\nabla u, \nabla v\big) + \big(w\cdot\nabla u, v\big) + \big(\nabla\cdot v, p\big) = (f,v) & \forall\, v \in V \\ \big(\nabla\cdot u, q\big) = 0 & \forall\, q \in \tilde{Q} \end{array} \tag{5.3}$$

5.0.4 Lemma: *If the domain Ω is two or three dimensional, the trilinear form $\big(w\cdot\nabla u, v\big)$ admits the estimate*

$$\big(w\cdot\nabla u, v\big) \leqslant C|w|_{W^{1,2}}\,|u|_{W^{1,2}}\,|v|_{W^{1,2}}. \tag{5.4}$$

Proof. Using Hölder inequality twice, we obtain

$$\big(w\cdot\nabla u, v\big) \leqslant |w|_{W^{1,2}}\big\|uv\big\|_{L^2} \leqslant |w|_{W^{1,2}}\big\|u\big\|_{L^4}\big\|v\big\|_{L^4}.$$

Applying the embedding Theorem 1.2.2, the estimate (5.4) follows with C being the square of the constant in (1.9). □

5.0.5 Lemma: *Assume that $w \in W_0^{1,2}(\Omega;\mathbb{R}^d)$ and*

$$\big(\nabla\cdot w, |v|^2\big) \leqslant 0 \quad \forall v \in L^4(\Omega). \tag{5.5}$$

Then the Oseen problem (5.3) *has a unique solution $(u,p) \in V \times \tilde{Q}$ and*

$$\big\|u\big\|_{W^{1,2}} \leqslant \frac{C_p}{\nu}\|f\|_{L^2(\Omega)}, \tag{5.6}$$

Proof. By partial integration and condition (5.5), we have for any $w, v \in V$

$$\big(w\cdot\nabla v, v\big) = -\big(v, w\cdot\nabla v\big) - \big(\nabla\cdot wv, v\big) \geqslant -\big(v, w\cdot\nabla v\big),$$

and thus $\big(w\cdot\nabla v, v\big) \geqslant 0$. Therefore, for any $v \in V$

$$\big(\nu\nabla v, \nabla v\big) + \big(w\cdot\nabla v, v\big) \geqslant \nu|v|^2_{1,2;\Omega} \tag{5.7}$$

independent of w. Now, let

$$\tilde{V} = \big\{v \in V \big| \nabla\cdot v = 0\big\}.$$

Then, the reduced problem

$$\big(\nu\nabla u, \nabla v\big) + \big(w\cdot\nabla u, v\big) = \big(f, v\big) \quad \forall v \in \tilde{V} \tag{5.8}$$

admits a unique solution $u \in \tilde{V}$ admitting estimate (5.6) due to (5.7) and the Lax-Milgram Lemma 1.3.6. □

5.0.6 Theorem: *If the right hand side f is divergence free and admits the estimate*

$$\|f\| \leqslant \frac{\nu^2}{CC_\Omega}, \tag{5.9}$$

where C_Ω is the diameter of Ω and C the constant in (5.4)*, then the Navier-Stokes equations* (5.1) *in weak form have a unique solution $(u, p) \in V \times \tilde{Q}$.*

Proof. By contraction property of Picard iteration; see e. g. [QV94, Ch. 10]. □

5.0.7 We will approach the discretization of the Navier-Stokes equations (5.1) by combining the techniques of the previous chapters step by step. Therefore, we start by adding linear advection to the Laplacian and Stokes operators in order to end up using the discretization obtained in Picard iteration for the complete nonlinear problem.

5.1 Advection-Diffusion-Reaction Equation

5.1.1 The simplest model for diffusive transport problems is the linear advection-diffusion-reaction equation

$$-\nu\Delta u + b\cdot\nabla u + \varrho u = f, \tag{5.10}$$

with non-negative parameters ν and ϱ. The vector field $b \in L^\infty(\Omega; \mathbb{R}^d)$ fulfils the stability condition

$$\nabla\cdot b - \varrho \leqslant 0 \quad \text{a.e. in } \Omega. \tag{5.11}$$

5.1.2 We derive a discontinuous Galerkin discretization* of equation (5.10) by simply adding the stable advection discretization (1.50), the interior penalty discretization of the Laplacian in (2.12) and an additional mass matrix, yielding the weak formulation: find $u_h \in V_h$

$$\nu a_h^{\text{IP}}\big(u_h, v\big) + \beta_h\big(u_h, v\big) + \big(\varrho u_h, v\big)_{\mathbb{T}_h} = (f, v)_{\mathbb{T}_h} \quad \forall v \in V_h. \tag{5.12}$$

Here, the space V_h is chosen as in Section 2.2. Accordingly, the energy norm is chosen as

$$\big\|u\big\|_{\text{ADR}}^2 = \nu\big\|u\big\|_{1;h}^2 + \big\|u\big\|_\beta^2 + \varrho\big\|u\big\|^2. \tag{5.13}$$

5.1.3 Remark: The stability and convergence analysis of this discretization are conducted by combining the results of sections 1.4 and 2.2. If the test functions used for the stability analysis in both sections would be the same, we could just add the results and yield the combined estimates. As it is, the test functions for the advection equation had to be augmented by the term $\gamma hb\cdot\nabla u_h$ in Lemma 1.4.11 on page 30. Therefore, we will have to show that this term does not spoil stability and boundedness of the other operators involved.

5.1.4 Lemma: *For any function $v \in W^{s,2}(\mathbb{T}_h)$, the energy norm admits the L^2-projection estimate*

$$\big\|v - \Pi_h v\big\|_{ADR} \leqslant \big|\alpha v\big|_{s,2;\mathbb{T}_h}, \tag{5.14}$$

with $\alpha = \max(\sqrt{\nu}h^{s-1}, \sqrt{|b|}h^{s-1/2}, \sqrt{\varrho}h^s)$.

Proof. By summing up the projection estimates for each term of the energy norm separately, namely (1.57), (2.17) and (1.21). □

5.1.5 Lemma: *There exist constants c and γ such that the following stability estimate holds for any $v \in V_h$ (abbreviating $\tilde{v} = \gamma hb\cdot\nabla v$):*

$$\big\|v\big\|_{ADR}^2 \leqslant c\Big(\nu a_h^{IP}\big(v, v+\tilde{v}\big) + \beta_h\big(v, v+\tilde{v}\big) + \big(\varrho v, v+\tilde{v}\big)_{\mathbb{T}_h}\Big). \tag{5.15}$$

Proof. First, we apply Lemma 1.4.11 to the advection term and obtain constants c_a and γ_a such that

$$\big\|v + h\gamma_a b\cdot\nabla v\big\|_\beta^2 \leqslant c_a\beta_h\big(v, v + h\gamma_a b\cdot\nabla v\big).$$

We observe that this estimate holds with a different constant c for any positive $\gamma \leqslant \gamma_a$

The diffusion and reaction terms are both handled using inverse estimate, e. g.

$$\big(\varrho v, \gamma_r hb\cdot\nabla v\big) \leqslant c\gamma_r\big(\varrho v, v\big),$$

and choose γ_r such that $c\gamma_r \leqslant 1/2$. The terms in $\nu a_h^{\text{IP}}\big(v, \gamma_d hb\cdot\nabla v\big)$ are estimated accordingly. Finally, we chose γ as the minimum of γ_a, γ_r and γ_d to obtain the stability estimate. □

*An *hp*-analysis of a formulation with a nonsymmetric interior penalty discretization can be found in [HSS02]

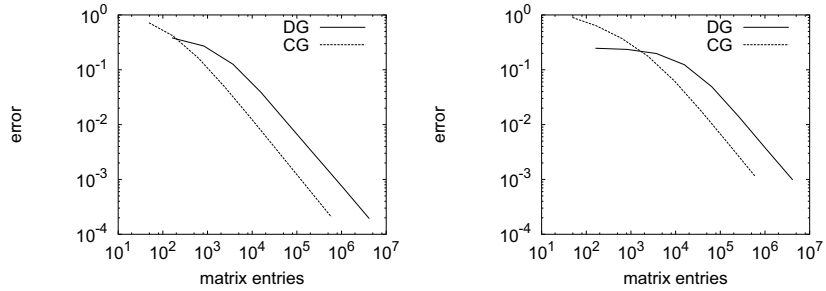

Figure 5.1: Comparison of DGFEM and CGFEM (both $\mathbb{Q}_1$ for a reaction dominated diffusion problem with $\nu = \frac{1}{100}$ (left) and $\nu = \frac{1}{1000}$ (right)

5.1.6 Corollary: *By the same inverse estimates as used in the last proof, we obtain a constant c independent of h and the parameters such that for any function $v \in V_h$*

$$\big\|v + \gamma h b\cdot\nabla v\big\|_{ADR} \leqslant c\big\|v\big\|_{ADR}. \tag{5.16}$$

5.1.7 Theorem: *Let $u \in W^{s,2}(\mathbb{T}_h)$ with $3/2 \leqslant s \leqslant k+1$ and $u_h \in V_h$ be solutions of equations* (5.10) *and* (5.12)*, respectively. Then, the error $e_u = u - u_h$ admits the estimate*

$$\big\|e_u\big\|_{ADR} \leqslant c\big|\alpha u\big|_{s,2;\mathbb{T}_h}, \tag{5.17}$$

where $\alpha = \max(\sqrt{\nu}h^{s-1}, \sqrt{|b|}h^{s-1/2}, \sqrt{\varrho}h^s)$.

Proof. The proof follows the generic proof of Theorem 1.3.9. Assumptions 1.3.5 to 1.3.8 hold true as a consequence of the preceding lemmas and corollaries. □

5.1.8 In Figure 5.1, we compare the accuracy of solutions to the discontinuous Galerkin scheme (5.12) and continuous finite elements for discretization of example A.3.1. We observe that the discontinuous method is better on coarse meshes, as long as the boundary layer is not resolved properly† An explanation for this behavior is that the discontinuous method does not obey the boundary values as long as the layer is not resolved, therefore being much closer to the limit solution. On the other hand, the continuous method is known to produce oscillations (see Figure 5.2) unless special measures like mass lumping or stabilization are taken. These on the other hand must be implemented carefully. Mass lumping for instance spoils Galerkin orthogonality of the error. The discontinuous scheme shows that the additional freedom introduced through the flux functions can be used to adapt the accuracy of boundary values to the discretization accuracy.

For mesh parameters h of the size of the boundary layer and smaller, we observe that the asymptotic behavior of the methods is similar, but the continuous method needs less matrix entries to

†The error was integrated using a 40^2 point Gauß quadrature rule. This way, we asserted that the error was measured correctly even on coarse meshes.

Figure 5.2: CGFEM (left) and DGFEM (right) solution of a reaction-dominated ($\nu = \frac{1}{1000}$) diffusion problem

achieve the same accuracy as the DG scheme. A quantitative investigation reveils that the error on fine meshes for these schemes is about the same if cell sizes are equal. But in this case, the DG scheme has more degrees of freedom and a denser matrix stencil, such that its numerical effort is larger.

5.2 Oseen Equations

5.2.1 Here, we extend the Stokes discretization of Chapter 4 to the case, where a linear advection operator of the type in equation (1.46) is added to the Stokes form. Similarly to advection-diffusion equations, the aim of the analysis is the derivation of a discretization and error estimates not deteriorating when the elliptic part becomes small. To this end, we again add the stable discontinuous Galerkin discretization for the advection problem in equation (1.50). The method and its analysis in this section were developed in [CKS03b, CKS02].

5.2.2 The mixed formulation of Oseen equations derived from Stokes equations (4.9), augmented by an advection term and an additional reaction term to account for instance for time stepping schemes reads

$$\begin{array}{rcccccll} ac1\nu\sigma & - & \nabla u & & & = & 0 & \\ -\nabla\cdot\sigma & + & b\cdot\nabla u + \gamma u & + & \nabla p & = & f & \text{in } \Omega. \\ & & \nabla\cdot u & & & = & 0 & \end{array} \tag{5.18}$$

We make the rather weak stability assumption that

$$0 \leqslant \gamma_0(x) := \gamma(x) - \frac{1}{2}\nabla\cdot b(x), \tag{5.19}$$

almost everywhere in Ω. In particular, $\gamma \equiv 0$ and a divergence free vector field b are permitted. Boundary conditions are chosen as for Stokes equations.

5.2.3 Remark: We are not assuming that γ_0 is uniformly bounded from zero by a positive constant. This assumption is quite common since it allows a simpler analysis (see [GR86, HSS02]), but it is not valid for the case of stationary, incompressible Navier-Stokes problems.

5.2.4 We apply the discretizations of the previous chapter and the previous section to the weak form of (5.18) to obtain the discontinuous Galerkin bilinear form

$$\begin{aligned} a_h(\sigma_h,\tau) + \quad b_h(u_h,\tau) \quad &= f_h^\sigma(\tau) \quad \forall \tau \in \Sigma_h, \\ -b_h(v,\sigma_h) + c_h(u_h,v) + \beta_h\big(u_h,v\big) + d_h(v,p_h) &= f_h^u(v) \quad \forall v \in V_h, \\ - \quad d_h(u_h,q) \quad + e_h(p_h,q) &= f_h^p(q) \quad \forall q \in Q_h, \end{aligned} \tag{5.20}$$

where the particular bilinear forms are defined in equations (4.17) to (4.21); in particular, $b_h(.,\tau)$ is the gradient on the velocity space tested with stress functions τ and $d_h(.,q)$ is the divergence of the velocity tested with a pressure test function. The form $\beta_h\big(u_h,v\big)$ defined in (1.50) is applied to each component of u separately, i. e., $\beta_h\big(u,v\big) = \sum \beta_h\big(u_i,v_i\big)$. The bilinear form of the complete Oseen problem is then

$$\begin{aligned} \mathcal{O}_h\big(\sigma,u,p;\tau,v,q\big) := a_h(\sigma,\tau) + b_h(u,\tau) - b_h(v,\sigma) + c_h(u,v) + \beta_h\big(u,v\big) \\ + d_h(v,p) - d_h(u,q) + e_h(p_h,q) \end{aligned}$$

Like in the previous section, we define the energy seminorm for this problem by adding the energy seminorm of the Stokes problem and of the advection problem, yielding

$$\big|(\sigma,u,p)\big|_{\mathrm{O}}^2 = \big|(\sigma,u,p)\big|_{\mathrm{S}}^2 + \sum_{i=1}^{d} \big\|u_i\big\|_\beta^2.$$

Since both forms are (semi)definite, so is their sum. Even more, we have the following stability result.

5.2.5 Lemma: *There exist constants c and γ, such that the Oseen bilinear form defined in* (5.20) *admits the stability estimates*

$$0 \leqslant \mathcal{O}_h\big(\sigma,u,p;\sigma,u,p\big) \tag{5.21}$$

$$\big|(\sigma,u,p)\big|_O^2 \leqslant c\mathcal{O}_h\big(\sigma,u,p;\sigma,u+h\gamma b\cdot\nabla u,p\big). \tag{5.22}$$

Proof. This result is obtained by combining the proofs of stability of the Stokes discretization in Chapter 4 and Lemma 5.1.5. Since the term $\mathcal{O}_h\big(0,u,0;0,h\gamma b\cdot\nabla u,0\big)$ can be controlled by inverse estimates exactly like in Lemma 5.1.5, we leave the details to the reader. □

5.2.6 Theorem: *Under the assumptions of Theorem 4.1.10 and those made in paragraph 5.2.4, the error (e_σ,e_u,e_p) between the discrete solution (σ_h,u_h,p_h) of* (5.20) *and the continuous solution (σ,u,p) of* (5.18) *admits the estimates*

$$\big|(e_\sigma,e_u,e_p)\big|_O \leqslant c\Big(\big|\alpha u\big|_{s,2;\mathbb{T}_h} + \big|h^{s-1}p\big|_{s-1,2;\mathbb{T}_h}\Big), \tag{5.23}$$

where $\alpha = \max(\sqrt{\nu}h^{s-1}, |b|h^{s-1/2})$.

Proof. Replacing the Stokes bilinear form by the Oseen bilinear form, estimates (4.33) and (4.34) are transformed into

$$\big|(\sigma,u,p)\big|_{\mathrm{O}} \leqslant c\Big(\big|\nu^{-\frac{1}{2}}\sigma\big|^2_{s-1,2;\mathbb{T}_h} + \big|\alpha u\big|^2_{s,2;\mathbb{T}_h} + \big|\nu^{-\frac{1}{2}}p\big|^2_{s-1,2;\mathbb{T}_h}\Big),$$

and

$$\mathcal{O}_h\big(\eta_\sigma,\eta_u,\eta_p;\tau,v,q\big) \leqslant c\big|(\tau,v,q)\big|_{\mathrm{O}}\Big(\big|\nu^{-\frac{1}{2}}\sigma\big|_{s-1,2;\mathbb{T}_h} + \big|\alpha u\big|_{s,2;\mathbb{T}_h} + \big|\nu^{-\frac{1}{2}}p\big|_{s-1,2;\mathbb{T}_h}\Big).$$

Combined with stability estimate (5.22) and Galerkin orthogonality, these two estimates allow us to obtain an energy norm estimate following the generic proof of Theorem 1.3.9. Estimate (5.23) follows by letting $\sigma = \nu\nabla u$. □

5.2.7 Remark: Again, like in Remark 4.1.19, the stability result holds if the additional stress tensor σ is eliminated from the system by static condensation. Then, combining the stability analysis for the Stokes problem with the analysis for the advection-reaction-diffusion problem in (5.12), we obtain stability and convergence results in the norm consisting of the L^2-norm of the pressure and the natural norm for the advection probelm defined in (1.53)

$$\big\|u\big\|^2_{\mathrm{O;IP}} = \big\|u\big\|^2_{\mathrm{ADR}} + \big\|p\big\|^2.$$

5.2.8 We show results for Kovasznay flow (example A.4.3 on page 173) with different Reynolds numbers $\mathrm{Re} = 1/\nu$ in Figure 5.3, where we used $\mathbb{Q}_2$ elements for V_h and Q_h. We observe h^3-convergence for u and h^2 for σ over the whole range of meshes and viscosities. Again, this is the superconvergence effect of the advection discretization observed earlier in this chapter. For the pressure, we observe the predicted convergence rates on fine grids. On coarser grids, the pressure approximation seems to suffer, resulting in a higher convergence rate.

The figure also shows that the method is robust with respect to the Reynolds number, as the errors do not increase. The scaling of the errors in the figure is according to the theoretical results. The errors in σ and p are in fact decreasing, which is due to the fact that the functions themselves are of order ν (see Figure A.9 on page 174). Concluding, we observe that the relative errors in u, p and σ are nearly independent of the viscosity.

5.3 Navier-Stokes Equations

5.3.1 After an accurate and stable discretization for the Oseen problem has been found, we would like to use this discretization in the Picard iteration (5.2). We cannot do this in a straight forward fashion, since the discrete velocities are only weakly divergence free, when tested with the pressure space (see [CKS05]). We will shortly discuss the problems arising by considering the continuous form of the equations (5.1).

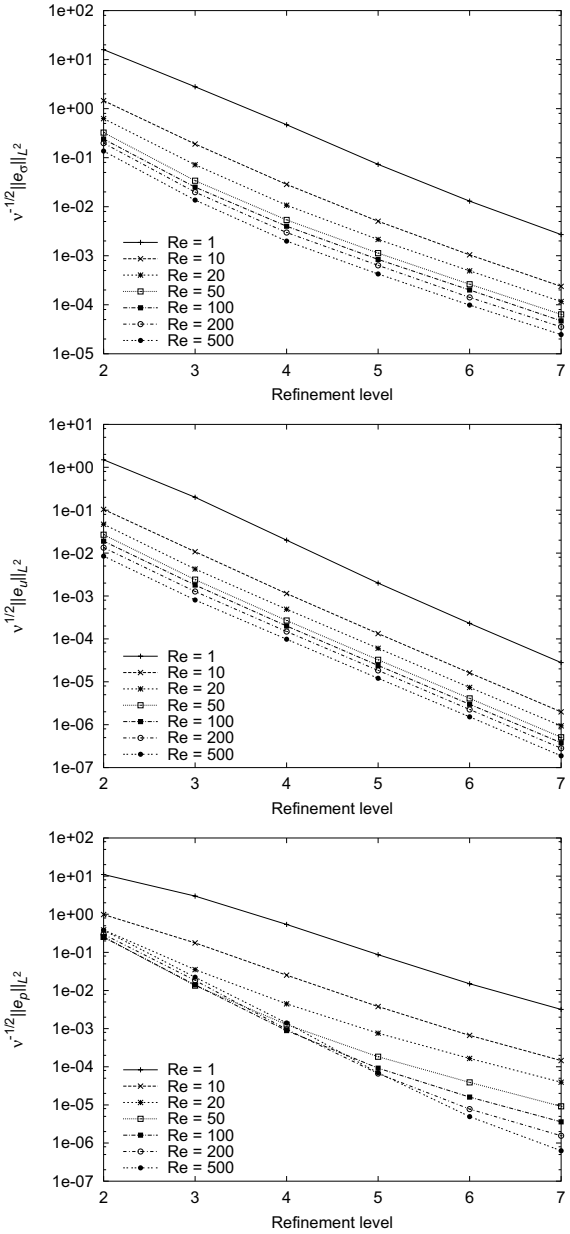

Figure 5.3: Approximation of Kovasznay flow (Oseen equations) with LDG and homogeneous $\mathbb{Q}_2$-elements

5.3.2 If we multiply the first equation of (5.1) by u, integrate by parts and use the boundary conditions, we get

$$\big(\nu\nabla u, \nabla u\big)_\Omega + \frac{1}{2}\big(|u|^2, \nabla\cdot u\big)_\Omega - \big(p, \nabla\cdot u\big)_\Omega = \big(f, u\big)_\Omega.$$

We see that we must use the incompressibility condition to obtain the equation

$$\big(\nu\nabla u, \nabla u\big)_\Omega = \big(f, u\big)_\Omega$$

from which the stability estimate (5.6) immediately follows.

In general, since exact incompressibility is very difficult to achieve after discretization, it is usually only enforced weakly. This weak incompressibility is enough, in a wide variety of cases, to guarantee that the discrete version of the term

$$\big(p, \nabla\cdot u\big)$$

is exactly zero, as for most mixed methods for the Stokes and Navier-Stokes equations, or non-negative, as for the DG methods considered for the Stokes and Oseen problems seen above. Unfortunately, this is not true for the discrete version of the term

$$\frac{1}{2}\big(|u|^2, \nabla\cdot u\big). \tag{5.24}$$

because the square of the modulus of the approximate velocity does *not* necessarily (and usually does not) belong to the space of the approximate pressure.

5.3.3 Remark: If the viscosity ν is not too small and continuous finite elements are used, it is possible to control the divergence through the energy norm estimate. Actually, we have

$$\|\nabla\cdot u_h\| = \|\nabla\cdot u_h - \nabla\cdot u\| \leqslant \|\nabla(u_h - u)\| \leqslant ch^k|u|_{k+1}$$

for polynomials of degree k. Furthermore, u_h is bounded in $L^4(\Omega)$ Therefore, the term (5.24) will become negligible on fine meshes. It is important to note though that the solvability of the Oseen problem will depend on the mesh size and the current iterate of the Picard iteration, a situation hardly acceptable for practical applications.

For a general discontinuous finite element method, this situation is even worse, since $\nabla\cdot u_h$ is not even in $L^2(\Omega)$ and the jumps of u_h across cell edges have to be controlled as well; furthermore, the integration by parts leading to (5.3.2) can only be performed cell by cell, yielding another possible source of instability.

5.3.4 In the sixties, a modification of the nonlinearity was suggested to solve this problem; see [Tém66] and [Tém68]. It consists of adding a consistent term (since for the continuous solution holds $\nabla\cdot u = 0$) as a kind of divergence correction:

$$\begin{aligned} -\nu\Delta u + \nabla\cdot(u\otimes u) - \frac{1}{2}(\nabla\cdot u)\,u + \nabla p &= f && \text{in } \Omega, \\ \nabla\cdot u &= 0 && \text{in } \Omega, \\ u &= 0 && \text{on } \Gamma. \end{aligned} \tag{5.25}$$

Multiplying the first equation by u, integrating by parts and using the boundary conditions, we get

$$\big(\nu\nabla u, \nabla u\big)_\Omega - \big(p, \nabla\cdot u\big)_\Omega = \big(f, u\big)_\Omega.$$

As a consequence, stability can follow from weak incompressibility. It is also the case for the first discontinuous Galerkin method for the incompressible Navier-Stokes equations [KJ98], a method which uses *locally* divergence-free polynomial approximations of the velocity, and for the more recent discontinuous Galerkin method developed in [GRW02] (see 5.3.8 below).The only problem with this approach is that local conservativity cannot be achieved because the first equation is not written in divergence form; we will discuss the construction of conservative methods at the end of this chapter.

5.3.5 Remark: In order to keep the presentation of the following sections simple, we focus on the discretization of the Laplacian operator by the interior penalty method. The analysis here relies on the ellipticity of the discrete form stated in Lemma 2.2.5 and its boundedness. Therefore, the form $a_h^{\mathrm{IP}}\big(.,.\big)$ can by Corollary 3.2.13 be replaced for instance by the Schur complement of the LDG method.

5.3.6 Determining a DG formulation for the nonlinear advection term is a topic of ongoing research. Since the numerical solution usually is not pointwise divergence free, special care has to be taken for the possible destabilizing effect of its divergence. This problem was circumvented by Karakashian and Jureidini in [KJ98] by using solenoidal shape functions. A broader view on the piecewise solenoidal approach is given in the recent article [HL06]

5.3.7 Since we do not use solenoidal ansatz spaces, we present two possible techniques for eliminating the destabilizing effect of the divergence. One consists in subtracting the discrete divergence of the formulation and was suggested by Girault et al. in [GRW02]. The other option is forcing the divergence of the discrete velocity u_h to zero by choosing suitable combinations of function spaces. Two different versions of this approach were suggested in [CKS05] and [CKS07] and will be discussed at the end of this chapter.

5.3.8 In the following, we will show results produced with the trilinear form

$$\tilde{c}_h(u;v,w) = \big(u\cdot\nabla v, w\big)_{\mathbb{T}_h} + \sum_{T\in T_h}\big\langle|\{\!\{u\}\!\}\cdot\mathbf{n}|\,(v^\downarrow - v^\uparrow), w^\downarrow\big\rangle_{\partial T_-} + \frac{1}{2}\big(\nabla\cdot uv, w\big)_{\mathbb{T}_h} - \frac{1}{2}\big\langle[\![u\cdot\mathbf{n}]\!], \{\!\{v\cdot w\}\!\}\big\rangle_{\mathbb{E}_h^i}, \quad (5.26)$$

suggested in [GRW02] with the same rationale as in (5.25). Here,

$$\partial T_- = \big\{x\in\partial\mathbb{T}\big|u_T(x)\cdot\mathbf{n}(x) < 0\big\}. \quad (5.27)$$

5.3.9 Remark: In the formulation in the previous paragraph, the inflow boundary ∂T_- and the notion of upwind and downwind values are with respect to the value of the possibly discontinuous velocity u *inside* the mesh cell T. In particular, if $E = T^1\cap T^2$ is the edge between two grid cells, the union $\partial T_-^1\cup(\partial T^2\setminus\partial T_-^2)$ may not contain the whole edge E as a subset and the intersection $\partial T_-^1\cap(\partial T^2\setminus\partial T_-^2)$ may be nonempty. For this reason, the classical stability analysis using integration by parts (see Lemma 1.4.7) cannot be used.

L	$\|\|e_\sigma\|\|$	order	$\|\|e_u\|\|$	order	$\|\|e_p\|\|$	order	steps
2	7.1e-1	—	7.2e-1	—	1.3e+0	—	38
3	4.3e-1	0.73	2.7e-1	1.43	4.8e-1	1.44	35
4	2.5e-1	0.75	6.8e-2	1.97	2.1e-1	1.18	25
5	1.4e-1	0.86	1.9e-2	1.88	9.7e-2	1.12	19
6	7.3e-2	0.94	5.1e-3	1.88	4.7e-2	1.04	18

Table 5.1: Convergence for Navier-Stokes equations (Kovasznay flow, $Re = 10$, $\mathbb{Q}_1$)

5.3.10 The DG discretization of the stationary Navier-Stokes equations is obtained by replacing the advection form $\beta_h\big(u_h, v\big)$ in (5.20) by $\tilde{c}_h(u_h; u_h, v)$ (and here replacing the LDG method by interior penalty), yielding the nonlinear system of equations

$$\begin{aligned} a_h^{\text{IP}}\big(u_h, v\big) + \tilde{c}_h(u_h; u_h, v) + d_h(v, p_h) &= f_h^u(v) \;\; \forall v \in V_h, \\ -d_h(u_h, q) \qquad\quad + e_h(p_h, q) &= f_h^p(q) \;\; \forall q \in Q_h. \end{aligned} \tag{5.28}$$

5.3.11 The nonlinear problem (5.28) is solved successively by the Picard iteration in Definition 5.0.2. For the DG method, it reads: given the previous approximation $(u^{(n)}, p^{(n)})$, compute (u, p) by solving the linear system

$$\begin{aligned} a_h^{\text{IP}}\big(u, v\big) + \tilde{c}_h(u^{(n)}; u, v) + d_h(v, p) &= f^u(v) \;\; \forall v \in V, \\ -d_h(u, q) \qquad\quad + e_h(p, q) &= f_h^p(q) \;\; \forall q \in Q_h, \end{aligned} \tag{5.29}$$

and let $(u^{(n+1)}, p^{(n+1)}) = (u, p)$

5.3.12 In order to demonstrate the convergence properties of this discretization (here with the LDG method for the elliptic operator), we solve for Kovasznay's analytical solution to the Navier-Stokes equations (see example A.4.3). Table 5.1 summarizes the results. First, we see that the convergence orders are as expected: one for σ and p and two for the velocity ($\mathbb{Q}_1$-elements have been used).

The rightmost column shows the number of Picard iteration steps needed to obtain a final residual below 10^{-10} (start residuals were all between 3 and 10). Here, the linear solver for the system (5.29) was iterated until the residual reached one hundredth of the start residual. The falling numbers suggest that Picard iteration converges better if the problem is approximated better.

5.3.13 The remainder of this chapter is devoted to schemes using the original form (5.1) of the Navier-Stokes equations. They are based on enforcing *exact* incompressibility in the space $H^{\text{div}}(\Omega)$ defined in (1.7). It is clear from the reasoning in Paragraph 5.3.2, that this yields stable Oseen problems. We consider two versions of this method: the first relies on a stable DG scheme with mixed order $\mathbb{Q}_k/\mathbb{P}_{k-1}$ discretization for V_h and Q_h with a postprocessing of the velocity into a divergence conforming space. Later on, we present a DG scheme directly formulated in a subspace of $H^{\text{div}}(\Omega)$.

5.3.14 In order to derive the first method, we have a closer look at the Picard iteration (5.29). By Lemma 5.2.5 (resp. Lemma 1.4.11), we see that only the divergence of $w = u^{(n)}$ enters the stability estimate for the form $O(w; u, v)$.

Therefore, we consider a method with *two* different approximations to the velocity: One approximation for u and another for w. For stability of the method, it is sufficient, that w is strongly incompressible, while we still compute a solution u which only weakly incompressible. Local conservativity can be achieved for such a DG method: even if the equations are not written in conservative form, we can be compensate for this by the fact that the approximation w is globally divergence-free.

We construct the second approximation $w \in W_h$, where W_h is the $H^{\mathrm{div}}(\Omega)$-conforming space constructed with BDM elements (see Definition 4.2.10), applying a simple (and local) postprocessing operator $\mathbb{P}$ to the computed approximation u.

5.3.15 Definition: For a piecewise smooth velocity field u, we define the BDM postprocessing operator $\mathbb{P}$ cellwise by

$$\mathbb{P}u|_T = \mathbb{P}_T\big(u|_T, \widehat{u}^p\big), \qquad T \in \mathbb{T}_h,$$

where $\widehat{u}^p$ is the numerical flux (4.14) related to the incompressibility constraint.

The cell wise postprocessing operator uses the BDM projection of Definition 4.2.12, where we apply the cell moments (4.51) to the argument u and the face moments (4.50) to the argument $\widehat{u}^p$.

5.3.16 Lemma: *The operator $\mathbb{P}u$ is well-defined and maps the space V_h into V_h^{div}. Furthermore, if $u \in V_h$ satisfies* (4.48)*, then*

$$\nabla \cdot \mathbb{P}u = 0 \qquad \textit{a. e. in } \Omega. \tag{5.30}$$

Proof. The first statements are due to the properties of the BDM projection in Remark 4.2.13.

To prove (5.30), we first observe that $\nabla\cdot\mathbb{P}u \in Q_h$. This is due to the fact that $\nabla\cdot\mathbb{P}u|_T \in P_{k-1}(T)$ for all $T \in \mathbb{T}_h$ and

$$\int_\Omega \nabla \cdot \mathbb{P}u \, dx = \int_\Gamma \mathbb{P}u \cdot n \, ds = \int_\Gamma \widehat{u}^p \cdot n \, ds = 0,$$

in view of the definitions of $\mathbb{P}$ and $\widehat{u}^p$.

Now, let $u \in V_h$ satisfy (4.48). For $q \in Q_h$, we obtain

$$\begin{aligned}\big(\nabla \cdot \mathbb{P}u, q\big)_\Omega &= \sum_{T\in\mathbb{T}_h} \big(-\big(\mathbb{P}u, \nabla q\big)_T + \big\langle \mathbb{P}u \cdot \mathbf{n}_T, q\big\rangle_{\partial T}\big)\\ &= \sum_{T\in\mathbb{T}_h} \big(-\big(u, \nabla q\big)_T + \big\langle \widehat{u}^p \cdot \mathbf{n}_T, q\big\rangle_{\partial T}\big)\\ &= 0.\end{aligned}$$

Here, we have used integration by parts, the properties of $\mathbb{P}$ and (4.48). Thus, we have $\nabla\cdot\mathbb{P}u \equiv 0$ in Ω. □

5.3.17 Lemma: *Let $v \in V_h$. Then we have*

$$\big\|\mathbb{P}v\big\|_{1;h} \leqslant C_{\text{stab}}\big\|v\big\|_{1;h}, \tag{5.31}$$

with a stability constant $C_{\text{stab}} > 0$ that is independent of the mesh-size. The proof is rather technical and lengthy, so we refer the reader to [CKS05].

5.3.18 Definition: Using the postprocessing operator $\mathbb{P}$, we can define the weak form of our discretization of the Navier-Stokes equations as

$$\begin{aligned} a_h^{\text{IP}}\big(u_h, v\big) + O_h(\mathbb{P}u_h; u_h, v) + d_h(v, p_h) &= f_h^u(v) \ \forall v \in V_h, \\ -d_h(u_h, q) \qquad\qquad + e_h(p_h, q) &= f_h^p(q) \ \forall q \in Q_h, \end{aligned} \tag{5.32}$$

where the forms are the same as in (5.20) and $O_h(\mathbb{P}w; u, v)$ is simply the advection form (1.50) with b replaced by $\mathbb{P}w$.

The Picard iteration for this problem is again obtained by replacing the first argument to O_h by the solution of the previous step.

5.3.19 Lemma: *Let $u \in V + V_h$ and $w_1, w_2, v \in V_h$. Then there is a Lipschitz constant C_o independent of the mesh-size, such that*

$$|O_h(\mathbb{P}w_1; u, v) - O_h(\mathbb{P}w_2; u, v)| \leqslant C_o\big\|w_1 - w_2\big\|_{1;h}\big\|u\big\|_{1;h}\big\|v\big\|_{1;h}.$$

Proof. According to the definition of $\beta_h\big(.,.\big)$ in (1.50), we must estimate the form

$$\begin{aligned} O_h(\mathbb{P}w_1; u, v) - O_h(\mathbb{P}w_2; u, v) &= \big(\mathbb{P}w_1 - \mathbb{P}w_2\big)\cdot\nabla u, v\big)_{\mathbb{T}_h} \\ &+ \big\langle|\mathbb{P}w_1\cdot\mathbf{n}|u^{\downarrow_1} - u^{\uparrow_1}, v^{\downarrow_1}\big\rangle_{\mathbb{E}_h^i} - \big\langle|\mathbb{P}w_2\cdot\mathbf{n}|u^{\downarrow_2} - u^{\uparrow_2}, v^{\downarrow_1}\big\rangle_{\mathbb{E}_h^i} \\ &+ \big\langle|\mathbb{P}w_1\cdot\mathbf{n}|u, v\big\rangle_{\mathbb{E}_h^{-1}} - \big\langle|\mathbb{P}w_2\cdot\mathbf{n}|u, v\big\rangle_{\mathbb{E}_h^{-2}}, \end{aligned}$$

where the indices indicate upstream and downstream directions for w_1 and w_2, respectively. By the discrete Sobolev embedding (1.19) and Lemma 5.3.17, the first term can immediately be estimated through

$$\begin{aligned} \big((\mathbb{P}w_1 - \mathbb{P}w_2)\cdot\nabla u, v\big)_{\mathbb{T}_h} &\leqslant \big\|\mathbb{P}w_1 - \mathbb{P}w_2\big\|_{L^4}\big\|u\big\|_{1;h}\big\|v\big\|_{L^4} \\ &\leqslant c\big\|\mathbb{P}w_1 - \mathbb{P}w_2\big\|_{1;h}\big\|u\big\|_{1;h}\big\|v\big\|_{1;h}. \end{aligned}$$

For the second term, we observe that for any vector field b

$$\big\langle|b\cdot\mathbf{n}|u^{\downarrow} - u^{\uparrow}, v^{\downarrow}\big\rangle_{\mathbb{E}_h^i} = -\big\langle b\cdot\mathbf{n}^+(u^+ - u^-), v^+\big\rangle_{\mathbb{E}_h^i},$$

where the adjacent cells T_+ and T_- of an edge are chosen arbitrarily. Therefore,

$$\begin{aligned} \big|\big\langle|\mathbb{P}w_1\cdot\mathbf{n}|u^{\downarrow_1} - u^{\uparrow_1}, v^{\downarrow_1}\big\rangle_{\mathbb{E}_h^i} - \big\langle|\mathbb{P}w_2\cdot\mathbf{n}|u^{\downarrow_2} - u^{\uparrow_2}, v^{\downarrow_1}\big\rangle_{\mathbb{E}_h^i}\big| \\ = \big|\big\langle(\mathbb{P}w_1 - \mathbb{P}w_2)\cdot\mathbf{n}^+(u^+ - u^-), v^+\big\rangle_{\mathbb{E}_h^i}\big| \\ \leqslant \sum_{E\in\mathbb{E}_h^i}\big\|\mathbb{P}w_1 - \mathbb{P}w_2\big\|_{L^4(E)}\big\|[\![u\otimes\mathbf{n}]\!]\big\|_{L^2(E)}\big\|v\big\|_{L^4(E)}. \end{aligned}$$

Using inverse estimate and the fact that by definition $\|[\![u]\!]\|_{\mathbb{E}_h^i} \leqslant ch\|u\|_{1;h}$ yields

$$|\langle|\mathbb{P}w_1\cdot\mathbf{n}|u^{\downarrow_1}-u^{\uparrow_1},v^{\downarrow_1}\rangle_{\mathbb{E}_h^i}-\langle|\mathbb{P}w_2\cdot\mathbf{n}|u^{\downarrow_2}-u^{\uparrow_2},v^{\downarrow_1}\rangle_{\mathbb{E}_h^i}| \leqslant c\|\mathbb{P}w_1-\mathbb{P}w_2\|_{1;h}\|u\|_{1;h}\|v\|_{1;h}.$$

Finally, in order to estimate the boundary terms, we observe that for any vector field b, we obviously have

$$\langle|b\cdot\mathbf{n}||u|,|v|\rangle_{\mathbb{E}_h^-}\leqslant\langle|b\cdot\mathbf{n}||u|,|v|\rangle_{\mathbb{E}_h^\Gamma}.$$

Therefore, applying triangle inequality to $|\mathbb{P}w_1\cdot\mathbf{n}|-|\mathbb{P}w_2\cdot\mathbf{n}|$, we obtain

$$|\langle|\mathbb{P}w_1\cdot\mathbf{n}|u,v\rangle_{\mathbb{E}_h^{-_1}}-\langle|\mathbb{P}w_2\cdot\mathbf{n}|u,v\rangle_{\mathbb{E}_h^{-_2}}| \leqslant\sum_{E\in\mathbb{E}_h^\Gamma}\|\mathbb{P}w_1-\mathbb{P}w_2\|_{L^4(E)}\|u\|_{L^2(E)}\|v\|_{L^4(E)}$$

and we continue like on interior edges. Summing up the results proves the estimate. □

5.3.20 Theorem: *Assume that*

$$\mu:=\frac{C_oC_p\|f\|_0}{\nu^2\alpha^2}<1, \tag{5.33}$$

where α is the constant of the stability estimate for the interior penalty method (2.15)*, C_o the Lipschitz constant of Lemma 5.3.19 and C_p the constant from Friedrichs inequality* (1.18)*. Then the DG method* (5.32) *defines a unique solution $(u_h,p_h)\in V_h\times Q_h$. It satisfies the bounds*

$$\|u_h\|_{1;h}\leqslant\frac{C_p\|f\|_0}{\nu\,\alpha}, \tag{5.34}$$

$$\|p_h\|_0\leqslant\gamma^{-1}\left(2+\frac{C_a}{\alpha}\alpha\right)C_p\|f\|_0, \tag{5.35}$$

where C_a is the constant in the boundedness of $a_h^{IP}(.,.)$. Finally,

$$\sum_{E\in E_h^i}\int_E|\mathbb{P}u_h\cdot\mathbf{n}|\,|[\![u_h\otimes\mathbf{n}]\!]|^2\,ds+\int_\Gamma|\mathbb{P}u_h\cdot\mathbf{n}||u_h|^2\,ds\leqslant\frac{C_p^2\|f\|_0^2}{\nu\alpha}. \tag{5.36}$$

Proof. First, we eliminate the pressure from the problem by restricting ourselves to the weakly divergence-free subspace of V_h,

$$Z_h=\{v\in V_h:d_h(v,q)=0\quad\forall q\in Q_h\}. \tag{5.37}$$

The approximate velocity is thus characterized as the only function $u_h\in Z_h$ such that

$$a_h^{\mathrm{IP}}(u_h,v)+O_h(\mathbb{P}u_h;u_h,v)=(f,v)\qquad\forall v\in Z_h. \tag{5.38}$$

Let us introduce the solution operator S of the Oseen problem. For $\overline{u} \in Z_h$, $u = S(\overline{u})$ denotes the solution of the problem: find $u \in Z_h$ such that

$$a_h^{\mathrm{IP}}\big(u, v\big) + O_h(\mathbb{P}\overline{u}; u, v) = \big(f, v\big) \qquad \forall v \in Z_h.$$

Note that since $\overline{u} \in Z_h$ we have, by Lemma 5.3.16, that $\nabla \cdot \mathbb{P}\overline{u} = 0$. As a consequence, this problem is uniquely solvable. Furthermore, by the coercivity of the form $a_h^{\mathrm{IP}}\big(.,.\big)$ in (2.15) and O_h in Lemma 5.2.5,

$$\nu\alpha\big\|u\big\|_{1;h}^2 \leqslant a_h^{\mathrm{IP}}\big(u, u\big) + O_h(\mathbb{P}\overline{u}; u, u) = \big(f, u\big) \leqslant \big\|f\big\|_0\big\|u\big\|_0.$$

By Friedrichs inequality (1.18), we obtain

$$\nu\alpha\big\|u\big\|_{1;h}^2 \leqslant C_p\big\|f\big\|_0\big\|u\big\|_{1;h}.$$

Hence, the solution u to the above problem satisfies

$$\big\|u\big\|_{1;h} \leqslant \frac{C_p\big\|f\big\|_0}{\nu\alpha}, \tag{5.39}$$

and S maps

$$\mathcal{K}_h = \left\{ v \in Z_h \;\middle|\; \big\|v\big\|_{1;h} \leqslant \frac{C_p\big\|f\big\|_0}{\nu\alpha} \right\}$$

into itself.

Next, we show that S is a contraction on $\mathcal{K}_h$ under the smallness condition (5.33). To do so, let $\overline{u}_1, \overline{u}_2$ be in $\mathcal{K}_h$, and set $u_1 = S(\overline{u}_1)$, $u_2 = S(\overline{u}_2)$. Then

$$\nu\alpha\big\|u_1 - u_2\big\|_{1;h}^2 \leqslant a_h^{\mathrm{IP}}\big(u_1 - u_2, u_1 - u_2\big).$$

Since

$$a_h^{\mathrm{IP}}\big(u_1 - u_2, v\big) + O_h(\mathbb{P}\overline{u}_1; u_1, v) - O_h(\mathbb{P}\overline{u}_2; u_2, v) = 0,$$

for any $v \in Z_h$, taking $v = u_1 - u_2$ we get

$$\begin{aligned}\nu\alpha\big\|u_1 - u_2\big\|_{1;h}^2 \leqslant & \underbrace{-O_h(\mathbb{P}\overline{u}_2; u_1 - u_2, u_1 - u_2)}_{\leqslant 0} \\ & + O_h(\mathbb{P}\overline{u}_2; u_1, u_1 - u_2) - O_h(\mathbb{P}\overline{u}_1; u_1, u_1 - u_2).\end{aligned}$$

By the continuity property of O_h in Lemma 5.3.19, the bound (5.39) and the continuity of the post-processing operator $\mathbb{P}$ in Lemma 5.3.17,

$$\begin{aligned}\nu\alpha\big\|u_1 - u_2\big\|_{1;h}^2 \leqslant & C_o\big\|\overline{u}_1 - \overline{u}_2\big\|_{1;h}\big\|u_1\big\|_{1;h}\big\|u_1 - u_2\big\|_{1;h} \\ \leqslant & \frac{C_o C_p\big\|f\big\|_0}{\nu\alpha}\big\|\overline{u}_1 - \overline{u}_2\big\|_{1;h}\big\|u_1 - u_2\big\|_{1;h} \\ = & \nu\alpha\mu\,\big\|\overline{u}_1 - \overline{u}_2\big\|_{1;h}\big\|u_1 - u_2\big\|_{1;h}.\end{aligned}$$

This implies that

$$\big\|u_1 - u_2\big\|_{1;h} \leqslant\ \mu\,\big\|\overline{u}_1 - \overline{u}_2\big\|_{1;h}, \tag{5.40}$$

and so, if $\mu < 1$, that is, if the smallness condition (5.33) is satisfied, the mapping S is a contraction. Hence, S has a unique fixed point $u_h \in \mathcal{K}_h$, which is the solution to problem (5.38).

Now that the velocity u_h has been computed, the pressure is the solution $p_h \in Q_h$ of

$$d_h(v, p_h) = \big(f, v\big) - a_h^{\rm IP}\big(u_h, v\big) - O_h(\mathbb{P}u_h; u_h, v) \qquad \forall v \in V_h. \tag{5.41}$$

Due to boundedness of the involved forms and the Friedrichs inequality in Lemma1.2.3, the right-hand side defines a continuous linear functional on V_h and Z_h. The inf-sup condition in Lemma 4.2.2 then guarantees the existence of a unique solution p_h to the above problem. It can then easily be seen that the tuple (u_h, p_h) is the unique solution to the DG method in 5.20 with velocity $b = \mathbb{P}u_h$.

Next, we show the stability bounds for (u_h, p_h). The bound for $\big\|u_h\big\|_{1;h}$ in (5.34) follows since $u_h \in \mathcal{K}_h$. To obtain the bound for the upwind term in (5.36), note that

$$\nu\alpha\big\|u_h\big\|_{1;h}^2 + O_h(\mathbb{P}u_h; u_h, u_h) \leqslant C_p\big\|f\big\|_0\big\|u_h\big\|_{1;h}^2 \leqslant \frac{1}{2}\frac{C_p^2\big\|f\big\|_0^2}{\nu\alpha} + \frac{1}{2}\nu\alpha\big\|u_h\big\|_{1;h}^2.$$

Similarly to the previous arguments, here we have used the coercivity of $a_h^{\rm IP}\big(.,.\big)$, equation (5.38) with test function $v = u_h$, and the Friedrichs inequality. Bringing the term $\frac{1}{2}\nu\alpha\big\|u_h\big\|_{1;h}^2$ to the left-hand side and observing the coercivity of O_h give the stability bound in (5.36).

Moreover, using the inf-sup condition in Lemma 4.2.2, the Friedrichs inequality in Lemma 1.2.3, the continuity of the bilinear forms, Lemma 5.3.19, and the stability of $\mathbb{P}$ in Lemma 5.3.17, we have from (5.41)

$$\gamma\big\|p_h\big\|_0 \leqslant \sup_{0\neq v\in V_h} \frac{d_h(v, p_h)}{\big\|v\big\|_{1;h}} \leqslant C_p\big\|f\big\|_0 + \nu\, C_a\big\|u_h\big\|_{1;h} + C_o\big\|u_h\big\|_{1;h}^2.$$

Taking into account the stability bound for u_h and Assumption (5.33) gives

$$\begin{aligned}\big\|p_h\big\|_0 &\leqslant \gamma^{-1}\bigg(C_p\big\|f\big\|_0 + \frac{C_aC_p\big\|f\big\|_0}{\alpha} + \frac{C_oC_p\big\|f\big\|_0}{\nu^2\alpha^2}C_p\big\|f\big\|_0\bigg)\\ &\leqslant \gamma^{-1}C_p\big\|f\big\|_0\left(2 + \frac{C_a}{\alpha}\right).\end{aligned}$$

This gives the desired bound (5.35) for p_h. □

5.3.21 Corollary: *If $(u_h^{\ell+1}, p_h^{\ell+1})_{\ell=0,1,\ldots}$ is the sequence produced by the Picard iteration for the postprocessed scheme, then*

$$\begin{aligned}\big\|u_h - u_h^{\ell+1}\big\|_{1;h} &\leqslant 2\left(\frac{C_p\big\|f\big\|_0}{\nu\alpha}\right)\frac{\mu^\ell}{(1-\mu)},\\ \big\|p_h - p_h^{\ell+1}\big\|_0 &\leqslant 2\,\gamma^{-1}\left(\frac{C_a + 2\alpha}{\alpha}\right)C_p\big\|f\big\|_0\frac{\mu^\ell}{(1-\mu)},\end{aligned}$$

for any initial guess $(u_h^0, p_h^0) \in V_h \times Q_h$.

Proof. Since $u_h^{\ell+1} = S(u_h^\ell)$ and S is a contraction with Lipschitz constant μ, we immediately get y the Banach fixed point theorem

$$\left\|u_h - u_h^{\ell+1}\right\|_{1;h} \leqslant \frac{\mu^\ell}{(1-\mu)}\left\|u_h^2 - u_h^1\right\|_{1;h}.$$

The result now follows from the fact that, by the stability bound (5.39),

$$\left\|u_h^m\right\|_{1;h} \leqslant \frac{C_p\left\|f\right\|_0}{\nu\alpha},$$

for $m \geqslant 1$.

To obtain the estimate for the pressure, we proceed as follows. First, we note that, from the momentum equation, we have

$$d_h(v, p_h^{\ell+1}) = \big(f, v\big) - a_h^{\text{IP}}\big(u_h^{\ell+1}, v\big) - O_h(\mathbb{P}u_h^\ell; u_h^{\ell+1}, v) \qquad \forall v \in V_h.$$

This implies that

$$\begin{aligned} d_h(v, p_h - p_h^{\ell+1}) &= -a_h^{\text{IP}}\big(u_h - u_h^{\ell+1}, v\big) - O_h(\mathbb{P}u_h; u_h, v) + O_h(\mathbb{P}u_h^\ell; u_h, v) \\ &\quad -O_h(\mathbb{P}u_h^\ell; u_h, v) + O_h(\mathbb{P}u_h^\ell; u_h^{\ell+1}, v) \\ &= -a_h^{\text{IP}}\big(u_h - u_h^{\ell+1}, v\big) - O_h(\mathbb{P}u_h; u_h, v) + O_h(\mathbb{P}u_h^\ell; u_h, v) \\ &\quad O_h(\mathbb{P}u_h^\ell; u_h^{\ell+1} - u_h, v). \end{aligned}$$

We insert this expression in the inf-sup condition for d_h, use the stability properties of $a_h^{\text{IP}}\big(.,.\big)$, $O_h(.;.,.)$, and $\mathbb{P}$, take into account the bounds for $\left\|u_h\right\|_{1;h}$, $\left\|u_h^\ell\right\|_{1;h}$, and the contraction property of S to obtain

$$\begin{aligned} \left\|p_h - p_h^{\ell+1}\right\|_0 &\leqslant \gamma^{-1}\left(\nu C_a\mu + \frac{C_oC_p\left\|f\right\|_0}{\nu\alpha}(1+\mu)\right)\left\|u_h - u_h^\ell\right\|_{1;h} \\ &\leqslant \gamma^{-1}\mu\left(\nu C_a + 2\nu\alpha\right)\left\|u_h - u_h^\ell\right\|_{1;h} \end{aligned}$$

The desired bound for $\left\|p_h - p_h^{\ell+1}\right\|_0$ then follows from the bound for $\left\|u_h - u_h^\ell\right\|_{1;h}$. □

5.3.22 Remark: If we set

$$\hat{\alpha} = \min\{1, \alpha\}, \qquad \hat{C}_p = \max\{C_\Omega, C_p\}, \qquad \hat{C}_o = \max\{C_\Omega, C_o\}, \tag{5.42}$$

then both the smallness assumptions in (5.9) and (5.33) are satisfied if we have that

$$\frac{\hat{C}_o\hat{C}_p\left\|f\right\|_0}{\nu^2\,\hat{\alpha}^2} < 1.$$

Hence, both the Navier-Stokes equations and their DG approximation are uniquely solvable. Under a smallness condition that is slightly more restrictive, we obtain the following estimates.

5.3.23 Theorem: *Assume that*

$$\frac{\hat{C}_o\hat{C}_p\big\|f\big\|_0}{\nu^2\,\hat{\alpha}^2} \leqslant \frac{1}{2}, \tag{5.43}$$

and that the exact solution (u,p) *of the Navier-Stokes equations* (5.1) *satisfies*

$$u \in H^{s+1}(\Omega)^2, \qquad p \in H^s(\Omega), \qquad s \geq 1. \tag{5.44}$$

Then

$$\begin{aligned}
\big\|u-u_h\big\|_{1;h} &\leqslant Ch^{\min\{k,s\}}\big(\big\|u\big\|_{s+1} + \nu^{-1}\,\big\|p\big\|_s\big),\\
\big\|u-\mathbb{P}u_h\big\|_{1;h} &\leqslant Ch^{\min\{k,s\}}\big(\big\|u\big\|_{s+1} + \nu^{-1}\,\big\|p\big\|_s\big),\\
\big\|p-p_h\big\|_0 &\leqslant Ch^{\min\{k,s\}}\big(\nu\big\|u\big\|_{s+1} + \big\|p\big\|_s\big),
\end{aligned}$$

where the constants are independent of h.

Proof. We modify the approach used in the previous section to get error estimates for the DG method for the Oseen problem in two ways. First, we use the *non-conforming* approach introduced in [PS02] and later used in [SST03], and consider the expression

$$\mathcal{R}_h(u,p) := \sup_{0\neq v\in V_h} \frac{|R_h(u,p;v)|}{\big\|v\big\|_{1;h}},$$

where

$$R_h(u,p;v) := \nu a_h^{\mathrm{IP}}\big(u,v\big) + O_h(u;u,v) + d_h(v,p) - \big(f,v\big), \qquad v\in V_h. \tag{5.45}$$

The second modification is, of course, due to the presence of the convective non-linearity.

Let us begin with the estimate for the error $\big\|u-u\big\|_{1;h}$, for which we are going to prove an estimate in terms of best approximation, namely

$$\begin{aligned}
\big\|u-u_h\big\|_{1;h} \leqslant C\Bigg(&\inf_{v\in V_h}\big\|u-v\big\|_{1;h} + \inf_{v\in V_h^{\mathrm{div}}}\big\|u-v\big\|_{1;h}\\
&+ \inf_{q\in Q_h}\frac{1}{\nu}\big\|p-q\big\|_0 + \frac{1}{\nu}\mathcal{R}_h(u,p)\Bigg).
\end{aligned} \tag{5.46}$$

To prove this estimate, we proceed as in the error analysis of standard mixed methods, see e.g. [BF91], and consider first an element $v\in Z_h$. Using a standard triangle inequality

$$\big\|u-u_h\big\|_{1;h} \leqslant \big\|u-v\big\|_{1;h} + \big\|v-u_h\big\|_{1;h},$$

we focus on the second term. Using the definition of R_h in (5.45), we obtain

$$\begin{aligned}
\nu\alpha\big\|v-u_h\big\|_{1;h} \leqslant& a_h^{\mathrm{IP}}\big(v-u_h, v-u_h\big)\\
=& a_h^{\mathrm{IP}}\big(v-u, v-u_h\big) + R_h(u,p;v-u_h)\\
&- d_h(v-u_h, p-p_h)\\
&\underbrace{+O_h(\mathbb{P}u_h;u_h,v-u_h) - O_h(u;u,v-u_h)}_{=:\mathcal{O}}
\end{aligned}$$

While the first two terms in this sum can be estimated immediately by ellipticity and the definition of $\mathcal{R}_h(u,p)$, respectively, the nonlinearity and incompressibility need some more thought. Since $v - u_h \in Z_h$, we have

$$d_h(v-u_h, p-p_h) = d_h(v-u_h, p) = d_h(v-u_h, p-q),$$

for any $q \in Q_h$. From the boundedness of the forms follows

$$|d_h(v-u_h, p-p_h)| \leqslant C\big\|v-u_h\big\|_{1;h}\big\|p-q\big\|_0 \qquad \forall\, q \in Q_h.$$

Finally,

$$\begin{aligned}\mathcal{O} =& O_h(\mathbb{P}u_h; u, v-u_h) - O_h(\mathbb{P}v; u, v-u_h)\\ &- O_h(\mathbb{P}u_h; v-u_h, v-u_h)\\ &+ O_h(\mathbb{P}v; u, v-u_h) - O_h(u; u, v-u_h)\\ &+ O_h(\mathbb{P}u_h; v-u, v-u_h)\\ =:& T_1 + T_2 + T_3 + T_4.\end{aligned}$$

Note that due to Lemma 5.3.19, Lemma 5.3.17, the stability bound for u in (5.6), and the definitions of the parameters in (5.42), we have

$$\begin{aligned}T_1 &\leqslant C_o\big\|u\big\|_1\big\|v-u_h\big\|_{1;h}^2\\ &\leqslant \frac{C_o C_O mega\,\big\|f\big\|_0}{\nu}\big\|v-u_h\big\|_{1;h}^2\\ &\leqslant \frac{\hat{C}_o\hat{C}_p\big\|f\big\|_0}{\nu\hat{\alpha}}\big\|v-u_h\big\|_{1;h}^2\\ &\leqslant \frac{1}{2}\nu\hat{\alpha}\,\big\|v-u_h\big\|_{1;h}^2,\end{aligned}$$

by the smallness condition (5.43).

Next, by Lemma 5.2.5, $T_2 \leqslant 0$ and by Lemma 5.3.19, Lemma 5.3.17, and the bound for u_h in Theorem 5.3.20,

$$\begin{aligned}T_4 &\leqslant \frac{C_o C_p\big\|f\big\|_0}{\nu\alpha}\big\|u-v\big\|_{1;h}\big\|v-u_h\big\|_{1;h}\\ &\leqslant \frac{\hat{C}_o\hat{C}_p\big\|f\big\|_0}{\nu\hat{\alpha}}\big\|u-v\big\|_{1;h}\big\|v-u_h\big\|_{1;h}\\ &\leqslant \frac{1}{2}\nu\hat{\alpha}\,\big\|u-v\big\|_{1;h}\big\|v-u_h\big\|_{1;h},\end{aligned}$$

by the smallness condition (5.43).

Finally, by the Lipschitz property of the form O_h in Lemma 5.3.19, we have

$$\begin{aligned}T_3 &\leqslant C_o\big\|u-\mathbb{P}v\big\|_{1;h}\big\|u\big\|_1\big\|v-u_h\big\|_{1;h}\\ &\leqslant \frac{C_o C_\Omega\big\|f\big\|_0}{\nu}\big\|u-\mathbb{P}v\big\|_{1;h}\big\|v-u_h\big\|_{1;h}\\ &\leqslant C\big\|u-\mathbb{P}v\big\|_{1;h}\big\|v-u_h\big\|_{1;h},\end{aligned}$$

by the bound for u in (5.6) and the smallness condition (5.43). Now, take an arbitrary function $\widetilde{v}$ in V_h^{div}. Since $\mathbb{P}$ reproduces functions in V_h^{div}, we have $\mathbb{P}\widetilde{v} = \widetilde{v}$, and so

$$\begin{aligned}\big\|u - \mathbb{P}v\big\|_{1;h} &\leqslant \big\|u - \widetilde{v}\big\|_{1;h} + \big\|\mathbb{P}\widetilde{v} - \mathbb{P}v\big\|_{1;h} \\ &\leqslant \big\|u - \widetilde{v}\big\|_{1;h} + C_{\text{stab}}\big\|\widetilde{v} - v\big\|_{1;h} \\ &\leqslant (1 + C_{\text{stab}})\big\|u - \widetilde{v}\big\|_{1;h} + C_{\text{stab}}\big\|u - v\big\|_{1;h},\end{aligned}$$

by Lemma 5.3.17. This implies that

$$T_3 \leqslant \frac{\nu\hat{\alpha}}{2}\left(\left(\frac{1 + C_{\text{stab}}}{C_{\text{stab}}}\right)\big\|u - \widetilde{v}\big\|_{1;h} + \big\|u - v\big\|_{1;h}\right)\big\|v - u_h\big\|_{1;h}.$$

Thus, gathering all the estimates above, we obtain

$$\big\|v - u_h\big\|_{1;h} \leqslant C\big(\big\|u - v\big\|_{1;h} + \big\|u - \widetilde{v}\big\|_{1;h} + \big\|p - q\big\|_0 + \mathcal{R}_h(u,p)\big),$$

and by triangle inequality

$$\big\|u - u_h\big\|_{1;h} \leqslant C\big(\big\|u - v\big\|_{1;h} + \big\|u - \widetilde{v}\big\|_{1;h} + \big\|p - q\big\|_0 + \mathcal{R}_h(u,p)\big), \tag{5.47}$$

for any $v \in Z_h$, $\widetilde{v} \in V_h^{\text{div}}$, and $q \in Q_h$.

It remains to replace $v \in Z_h$ by an arbitrary function in $v \in V_h$. To this end, fix $v \in V_h$ and consider the problem: Find $r \in V_h$ such that

$$d_h(r,q) = d_h(u - v, q) \qquad \forall q \in Q_h.$$

The inf-sup condition in Lemma 4.2.2 guarantees that such a solution r exists. Furthermore, it can be easily seen that we have

$$\big\|r\big\|_{1;h} \leqslant \gamma^{-1}C\big\|u - v\big\|_{1;h},$$

in view of the inf-sup condition for d_h and the continuity of the form d_h. By construction and since $d_h(u,q) = 0$ for any $q \in Q_h$, we further have that $r + v \in Z_h$. Inserting $r + v$ in (5.47), employing the triangle inequality, and taking into account the above bound for r yield the abstract error estimate (5.46) for the velocity.

As a consequence of the approximation result (5.46), we obtain the following estimate of the error between u and its globally solenoidal approximation $\mathbb{P}u_h$

$$\big\|u - \mathbb{P}u_h\big\|_{1;h} \leqslant (1 + C_{\text{stab}}) \inf_{\widetilde{v} \in V_h^{\text{div}}} \big\|u - \widetilde{v}\big\|_{1;h} + C_{\text{stab}}\big\|u - u_h\big\|_{1;h}. \tag{5.48}$$

To see this, note that

$$\begin{aligned}\big\|u - \mathbb{P}u_h\big\|_{1;h} &\leqslant \big\|u - \widetilde{v}\big\|_{1;h} + \big\|\widetilde{v} - \mathbb{P}u_h\big\|_{1;h} \\ &\leqslant \big\|u - \widetilde{v}\big\|_{1;h} + C_{\text{stab}}\big\|\widetilde{v} - u_h\big\|_{1;h} \\ &\leqslant (1 + C_{\text{stab}})\big\|u - \widetilde{v}\big\|_{1;h} + C_{\text{stab}}\big\|u - u_h\big\|_{1;h},\end{aligned}$$

where $\widetilde{v}$ is any element of V_h^{div}. Here, we have used the stability bound in Lemma 5.3.17 and the fact that $\mathbb{P}$ reproduces polynomials in V_h^{div}. This shows the inequality (5.48).

Now, let us obtain the estimate for the pressure. We claim that the error in the pressure satisfies

$$\begin{aligned}\big\|p-p_h\big\|_0 \leqslant C\Bigg[&\inf_{v\in V_h} \nu \,\big\|u-v\big\|_{1;h} + \inf_{\widetilde{v}\in V_h^{\mathrm{div}}} \nu \,\big\|u-\widetilde{v}\big\|_{1;h} \\ &+ \inf_{q\in Q_h} \big\|p-q\big\|_0 + \mathcal{R}_h(u,p)\Bigg].\end{aligned} \tag{5.49}$$

To see this, we proceed in a way similar to the one used to deal with the velocity. Thus, we begin by noting that for $q \in Q_h$

$$\begin{aligned}\big\|p-p_h\big\|_0 &\leqslant \big\|q-p_h\big\|_0 + \big\|p-q\big\|_0 \\ &\leqslant \gamma^{-1} \sup_{v\in V_h} \frac{d_h(v,q-p_h)}{\big\|v\big\|_{1;h}} + \big\|p-q\big\|_0 \\ &\leqslant \gamma^{-1} \sup_{v\in V_h} \frac{d_h(v,p-p_h)}{\big\|v\big\|_{1;h}} + \gamma^{-1} \sup_{v\in V_h} \frac{d_h(v,q-p)}{\big\|v\big\|_{1;h}} + \big\|p-q\big\|_0,\end{aligned}$$

where we have use the inf-sup condition in Lemma 4.2.2. Therefore,

$$\big\|p-p_h\big\|_0 \leqslant \gamma^{-1} \sup_{v\in V_h} \frac{d_h(v,p-p_h)}{\big\|v\big\|_{1;h}} + (1+\gamma^{-1}C_b)\,\big\|p-q\big\|_0, \tag{5.50}$$

by the continuity of the form d_h.

To bound the first term on the right-hand side of (5.50), we note that

$$d_h(v,p-p_h) = -a_h^{\mathrm{IP}}\big(u-u_h,v\big) - O_h(u;u,v) + O_h(\mathbb{P}u_h;u_h,v) + R_h(u,p;v),$$

for any $v \in V_h$, and proceed as in the previous step to obtain

$$\begin{aligned}d_h(v,p-p_h) &= a_h^{\mathrm{IP}}\big(u-u_h,v\big) + O_h(\mathbb{P}u_h;u_h-u,v) \\ &\quad + O_h(\mathbb{P}u_h;u,v) - O_h(u;u,v) + R_h(u,p;v) \\ &\leqslant \Bigg[(\nu\, C_a + \frac{1}{2}\nu\,\hat{\alpha})\,\big\|u-u_h\big\|_{1;h} \\ &\quad + \frac{\nu\hat{\alpha}}{2\,C_{\mathrm{stab}}}\,\big\|u-\mathbb{P}u_h\big\|_{1;h} + \mathcal{R}_h(u,p)\Bigg]\,\big\|v\big\|_{1;h},\end{aligned}$$

and since

$$\big\|u-\mathbb{P}u_h\big\|_{1;h} \leqslant (1+C_{\mathrm{stab}})\big\|u-\widetilde{v}\big\|_{1;h} + C_{\mathrm{stab}}\big\|u-u_h\big\|_{1;h},$$

from (5.40), we get

$$\begin{aligned}d_h(v,p-p_h) \leqslant &(C_a+\hat{\alpha})\,\nu\,\big\|u-u_h\big\|_{1;h} \\ &+ \frac{\hat{\alpha}}{2}\,\left(\frac{1+C_{\mathrm{stab}}}{C_{\mathrm{stab}}}\right)\,\nu\,\big\|u-\widetilde{v}\big\|_{1;h} + \mathcal{R}_h(u,p).\end{aligned}$$

Inserting this inequality in (5.50) and using the bound (5.46) for $\big\|u-u_h\big\|_{1;h}$ from Step 1, we immediately obtain the abstract estimate (5.49) for the pressure.

Under the regularity Assumption (5.44), the following standard approximation property holds

$$\inf_{v\in V_h} \nu \left\|u-v\right\|_{1;h} + \inf_{q\in Q_h} \left\|p-q\right\|_0 \leqslant C_{\rm app} h^{\min\{k,s\}} \left[\nu \left\|u\right\|_{s+1} + \left\|p\right\|_s\right].$$

Moreover, from the results in [BF91], see also [HL02, Section 3], we have

$$\inf_{\widetilde{v}\in V_h^{\rm div}} \left\|u-\widetilde{v}\right\|_{1;h} \leqslant C_{\rm app} h^{\min\{k,s\}} \left\|u\right\|_{s+1}.$$

Finally, we have that

$$\mathcal{R}_h(u,p) \leqslant C_{\rm app} h^{\min\{k,s\}} \left[\nu \left\|u\right\|_{s+1} + \left\|p\right\|_s\right],$$

with a constant $C_{\rm app}$ independent of the mesh-size.

To see the estimate of the residual, we proceed as follows. For $v \in V_h$, it is easy to see that $R_h(u,p;v)$ is given by

$$R_h(u,p;v) = \sum_{E\in\mathbb{E}_h} \int_E \{\!\{\nu\nabla u - P_h(\nu\nabla u)\}\!\} : [\![v\otimes n]\!]\, ds - \sum_{E\in\mathbb{E}_h} \int_E \{\!\{p - P_h p\}\!\} [\![v\cdot n]\!]\, ds,$$

with $P_h : L^2(\Omega)^{2\times 2} \to \Sigma_h$ and $P_h : L^2(\Omega)/\mathbb{R} \to Q_h$ denoting the L^2-projections onto Σ_h and Q_h, respectively. The desired estimate follows then by proceeding as the proof of [SST03, Proposition 8.1] and using standard approximation results for P_h and P_h.

It is now a simple matter to see that the error estimates of Theorem 5.3.23 follow by inserting the approximation estimates obtained in the previous step into the abstract bounds for the velocity, (5.46), for its globally solenoidal post-processing (5.48) and the pressure (5.49). This completes the proof of Theorem 5.3.23. □

5.3.24 In Table 5.2 we show the errors and convergence rates in p, u and σ obtained for $\nu = 0.1$. The errors in p and σ are measured in the L^2-norm while $u - u_h$ and $u - \mathbb{P}u_h$ are evaluated in the broken H^1-norm. We observe the predicted first order convergence for all the error components, in full agreement with the results of Theorem 5.3.23. Notice that we have scaled the L^2-error in σ by ν^{-1} so that this error can be directly compared to the H^1-errors in $u - u_h$ and $u - \mathbb{P}u_h$. These three errors are all of the same magnitude, with a slight advantage for the post-processed solution.

In Table 5.3 we show the seminorm of the errors $u-u_h$ and $u-\mathbb{P}u_h$ which measures their jumps. Notice that it superconverges with order $3/2$. This means that the relative contribution of this seminorm to the $\left\|\cdot\right\|_{1;h}$−norm diminishes as h decreases. An analysis of this phenomenon remains to be carried out.

In Table 5.4, we show the L^2-errors in the velocities and their corresponding convergence orders. In the first column, we observe that the velocities converge with second order. In the second column, we notice that by post-processing the error is reduced by a factor of roughly 3/2. Therefore, the post-processed solution should be used as the best approximation obtained by our scheme. Furthermore, we show the L^∞-norms of the divergence of $\mathbb{P}u_h$ (evaluated at the points of a 4-by-4 Gauss formula on each cell). These are of the order of the residual of the non-linear iteration, confirming that the post-processed solution is indeed divergence-free.

L	$\|p-p_h\|_0$		$\|\|u-u_h\|\|_{1;h}$		$\|\|u-\mathbb{P}u_h\|\|_{1;h}$		$\nu^{-1}\|\sigma-\sigma_h\|_0$	
3	2.2e+0	—	1.2e+1	—	8.1e+0	—	7.0e-0	—
4	1.0e+0	1.12	5.4e+0	1.11	3.2e+0	1.33	3.4e-0	1.05
5	4.8e-1	1.10	2.4e+0	1.16	1.4e+0	1.18	1.6e-0	1.07
6	2.3e-1	1.04	1.1e+0	1.18	6.8e-1	1.06	7.8e-1	1.04
7	1.2e-1	1.01	4.7e-1	1.17	3.4e-1	1.02	3.9e-1	1.02
8	5.8e-2	1.00	2.2e-1	1.13	1.7e-1	1.01	1.9e-1	1.02

Table 5.2: Errors and orders of convergence for $\nu = 0.1$.

L	$\|u-u_h\|_h$		$\|u-\mathbb{P}u_h\|_h$	
3	9.1e+0	—	4.8e+0	—
4	4.2e+0	1.11	1.5e+0	1.72
5	1.8e+0	1.23	4.7e-1	1.64
6	7.2e-1	1.32	1.6e-1	1.54
7	2.8e-1	1.39	5.6e-2	1.52
8	1.0e-1	1.44	2.0e-2	1.51

Table 5.3: Errors and orders of convergence for $\nu = 0.1$ in the jump seminorm $|v|_h := \left\{\sum_{E\in\mathbb{E}_h}\int_E \kappa_0 h^{-1}\,|[\![v\otimes n]\!]|^2\,ds\right\}^{1/2}$.

L	$\|u-u_h\|_0$		$\|u-\mathbb{P}u_h\|_0$		$\|\nabla\cdot\mathbb{P}u_h\|_\infty$
3	6.4e-1	—	4.9e-1	—	1.4e-12
4	1.6e-1	2.03	1.1e-1	2.22	1.4e-12
5	3.3e-2	2.22	2.0e-2	2.37	3.2e-12
6	7.1e-3	2.24	4.2e-3	2.27	1.5e-11
7	1.6e-3	2.19	9.8e-4	2.12	1.8e-12
8	3.5e-4	2.13	2.4e-4	2.04	2.9e-11

Table 5.4: L^2-errors and orders of convergence in the velocity and L^∞-norm of the divergence of the post-processed solution $\mathbb{P}u_h$ for $\nu = 0.1$.

L	$\nu = 1$	$\nu = 0.1$	$\nu = 0.01$
3	14	33	865
4	10	21	106
5	8	16	54
6	6	12	29
7	5	10	18
8	6	10	13

Table 5.5: Number of iterations for convergence of the non-linear iteration.

The convergence of the non-linear iteration under consideration is illustrated in Table 5.5. Displayed is the number of steps required to reduce the start residual by a factor of 10^7. The initial guess for the iterations is the vector $u_h^0 = 0$. The linear system in each step is solved by a preconditioned GMRES method up to a relative accuracy of 10^{-4}. Therefore, the error of the linear iterations is small enough to be neglected. Table 5.5 shows that the number of iteration steps is not only bounded independently of the mesh-size, but in fact decreasing. This is in perfect agreement with our theoretical results in Theorem 5.3.20. If we decrease the viscosity, the increase of the number of iteration steps for convergence is quite moderate on fine grids. Of course, this only holds as long as there is convergence. With $\nu = 10^{-3}$, the non-linear iteration does not converge anymore, probably because the stationary solution is not stable in this case.

5.3.25 Remark: For a DG method using the div-conforming space V_h^{div} in (4.45), it can be readily seen that a field $u \in V_h^{\text{div}}$ satisfying (4.49) already *is* exactly incompressible. Hence, for such a DG method, we can take $\mathbb{P}$ as the identity; we will study this approach in the following pargraphs.

We first notice that, since the space of velocities V_h is included in $H^{\text{div}}(\Omega)$, the jump of the normal component of v vanishes over edges. Therefore, unlike for the case treated before, there is no need to introduce a numerical flux associated with p_h.

5.3.26 Theorem: *Under the above assumptions, theorems 5.3.20 and 5.3.23 hold for the divergence conforming method. In particular, we have the optimal error bound*

$$\left\|u - u_h\right\|_{1;h} + \left\|p - p_h\right\|_0 \leqslant Ch^k \left(\left\|u\right\|_{k+1} + \left\|p\right\|_k\right),$$

with a constant C independent of the mesh size. Moreover, the approximate velocity u_h is exactly divergence-free.

5.3.27 To carry out our numerical experiments, we use the symmetric interior penalty method. We consider quadratic mesh cells and the local spaces are, since this method does not use a separate space for σ,

$$V(T) \times Q(T) = RT_k(T) \times \mathbb{Q}_k(T),$$

5.3.28 Again, we use example A.4.3. Since this function has boundary values different from zero, special precautions must be taken. As pointed out in Section 2, in order to obtain a divergence-free solution in the elements adjacent to the boundary, the Dirichlet boundary condition for the normal component must be implemented in a *strong* way, while the tangential components obtain their boundary values weakly through the form (5.20). The strong boundary values are obtained by interpolating the normal component of the prescribed boundary function in the set of Gauss points required to integrate RT_k exactly on a face and then initializing the start value of the nonlinear iteration to this value. In all subsequent iteration steps, the residual and update vectors are set to zero in the corresponding components. The tangential component of the velocity field is prescribed in weak form, as with standard DG schemes.

5.3.29 Absolute errors and convergence rates for several pairs of $\text{RT}_k/\mathbb{Q}_k$ polynomials are listed in Table 5.6. It exhibits clearly the expected convergence orders for error of the velocity in $H^1(\mathbb{T}_h)$ and $L^2(\Omega)$. The pressure errors in this table converge much faster than expected; a fact we cannot explain right now.

k	L	$\|\nabla e_u\|$	ord.	$\|e_u\|$	ord.	$\|p\|$	ord.	$\|\nabla \cdot u\|_\infty$
	4	9.7e+0	1.01	2.3e-1	1.82	3.9e+0	1.63	8.7e-10
	5	4.9e+0	1.00	6.2e-2	1.89	1.2e+0	1.75	2.8e-09
1	6	2.4e+0	1.00	1.6e-2	1.93	3.3e-1	1.81	6.7e-09
	7	1.2e+0	1.00	4.2e-3	1.96	9.3e-2	1.84	3.3e-08
	8	6.1e-1	1.00	1.1e-3	1.98	2.6e-2	1.86	2.8e-09
	3	3.9e+0	2.22	9.3e-2	3.12	2.6e+0	2.01	4.8e-10
	4	9.9e-1	1.97	1.2e-2	2.96	4.5e-1	2.53	1.7e-09
2	5	2.5e-1	2.00	1.5e-3	3.00	6.5e-2	2.79	4.1e-09
	6	6.2e-2	2.00	1.9e-4	3.00	9.2e-3	2.82	2.0e-08
	7	1.5e-2	2.00	2.3e-5	3.00	1.4e-3	2.77	7.3e-08
	4	6.5e-2	2.99	6.8e-4	3.78	4.2e-2	3.46	1.9e-09
3	5	8.0e-3	3.03	4.6e-5	3.87	3.2e-3	3.71	7.9e-09
	6	9.8e-4	3.02	3.0e-6	3.93	2.5e-4	3.72	2.0e-08
	3	5.3e-2	4.27	6.8e-4	5.31	6.3e-2	3.91	7.6e-10
4	4	3.4e-3	3.97	2.1e-5	5.00	2.7e-3	4.52	3.6e-09
	5	2.1e-4	4.00	6.6e-7	5.01	9.5e-5	4.85	6.7e-09

Table 5.6: Errors for Kovasznay flow ($\nu = 1$) and pairs $\mathrm{RT}_k/\mathbb{Q}_k$.

Chapter 6

Linear Solvers

The solution of the linear systems arising from discontinuous Galerkin discretizations has been neglected for a long time. Time dependent hyperbolic problems are solved mostly by explicit time-stepping schemes, while a Gauß-Seidel method with suitable ordering of the degrees of freedom yields sufficient results in the stationary case. In the presence of a second order elliptic term, these methods fail to be efficient. Implicit time-stepping is required to avoid unreasonably small space grids due to the Courant-Friedrichs-Lewy condition. Furthermore, the condition number of the linear system in the stationary case or for large time steps increases with mesh refinement, resulting in very slow iterative solvers on fine grids.

Therefore, preconditioners are required to counter the effect of the increasing condition number. Before our work in [GK03c], very few results in this direction were known. In [RVW96] Rusten et al. had used an inconsistent multilevel interior penalty method to precondition the conforming mixed discretization of Poisson equation. Furthermore, Ewing et al. had investigated domain decomposition [ELPV93] and multilevel [BEPS96] methods for the cell-centered finite difference method, which corresponds to the interior penalty method in Section 2.2 with constant shape functions. Recently, a domain decomposition preconditioner for the interior penalty method was presented by Lasser and Toselli in [LT01].

After summarizing the most important results on Krylov-space solvers, we present the multilevel method of Gopalakrishnan and Kanschat in [GK03c, GK03b], which applies to the interior penalty method as well as to the Schur complement of the matrix produced by the LDG method, as shown in [GK03a]. We demonstrate how data structures can be set up for a discontinuous multilevel method, so that locally refined grids can be handled in an efficient way without additional coding of the particular application.

A general framework for preconditioning of saddle point problems was presented by Murphy et al. in [MGW00]. It successfully applied to various discretizations of Stokes and Oseen equations in [ESW02, KLW02, SEKW01]. Here, we apply this concept successfully to the LDG saddle point problem, where we obtain an analytical eigenvalue estimate, and to the Stokes system. Computational results show its feasibility in a DG context.

The remaining sections of this chapter are devoted to the application of the preconditioning concepts to advection dominated problems. In the section on advection, we show results from [GK03c], indicating that the multilevel scheme combined with a Gauß-Seidel smoother with downwind ordering improves when advection becomes dominant and can therefore be considered robust with respect to the Peclet number. Furthermore, we show new results exhibiting this robustness even for flow with vortices. Finally, we show that a

preconditioner following the ideas by Kay et al. in [KLW02] yields good results for Oseen equations.

6.1 Krylov space methods

6.1.1 Krylov space methods —in particular preconditioned versions— are the most efficient iterative schemes known for solving sparse linear systems of equations

$$\mathsf{A}\mathsf{u} = \mathsf{f}.$$

They exist in several variants for linear systems that are symmetric, nonsymmetric, positive definite or not.

Still, their performance usually depends on a good preconditioner P for the matrix A. It is those preconditioners the following sections deal with after we presented the most important facts on the iterative schemes.

6.1.2 Algorithm (Preconditioned conjugate gradient): With a given start vector u_0, compute the initial residual $\mathsf{r}_0 := \mathsf{f} - \mathsf{A}\mathsf{u}_0$ and the preconditioned residual $\mathsf{z}_0 := \mathsf{P}\mathsf{r}_0$. Let $\mathsf{p}_0 := \mathsf{z}_0$.

While $\|\mathsf{r}_k\| > \varepsilon$ do for $k = 0, 1, \ldots$

$$\begin{aligned}
\alpha_k &:= \frac{\mathsf{r}_k^T \mathsf{z}_k}{\mathsf{p}_k^T \mathsf{A}\mathsf{p}_k} \\
\mathsf{u}_{k+1} &:= \mathsf{u}_k + \alpha_k \mathsf{p}_k \\
\mathsf{r}_{k+1} &:= \mathsf{r}_k - \alpha_k \mathsf{A}\mathsf{p}_k \\
\mathsf{z}_{k+1} &:= \mathsf{P}\mathsf{r}_{k+1} \\
\beta_k &:= \frac{\mathsf{r}_{k+1}^T \mathsf{z}_{k+1}}{\mathsf{r}_k^T \mathsf{z}_k} \\
\mathsf{p}_{k+1} &:= \mathsf{p}_k + \beta_k \mathsf{p}_k
\end{aligned}$$

6.1.3 Algorithm (GMRES with right preconditioning): Compute the initial residual $\mathsf{r}_0 := \mathsf{f} - \mathsf{A}\mathsf{u}_0$, $\beta := \|\mathsf{r}_0\|$ and $\mathsf{v}_1 := \mathsf{r}_0/\beta$.

While $\|\mathsf{y}_k\| > \varepsilon$ do for $k = 1, 2, \ldots$

1. Compute $\mathsf{w} := \mathsf{A}\mathsf{P}\mathsf{v}_k$.
2. Compute column k of the matrix $\mathsf{H} \in \mathbb{R}^{k+1,k}$ and orthogonalize w by

$$\begin{aligned}
h_{ik} &:= \mathsf{w}^T \mathsf{v}_i \qquad i = 1, \ldots, k \\
\mathsf{w} &:= \mathsf{w} - h_{ik} \mathsf{v}_i.
\end{aligned}$$

3. Let

$$h_{k+1,k} := \|\mathsf{w}\|$$
$$\mathsf{v}_{k+1} := \frac{\mathsf{w}}{h_{k+1,k}}$$

4. Compute $\mathsf{y}_k := \operatorname{argmin}\|\beta\mathsf{e}_1 - \mathsf{H}\mathsf{y}\|$.

Finally, let

$$\mathsf{u} = \mathsf{u}_0 + \mathsf{P}\sum_j y_j \mathsf{v}_j$$

6.1.4 Remark: The GMRES algorithm requires an auxiliary vector in each iteration step. Therefore, it is customary to restart the iteration after a certain number of steps.

6.1.5 Lemma: *Let* $\|\mathsf{u}\|_\mathsf{A}^2 = \mathsf{u}^T\mathsf{A}\mathsf{u}$. *The conjugate gradient method minimizes the error in step* k *in the sense that there is a polynomial* $p \in \mathbb{P}_{k-1}$ *such that*

$$\mathsf{u}_k = \mathsf{u}_0 + p(\mathsf{A})\mathsf{r}_0,$$

and p *minimizes the expression*

$$\|(\mathsf{I} - \mathsf{A}p(\mathsf{A}))(\mathsf{u} - \mathsf{u}_0)\|_\mathsf{A} \tag{6.1}$$

in the space $\mathbb{P}_{k-1}$. *The GMRES method minimizes the norm of the residual in each step such that*

$$\|(\mathsf{I} - \mathsf{A}p(\mathsf{A}))\mathsf{r}_0\| \tag{6.2}$$

is minimal in $\mathbb{P}_{k-1}$.

Proof. see [Saa00]. □

6.1.6 Definition: The *spectral condition number* of a symmetric positive definite matrix A is defined by

$$\operatorname{cond}_2 A = \frac{\lambda_{\max}(\mathsf{A})}{\lambda_{\min}(\mathsf{A})},$$

where $\lambda_{\max}(\mathsf{A})$ and $\lambda_{\min}(\mathsf{A})$ are the largest and smallest eigenvalues of A, respectively.

6.1.7 Corollary: *The error after the* n*th conjugate gradient iteration step admits the estimate*

$$\|\mathsf{u}^n - \mathsf{u}\|_\mathsf{A} \leqslant \frac{\sqrt{\operatorname{cond}_2 \mathsf{A}} - 1}{\sqrt{\operatorname{cond}_2 \mathsf{A}} + 1}\|\mathsf{u}^{n-1} - \mathsf{u}\|_\mathsf{A} \leqslant \left(\frac{\sqrt{\operatorname{cond}_2 \mathsf{A}} - 1}{\sqrt{\operatorname{cond}_2 \mathsf{A}} + 1}\right)^n \|\mathsf{u}^0 - \mathsf{u}\|_\mathsf{A}. \tag{6.3}$$

Proof. The second inequality follows immediately by iteration. The first inequality follows from (6.1) by using Chebyshev polynomials for the interval $[\lambda_{\min}(\mathsf{A}), \lambda_{\max}(\mathsf{A})]$ (cf. [Saa00]). □

6.1.8 Corollary: *Let the matrix* A *have real eigenvalues and be positive definite (but not necessarily symmetric). Then, the residual of the* n*th step of the GMRES iteration admits the estimate*

$$\left\|\mathsf{A}\mathsf{u}^n - \mathsf{f}\right\| \leqslant \left(\frac{\sqrt{\operatorname{cond}_2 \mathsf{A}} - 1}{\sqrt{\operatorname{cond}_2 \mathsf{A}} + 1}\right)^n \left\|\mathsf{A}\mathsf{u}^0 - \mathsf{f}\right\|. \tag{6.4}$$

Proof. This estimate follows from (6.2) in the same way as (6.3) from (6.1). □

6.2 Interior Penalty

6.2.1 Using the basis of V_h chosen in paragraph 1.1.12, equation (2.12) is equivalent to the linear system of equations

$$\mathsf{A}\mathsf{u} = \mathsf{f}, \tag{6.5}$$

where $\mathsf{u} \in \mathbb{R}^n$ is the coefficient vector of u with respect to the basis functions and the entries of A and f are defined by (we continue using double indices for the basis functions)

$$\mathsf{A}_{S,i;T,j} = a_h^{\mathrm{IP}}\big(\varphi_{S,i}, \varphi_{T,j}\big), \qquad \mathsf{f}_{T,j} = \big(f, \varphi_{T,j}\big)_T. \tag{6.6}$$

6.2.2 Lemma: *Assume that the shape regularity Assumption 1.1.13 holds. Then, the spectral condition number* $\operatorname{cond}_2(A_h)$ *of the operator* A_h *defined by*

$$\big(A_h u, v\big) = a_h^{IP}\big(u, v\big) \quad \forall u, v \in V_h, \tag{6.7}$$

behaves asymptotically like

$$\begin{aligned} \operatorname{cond}_2(A_h) &= \mathcal{O}(h_{\min}^{-2}) && h \to 0, \\ \operatorname{cond}_2(A_h) &= \mathcal{O}(\kappa) && \kappa \to \infty. \end{aligned}$$

Here, $h_{\min}$ *is the diameter of the smallest grid cell. The condition number of the matrix* A *behaves asymptotically like the condition number of the operator* A_h*.*

Proof. Let $u^0 \in W^{1,2}(\Omega)$ be an eigenfunction to the smallest eigenvalue λ_0 of the continuous Poisson operator. It depends on the domain Ω only. Furthermore, it is known to be a smooth function. Let u_h^0 be its interior penalty approximation, i. e.,

$$a_h^{\mathrm{IP}}\big(u_h^0, v\big) = a_h^{\mathrm{IP}}\big(u^0, v\big) \qquad \forall v \in V_h.$$

Due to the consistence of $a_h^{\mathrm{IP}}\big(.,.\big)$, we have

$$\lambda_0 \big\|u^0\big\|^2 = a_h^{\mathrm{IP}}\big(u^0, u^0\big) = a_h^{\mathrm{IP}}\big(u_h^0, u_h^0\big) + \mathcal{O}(h^{2k}).$$

Therefore, we conclude that independent of $\kappa > \kappa_0$,

$$\lambda_{\min}(A_h) \leqslant \lambda_0 + \mathcal{O}(h^{2k}). \tag{6.8}$$

If $\kappa > \kappa_0$, a lower bound of the minimal eigenvalue follows from the stability of the method with respect to the energy norm and Friedrichs inequality in Lemma 1.2.3.

We estimate the maximal eigenvalue by using the fact that due to symmetry of A_h for any $u \in V_h$

$$a_h(u,u) \leqslant \lambda_{\max}(A_h)\big(u,u\big).$$

Choosing u such that $u = 1$ on a cell T such that the diameter of T is minimal and zero everywhere else, we conclude that

$$\lambda_{\max}(A_h) \geqslant \frac{a_h(u,u)}{\big\|u\big\|^2} = c\frac{\kappa}{h_{\min}^2}.$$

An upper bound for the maximal eigenvalue with the same asymptotic behavior follows from the boundedness of the energy norm and inverse estimates. □

6.2.3 Corollary: *Solution of the discrete problem* (6.5) *by the conjugate gradient method requires* $\mathcal{O}(h_{\min}^{-1})$ *iteration steps, where* $h_{\min}$ *is the diameter of the smallest grid cell.*

6.2.4 The previous lemma and corollary imply that the solution of the discrete linear system (6.5) will be very time-consuming if performed on fine grids without a suitable preconditioner. The remainder of this section is devoted to develop such a preconditioner. We will show, that the suggested preconditioner is *uniform*, that is, the condition number of the product of preconditioner and matrix is bounded independent of the mesh size.

6.2.5 We recall that we assume a hierarchy of triangulations

$$\mathbb{T}_0 \sqsubseteq \ldots \sqsubseteq \mathbb{T}_\ell \sqsubseteq \ldots \sqsubseteq \mathbb{T}_L.$$

Since, due to Assumption 1.1.4, this hierarchy corresponds to a nesting of the finite element spaces,

$$V_0 \subset \cdots \subset V_\ell \subset \cdots \subset V_L. \tag{6.9}$$

We remark that the L^2-projection $\Pi_\ell : L^2(\Omega) \to V_\ell$ is obviously an operator also from $V_{\ell+1}$ to V_ℓ. By its projection properties we have for any $u \in V_\ell$ and $v \in V_{\ell+1}$

$$\big(u, v\big) = \big(u, \Pi_\ell v\big) = \big(\Pi_\ell^* u, v\big).$$

Therefore, the adjoint operator Π_ℓ^* or *prolongation* is simply the embedding of V_ℓ into $V_{\ell+1}$.

6.2.6 Definition: The matrix R_ℓ associated with the projection Π_ℓ is called *restriction*, and its transpose R_ℓ^T, the matrix representing the embedding operator, *prolongation*

On each grid cell T, we define the *cell matrix* A_T by

$$(\mathsf{A}_T)_{ij} = a_h^{\mathrm{IP}}\big(\varphi_{T,i}^\ell, \varphi_{T,j}^\ell\big). \tag{6.10}$$

Here, ℓ is the grid level on which the cell T is located. Analogous to (6.6), we define level matrices $\mathsf{A}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$ by using shape functions $S, T \in \mathbb{T}_\ell$.

6.2.7 Definition: The *mass matrix* M_ℓ on level ℓ is the matrix with entries

$$(\mathsf{M}_\ell)_{S,i;T,j} = \big(\varphi_{S,i}, \varphi_{T,j}\big).$$

In the context of discontinuous finite elements, this matrix is always block diagonal, that is, $(\mathsf{M}_\ell)_{S,i;T,j} = 0$ if $S \neq T$. Furthermore, if the shape functions are Legendre polynomials, it is a diagonal matrix. Legendre polynomials for the actual grid cells would yield the identity matrix.

6.2.8 Definition: A main ingredient of multilevel methods is the smoother S_ℓ. We consider four simple smoothers here, point and block versions of Jacobi and Gauß-Seidel methods.

The *diagonal* and *block diagonal* of the matrix A_ℓ are

$$\mathsf{D}_\ell = \operatorname{diag}(\mathsf{A}_\ell)_{T,i;T,i} \qquad\qquad \mathsf{D}_{B;\ell} = \operatorname{diag}_{T\in\mathbb{T}_\ell}(\mathsf{A}_T)$$

Given any ordering "$<$" of the grid cells, we define an ordering of the index set $\{(T,i)\}$ by

$$(S,i) < (T,j) \;\;\Leftrightarrow\;\; S < T \;\vee\; (S = T \wedge i < j).$$

Then, the *lower triangle* and *lower block triangle* of the matrix are the matrices

$$(\mathsf{L}_\ell)_{S,i;T,j} = \begin{cases} (\mathsf{A}_\ell)_{S,i;T,j} & \text{if } (S,i) < (T,j) \\ 0 & \text{else,} \end{cases}$$

$$(\mathsf{L}_{B;\ell})_{ST} = \begin{cases} (\mathsf{A}_\ell)_{ST} & \text{if } S < T \\ 0 & \text{else,} \end{cases}$$

The two variants of the Jacobi smoother, namely *point Jacobi* and *block Jacobi*, are

$$\mathsf{S}_{J;\ell} = \omega\mathsf{D}_\ell^{-1}, \qquad\qquad \mathsf{S}_{J_B;\ell} = \omega\mathsf{D}_{B;\ell}^{-1}$$

with *relaxation parameter* ω and the point and block Gauß-Seidel smoothers are

$$\mathsf{S}_{G;\ell} = (\mathsf{D}_\ell + \mathsf{L}_\ell)^{-1}, \qquad\qquad \mathsf{S}_{G_B;\ell} = (\mathsf{D}_{B;\ell} + \mathsf{L}_{B;\ell})^{-1}.$$

6.2.9 Algorithm: We define the action of the *variable V-cycle* operator $\breve{\mathsf{P}}_\ell$ on a vector d_ℓ recursively. With S_ℓ any of the smoothers in Definition 6.2.8, let

$$\mathsf{S}_\ell^{(i)} = \begin{cases} \mathsf{S}_\ell & i \text{ odd} \\ \mathsf{S}_\ell^T & i \text{ even.} \end{cases}$$

Let the numbers smoothing steps m_ℓ be positive integers such that

$$\beta_0 m_\ell \leqslant m_{\ell-1} \leqslant \beta_1 m_\ell, \tag{6.11}$$

with $1 < \beta_0 \leqslant \beta_1$.

Then, let $\breve{\mathsf{P}}_0 = \mathsf{A}_0^{-1}$. If $\ell \neq 0$, set $x^{(0)} = 0$ and compute $\breve{\mathsf{P}}_\ell \mathsf{d}_\ell$ (assuming that $\breve{\mathsf{P}}_{\ell-1}$ is already defined) by the following steps:

1. (Pre-smoothing) Compute $\mathsf{x}^{(m_\ell)}$ iteratively by

$$\mathsf{x}^{(i)} = \mathsf{x}^{(i-1)} + \mathsf{S}_\ell^{(i)}(\mathsf{d}_\ell - \mathsf{A}_\ell \mathsf{x}^{(i-1)}), \quad i = 1, \ldots, m_\ell.$$

2. (Coarse grid correction) Let

$$\mathsf{y}^{(0)} = \mathsf{x}^{(m_\ell)} + \mathsf{R}_{\ell-1}^T \breve{\mathsf{P}}_{\ell-1} \mathsf{R}_{\ell-1}(\mathsf{d}_\ell - \mathsf{A}_\ell \mathsf{x}^{(m_\ell)}).$$

3. (Post-smoothing) Compute $\mathsf{y}^{(m_\ell)}$ iteratively by

$$\mathsf{y}^{(i)} = \mathsf{y}^{(i-1)} + \mathsf{S}_\ell^{(m_\ell+i)}(\mathsf{d}_\ell - \mathsf{A}_\ell \mathsf{y}^{(i-1)}), \quad i = 1, \ldots, m_\ell.$$

4. Set $\breve{\mathsf{P}}_\ell \mathsf{d}_\ell = \mathsf{y}^{(m_\ell)}$.

6.2.10 Remark: A typical choice is $\beta_0 = \beta_1 = 2$ and $m_L = 1$. In this case, the number of smoothing steps on level ℓ is $2^{L-\ell}$. Then, the asymptotic complexity of the variable V-cycle is the same as that of the W-cycle.

Choosing $\beta_0 = \beta_1 = 1$ yields the standard V-cycle. This algorithm is not covered by our analysis, but numerical results later show that it is feasible, too.

6.2.11 Theorem: *Suppose that regularity Assumption* (2.5) *holds. Then,* $\breve{\mathsf{P}}_L$ *is symmetric and positive definite and there is a constant* c_κ *independent of the level (or of* h_ℓ*) such that*

$$\frac{1}{\zeta} \leqslant \frac{\mathsf{x}^T \mathsf{A}_\ell \breve{\mathsf{P}}_\ell \mathsf{A}_\ell \mathsf{x}}{\mathsf{x}^T \mathsf{A}_\ell \mathsf{x}} \leqslant \zeta \quad \forall \mathsf{x} \in \mathbb{R}^{n_L}, \tag{6.12}$$

with $\zeta = (c_\kappa + m_L^{\alpha/2})/m_L^{\alpha/2}$ *and* α *from equation* (2.5)*. Consequently, the spectral condition number of* $\breve{\mathsf{P}}_L \mathsf{A}_L$ *satisfies* $\operatorname{cond}_2(\breve{\mathsf{P}}_L \mathsf{A}_L) \leqslant \zeta^2$.

6.2.12 Remark: This theorem is the main result of [GK03c]. In virtue of the abstract theory given in [BPX91], the proof reduces to verification of two conditions.

1. There exists $\tilde{\omega} \in (0, 2)$ such that for all $\ell = 1, \ldots, L$

$$\frac{\tilde{\omega}}{\lambda_\ell} \mathsf{x}^T \mathsf{M}_\ell \mathsf{x} \leqslant 2\mathsf{x}^T \mathsf{S}_\ell \mathsf{x} - \mathsf{x}^T \mathsf{S}_\ell^T \mathsf{A}_\ell \mathsf{S}_\ell \mathsf{x}, \quad \forall x \in \mathbb{R}^{n_\ell}.$$

2. Let $Q_{\ell-1} : V_\ell \to V_{\ell-1}$ be the Ritz projection defined by

$$a_{\ell-1}^{\mathrm{IP}}\big(Q_{\ell-1}u, v\big) = a_\ell^{\mathrm{IP}}\big(u, v\big) \quad \forall u \in V_\ell,\ v \in V_{\ell-1}. \tag{6.13}$$

   Then, there is $\eta \in (0, 1)$ and $C_P > 0$ such that

$$\big|a_\ell^{\mathrm{IP}}\big(u - Q_{\ell-1}u, u\big)\big| \leqslant C_P \left(\frac{\big\|A_\ell u\big\|_{1;h}^2}{\lambda_\ell}\right) a_\ell^{\mathrm{IP}}\big(u, u\big)^{1-\eta}. \tag{6.14}$$

The following lemmas present some intermediary results required in the proof of Theorem 6.2.11, which is deferred to paragraph 6.2.15.

6.2.13 Lemma: *Let S_ℓ be any of the smoothing operators in Definition 6.2.8. If the smoother is of Jacobi type, we assume further that ω is sufficiently small. Then, S_ℓ is a smoothing operator, that is, there exists an $\tilde{\omega} \in (0,2)$ such that*

$$\frac{\tilde{\omega}}{\lambda_\ell}\mathsf{x}^T\mathsf{M}_\ell\mathsf{x} \leqslant 2\mathsf{x}^T\mathsf{S}_\ell\mathsf{x} - \mathsf{x}^T\mathsf{S}_\ell^T\mathsf{A}_\ell\mathsf{S}_\ell\mathsf{x}, \quad \forall x \in \mathbb{R}^{n_\ell} \tag{6.15}$$

Proof. This result has been proven in a very general context in [BP92] □

6.2.14 Lemma: *Assume that regularity Assumption 2.5 holds. Then, for all $u \in V_\ell$, with $\ell = 1, \ldots, L$ holds*

$$\big\|u - Q_{\ell-1}u\big\|_\ell \leqslant C(\kappa)h_\ell^\alpha \big\|A_\ell u\big\|_{W^{-1+\alpha,2};\Omega}.$$

Proof. We start with

$$\big\|u - Q_{\ell-1}u\big\|_\ell \leqslant \big\|u - w\big\|_\ell + \big\|w - Q_{\ell-1}u\big\|_\ell, \tag{6.16}$$

where $w \in W^{1+\alpha,2}(\Omega)$ is the solution of the dual problem

$$-\Delta w = A_\ell u \quad \text{in } \Omega.$$

By consistency of the interior penalty method and the definition of A_ℓ, we have

$$a_\ell^{\text{IP}}\big(w, v\big) = \big(A_\ell u, v\big) = a_\ell^{\text{IP}}\big(u, v\big) \quad \forall v \in V_\ell.$$

Therefore, u is the Ritz projection of w in V_ℓ, and in virtue of estimate (2.5), we have

$$\big\|u - w\big\|_\ell \leqslant Ch_\ell^\alpha \big|w\big|_{W^{1+\alpha,2}(\Omega)}. \tag{6.17}$$

For the estimation of the second term in (6.16), we use the same arguments as above, yielding

$$a_{\ell-1}^{\text{IP}}\big(w, v\big) = \big(A_\ell u, v\big) = a_{\ell-1}^{\text{IP}}\big(Q_{\ell-1}u, v\big) \quad \forall v \in V_{\ell-1}.$$

by definition of $Q_{\ell-1}$. Therefore,

$$\big\|w - Q_{\ell-1}u\big\|_{\ell-1} \leqslant Ch_\ell^\alpha \big\|w\big\|_{W^{1+\alpha,2}(\Omega)}.$$

It remains to show, that the term $\big\|w - Q_{\ell-1}u\big\|_\ell$ is bounded sufficiently by the coarse level norm $\big\|w - Q_{\ell-1}u\big\|_{\ell-1}$. Observe that $[\![w - Q_{\ell-1}u]\!]$ is zero on each edge $E \in \mathbb{E}_\ell$ which is not a subset of an edge in $\mathbb{E}_{\ell-1}$.. Therefore,

$$\sum_{E\in\mathbb{E}_\ell} \frac{\kappa}{h_E}\big\|[\![w - Q_{\ell-1}u]\!]\big\|_E^2 = 2\sum_{E\in\mathbb{E}_{\ell-1}} \frac{\kappa}{h_E}\big\|[\![w - Q_{\ell-1}u]\!]\big\|_E^2,$$

and consequently

$$\big\|w - Q_{\ell-1}u\big\|_\ell^2 \leqslant 2\big\|w - Q_{\ell-1}u\big\|_{\ell-1} \leqslant Ch_\ell^\alpha \big\|w\big\|_{W^{1+\alpha,2}(\Omega)}. \tag{6.18}$$

Adding up (6.17) and (6.18) completes the proof. □

6.2.15 Proof of Theorem 6.2.11 We will prove (6.14) with $\eta = \alpha/2$ and α from estimate (2.5). Lemma 6.2.14 and interpolation of Sobolev norms yield

$$\|u - Q_{\ell-1}u\|_\ell \leqslant Ch_\ell^\alpha \|A_\ell u\|_{-1+\alpha,2} \leqslant Ch_\ell^\alpha \|A_\ell u\|_{-1,2}^{1-\alpha} \|A_\ell u\|^\alpha.$$

Assuming for the moment that

$$\|A_\ell u\|_{-1,2} \leqslant C\|u\|_\ell, \tag{6.19}$$

the proof of estimate (6.14) is completed as follows:

$$\begin{aligned} a_\ell^{\mathrm{IP}}\big(u - Q_{\ell-1}u, u\big) &\leqslant \|u - Q_{\ell-1}u\|_\ell \|u\|_\ell \\ &\leqslant Ch_\ell^\alpha \|A_\ell u\|^\alpha \|u\|_\ell^{2-\alpha} \\ &\leqslant C\left(\frac{\|A_\ell u\|^2}{\lambda_\ell}\right)^{\frac{\alpha}{2}} a_\ell^{\mathrm{IP}}\big(u,u\big)^{1-\frac{\alpha}{2}}, \end{aligned}$$

which is (6.14) with $\eta = \alpha/2$. Thus, it remains to prove (6.19).

$$\begin{aligned} \|A_\ell u\|_{-1,2} &= \sup_{\psi \in W^{1,2}(\Omega)} \frac{\big(A_\ell u, \psi\big)}{|\psi|_{1,2}} \\ &\leqslant \sup_{\psi \in W^{1,2}(\Omega)} \frac{\big(A_\ell u, \eta_\psi\big)}{|\psi|_{1,2}} + \sup_{\psi \in W^{1,2}(\Omega)} \frac{a_\ell^{\mathrm{IP}}\big(u, \Pi_h \psi\big)}{|\psi|_{1,2}}. \end{aligned}$$

For the first term, we use approximation estimate (1.21) to obtain

$$\frac{\big(A_\ell u, \eta_\psi\big)}{|\psi|_{1,2}} \leqslant Ch_\ell \|A_\ell u\|.$$

Stability of the projection yields

$$\frac{\big(A_\ell u, \Pi_h \psi\big)}{|\psi|_{1,2}} \leqslant C\|u\|_\ell.$$

Finally, since A_ℓ is symmetric positive definite, we have

$$h_\ell^2 \|A_\ell u\|^2 \leqslant h_\ell^2 \lambda_\ell\big(A_\ell u, u\big) \leqslant a_\ell^{\mathrm{IP}}\big(u,u\big),$$

which completes the proof. □

6.2.16 Remark: The preceding analysis focused on the case of the multigrid method as a preconditioner in an outer Krylov space iteration, where the condition number $\mathrm{cond}_2(\breve{\mathsf{P}}_L \mathsf{A}_L)$ is significant for efficiency.

Alternatively, the multilevel method can be used as a solver in an iteration like

$$\mathsf{u}^{(n+1)} = \mathsf{u}^n - \breve{\mathsf{P}}_L(\mathsf{b}_L - \mathsf{A}_L \mathsf{u}^n).$$

Here, the *contraction number* of the iteration matrix

$$\varrho(\mathsf{I} - \breve{\mathsf{P}}_L \mathsf{A}_L) = \max_i \{|\lambda_i(\mathsf{I} - \breve{\mathsf{P}}_L \mathsf{A}_L)|\}, \tag{6.20}$$

where $\lambda_i(\mathsf{I} - \breve{\mathsf{P}}_L \mathsf{A}_L)$ are the eigenvalues of $\mathsf{I} - \breve{\mathsf{P}}_L \mathsf{A}_L$. Without covering this case by our analysis, we will show some numerical results below.

		$m(\ell) = 2^{L-\ell}$		$m(\ell) = 1$	
L	$\kappa(\mathsf{A}_L)$	$\kappa(\bar{\mathsf{P}}_L\mathsf{A}_L)$	$\varrho(\mathsf{I} - \bar{\mathsf{P}}_L\mathsf{A}_L)$	$\kappa(\bar{\mathsf{P}}_L\mathsf{A}_L)$	$\varrho(\mathsf{I} - \bar{\mathsf{P}}_L\mathsf{A}_L)$
2	10	1.36	0.19	1.36	0.19
3	22	1.71	0.26	1.72	0.27
4	79	1.97	0.36	2.11	0.42
5	312	2.08	0.41	2.39	0.54
6	1246	2.11	0.42	2.56	0.62
7	4981	2.11	0.42	2.66	0.66
8	19921	2.12	0.42	2.73	0.69

Table 6.1: Condition numbers and contraction numbers, when Gauß-Seidel smoother, and $\mathbb{Q}_1$ elements are used. $\Omega = (-1,1)^2$, $\kappa = 3$

	$m(\ell) = 2^{L-\ell}$		$m(\ell) = 2^{L-\ell+1}$	
L	$\kappa(\bar{\mathsf{P}}_L\mathsf{A}_L)$	$\varrho(\mathsf{I} - \bar{\mathsf{P}}_L\mathsf{A}_L)$	$\kappa(\bar{\mathsf{P}}_L\mathsf{A}_L)$	$\varrho(\mathsf{I} - \bar{\mathsf{P}}_L\mathsf{A}_L)$
2	1.62	0.30	1.14	0.10
3	2.23	0.40	1.36	0.17
4	2.72	0.54	1.49	0.21
5	2.95	0.64	1.54	0.22
6	3.02	0.67	1.56	0.22
7	3.04	0.68	1.56	0.22

Table 6.2: Condition numbers and contraction numbers when Jacobi smoother and $\mathbb{Q}_1$ elements are used. $\Omega = (-1,1)^2$, $\kappa = 3$

6.2.17 The following numerical experiments are performed on hierarchies of Cartesian meshes generated by global refinement (see A.1.1 on page 167), unless stated otherwise. The right hand side was chosen according to example A.2.2 on page 170, since trigonometric functions are approximated by eigenfunctions of the matrix A_L and cause spurious convergence effects. The start vector for all iterations is $\mathsf{u}^{(0)} = 0$.

6.2.18 We start tests of the multilevel method with the standard situation of $\mathbb{Q}_1$ elements and the block-Gauß-Seidel smoother $\mathsf{S}_{G_B;\ell}$ of Definition 6.2.8 on page 136.*

Table 6.1 shows results for this case. In the first column, we display the condition number of the matrix A_L without preconditioning for comparison; they exhibit the expected growth behavior and increase by a factor of four with each refinement step. The next column shows the preconditioner according to the preceding analysis. Clearly, the condition numbers stay bounded by a very moderate number, resulting in a conjugate gradient iteration with contraction number 0.19. This number is much smaller than the contraction number of the multigrid method (6.20) itself, shown in the third column.

6.2.19 The block-Jacobi smoother $\mathsf{S}_{J_B;\ell}$ is subject of Table 6.2. Here, the relaxation parameter is $\omega = 1$. The left column shows that the condition numbers of the preconditioned system

*The condition numbers in the following paragraphs are taken from [GK03c], where the stabilization on the boundary was equal to the one in the interior ($\vartheta = 1$ in (2.19)). The results shown in Remark 6.2.22 reflect the finer stability analysis and are computed for the form in (2.12).

	$\mathbb{Q}_2$ elements ($\kappa = 8$)			$\mathbb{Q}_3$ elements ($\kappa = 22$)		
L	$\kappa(\mathsf{A}_L)$	$\kappa(\breve{\mathsf{P}}_L\mathsf{A}_L)$	$\varrho(\mathsf{I}-\breve{\mathsf{P}}_L\mathsf{A}_L)$	$\kappa(\mathsf{A}_L)$	$\kappa(\breve{\mathsf{P}}_L\mathsf{A}_L)$	$\varrho(\mathsf{I}-\breve{\mathsf{P}}_L\mathsf{A}_L)$
2	23	2.07	0.36	77	2.97	0.53
3	69	2.11	0.39	269	2.88	0.50
4	263	2.14	0.40	1061	2.90	0.52
5	1041	2.16	0.41	4235	2.92	0.52
6	4154	2.16	0.41	16934	2.92	0.52
7	16605	2.16	0.41	67731	2.92	0.52

Table 6.3: Condition numbers of A_L and $\breve{\mathsf{P}}_L\mathsf{A}_L$ when Gauß-Seidel smoother and biquadratic ($\mathbb{Q}_2$) and bicubic ($\mathbb{Q}_3$) shape functions are used. $\Omega = (-1,1)^2$.

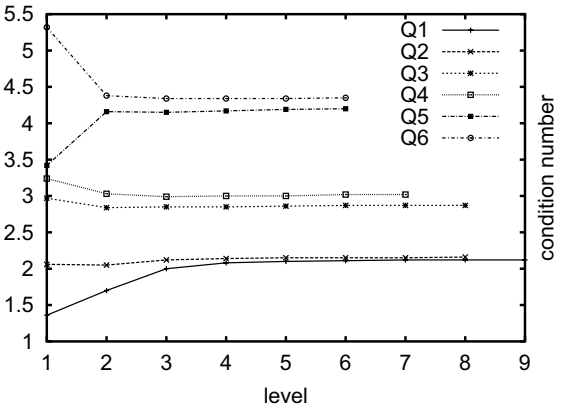

Figure 6.1: Condition numbers for different shape function spaces on a sequence of meshes

as well as the contraction number of the linear multigrid method are considerably higher than for the Gauß-Seidel smoother. Since both methods can be implemented with about the same numerical effort, Gauß-Seidel should be preferred. The right column shows that doubling the number of smoothing steps can improve the performance of the Jacobi smoother a lot.

6.2.20 Convergence for shape function spaces $\mathbb{Q}_2$ and $\mathbb{Q}_3$ is displayed in Table 6.3. We see that the condition number increases only slightly with polynomial degree, a result also evident from Figure 6.1, where the condition numbers on a sequence of meshes are displayed up to polynomials in $\mathbb{Q}_6$.

6.2.21 In the analysis, we assumed a mild regularity assumption (2.5), that holds for non-convex polygonal domains. We now investigate the performance of the preconditioner on the domains with reentrant corners of example A.2.3 on page 170. While (2.5) holds with $\alpha < 2/3$ on the L-shaped domain Ω_{270} with interior angle of $270°$, it does not hold for the slit domain Ω_{360}, where $\alpha = 1/2$. We conclude from Table 6.4 that the algorithm yields a good preconditioner in both cases.

	Ω_{270}		Ω_{360}	
L	$\kappa(\tilde{\mathsf{P}}_L\mathsf{A}_L)$	$\varrho(\mathsf{I}-\tilde{\mathsf{P}}_L\mathsf{A}_L)$	$\kappa(\tilde{\mathsf{P}}_L\mathsf{A}_L)$	$\varrho(\mathsf{I}-\tilde{\mathsf{P}}_L\mathsf{A}_L)$
2	1.70	0.28	1.70	0.31
3	1.96	0.38	1.92	0.35
4	2.08	0.41	2.06	0.40
5	2.11	0.42	2.10	0.41
6	2.11	0.42	2.11	0.42
7	2.12	0.42	2.12	0.42

Table 6.4: Condition numbers and contraction numbers for L-shaped and slit domains using bilinear shape functions and Gauß-Seidel smoothing

	$\mathbb{Q}_1$, $\kappa=2$			$\mathbb{Q}_2$, $\kappa=4$			$\mathbb{Q}_3$, $\kappa=8$		
L	n_{10}	$\overline{\varrho}$	$\overline{r}$	n_{10}	$\overline{\varrho}$	$\overline{r}$	n_{10}	$\overline{\varrho}$	$\overline{r}$
1	6	0.017	1.763	8	0.056	1.253	9	0.076	1.116
2	8	0.053	1.275	11	0.117	0.932	12	0.134	0.874
3	9	0.058	1.235	12	0.139	0.858	13	0.162	0.792
4	9	0.062	1.210	12	0.143	0.846	13	0.167	0.777
5	9	0.062	1.208	12	0.144	0.843	13	0.168	0.776
6	9	0.062	1.209	12	0.143	0.843	13	0.168	0.776
7	9	0.062	1.211	12	0.143	0.844	13	0.168	0.776

Table 6.5: Contraction and convergence rates for the variable V-cycle with block-Gauß-Seidel smoother

6.2.22 Remark: Condition numbers are more of theoretical interest and their relation to the number of steps required for a certain accuracy is only indirect. Furthermore, their meaning in this context is limited to symmetric positive definite matrices. To be able to compare with preconditioners for other discretizations later, we summarize the results of the preceding paragraphs showing values of immediate practical relevance[†]. First, there is n_{10}, the number of steps required to reduce the initial residual by 10^{10}. Since this value is discrete, we also show the average *contraction rate* $\overline{\varrho}$ and the decadic *convergence rate*

$$\overline{r} = -\log_{10}\overline{\varrho} \tag{6.21}$$

according to Varga [Var99].

6.2.23 The method can be applied to $\mathbb{P}_k$ shape function spaces, as shown in Table 6.6. The figures here are very similar to Table 6.6 and we conclude that the multilevel preconditioner 6.2.9 is as efficient for complete polynomial spaces as for tensor product polynomials.

6.2.24 Much like the condition number of the original matrix A_L itself, the condition number of the preconditioned system depends on the stabilization parameter κ. Figure 6.2 shows how the value $1/\bar{r}$ varies with κ. We observe a dependence like $\bar{r}^2 \approx 1/\kappa$ for large values of κ, which corresponds to condition numbers depending linearly on κ. In particular, this figure shows that κ must be chosen carefully to yield an efficient solver. We remark that the shape of the curve is in good correspondence with the results for incomplete Cholesky preconditioners in [Cas01, Cas02].

[†]From now on, values are computed with the improved boundary stabilization

	$\mathbb{P}_1$, $\kappa = 2$			$\mathbb{P}_2$, $\kappa = 4$			$\mathbb{P}_3$, $\kappa = 8$		
L	n_{10}	$\overline{\varrho}$	$\overline{r}$	n_{10}	$\overline{\varrho}$	$\overline{r}$	n_{10}	$\overline{\varrho}$	$\overline{r}$
1	5	0.000	4.382	7	0.024	1.617	8	0.042	1.378
2	9	0.054	1.268	11	0.100	0.999	13	0.145	0.838
3	9	0.070	1.158	12	0.130	0.885	14	0.172	0.765
4	9	0.072	1.142	12	0.135	0.870	14	0.179	0.748
5	9	0.071	1.150	12	0.135	0.868	14	0.179	0.747
6	9	0.070	1.155	12	0.135	0.871	14	0.179	0.747
7	9	0.069	1.162	12	0.133	0.875	14	0.179	0.747

Table 6.6: Reduction and convergence rates for the variable V-cycle with block-Gauß-Seidel smoother ($\mathbb{P}_k$ shape functions)

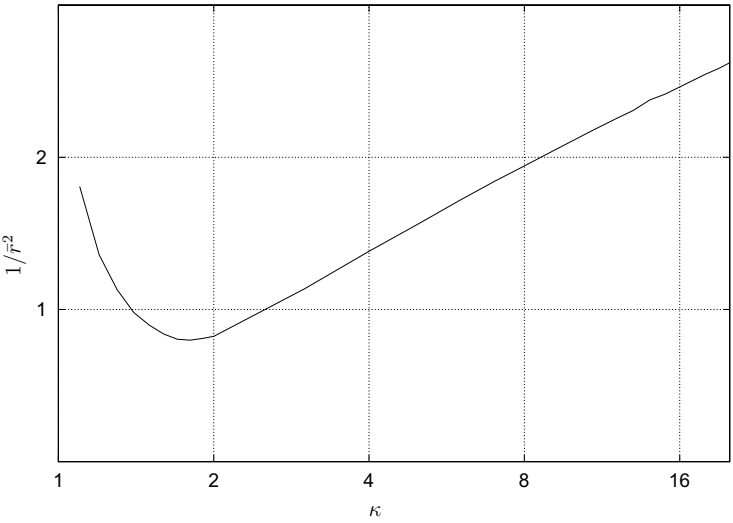

Figure 6.2: Solver performance depending on stabilization parameter κ.

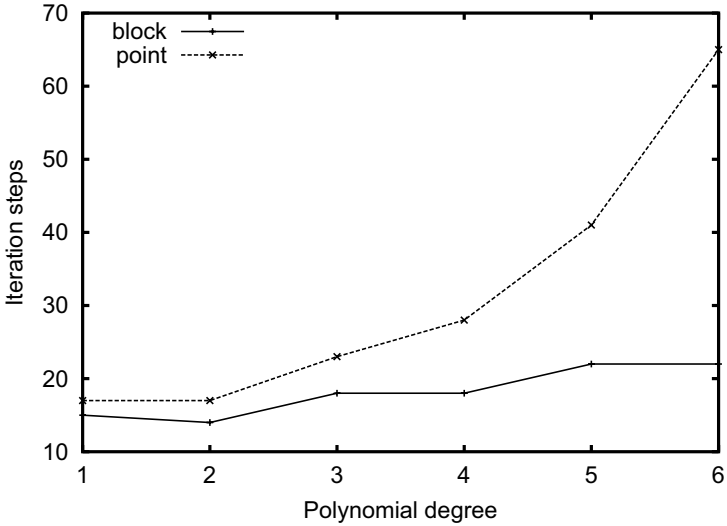

Figure 6.3: Comparison of point and block smoothers for different shape function spaces

	$\mathbb{Q}_1$, $\kappa = 2$			$\mathbb{P}_1$, $\kappa = 2$			$\mathbb{P}_2$, $\kappa = 5$		
L	n_{10}	$\overline{\varrho}$	$\overline{r}$	n_{10}	$\overline{\varrho}$	$\overline{r}$	n_{10}	$\overline{\varrho}$	$\overline{r}$
1	7	.025	1.598	6	.017	1.766	8	.040	1.395
2	9	.068	1.170	9	.064	1.194	12	.123	0.910
3	10	.098	1.008	10	.093	1.032	14	.180	0.744
4	11	.113	0.945	11	.111	0.957	15	.211	0.676
5	11	.121	0.915	11	.121	0.917	16	.227	0.644
6	12	.123	0.911	12	.123	0.909	16	.230	0.638

Table 6.7: Contraction and convergence rates on a three-dimensional cube

6.2.25 Comparison between the point Gauß-Seidel smoother $\mathsf{S}_{G;L}$ and the block smoother $\mathsf{S}_{G_B;L}$ in Figure 6.3 (from [GK03b]) exhibits the superiority of the block version. While the iteration counts n_{10} of the block-preconditioned system increase only moderately with the polynomial degree, the number of steps seems to increase exponentially with point-Gauß-Seidel.

6.2.26 Results for the three-dimensional case are shown in Table 6.7, where we observe the same behavior as in two dimensions.

6.2.27 All experiments above have been performed on equidistant Cartesian grids. In Table 6.8, we show that the method is robust with respect to the shape of the grid cells. First, we show results for the irregular coarse grid in Figure A.2 on page 168, refined regularly. In this case, the inclusion of the spaces V_ℓ still holds and our theory is applicable. The results confirm that the reduction factors remain bounded independent of the mesh width.

	Irregular, $\kappa = 3.5$			Distorted, $\kappa = 6$		
L	n_{10}	$\overline{\varrho}$	$\overline{r}$	n_{10}	$\overline{\varrho}$	$\overline{r}$
1	17	0.248	0.606	7	0.030	1.517
2	20	0.307	0.513	16	0.228	0.642
3	21	0.326	0.487	20	0.301	0.521
4	22	0.337	0.472	22	0.340	0.469
5	21	0.331	0.480	23	0.364	0.439
6	21	0.329	0.483	25	0.387	0.412

Table 6.8: Reduction and convergence rates for the variable V-cycle with block-Gauß-Seidel smoother on non-Cartesian grids ($\mathbb{Q}_1$ shape functions)

The second column shows results on refinements of a single square cell, where the mesh is distorted with a random vector field in each refinement step (see Figure A.1 on page 168). In this case, inclusion of the level spaces V_ℓ does not hold anymore. Furthermore, the shape of the cells degenerates more and more during refinement. Therefore, we had to choose $\kappa = 6$ Still, the results show reasonable convergence rates.

6.3 Local multigrid

6.3.1 This section explains the implementation of the multi-level preconditioner presented in Section 6.2 on locally refined grids. The goals of this implementation are

1. Optimal complexity: the method should use the same amount of operations per degree of freedom as in the globally refined case. This requires *local smoothing* to avoid repeated smoothing on a single cell. On the other hand, local smoothing may not reduce the quality as a preconditioner.
2. The multilevel method should use the same matrices for local smoothing and matrix-vector multiplications. Assembling of these matrices should be possible in a cell-by-cell fashion similar to the case of global refinement.

Following [Bra93], we perform a multilevel method on the complete level spaces V_ℓ, but we restrict the smoother to the part that is really refined to level ℓ.

6.3.2 Definition: We call a grid cell in a hierarchy of locally refined meshes $\{\mathbb{T}_\ell\}$ *active*, if it is not refined in any of the triangulations of the hierarchy. Active cells are the locally finest cells of the hierarchy. In Figure 6.4, active cells are shaded, while inactive ones are white. The shape functions of all active cells constitute the discretization space V_L of the finest level.

6.3.3 Definition: We define the level ℓ_T of a grid cell T in the triangulation hierarchy $\{\mathbb{T}_\ell\}$ recursively as follows: if a cell belongs to the coarse mesh, its level is zero. Otherwise, it was obtained by refinement of another grid cell T_p and we set

$$\ell_T = \ell_{T_p} + 1.$$

Since there is no coarsening in a single hierarchy $\{\mathbb{T}_\ell\}$[‡], this level is uniquely defined.

[‡]This does not rule out coarsening in the process of obtaining the hierarchy

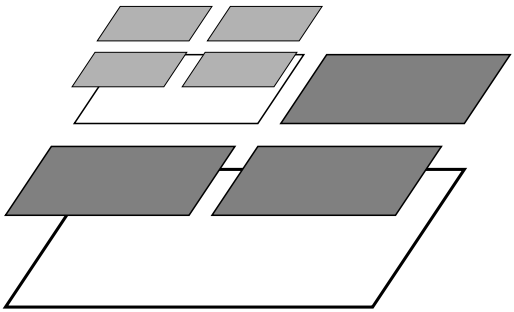

Figure 6.4: A hierarchy of three meshes with local refinement

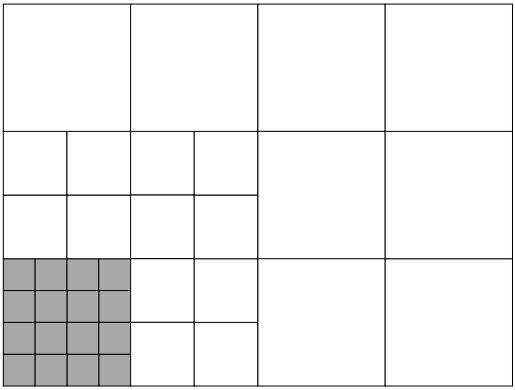

Figure 6.5: Splitting of $\mathbb{T}_\ell$ into $\mathbb{T}_\ell^S$ (shaded cells) and $\mathbb{T}_\ell^L$ (white)

6.3.4 Remark: Figure 6.4 shows that a cell of level ℓ may belong to several meshes $\mathbb{T}_\ell, \mathbb{T}_{\ell+1}$ and so on. For instance, the shaded cells on the intermediate level (cell level 1) belong to the meshes $\mathbb{T}_1$ and $\mathbb{T}_2$.

6.3.5 Definition: Each triangulation $\mathbb{T}_\ell$ is partitioned into the set of cells strictly on level ℓ

$$\mathbb{T}_\ell^S = \big\{T \in \mathbb{T}_\ell \big| \ell_T = \ell\big\},$$

and the set of cells on lower levels

$$\mathbb{T}_\ell^L = \big\{T \in \mathbb{T}_\ell \big| \ell_T < \ell\big\}.$$

This partitioning is explained in Figure 6.5. We remark that cells in $\mathbb{T}_\ell^L$ may contain grid cells of different lower levels. Obviously, there holds

$$\mathbb{T}_\ell = \mathbb{T}_\ell^S \cup \mathbb{T}_\ell^L.$$

We call the divider between those two sets the *refinement edge* $\mathbb{E}_l^E$ between levels ℓ and $\ell-1$, that is

$$\mathbb{E}_\ell^E = \left\{ E \in \mathbb{E}_\ell \middle| E \cap \bigcup_{\mathbb{T}_\ell^S} T = E \ \wedge \ E \cap \bigcup_{\mathbb{T}_\ell^L} T = E \right\}$$

6.3.6 The partitioning of the spaces V_ℓ into subspaces follows the splitting of $\mathbb{T}_\ell$: we let

$$V_\ell = V_\ell^S \oplus V_\ell^L, \tag{6.22}$$

where V_ℓ^S are the functions in V_ℓ with support in $\mathbb{T}_\ell^S$ and V_ℓ^L accordingly. Equation (6.22) holds for discontinuous finite elements. A basis for the subspaces is obtained by restricting the definition of the basis in paragraph 1.1.12 to the subsets of the triangulation. We will assume that the basis is ordered in a way that cells in $\mathbb{T}_\ell^S$ are before those in $\mathbb{T}_\ell^L$. Then, a function $u \in V_\ell$ is represented by a coefficient vector u_ℓ of the form $(\mathsf{u}_\ell^S, \mathsf{u}_\ell^L)$.

6.3.7 Remark: The ordering required in the previous paragraph does not impose any ordering of the basis functions inside V_ℓ^S and V_ℓ^L. Such an ordering may be chosen to optimize the smoother (see Remark 1.4.19).

6.3.8 A system of equations on V_ℓ written in coordinates with respect to this basis,

$$\mathsf{A}_\ell \mathsf{u}_\ell = \mathsf{f}_\ell, \tag{6.23}$$

is then split into

$$\begin{pmatrix} \mathsf{A}_\ell^{SS} & \mathsf{A}_\ell^{SL} \\ \mathsf{A}_\ell^{LS} & \mathsf{A}_\ell^{LL} \end{pmatrix} \begin{pmatrix} \mathsf{u}_\ell^S \\ \mathsf{u}_\ell^L \end{pmatrix} = \begin{pmatrix} \mathsf{f}_\ell^S \\ \mathsf{f}_\ell^L \end{pmatrix}. \tag{6.24}$$

6.3.9 Remark: If the matrix A in (6.23) was obtained by applying a bilinear form $a_\ell(u,v)$ on V_ℓ to the basis functions of V_ℓ, then the matrices A_ℓ^{SS} and A_ℓ^{LL} correspond to the restriction of $a_\ell(.,.)$ to the subspaces V_ℓ^S and V_ℓ^L, respectively.

6.3.10 Remark: The matrices A_ℓ^{SL} and A_ℓ^{LS} consist of boundary fluxes on the refinement edge $\mathbb{E}_\ell^E$ only. Since we expect the cells adjacent to these edges to be a small number compared to the whole mesh, these matrices have very few nonzero entries and can be stored very efficiently.

6.3.11 Definition: The restriction operator $\mathsf{R}_{\ell-1} : V_\ell \to V_{\ell-1}$ acts as L^2-projection $\mathsf{R}_{\ell-1}^S$ on V_ℓ^S and as identity on V_ℓ^L,

$$\mathsf{R}_{\ell-1}\mathsf{x}_\ell = \mathsf{R}_{\ell-1}^S \mathsf{x}_\ell^S + \mathsf{x}_\ell^L. \tag{6.25}$$

6.3.12 Algorithm: Let $\mathsf{S}_{\ell;S}^{(i)}$, $i = 1, \dots$ be a sequence of smoothers as in algorithm 6.2.9 on page 136, but acting on V_ℓ^S only. Then, the action of the *local variable V-cycle* operator P_ℓ on a vector d_ℓ is defined recursively by $\mathsf{P}_0 = \mathsf{A}_0^{-1}$ and ($\mathsf{x}^{(0)} = 0$):

1. (Pre-smoothing) Compute $\mathsf{x}_S^{(m_\ell)}$ iteratively by
$$\mathsf{x}_S^{(i)} = \mathsf{x}_S^{(i-1)} + \mathsf{S}_{\ell;S}^{(i)}(\mathsf{d}_\ell^S - \mathsf{A}_\ell^{SS}\mathsf{x}_S^{(i-1)}), \quad i = 1, \dots, m_\ell,$$
and let $\mathsf{x}^{(m_\ell)} = (\mathsf{x}_S^{(m_\ell)}, 0)^T$.
2. (Coarse grid correction) Let
$$\mathsf{y}^{(0)} = \mathsf{x}^{(m_\ell)} + \mathsf{R}_{\ell-1}^T\mathsf{P}_{\ell-1}\big(\mathsf{R}_{\ell-1}^S(\mathsf{d}_\ell^S - \mathsf{A}_\ell^{SS}\mathsf{x}_S^{(m_\ell)}) + d_\ell^L - \mathsf{A}_\ell^{LS}\mathsf{x}_S^{(m_\ell)}\big).$$
3. (Post-smoothing) Compute $\mathsf{y}^{(m_\ell)}$ iteratively by
$$\mathsf{y}_S^{(i)} = \mathsf{y}_S^{(i-1)} + \mathsf{S}_{\ell;S}^{(m_\ell+i)}(\mathsf{g}_\ell^S - \mathsf{A}_\ell^{SS}\mathsf{y}_S^{(i-1)}), \quad i = 1, \dots, m_\ell.$$
where $\mathsf{g}_\ell^S = d_\ell^S - \mathsf{A}_\ell^{SL}\mathsf{y}_L^{(0)}$
4. Set $\mathsf{P}_\ell\mathsf{d}_\ell = (\mathsf{y}_S^{(m_\ell)}, \mathsf{y}_L^{(0)})$.

6.3.13 Theorem: *Algorithm 6.3.12 is equivalent to algorithm 6.2.9 if the smoothers there are taken of the form*
$$\mathsf{S}_\ell^{(i)} = \begin{pmatrix} \mathsf{S}_{\ell;S}^{(i)} & 0 \\ 0 & 0 \end{pmatrix}.$$

Proof. First, due to the structure of the smoother and due to $x^{(0)} = 0$, we have $x_L^{(0)} = 0$ during the whole pre-smoothing iteration. Therefore,
$$\begin{aligned}\mathsf{S}_\ell^{(i)}(\mathsf{d}_\ell - \mathsf{A}_\ell x^{(i)}) &= \begin{pmatrix} \mathsf{S}_{\ell;S}^{(i)} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \mathsf{d}_\ell^S - \mathsf{A}_\ell^{SS}x_S^{(i)} \\ \mathsf{d}_\ell^L - \mathsf{A}_\ell^{LS}x_S^{(i)} \end{pmatrix} \\ &= \begin{pmatrix} \mathsf{S}_{\ell;S}^{(i)}(\mathsf{d}_\ell^S - \mathsf{A}_\ell^{SS}\mathsf{x}_S^{(i-1)}) \\ 0 \end{pmatrix}\end{aligned}$$
The form of the restricted residual in the coarse grid correction is an immediate consequence of the splitting of the restriction operator in (6.25). Finally, $y_L^{(i)}$ remains constant during post-smoothing and therefore
$$\begin{aligned}\mathsf{S}_\ell^{(m_\ell+i)}(\mathsf{d}_\ell - \mathsf{A}_\ell y^{(i)}) &= \begin{pmatrix} \mathsf{S}_{\ell;S}^{(m_\ell+i)} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \mathsf{d}_\ell^S - \mathsf{A}_\ell^{SS}y_S^{(i)} - \mathsf{A}_\ell^{SL}y_L^{(0)} \\ \mathsf{d}_\ell^L - \mathsf{A}_\ell^{LS}y_S^{(i)} - \mathsf{A}_\ell^{LL}y_L^{(0)} \end{pmatrix} \\ &= \begin{pmatrix} \mathsf{S}_{\ell;S}^{(m_\ell+i)}(\mathsf{g}_\ell^S - \mathsf{A}_\ell^{SS}\mathsf{y}_S^{(i-1)}) \\ 0 \end{pmatrix},\end{aligned}$$
which concludes the proof. □

6.3.14 Remark: For $\mathsf{S}_{\ell;S}^{(i)}$ any of the smoothing methods in Definition 6.2.8 applied to A_ℓ^{SS} may be used. Here, it is important that the matrix A_ℓ^{SS} used by the relaxation method is the same matrix used in the computation of residuals. Therefore, no additional matrix has to be assembled or stored.

6.3.15 Remark: The splitting of operators in the algorithm renders it possible, that shape functions of an active grid cell on a lower level are used *only* on this level, not on finer ones. In fact, the function corresponding to the vector d_ℓ on an active cell equals the function represented by d_L on the finest level, possibly modified by boundary contributions from neighboring refined cells. Furthermore, this means that matrices A_ℓ^{LL} are neither assembled, not used.

6.3.16 In the following paragraphs, we explain how this algorithm can be implemented using only the routines used for integrating the bilinear and linear forms on grid cells and edges. Since these are used for assembling the global system matrix as well, we assure that no additional work is required for the local multigrid method.

6.3.17 We assume that there are functions `CellMatrix(T)` and `EdgeFlux(E)`, computing the local contributions of the bilinear form $a_h(.,.)$ by some integration method. The result of the function `CellMatrix(T)` is a single matrix $\mathsf{C}^{(T)}$ of dimension $n_T \times n_T$. The result of `EdgeFlux(E)` is a set of four matrices $(\mathsf{E}^{(T_i,T_j)})_{i,j=1,2}$ of dimensions $n_{T_i} \times n_{T_j}$. Here, T_1 and T_2 are the two cells adjacent to E. For simplicity, we assume that this function can handle boundary fluxes as well; then, the parameter T_2 is void and only $\mathsf{E}^{(T_1,T_1)}$ is returned.

6.3.18 Algorithm: With the matrices defined in the previous paragraph, the system matrix A_h is assembled in a two-step summation process: start with $\mathsf{A} = 0$ (omitting the index h for obvious reasons).

1. For each (active) cell $T \in \mathbb{T}_h$ let

$$\mathsf{A}_{T,i;T,j} = \mathsf{A}_{T,i;T,j} + \mathsf{C}_{ij}^{(T)} \qquad i,j = 1 \ldots n_T.$$

2. For each interior edge $E \in \mathbb{E}_h^i$ let T_1 and T_2 be the adjacent cells and let

$$\mathsf{A}_{T_\mu,i;T_\nu,j} = \mathsf{A}_{T_\mu,i;T_\nu,j} + \mathsf{E}_{ij}^{(T_\mu,T_\nu)} \qquad \begin{matrix} i = 1 \ldots n_{T_\mu} \\ j = 1 \ldots n_{T_\nu} \end{matrix} \quad \mu,\nu = 1,2$$

   On each boundary edge do the same with $\mu,\nu = 1$ only.

6.3.19 Remark: Algorithm 6.3.18, which is the usual way of building matrices in finite element codes, applies to locally refined meshes with hanging nodes. In this case, the edges E in the second step must be the edges of the finer cells (see Figure 6.6). Here, the larger cell appears twice in the summation process, as T_2^a and T_2^b. Stabilization parameters must be chosen properly in this case; see the discretization chapters for details. This algorithm also applies to hp-adaptive discretizations, where the number of degrees of freedom varies between grid cells.

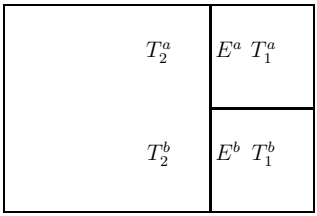

Figure 6.6: Assembling on locally refined meshes

6.3.20 Algorithm: A scheme similar to algorithm 6.3.18 is used to assemble the matrices A_ℓ^{SS}, A_ℓ^{LS} and A_ℓ^{SL}. Start with all matrices zero and,

1. for each cell T of the triangulation hierarchy $\{\mathbb{T}_\ell\}$, let ℓ_T be the level of the cell (see 6.3.3). Let
$$(\mathsf{A}_{\ell_T}^{SS})_{T,i;T,j} = (\mathsf{A}_{\ell_T}^{SS})_{T,i;T,j} + \mathsf{C}_{ij}^{(T)} \qquad i,j = 1 \dots n_T.$$
2. For each interior edge E of the hierarchy let T_1 and T_2 be the adjacent cells and ℓ_E the level of the edge. If there is no neighbor cell on level ℓ_E on one side, the cell on the coarser level is taken.
   (a) If $\ell_E = \ell_{T_1} = \ell_{T_2}$, let
   $$(\mathsf{A}_{\ell_E}^{SS})_{T_\mu,i;T_\nu,j} = (\mathsf{A}_{\ell_E}^{SS})_{T_\mu,i;T_\nu,j} + \mathsf{E}_{ij}^{(T_\mu,T_\nu)} \qquad \begin{matrix} i = 1 \dots n_{T_\mu} \\ j = 1 \dots n_{T_\nu} \end{matrix} \quad \mu,\nu = 1,2.$$
   The same is done on the boundary, where $\mu,\nu = 1$ only.
   (b) Otherwise, $E \in \mathbb{E}_{\ell_E}^E$ and one of the neighbors T_1 and T_2, say T_1 is on level $\ell_E - 1$. Then, let
   $$\begin{aligned}
   (\mathsf{A}_{\ell_E}^{SS})_{T_2,i;T_2,j} &= (\mathsf{A}_{\ell_E}^{SS})_{T_2,i;T_2,j} + \mathsf{E}_{ij}^{(T_2,T_2)} \\
   (\mathsf{A}_{\ell_E-1}^{SS})_{T_1,i;T_1,j} &= (\mathsf{A}_{\ell_E-1}^{SS})_{T_1,i;T_1,j} + \mathsf{E}_{ij}^{(T_1,T_1)} \\
   (\mathsf{A}_{\ell_E}^{LS})_{T_1,i;T_2,j} &= (\mathsf{A}_{\ell_E}^{LS})_{T_1,i;T_2,j} + \mathsf{E}_{ij}^{(T_1,T_2)} \\
   (\mathsf{A}_{\ell_E}^{SL})_{T_2,i;T_1,j} &= (\mathsf{A}_{\ell_E}^{SL})_{T_2,i;T_1,j} + \mathsf{E}_{ij}^{(T_2,T_1)}
   \end{aligned}$$

6.3.21 Remark: The algorithms for assembling the system matrix and for assembling the level matrices use the same cell matrices, which simplifies the coding for complex systems of equations considerably.

6.3.22 Remark: Since the matrix A_L^{LL} is never assembled and since the matrices A_ℓ^{SS} are not split into an active part and an overlapped part, it is impossible to use the level matrices for the multiplication with the system matrix A_h. The resulting disadvantage of doubling the memory requirements may be overcome by two remedies: first, low accuracy is sufficient for the level matrices, since they are used in a preconditioner. Second, assembling of the system matrix may be replaced by integrating the residual immediately; this approach may be even more time efficient inside a nonlinear iteration, where only few linear residuals with the same matrix are needed.

	$\mathbb{Q}_1, \kappa = 2$			$\mathbb{Q}_2, \kappa = 4$			$\mathbb{P}_1, \kappa = 2$		
L	n_{10}	$\overline{\varrho}$	$\overline{r}$	n_{10}	$\overline{\varrho}$	$\overline{r}$	n_{10}	$\overline{\varrho}$	$\overline{r}$
2	8	0.053	1.275	11	0.117	0.932	9	0.054	1.268
3	9	0.058	1.235	12	0.139	0.858	9	0.070	1.158
4	9	0.062	1.210	12	0.143	0.846	9	0.072	1.142
5	9	0.062	1.208	12	0.144	0.843	9	0.071	1.150
6	9	0.062	1.209	12	0.143	0.843	9	0.070	1.155
7	9	0.062	1.211	12	0.143	0.844	9	0.069	1.162
8	9	0.061	1.213	12	0.143	0.845	9	0.068	1.169

Table 6.9: Contraction and convergence rates on locally refined grids

6.3.23 In Table 6.9 we show that the preconditioner is well suited for locally refined meshes and local smoothing. The contraction numbers displayed in are exactly the same as in tables 6.5 and 6.6. Results were obtained with the meshes in Figure A.3.

6.4 LDG

6.4.1 In this section, we will continue the investigations of Section 6.2 and develop preconditioning methods for the LDG method. This will be achieved in two steps. First, a preconditioner for the Schur complement of the mixed linear system is derived. In a second step, a block preconditioner is constructed for the complete system.

6.4.2 Using bases according to Definition 1.1.11 on page 17 for the spaces Σ_h and V_h, the discrete system (3.10) becomes a linear system of equations of the type

$$\mathfrak{A}\mathfrak{u} = \mathfrak{f}, \tag{6.26}$$

with

$$\mathfrak{A} = \begin{pmatrix} \mathsf{M}_\nu & -\mathsf{B}^T \\ \mathsf{B} & \mathsf{C} \end{pmatrix}, \quad \mathfrak{u} = \begin{pmatrix} \sigma \\ \mathsf{u} \end{pmatrix}, \quad \mathfrak{f} = \begin{pmatrix} \mathsf{f}_\sigma \\ \mathsf{f}_u \end{pmatrix}. \tag{6.27}$$

Here, like in previous chapter, a matrix entry at location $T, i; S, j$ is computed by applying the corresponding bilinear form to basis functions with these indices. For instance, let $\varphi_{T,i}$ be a basis function of V_h and $\psi_{S,j}$ a (vector-valued) basis function of Σ_h, then

$$B_{T,i;S,j} = b_h(\varphi_{T,i}, \psi_{S,j}).$$

6.4.3 Definition: Associated with the matrix $\mathfrak{A}$ is its *Schur complement* S, which is the matrix

$$\mathsf{S} = \mathsf{B}\mathsf{M}_\nu^{-1}\mathsf{B}^T + \mathsf{C}, \tag{6.28}$$

which is generated from $\mathfrak{A}$ by block-elimination of the variable σ.

6.4.4 Remark: By definition, S is symmetric if M_ν and C are symmetric, which is the case for both version of the LDG method. Furthermore, (σ, u) is a solution of (6.26) if and only if u is a solution of

$$\mathsf{S}\mathsf{u} = \mathsf{f}_S \qquad \mathsf{f}_S = \mathsf{f}_\sigma - \mathsf{B}\mathsf{M}_\nu^{-1}\mathsf{f}_u, \tag{6.29}$$

and $\sigma = -\mathsf{M}_\nu^{-1}\mathsf{B}^T\mathsf{u}$. Finally, since M_ν is positive definite, C positive semi-definite and the system is uniquely solvable (Lemma 3.1.6), the matrix S is positive definite.

6.4.5 The bilinear form $s_h(.,.)$ associated with the matrix S corresponds to a primal formulation of the LDG method. This primal formulation is described and analyzed in [ABCM00, ABCM01]. We will use the equivalence of $s_h(.,.)$ and $a_h^{\mathrm{IP}}(.,.)$ stated in Corollary 3.2.13 to construct efficient preconditioners for S.

6.4.6 Remark: The matrix M_ν consists of d diagonal blocks which are both scaled versions of the mass matrix. Since the discretization is discontinuous, these blocks decouple further into smaller blocks consisting of scaled cell mass matrices for each mesh cell. If we chose orthogonal polynomials on T (with respect to the weight $1/\nu$), M_ν would be a diagonal matrix (see [Cas01]).

Therefore, evaluation of the Schur complement is particularly simple; M_ν can be inverted in a cheap way before the iterative solution and multiplication with M_ν^{-1} becomes about as expensive as a multiplication with C.

6.4.7 Definition: Two symmetric, positive definite bilinear forms $a(.,.)$ and $b(.,.)$ on the space V_h are called *spectrally equivalent*, if the estimate

$$c_1 a(v,v) \leqslant b(v,v) \leqslant c_2 a(v,v), \tag{6.30}$$

holds with constants c_1 and c_2 for any $v \in V_h$.

6.4.8 Lemma: *Let $a_h(.,.)$ and $s_h(.,.)$ be two symmetric, positive definite and spectrally equivalent bilinear forms with associated matrices A and S (both with respect to the same basis). Let furthermore P be a uniform preconditioner for A. Then, P is also a uniform preconditioner for S.*

Proof. If P is a uniform preconditioner for A, then there are constants c_3 and c_4 such that for any $\mathsf{v} \in \mathbb{R}^n$ ($n = \dim V$)

$$c_3 \mathsf{v}^T\mathsf{A}\mathsf{v} \leqslant \mathsf{v}^T\mathsf{A}\mathsf{P}\mathsf{A}\mathsf{v} \leqslant c_4 \mathsf{v}^T\mathsf{A}\mathsf{v},$$

or equivalently,

$$c_3 \mathsf{v}^T\mathsf{A}^{-1}\mathsf{v} \leqslant \mathsf{v}^T\mathsf{P}\mathsf{v} \leqslant c_4 \mathsf{v}^T\mathsf{A}^{-1}\mathsf{v}.$$

Since from (6.30) follows

$$c_1 \mathsf{v}^T\mathsf{S}^{-1}\mathsf{v} \leqslant \mathsf{v}^T\mathsf{A}^{-1}\mathsf{v} \leqslant c_2 \mathsf{v}^T\mathsf{S}^{-1}\mathsf{v},$$

the lemma follows by combining the last two estimates. □

	$\mathbb{Q}_1$		$\mathbb{Q}_2$		$\mathbb{Q}_3$		$\mathbb{Q}_4$	
L	n_{10}	$\bar{r}$	n_{10}	$\bar{r}$	n_{10}	$\bar{r}$	n_{10}	$\bar{r}$
1	10	1.363	12	.864	13	.794	16	.648
2	14	0.763	14	.758	15	.702	17	.606
3	14	0.715	14	.752	16	.653	17	.592
4	15	0.708	14	.743	16	.645	17	.592
5	15	0.701	14	.742	16	.643	17	.591
6	15	0.701	14	.741	16	.642	17	.592
7	15	0.701	14	.741	16	.642		

Table 6.10: Conjugate gradient convergence rates for standard LDG with preconditioner $\breve{\mathsf{P}}$.

	γ_{uu}	κ
$\mathbb{Q}_1$	2	4
$\mathbb{Q}_2$	4	9
$\mathbb{Q}_3$	4	13
$\mathbb{Q}_4$	4	13

Table 6.11: Values of stabilization parameters yielding optimal convergence

6.4.9 Lemma: *The multilevel preconditioner $\breve{\mathsf{P}}_L$ as defined in Algorithm 6.2.9 on page 136 is a uniform preconditioner for the LDG Schur complement matrix S_L on level L.*

Proof. This result is an immediate consequence of Lemma 6.4.8 the equivalence (3.31) of the bilinear forms $a_h^{\mathrm{IP}}(.,.)$ and $s_h(.,.)$. □

6.4.10 The following numerical results give evidence to the fact that the preconditioner $\breve{\mathsf{P}}_L$ for the matrix S_L is not only uniform , but that the resulting condition numbers are in fact sufficiently small to recommend this method for practical computation.

6.4.11 Like in the previous chapter, we begin by investigating the convergence properties on Cartesian grids (see A.1.1 on page 167), $\nu = 1$ and the right hand side chosen according to example A.2.2 on page 170. Table 6.10 shows convergence rates of the cg method for the LDG Schur complement S_L using the preconditioner $\breve{\mathsf{P}}_L$ for different polynomial degrees on Cartesian grids. As for the interior penalty method, the convergence rate is independent of the mesh level as soon as h is sufficiently small. Furthermore, it deteriorates only slowly with the polynomial degree. The choice of parameters can be found in Table 6.11. These parameters were found optimal by experiment. A theoretical justification for their unexpected dependence on the polynomial degree is still lacking.

6.4.12 The preconditioner $\breve{\mathsf{P}}$ has two free parameters, namely $\widehat{\gamma_{uu}}$ and κ. In paragraph 3.2.10 on page 75, we observed that $\widehat{\gamma_{uu}}$ does not influence the discretization accuracy over a wide range of values. Figure 6.7 suggests, that its influence on the preconditioner is much more important. The graphs suggest that $\widehat{\gamma_{uu}}$ should be chosen close to the optimal value to achieve good convergence. Still, the peak of the convergence rates is quite flat and a factor of 4 changes the convergence rates by less than 25%. Furthermore, the optimal value is larger than the point

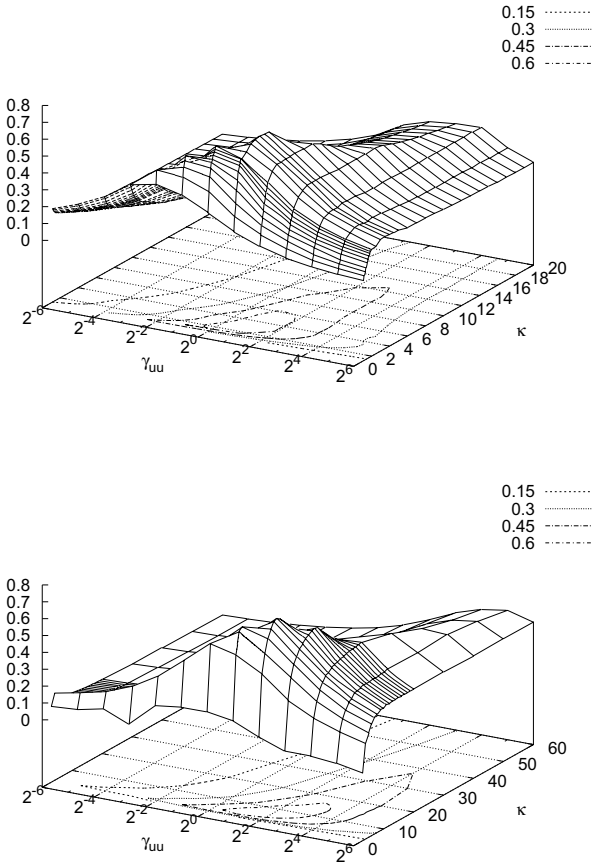

Figure 6.7: Multigrid convergence rates for LDG with IP preconditioning, $\mathbb{Q}_1$ (left) and $\mathbb{Q}_2$ (right) elements

L	$\breve{\mathsf{P}}_L A_L$	$\mathsf{P}_L A_L$
1	12	6
2	27	14
3	39	20
4	59	29
5	82	50
6	114	85
7	155	143

Table 6.12: Iteration counts n_{10} for superconvergent LDG

	$\mathbb{Q}_1$		$\mathbb{Q}_2$		$\mathbb{Q}_3$	
L	n_{10}	$\bar{r}$	n_{10}	$\bar{r}$	n_{10}	$\bar{r}$
1	17	.874	28	.548	38	.408
2	31	.497	38	.403	46	.333
3	37	.418	41	.376	50	.309
4	38	.402	43	.363	52	.301
5	39	.394	44	.358	53	.299
6	40	.392	45	.357	54	.298
7	40	.393	45	.357	54	.298

Table 6.13: Performance of cg method for modified superconvergent LDG

where the accuracy of the discretization levels off in Figure 3.1 on page 76. Similarly, the convergence rates do not deteriorate fast if κ is chosen larger than the optimal value.

Preconditioning the superconvergent scheme

6.4.13 We apply the same preconditioning techniques to the superconvergent method, even if (3.31) is not proven for this case. The results with $\widehat{\gamma_{uu}} = 2$ are in Table 6.12. They exhibit a growth of the iteration counts with about $\sqrt{h_L}$ with the interior penalty preconditioner $\breve{\mathsf{P}}$ ($\kappa = 3$). We compare this preconditioner with the genuine LDG preconditioner P, where the smoothing is performed on the matrix S itself and conclude that both solvers are not satisfactory.

6.4.14 The performance of the superconvergent LDG method can be improved by using the flux (3.19) on the Dirichlet boundary (see also Remark 3.3.14). Indeed, Table 6.13 shows that the multilevel interior penalty scheme is a uniform preconditioner for the resulting method.

Solving the saddle-point system

6.4.15 Preconditioners for (stabilized) saddle-point systems can be derived using the following result due to [MGW00]:

$$\begin{pmatrix} \mathsf{A} & 0 \\ \mathsf{B} & \mathsf{S} \end{pmatrix}^{-1} \begin{pmatrix} \mathsf{A} & -\mathsf{B}^T \\ \mathsf{B} & \mathsf{C} \end{pmatrix} = \begin{pmatrix} \mathsf{I} & -\mathsf{A}^{-1}\mathsf{B}^T \\ 0 & \mathsf{I} \end{pmatrix}.$$

If we could compute the inverse on the left, the spectrum of the matrix would consist of the single value 1, and Krylov-space methods would converge in very few steps. Unfortunately, this includes the exact inversion of S, which we have to replace by a reasonable approximate, the uniform preconditioner P, either $\breve{\mathsf{P}}_L$ or P_L. Thus, we define the preconditioner for the saddle-point system

$$\mathfrak{P} := \begin{pmatrix} \mathsf{A} & 0 \\ \mathsf{B} & \gamma_u^{-1}\mathsf{P}^{-1} \end{pmatrix}, \tag{6.31}$$

where we introduced the additional tuning parameter $\gamma_u > 0$.

6.4.16 Lemma: *The matrix $\mathfrak{A}\mathfrak{P}^{-1}$ has only real eigenvalues and its spectrum is*

$$\sigma(\mathfrak{A}\mathfrak{P}^{-1}) = \sigma(\gamma_u \breve{\mathsf{P}} S) \cup \{1\}. \tag{6.32}$$

Proof. (from [Kan03b]) The proof is based on the proof to [Kla98, Lemma 3.1]. Instead of computing the spectrum of $\mathfrak{A}\mathfrak{P}^{-1}$, we consider the transpose system and solve the generalized eigenvalue problem

$$\mathfrak{A}^T x = \lambda \mathfrak{P}^T x,$$

or, splitting into components,

$$\begin{aligned} \mathsf{A}\sigma \ + \mathsf{B}^T\mathsf{u} &= \lambda\mathsf{A}\sigma + \ \ \lambda\mathsf{B}^T\mathsf{u}, \\ -\mathsf{B}\sigma + \ \mathsf{C}\mathsf{u} \ &= \qquad \lambda\gamma_u^{-1}\breve{\mathsf{P}}_L^{-1}\mathsf{u}. \end{aligned} \tag{6.33}$$

If $\lambda = 1$, the first equation is void. In this case, a non-trivial solution to

$$-\mathsf{B}\sigma + \underbrace{(\mathsf{C} - \gamma_u^{-1}\breve{\mathsf{P}}_L^{-1})}_{=:\hat{\mathsf{C}}}\mathsf{u} = 0,$$

is an eigenvector to the eigenvalue 1. If $\hat{\mathsf{C}}$ is singular, choose a non-zero pair $(\sigma, \mathsf{u}) \in \ker B \times \ker \hat{\mathsf{C}}$ arbitrarily. If $\hat{\mathsf{C}}$ is regular, choose $\sigma \neq 0$ arbitrarily and $\mathsf{u} := \hat{\mathsf{C}}^{-1}\mathsf{B}\sigma$. Therefore, $\lambda = 1$ is eigenvalue.

For $\lambda \neq 1$, we compute the Schur complement of (6.33), namely,

$$\mathsf{S}\mathsf{u} = (\mathsf{B}\mathsf{A}^{-1}\mathsf{B}^T + \mathsf{C})\mathsf{u} = \lambda\gamma_u^{-1}\breve{\mathsf{P}}_L^{-1}\mathsf{u}.$$

Multiplying with $\gamma_u\breve{\mathsf{P}}_L$ from the left, we see that λ is an eigenvalue of $\gamma_u\breve{\mathsf{P}}_L S$ and u its associated eigenvector. Finally, we remark that the eigenvalues of $\mathfrak{A}\mathfrak{P}^{-1}$ are equal to those of $\mathfrak{P}^{-T}\mathfrak{A}^T$. □

6.4.17 Remark: Since the matrix $\mathfrak{A}\mathfrak{P}^{-1}$ is not normal, convergence results for Krylov space methods cannot be applied solely based on the previous lemma. In [Kla98], normality of the matrix is established with respect to another symmetric and positive definite matrix. Unfortunately, the theoretical results yield a number of iteration steps growing like $\mathcal{O}(2^L)$, a growth not observed in the numerical computations below.

	Schur		system	
L	n_{10}	$\bar{r}$	n_{10}	$\bar{r}$
1	10	1.438	11	1.226
2	14	0.768	15	0.717
3	14	0.721	15	0.670
4	15	0.710	16	0.665
5	15	0.703	16	0.660
6	15	0.703	16	0.660
7	15	0.702	16	0.660

Table 6.14: GMRES performance for the preconditioned LDG Schur complement and the preconditioned system ($\mathbb{Q}_1$-elements)

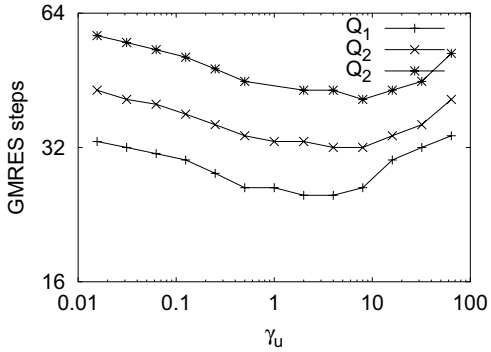

Figure 6.8: GMRES steps for the standard LDG system depending on the scaling factor γ_u.

6.4.18 Since the preconditioned matrix $\mathfrak{A}\mathfrak{P}^{-1}$ is non-symmetric, we apply GMRES iteration and compare to GMRES applied to the preconditioned Schur complement in Table 6.14. The setup of the test calculations is the same as in the previous sections. The table shows that the GMRES convergence rates on the Schur complement are exactly as the corresponding values for the cg iteration in Table 6.10 on page 153. Applied to the saddle point system, the GMRES method needs exactly one more iteration than for the Schur complement, possibly due to the additional eigenvalue according to Lemma 6.4.16.

We conclude that the GMRES method on the saddle point system is an efficient alternative if computation of the Schur complement is not feasible. Still, we remark that this approach suffers from vector lengths being 3 times (4 times in three dimensions) as large as for the Schur complement.

6.4.19 The tuning parameter γ_u serves to shift the spectrum of $\breve{\mathsf{P}}_L S$ with respect to the additional eigenvalue 1 of the system. Figure 6.8 exhibits that the convergence speed of the preconditioned GMRES method is rather insensitive to γ_u within a wide range. This is expected from Lemma 6.4.16, since the single eigenvalue 1 does not influence the convergence of Krylov space methods critically, even if isolated. Therefore, we will choose

$$\gamma_u = 1,$$

	ν										
L	4^0	4^{-1}	4^{-2}	4^{-3}	4^{-4}	4^{-5}	4^{-6}	4^{-7}	4^{-8}	4^{-9}	4^{-10}
2	1.6	1.5	1.6	2.3	3.2	4.2	5.4	6.6	7.8	9.0	10.1
3	1.4	1.4	1.4	1.7	2.3	3.2	4.4	5.6	6.8	8.0	9.2
4	1.3	1.3	1.3	1.5	1.7	2.4	3.4	4.5	5.9	7.1	8.3
5	1.2	1.3	1.3	1.3	1.5	1.6	2.4	3.5	4.7	6.1	7.3
6	1.2	1.2	1.3	1.3	1.5	1.3	1.6	2.5	3.7	5.2	6.3
7	1.2	1.2	1.3	1.3	1.3	1.5	1.2	1.7	2.6	3.8	5.4
8	1.2	1.2	1.2	1.3	1.3	1.5	1.3	1.1	1.7	2.7	4.0

Table 6.15: Convergence rates for GMRES with multilevel preconditioner using downwind block-Gauß-Seidel smoothing

in all further computations.

6.5 Advection Diffusion

6.5.1 Assume that the advection field b in equation (5.10) does not contain loops (see paragraph 6.5.5 below for the other case). In paragraph 1.4.19, we remarked that the mesh cells can be ordered in such a way that the matrix associated with the form $\beta_h(.,.)$ is upper block-triangular and the diagonal blocks correspond to the cell matrices. Therefore, the block-Gauß-Seidel smoother defined in 6.2.8 is a direct solver.

6.5.2 Based on these observations, we construct a multilevel preconditioner for the discretized advection-diffusion equation (1.50) using algorithm 6.2.9 with the modification that we do *not* use transposed smoothing operators. We also will allow different numbers of pre and post smoothing steps.

6.5.3 We claim that the preconditioner constructed this way is robust with respect to the diffusion parameter, that is, a GMRES iteration applied to the preconditioned matrix converges

- in one step, if the diffusion coefficient ν vanishes,
- like in the self-adjoint case, if $\nu \gg 1$ and
- the transition between the two extremal states is monotonous.

Indeed, Table 6.15 shows this behavior. The problem solved was example A.3.2 with $\mathbb{Q}_1$-elements. On the left side of the table, we see the same convergence rates as in Table 6.5 on page 142. They correspond to gaining more than one decimal in each iteration. To the right, these rates increase monotonously until up to 10 decimals are gained in a single step. This is the case where the smoother is nearly an exact solver. The convergence rates remain constant on the diagonals with $\nu/h^2 = c$.

	ν										
L	4^0	4^{-1}	4^{-2}	4^{-3}	4^{-4}	4^{-5}	4^{-6}	4^{-7}	4^{-8}	4^{-9}	4^{-10}
2	1.6	1.5	1.3	1.2	1.1	1.1	1.2	1.2	1.1	1.1	1.1
3	1.4	1.3	1.1	0.9	0.8	0.7	0.7	0.7	0.7	0.7	0.7
4	1.3	1.3	1.2	0.9	0.6	0.5	0.4	0.4	0.4	0.4	0.4
5	1.3	1.3	1.2	1.0	0.6	0.3	0.2	0.2	0.2	0.2	0.2
6	1.3	1.3	1.2	1.1	0.8	0.4	0.2	0.2	0.1	0.1	0.1
7	1.3	1.3	1.3	1.2	1.0	0.6	0.2	0.1	0.1	0.1	0.1
8	1.3	1.3	1.3	1.2	1.1	0.8	0.4	0.2	0.1	0.1	0.1

Table 6.16: Performance of GMRES with multilevel preconditioner using upwind block-Gauß-Seidel smoothing

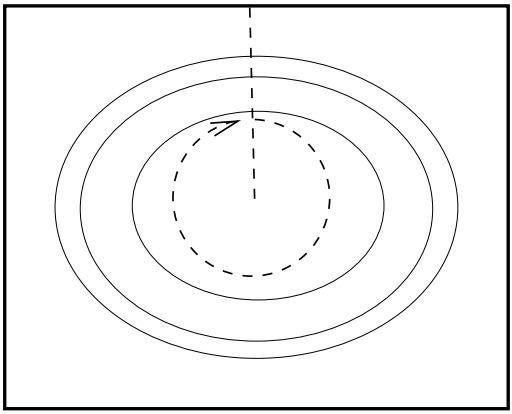

Figure 6.9: Numbering in case of a vortex

6.5.4 Since finding a downwind ordering is not a simple task in general, we have to check if it is really important. In Table 6.16, we solve the same problem as in Table 6.15, but this time sorting the degrees of freedom in opposite direction. Here, the convergence rates drop as ν tends to zero, reflecting a gain of accuracy of less than a decimal in 10 steps in the last column. Therefore, we conclude that at least a reasonable numbering algorithm is crucial for the performance of the preconditioner.

6.5.5 If the flow field has closed integral curves somewhere in the domain, a downwind numbering cannot be found anymore. In this case, a numbering scheme described in [BW97, HP97] can be employed. Assume that the vector field contains a vortex like in Figure 6.9. Then, this vortex is cut from its center to its boundary more or less perpendicular to the vector field. Afterwards, mesh cells are numbered starting at this cut-line in downwind direction until the cut-line is reached from the opposite side.

6.5.6 In Table 6.17, we show the effect of such a numbering in the case of the simple rotating flow field of example A.3.3. Since the matrix is not block-triangular, we do not expect solution

	ν										
L	4^0	4^{-1}	4^{-2}	4^{-3}	4^{-4}	4^{-5}	4^{-6}	4^{-7}	4^{-8}	4^{-9}	4^{-10}
2	1.4	1.5	1.4	1.7	1.7	1.9	2.2	2.4	2.8	3.1	2.2
3	1.3	1.4	1.3	1.3	1.4	1.4	1.4	1.4	1.4	1.4	1.3
4	1.3	1.3	1.3	1.3	1.1	1.0	1.0	1.0	1.0	1.1	1.0
5	1.3	1.3	1.3	1.3	1.2	1.0	0.7	0.6	0.6	0.6	0.6
6	1.3	1.3	1.3	1.3	1.2	1.0	0.8	0.5	0.4	0.4	0.3
7	1.3	1.3	1.3	1.3	1.3	1.1	0.9	0.6	0.3	0.2	—
8	1.3	1.3	1.3	1.3	1.3	1.3	1.0	0.7	0.4	0.2	0.1

Table 6.17: Convergence rates for GMRES with multilevel preconditioner using downwind block-Gauß-Seidel smoothing in a vortex

in a single step if $\nu \to 0$. Even more, we know from the solution obtained by the method of characteristics that the equation looses stability if $\nu = 0$ (see Remark 1.4.18). Therefore, we observe that the convergence rates on finer grids break down dramatically if $\nu < 4^{-6} \approx 2.4 \cdot 10^{-4}$. As long as the diffusion coefficient stays above this threshold, the performance of the preconditioner depends only very weakly on the diffusion coefficient. Therefore, we conclude that the proposed method is robust with respect to dominant advection in a wide range.

6.6 Stokes

6.6.1 Following paragraph 6.4.2, we derive the matrix formulation of the Stokes system 4.16 as

$$\mathfrak{A}\mathfrak{u} = \mathfrak{f}, \tag{6.34}$$

with

$$\mathfrak{A} = \begin{pmatrix} \mathbf{M}_\nu & -\mathbf{B}^T & \\ \mathbf{B} & \mathbf{C} & -\mathsf{D}^T \\ & \mathsf{D} & E \end{pmatrix}, \qquad \mathfrak{u} = \begin{pmatrix} \boldsymbol{\sigma} \\ \mathbf{u} \\ \mathsf{p} \end{pmatrix}, \qquad \mathfrak{f} = \begin{pmatrix} \mathbf{f}_\sigma \\ \mathbf{f}_u \\ \mathsf{f}_p \end{pmatrix}. \tag{6.35}$$

The matrices $\mathbf{M}_\nu$, $\mathbf{B}$ and $\mathbf{C}$ are block diagonal multiples of the LDG matrices M_ν, B and C, respectively in the sense that e.g. $\mathbf{B} = \operatorname{diag}(\mathsf{B}, \dots, \mathsf{B})$ with d diagonal blocks.

6.6.2 Definition: The *Schur complement* $\mathcal{S}$ of the matrix $\mathfrak{A}$ is

$$\mathcal{S} = \mathsf{E} + \mathsf{D}\mathbf{S}^{-1}\mathsf{D}^T, \tag{6.36}$$

where $\mathbf{S}$ is the block diagonal matrix $\operatorname{diag}(\mathsf{S}, \dots, \mathsf{S})$ with d diagonal blocks and S the LDG Schur complement defined in 6.4.3.

6.6.3 Lemma: *If the coefficients of the method are chosen as in 4.1.9, the condition number of $\mathcal{S}$ is bounded independent of the viscosity ν and the mesh size h.*

Proof. Following [Bec95], we estimate the minimal eigenvalue by the Rayleigh quotient for the Schur complement, namely

$$\begin{aligned}\lambda_{\min} &= \inf_{\mathsf{p}} \frac{\mathsf{p}^T\mathcal{S}\mathsf{p}}{\mathsf{p}^T\mathsf{M}\mathsf{p}} \\ &= \inf_{\mathsf{p}} \frac{\mathsf{p}^T\mathsf{D}\mathbf{S}^{-1}\mathsf{D}^T\mathsf{p} + \mathsf{p}^T\mathsf{E}\mathsf{p}}{\mathsf{p}^T\mathsf{M}\mathsf{p}} \\ &= \inf_{\mathsf{p}} \frac{\mathsf{p}^T\mathsf{D}\mathbf{S}^{-T}\mathbf{S}\mathbf{S}^{-1}\mathsf{D}^T\mathsf{p} + \mathsf{p}^T\mathsf{E}\mathsf{p}}{\mathsf{p}^T\mathsf{M}\mathsf{p}} \\ &= \inf_{\mathsf{p}} \sup_{\mathsf{u}\in\mathsf{B}_p} \frac{\mathsf{u}^T\mathsf{D}^T\mathsf{p} + \mathsf{p}^T\mathsf{E}\mathsf{p}}{\mathsf{p}^T\mathsf{M}\mathsf{p}} \\ &= \inf_{p\in Q_h} \sup_{u\in B_p} \frac{d_h(u,p) + e_h(p,p)}{\|p\|^2}.\end{aligned}$$

due to the symmetry of $\mathbf{S}$. Here, $\mathsf{B}_p = \{\mathsf{u} | \mathsf{u}^T\mathbf{S}\mathsf{u} = \mathsf{p}^T\mathsf{M}\mathsf{p}\}$ and B_p is the image of B_p in V_h. Since $\mathsf{u}^T\mathbf{S}\mathsf{u} \geqslant |(0,u)|_{\mathrm{LDG}}$, where u is the image of u in V_h, we have by the inf-sup condition (4.37)

$$\lambda_{\min} \geqslant \frac{c}{\nu}.$$

For the maximal eigenvalue, we follow the same path

$$\begin{aligned}\lambda_{\max} &= \sup_{\mathsf{p}} \frac{\mathsf{p}^T\mathcal{S}\mathsf{p}}{\mathsf{p}^T\mathsf{M}\mathsf{p}} \\ &= \sup_{p\in Q_h} \sup_{u\in B_p} \frac{d_h(u,p) + e_h(p,p)}{\|p\|^2}.\end{aligned}$$

By Friedrichs inequality,

$$e_h(p,p) \leqslant c\|\nu^{-\frac{1}{2}}p\|^2.$$

Furthermore by the equivalence (3.31),

$$d_h(u,p) \leqslant c\nu^{-1}\|u\|_{1;h}\|p\| \leqslant c\nu^{-1}\|p\|^2,$$

yielding

$$\lambda_{\max} \leqslant \frac{c}{\nu},$$

which concludes the proof, since $\lambda_{\max}/\lambda_{\min}$ does not depend on h and ν. □

6.6.4 Corollary: *An accurate solution of the Schur complement equation of* (6.34) *can be achieved by the conjugate gradient method preconditioned with the mass matrix in a number of iteration steps bounded independent of h and ν.*

6.6.5 Following the same path as in paragraph 6.4.15, we construct a block preconditioner for the Stokes saddle-point system using the preconditioner $\check{\mathsf{P}}$ for the LDG Schur complement and the mass matrix for the Stokes Schur complement

$$\mathfrak{P} = \begin{pmatrix} \mathbf{M}_\nu & & \\ \mathbf{B} & \check{\mathbf{P}}^{-1} & \\ & \mathsf{D} & \mathsf{M} \end{pmatrix} \tag{6.37}$$

	$\mathbb{Q}_1$		$\mathbb{Q}_2$	
L	n_{10}	$\bar{r}$	n_{10}	$\bar{r}$
1	27	.379	29	.350
2	31	.330	33	.305
3	34	.296	37	.271
4	35	.288	37	.274
5	35	.288	37	.271
6	35	.287	38	.270
7	36	.283	38	.269

Table 6.18: GMRES performance for the Stokes system

6.6.6 Since the accurate inversion of the Laplacian in each step of the outer iterative scheme is very time-consuming, we improve the efficiency of the Stokes preconditioner replacing it by the multi-level preconditioner $\tilde{\mathsf{P}}$ in each velocity component. The performance of this preconditioner can be seen in Table 6.18. We see that the number of iteration steps grows by about 50% compared to the "exact" preconditioner, but replacing about 20 inner iteration steps for solving the Poisson problem in each component by a single one.

Comparing these results with the previous chapter, we conclude that the solution of the Stokes problem requires less than twice as many iteration steps than the Poisson equation with the same amount of work per step and component. Therefore, it can be considered very efficient.

6.6.7 The condition number of the Schur complement and therefore the convergence rate of the saddle-point solver naturally depends on $\widehat{\gamma_{uu}}$ and $\widehat{\gamma_{uu}}$. In Figure 6.10, we show this dependence for solving the Poisseuille flow example A.4.1 with $\mathbb{Q}_1$ and $\mathbb{Q}_2$ elements. Displayed are isolines close to the optimal convergence rates $c = \bar{r}_{\max}$ as well as $c/2$, $c/4$ and $c/8$. We see a dependence on $\widehat{\gamma_{uu}}$ similar to the LDG method for Poisson's equation; the area with convergence rates better than $1/2$ of the optimum stretches by a factor of about 2^6. Remarkably, $\widehat{\gamma_{pp}}$ has a strong influence on the solver, even if it does not affect solution accuracy (see paragraph 4.1.17 on page 98). Here, convergence rates above $1/2$ of the optimum are obtained only in a span of 2^4.

6.7 Oseen equations

6.7.1 As soon as advection is added to the Stokes operator, the performance of the linear solver becomes depending on the viscosity ν in a crucial way. It turns out that the preconditioner developed in the previous section does not yield satisfactory results.

6.7.2 Analogous to equation (6.35), the matrix of the LDG discretization of Oseen equations has the form

$$\mathfrak{A} = \begin{pmatrix} \mathsf{M}_\nu & -\mathsf{B}^T & \\ \mathsf{B} & \mathsf{C} & -\mathsf{D}^T \\ & \mathsf{D} & \mathsf{E} \end{pmatrix},$$

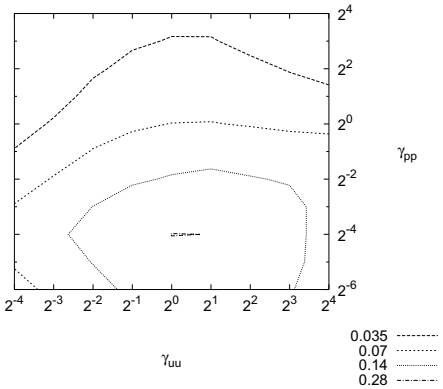

Figure 6.10: GMRES convergence rates depending on $\widehat{\gamma_{uu}}$ and $\widehat{\gamma_{pp}}$ (iso-lines). Poisseuille flow with $\mathbb{Q}_1$ elements

where now the matrix C consists of the jump terms of the Stokes operator (4.19) and the advection form $\beta_h(.,.)$ (cf. equation (5.20)). A preconditioner of the form

$$\mathfrak{P} = \begin{pmatrix} \mathsf{M}_\nu & & \\ \mathsf{B} & \check{\mathsf{P}}^{-1} & \\ & \mathsf{D} & \mathsf{P}_O \end{pmatrix}, \tag{6.38}$$

is constructed analogous to (6.37). Here, an efficient preconditioner for the Oseen Schur complement P_O must be found.

6.7.3 One scheme, suggested by Kay and Loghin (cf. [KLW02]), replaces the mass matrix as a preconditioner for the Schur complement by an operator more adapted to the advection dominated case. Assume that the "commutation" result

$$\mathsf{D}\mathsf{S}^{-1}\mathsf{D}^T \approx \mathsf{S}^{-1}\mathsf{D}\mathsf{D}^T$$

holds. Since D is the matrix associated with the DG divergence operator, $\mathsf{L} = \mathsf{D}\mathsf{D}^T$ is a matrix for a discrete Laplacian on the pressure space Q_h. Taking for instance the Stokes case ($b = 0$), we obtain ν^{-1} times identity. Since the mass matrix was shown to be a good preconditioner for the lower right block, we arrive at the following operator for preconditioning the Oseen Schur complement:

$$\mathsf{M}\mathsf{S}_p\mathsf{L},$$

where S_p is the matrix of a discretization of the operator $\nu\Delta + b \cdot \nabla$ on the space Q_h.

	ν						
L	10^{-1}	$10^{-1.5}$	10^{-2}	$10^{-2.5}$	10^{-3}	$10^{-3.5}$	10^{-4}
3	31	33	44	72	119	228	377
4	35	36	39	65	137	361	1183
5	38	39	39	49	110	290	945
6	40	45	42	43	71	217	774
7	41	51	48	45	50	106	397

Table 6.19: Convergence of exact Kay/Loghin preconditioner for two dimensional Poisseuille flow with $\mathbb{Q}_1$ elements

6.7.4 Remark: In [KLW02] it was shown that the commutation result postulated above holds indeed in the case of a constant advection field and if $\Omega = \mathbb{R}^d$.

6.7.5 We choose interior penalty discretizations to obtain the matrices S_p and L for simplicity (an LDG discretization would require additional variables for the pressure gradient just for preconditioning). Experiments showed that the method converged best with a stabilization parameter slightly larger than for the LDG preconditioner.

6.7.6 In Table 6.19, we present results for the Poisseuille flow problem (example A.4.1) on Cartesian grids. Following the results of Section 6.6, we chose $\widehat{\gamma_{uu}} = 2$ and $\widehat{\gamma_{pp}} = 1/16$ with $\mathbb{Q}_1$-elements. The table shows that on fine grids there is only a moderate increase of iteration steps up to a Reynolds number of 1000. In Particular, these numbers are not much higher than for the pure Stokes case with mass matrix preconditioning (see Table 6.18 on page 162). Then, for higher Reynolds numbers, the number of steps grows fast. Since this growth starts with higher Reynolds numbers on finer grids, it seems that a certain resolution of the possible boundary layers is required to make the method work. Here, the discontinuous Galerkin schemes may face a disadvantage compared to continuous methods, where a considerable, mesh dependent amount of artificial diffusion is induced by stabilization of the advection term.

6.7.7 In the previous paragraph, we considered the ideal but not feasible case of an exact solution of the Poisson and advection-diffusion problems arising in the block preconditioner (6.38). In fact, exact solution means to an accuracy of two orders of magnitude more accurate than the desired solution of the Oseen problem in *each* step unless a variant of GMRES for flexible preconditioning is used. Therefore, we replace the exact inversion by one or two V-cycles. The figures in the table show that the number of iteration steps is not much higher than for the exact solver using a single V-cycle only. In particular, the dependence on the viscosity ν is the similar to the exact preconditioner, so that no robustness is lost. Therefore, this simplification is the most efficient version of this preconditioner.

6.7.8 Finally, we test this preconditioner in the presence of vortices using the flow field of example A.3.3. Again, we use the clockwise ordering of grid cells introduced in Section 6.5. A more general numbering algorithm for arbitrary two-dimensional flow can be found in [HP97]. Solving Oseen equations yields the convergence shown in Table 6.21. Here, we used Bicgstab iteration instead of GMRES to avoid a nonfeasible number of auxiliary vectors. Therefore, the iteration counts with large viscosity are lower than above, since Bicgstab performs a double step on symmetric matrices. Again, we see a very moderate increase of iteration steps for lower

L	ν 10^{-1}	$10^{-1.5}$	10^{-2}	$10^{-2.5}$	10^{-3}	$10^{-3.5}$	10^{-4}
one V-cycle							
3	43	47	59	101	181	294	440
4	46	49	56	81	201	507	1129
5	50	53	55	74	134	388	1226
6	52	61	59	66	107	294	964
7	54	70	66	64	82	181	615
two V-cycles							
3	36	39	48	84	141	263	411
4	39	43	47	67	165	410	1151
5	41	46	47	60	121	323	1022
6	43	52	50	55	88	244	823
7	44	56	56	55	67	129	428

Table 6.20: Convergence of Kay/Loghin preconditioner for two dimensional Poisseuille flow with one and two V-cycles ($\mathbb{Q}_1$ elements)

L	ν 10^{-1}	$10^{-1.5}$	10^{-2}	$10^{-2.5}$	10^{-3}	$10^{-3.5}$
3	23	30	49	119	376	3689
4	24	30	47	114	365	—
5	26	28	41	92	308	4080
6	27	31	41	81	222	3055
7	28	31	46	83	215	1680

Table 6.21: Iteration steps n_{10} Kay/Loghin preconditioner for linear driven cavity flow with one variable V-cycles (Bicgstab, $\mathbb{Q}_1$ elements)

Reynolds numbers up to $\nu = 1/100$. Then, the system is still solvable in reasonable time up to $\nu = 1/1000$. A stronger deterioration of convergence with decrease of ν can be expected here, since the limit equation for $\nu = 0$ is not well-posed anymore.

Appendix A

Example problems

Many of the model problems used to demonstrate the feasibility of the methods and analysis in this book appear in several places. Therefore, we collect these test examples in this appendix.

A.1 Meshes and domains

A.1.1 Example: Most examples are computed on the unit square and unit cube $[0,1]^d$ in two and three dimensions, respectively, with Dirichlet boundary $\Gamma_D = \partial\Omega$. Unless superconvergence effects are to be cancelled, Cartesian grids may be used, originating from the refinement, local or global, of a single cell constituting the mesh $\mathbb{T}_0$.

A.1.2 Example: On the same grids as in example A.1.1, we fix the Dirichlet boundary

$$\Gamma_D = \big\{x \in \partial\Omega \big| x_1 = 0 \vee x_1 = 1\big\}.$$

On the Neumann boundary $\Gamma_N = \partial\Omega \setminus \Gamma_D$, we prescribe homogeneous or inhomogeneous boundary conditions, depending on the approximated solution.

A.1.3 Example: A variant of these globally refined meshes can be generated by shifting each vertex after refinement. An example, where the vertices are shifted by 15% of the minimal diameter of its adjacent cells is shown in Figure A.1.

A.1.4 Example: A second coarse mesh for the unit square is shown in Figure A.2. This mesh exhibits two topologically irregular points. Some superconvergence effects are cancelled there and optimality of L^∞ error estimates can be observed.

A.1.5 Example: Robustness with respect to local refinement is investigated on a hierarchy of meshes obtained by refining the positive quadrant of the square $[-1,1]^d$ in each step. Grid cells adjacent are refined if required to make the mesh one irregular. A sequence of such meshes is shown in Figure A.3. I norder to obtain a more pronounced locality of refinement, we also perform computations on meshes refined on a small circle around the origin in Figure A.4. The same types of local refinement are also tested in three dimensions (see Figure A.5.

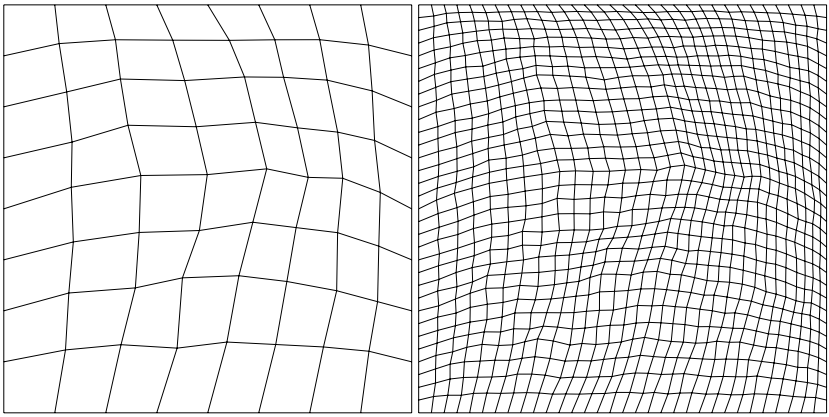

Figure A.1: Meshes with distorted vertices after 2 and 4 refinement steps

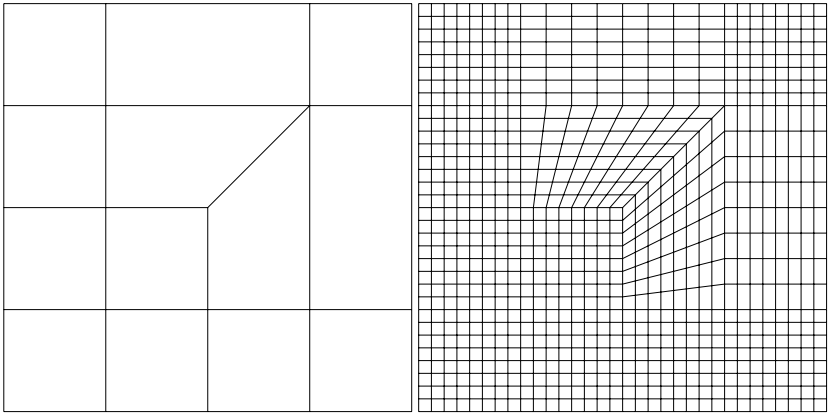

Figure A.2: Irregular subdivision of the unit square and second refinement

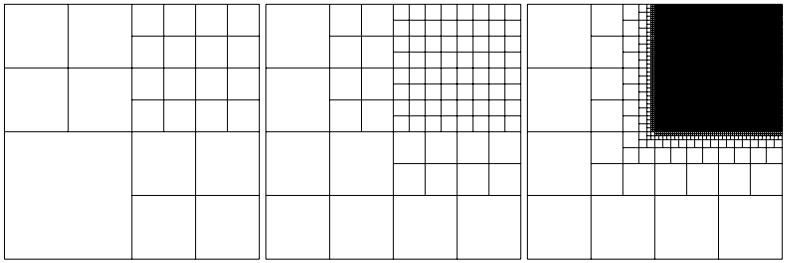

Figure A.3: Meshes with local refinement of the first quadrant, levels 2, 3 and 7

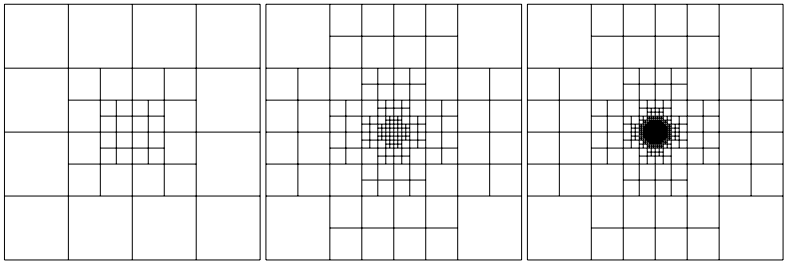

Figure A.4: Meshes with local refinement of a small circle, levels 2, 4 and 7

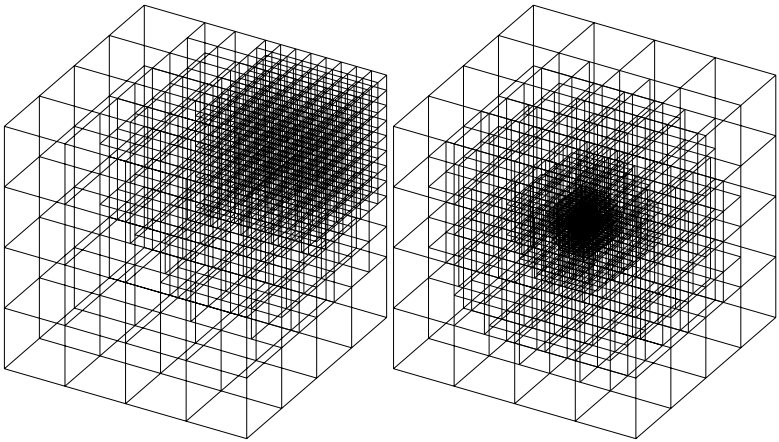

Figure A.5: Meshes with local refinement in three dimensions

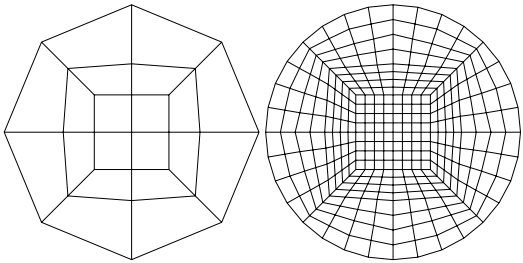

Figure A.6: Subdivisions of a circle. Coarse mesh with $\mathbb{Q}_1$-mapping, with $\mathbb{Q}_2$-mapping and refined mesh with $\mathbb{Q}_1$-mapping

A.1.6 Example: Even if we do not cover smoothly curved boundaries theoretically, we present some results for this case. As a prototype for such a case, we choose a hierarchy of triangulations of a circle as in Figure A.6. Remark that the shape of the computational domain (and the quality of approximation of the real domain) depends on the order of the mapping between reference cell and actual grid cell. The two coarse grids in Figure A.6 use bilinear and biquadratic mapping on the left and on the right, respectively.

A.2 Poisson equation

A.2.1 Example: Approximate the first eigenfunction of the Laplacian operator on the unit square with homogeneous boundary values, namely

$$u(x) = \prod_{i=1}^{d} \cos(\tfrac{\pi}{2} x_i), \qquad f(x) = \frac{d\pi^2}{4} \prod_{i=1}^{d} \cos(\tfrac{\pi}{2} x_i).$$

A.2.2 Example: A model problem with inhomogeneous boundary conditions is constructed using the exponential function.

$$u(x) = \prod_{i=1}^{d} \exp(x_i), \qquad f(x) = d \prod_{i=1}^{d} \exp(x_i).$$

A.2.3 Example: Solutions with lower regularity are obtained on domains with reentrant corners. Our examples include the leading singularities for the L-shaped and the slit domain (see Figure A.7) with right-hand side $f = 0$ and homogeneous boundary values on the reentrant edges. The values on the remaining parts of the boundary are set to fit the solution. The solutions can be found in [Kon67]. Introducing polar coordinates, we name r the distance from the origin and φ the angle computed counter-clockwise from the indentation into the unit square $[-1, 1]^2$, we have in the L-shaped domain

$$u(r, \varphi) = r^{\frac{2\pi}{3}} \sin(\tfrac{3}{2}\varphi). \qquad \text{(a)}$$

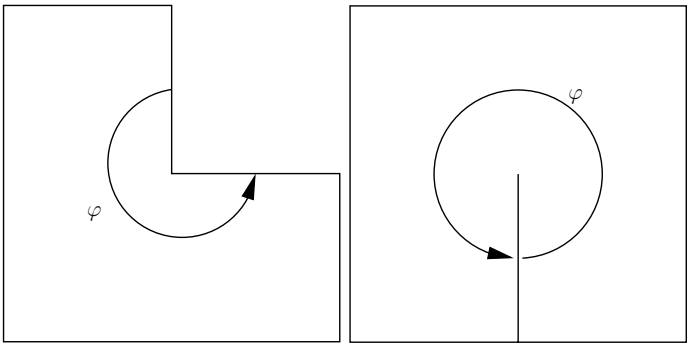

Figure A.7: L-shaped and slit domain

The solution in the slit domain is

$$u(r,\varphi) = r^{\frac{\pi}{2}} \sin(2\varphi). \tag{b}$$

A.3 Advection-diffusion-reaction

A.3.1 Example: In order to investigate the behavior of discretizations in the case of reaction boundary layers, we solve

$$\begin{aligned} -\nu\Delta u + u &= 1 \text{ in } \Omega = [-1,1]^2 \\ u &= 0 \text{ if } x = \pm 1 \\ \partial_n u &= 0 \text{ if } y = \pm 1. \end{aligned}$$

The solution of this problem is

$$u(x,y) = 1 - \frac{e^{\frac{x-1}{\sqrt{\nu}}} + e^{\frac{-x-1}{\sqrt{\nu}}}}{1 + e^{-\frac{2}{\sqrt{\nu}}}}. \tag{A.1}$$

This solution exhibits an exponential boundary layer of width $\sqrt{\nu}$ at the boundaries $x = \pm 1$.

A.3.2 Example: The simplest advection field in our investigation is a constant vector

$$b = \begin{pmatrix} \cos \pi/5 \\ \sin \pi/5 \end{pmatrix}.$$

While the vector field is constant, it is oblique with respect to the Cartesian grid cells in example A.1.1, avoiding a purely on-dimensional solution.

A.3.3 Example: Behavior of solvers and solution in flow fields with vortices is tested using the two-dimensional vector field (cf. [ESW02])

$$u(x,y) = \begin{pmatrix} 2(2y-1)(1-(2x-1)^2) \\ -2(2x-1)(1-(2y-1)^2) \end{pmatrix} \tag{A.2}$$

on the square $[0,1]^2$. It constitutes a single divergence free vortex rotating clockwise and vanishing at the boundaries.

A.4 Stokes and Oseen equations

A.4.1 Example (Poisseuille flow): A standard example for flow problems is Poisseuille flow through a straight pipe. Let the domain be the square $[-1,1]^2$ or the cylinder $[0,1] \times B_1(0)$. Then, the flow field

$$u^0(x,r) = \begin{pmatrix} 1-r^2 \\ 0 \end{pmatrix},$$

where $r = |y|$ and $r = \sqrt{y^2+z^2}$ in two and three dimensions, respectively, is a solution to the Stokes equations 4.9 with pressure

$$p(x,r) = -2(d-1)\nu x + \bar{p}.$$

Furthermore, it solves Oseen equations 5.18 with flow field $b = u^0$ and therefore it is a stationary solution to the incompressible Navier-Stokes equations.

We solve this problem with boundary conditions

$$\begin{aligned} u &= u^0 \quad \text{on } \partial\Omega \cap \{x = -1\} \\ u &= 0 \quad \text{on } \partial\Omega \cap \{r = 1\} \\ \partial_{\mathbf{n}} u + p\mathbf{n} &= 0 \quad \text{on } \partial\Omega \cap \{x = 1\}. \end{aligned}$$

The outflow boundary condition implies $p(1,r) = 0$, yielding $\bar{p} = -2(d-1)\nu$.

A.4.2 Example: A simple test for Stokes equations also suitable for higher order polynomials on the unit square $[0,1]^2$ in two dimensions are the functions

$$\begin{aligned} u_1(x,y) &= -e^x(y\cos y + \sin y) \\ u_2(x,y) &= e^x y \sin y \\ p(x,y) &= 2e^x \sin y. \end{aligned}$$

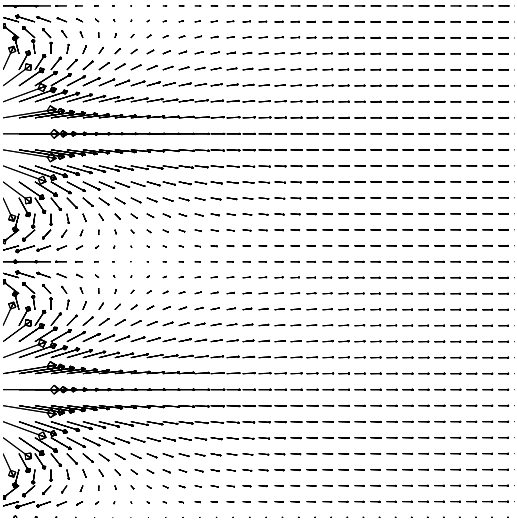

Figure A.8: Kovasznay flow field for Reynolds number 10

A.4.3 Example (2D Kovasznay flow): In [Kov48], a stationary solution to Navier-Stokes equations in the idealized wake behind a grid was computed for any Reynolds number. On the square $[-0.5, 1.5] \times [0, 2]$, this solution is

$$\begin{aligned} u_1(x,y) &= 1 - e^{\lambda x}\cos(2\pi y) \\ u_2(x,y) &= \frac{\lambda}{2\pi} e^{\lambda x}\sin(2\pi y) \\ p(x,y) &= -\frac{1}{2}e^{2\lambda x} + \bar{p}. \end{aligned}$$

Here,

$$\lambda = \frac{-8\pi^2}{\mathrm{Re} + \sqrt{\mathrm{Re}^2 + 64\pi^2}}.$$

The flow field u for $\mathrm{Re} = 1/\nu = 10$ is shown in Figure A.8.

Figure A.9 shows that the size of the solution is independent of the viscosity, while σ and p behave linearly with respect to ν.

In computations, we impose the Dirichlet boundary value u at the whole boundary of Ω.

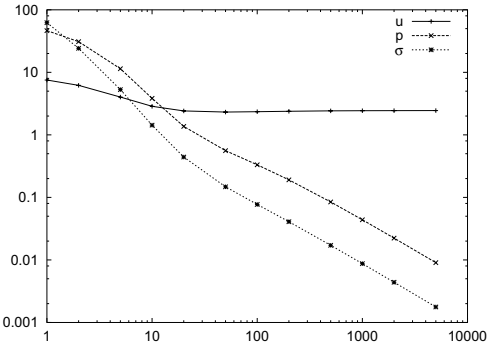

Figure A.9: L^2-norms of Kovasznay solutions σ, u and p depending on the Reynolds number

Bibliography

[ABCM00] D. N. Arnold, F. Brezzi, B. Cockburn, and D. Marini. Discontinuous Galerkin methods for elliptic problems. In B. Cockburn, G. E. Karniadakis, and C.-W. Shu, editors, *Discontinuous Galerkin Methods*, volume 11 of *LNCSE*, pages 89–101. Springer, 2000.

[ABCM01] D. N. Arnold, F. Brezzi, B. Cockburn, and D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779, 2001.

[ABF02] D. N. Arnold, D. Boffi, and R. S. Falk. Approximation by quadrilateral finite elements. *Math. Comput.*, 71:909–922, 2002.

[Ada75] R. A. Adams. *Sobolev Spaces*. Academic Press, 1975.

[Arn82] D. N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19(4):742–760, 1982.

[BDFM87] F. Brezzi, jr. Douglas, J., M. Fortin, and D. Marini. Efficient mixed finite element methods in two and three space variables. *RAIRO Modél. Math. Anal.Numér.*, 21:581–604, 1987.

[BDM85] F. Brezzi, jr. Douglas, J., and D. Marini. Two families of mixed finite elements for second order elliptic problems. *Numer. Math.*, 47:217–235, 1985.

[Bec95] R. Becker. *An Adaptive Finite Element Method for the Incompressible Navier-Stokes Equations on Time-Dependent Domains*. Dissertation, Universität Heidelberg, 1995.

[BEPS96] J. H. Bramble, R. E. Ewing, J. E. Pasciak, and J. Shen. The analysis of multigrid algorithms for cell centered finite difference methods. *Adv. Comput. Math.*, 5(1):15–29, 1996.

[Ber01] S. Berrone. Adaptive discretization of stationary and incompressible Navier–Stokes equations by stabilized finite elements. *Comput. Methods Appl. Mech. Engrg.*, 190(34):4435–4455, 2001.

[BF91] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*. Springer, 1991.

[BHK07] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II — a general purpose object oriented finite element library. *ACM Trans. Math. Softw.*, 33(4), 2007.

[BHL03] R. Becker, P. Hansbo, and M. G. Larson. Energy norm a posteriori error estimation for discontinuous Galerkin methods. *Comput. Methods Appl. Mech. Engrg.*, 192:723–733, 2003.

[BHS03] R. Becker, P. Hansbo, and R. Stenberg. A finite element method for domain decomposition with non-matching grids. *M2AN Math. Model. Numer. Anal.*, 37(2):209–225, 2003.

[BJK90] G. A. Baker, W. N. Jureidini, and O. Karakashian. Piecewise solenoidal vector fields and the Stokes problem. *SIAM J. Numer. Anal.*, 27:1466–1485, 1990.

[BP92] J. H. Bramble and J. E. Pasciak. The analysis of smoothers for multigrid algorithms. *Math. Comput.*, 58(198):467–488, 4 1992.

[BPX91] J. H. Bramble, J. E. Pasciak, and J. Xu. The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms. *Math. Comput.*, 56(193):1–34, 1991.

[BR01] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica*, 10:1–102, 2001.

[Bra93] J. H. Bramble. *Multigrid Methods*. Number 294 in Pitman research notes in mathematics series. Longman Scientific, 1993.

[BS02] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, 2nd edition, 2002.

[BW97] J. Bey and G. Wittum. Downwind numbering: Robust multigrid for convection-diffusion problems. *Appl. Numer. Math.*, 23(1):177–192, 1997.

[Cas01] P. Castillo. *Local Discontinuous Galerkin Methods for Convection-Diffusion and Elliptic Problems*. PhD thesis, University of Minnesota, Minneapolis, 2001.

[Cas02] P. Castillo. Performance of discontinuous Galerkin methods for elliptic PDE's. *SIAM J. Sci. Comput.*, 24:524–547, 2002.

[CCPS00] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau. An a priori error estimate of the local discontinuous Galerkin method for elliptic problems. *SIAM J. Numer. Anal.*, 38(5):1676–1706, 2000.

[CD00] B. Cockburn and C. N. Dawson. Some extensions of the local discontinuous Galerkin method for convection-diffusion equations in multidimensions. In J. R. Whiteman, editor, *The Proceedings of the Conference on the Mathematics of Finite Elements and Applications: MAFELAP X*, pages 225–238. Elsevier, 2000.

[Cia78] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, 1978.

[CKPS01] B. Cockburn, G. Kanschat, I. Perugia, and D. Schötzau. Superconvergence of the local discontinuous Galerkin method for elliptic problems on cartesian grids. *SIAM J. Numer. Anal.*, 39(1):264–285, 2001.

[CKS02] B. Cockburn, G. Kanschat, and D. Schötzau. The local discontinuous Galerkin method in incompressible fluid flow. In H. A. Mang, F. G. Rammerstorfer, and J. Eberhardsteiner, editors, *Proceedings of the Fifth World Congress on Computational Mechanics (WCCM V)*. Vienna University of Technology, 2002.

[CKS03a] B. Cockburn, G. Kanschat, and D. Schötzau. LDG methods for Stokes flow problems. In F. Brezzi, A. Buffa, S. Corsaro, and A. Murli, editors, *Numerical Mathematics and Advanced Applications: ENUMATH 2001*, pages 755–764, Milano, 2003. Springer Italia.

[CKS03b] B. Cockburn, G. Kanschat, and D. Schötzau. The local discontinuous Galerkin method for the Oseen equations. *Math. Comput.*, 73(246):569–593, 2003.

[CKS05] B. Cockburn, G. Kanschat, and D. Schötzau. A locally conservative LDG method for the incompressible Navier- Stokes equations. *Math. Comput.*, 74:1067–1095, 2005.

[CKS07] B. Cockburn, G. Kanschat, and D. Schötzau. A note on discontinuous Galerkin divergence-free solutions of the Navier-Stokes equations. *J. Sci. Comput.*, 31(1–2):61–73, 2007.

[CKSS02] B. Cockburn, G. Kanschat, D. Schötzau, and C. Schwab. Local discontinuous Galerkin methods for the Stokes system. *SIAM J. Numer. Anal.*, 40(1):319–343, 2002.

[CS98] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, 35:2440–2463, 1998.

[ELPV93] R. E. Ewing, R. D. Lazarov, J. E. Pasciak, and P. S. Vassilevski. Domain decomposition type iterative schemes for parabolic problems on locally refined grids. *SIAM J. Numer. Anal.*, 30:1537–1557, 1993.

[ESW02] H. C. Elman, D. J. Silvester, and A. J. Wathen. Performance and analysis of saddle point preconditioners for the discrete steady-state Navier-Stokes equations. *Numer. Math.*, 90(4):665 –688, 2002.

[FR76] J. Frehse and R. Rannacher. Eine l^1-fehlerabschätzung für diskrete grundlösungen in der methode der finiten elemente. *Bonner Math. Schriften*, 89:92–114, 1976. Tagungsband Finite Elemente.

[FS91] L. P. Franca and R. Stenberg. Error analysis of some Galerkin least squares methods for the elasticity equations. *SIAM J. Numer. Anal.*, 28:1680–1697, 1991.

[GK03a] J. Gopalakrishnan and G. Kanschat. Application of unified DG analysis to preconditioning DG methods. In K. J. Bathe, editor, *Computational Fluid and Solid Mechanics 2003*, pages 1943–1945. Elsevier, 2003.

[GK03b] J. Gopalakrishnan and G. Kanschat. Multi-level preconditioners for the interior penalty method. In F. Brezzi, A. Buffa, S. Corsaro, and A. Murli, editors, *Numerical Mathematics and Advanced Applications: ENUMATH 2001*, pages 795–804, Milano, 2003. Springer Italia.

[GK03c] J. Gopalakrishnan and G. Kanschat. A multilevel discontinuous Galerkin method. *Numer. Math.*, 95(3):527–550, 2003.

[GR86] V. Girault and P.-A. Raviart. *Finite element approximations of the Navier-Stokes equations*. Springer, 1986.

[Gri85] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Number 24 in Monographs and Studies in Mathematics. Pitman, 1985.

[GRW02] V. Girault, B. Rivière, and M. F. Wheeler. A discontinuous Galerkin method with non-overlapping domain decomposition for the Stokes and Navier-Stokes problems. Report 02-08, TICAM, Austin, 2002.

[GT98] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, 1998.

[HL02] P. Hansbo and M. G. Larson. Discontinuous Galerkin methods for incompressible and nearly incompressible elasticity by Nitsche's method. *Comput. Methods Appl. Mech. Engrg.*, 191:1895–1908, 2002.

[HL06] P. Hansbo and M. Larson. Piecewise divergence free discontinuous galerkin methods. Technical report, Chalmers Technical University, Göteborg, 2006.

[HN01] B. Heinrich and S. Nicaise. Nitsche mortar finite element method for transmission problems with singularities. Technical Report 01-10, SFB 393, Chemnitz, 2001. 3.

[HP97] W. Hackbusch and T. Probst. Downwind Gauss-Seidel smoothing for convection dominated problems. *Numer. Linear Algebra Appl.*, 4(2):85–102, 1997.

[HSS02] P. Houston, C. Schwab, and E. Süli. Discontinuous hp-finite element methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 39(6):2133–2163, 2002.

[JP86] C. Johnson and J. Pitkäranta. An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation. *Math. Comput.*, 46:1–26, 1986.

[Kan03a] G. Kanschat. Preconditioning discontinuous Galerkin saddle point systems. In K. J. Bathe, editor, *Computational Fluid and Solid Mechanics 2003*, pages 2016–2018. Elsevier, 2003.

[Kan03b] G. Kanschat. Preconditioning methods for local discontinuous Galerkin discretizations. *SIAM J. Sci. Comput.*, 25(3):815–831, 2003.

[KJ98] O. Karakashian and W. N. Jureidini. A nonconforming finite element method for the stationary Navier-Stokes equations. *SIAM J. Numer. Anal.*, 35(1):93–120, 1998.

[Kla98] A. Klawonn. Block-triangular preconditioners for saddle point problems with a penalty term. *SIAM J. Sci. Comput.*, 19(1):172–184, 1998.

[KLW02] D. Kay, D. Loghin, and A. J. Wathen. A preconditioner for the steady-state Navier-Stokes equations. *SIAM J. Sci. Comput.*, 24(1):237–256, 2002.

[Kon67] V. A. Kondrat'ev. Boundary value problems for elliptic equations in domains with conical or angular points. *Trans. Moscow Math. Soc.*, 16:227–313, 1967.

[Kov48] L. I. G. Kovasznay. Laminar flow behind a two-dimensional grid. *Proc. Camb. Philos. Soc.*, 44:58–62, 1948.

[KR02] G. Kanschat and R. Rannacher. Local error analysis of the interior penalty discontinuous Galerkin method for second order elliptic problems. *J. Numer. Math.*, 10(4):249–274, 2002.

[LR74] P. LeSaint and P.-A. Raviart. On a finite element method for solving the neutron transport equation. In C. de Boor, editor, *Mathematical aspects of finite elements in partial differential equations*, pages 89–123, New York, 1974. Academic Press.

[LT01] C. Lasser and A. Toselli. Overlapping preconditioners for discontinuous Galerkin approximations of second order problems. In N. Debit, M. Garbey, R. H. W. Hoppe, J. Périaux, D. Keyes, and Y. A. Kuznetsov, editors, *Thirteenth International Conference on Domain Decomposition Methods*. DDM.org, 2001.

[MGW00] M. F. Murphy, G. H. Golub, and A. J. Wathen. A note on preconditioning for indefinite linear systems. *SIAM J. Sci. Comput.*, 21(6):1969–1972, 2000.

[Mon03] P. Monk. *Finite Element Methods for Maxwell's Equations*. Oxford University Press, 2003.

[Nit71] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei der Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abh. Math. Sem. Univ. Hamburg*, 36:9–15, 1971.

[PS02] I. Perugia and D. Schötzau. An hp-analysis of the local discontinuous Galerkin method for diffusion problems. *J. Sci. Comput.*, 17:561–571, 2002. Special Issue: Proceedings of the ICOSAHOM-01.

[QV94] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*, volume 23 of *Series in Computational Mathematics*. Springer, Heidelberg, 1994.

[RT77] P.-A. Raviart and J. M. Thomas. A mixed method for second order elliptic problems. In I. Galligani and E. Magenes, editors, *Mathematical Aspects of the Finite Element Method*. Springer, New York, 1977.

[RVW96] T. Rusten, P. S. Vassilevski, and R. Winther. Interior penalty preconditioners for mixed finite element approximations of elliptic problems. *Math. Comput.*, 65(214):447–466, 4 1996.

[Saa00] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Oxford University Press, 2nd edition, 2000.

[SEKW01] D. J. Silvester, H. C. Elman, D. Kay, and A. J. Wathen. Efficient preconditioning of the linearized Navier-Stokes equations. *J. Comput. Appl. Math.*, 128:261–279, 2001.

[SST03] D. Schötzau, C. Schwab, and A. Toselli. hp-DGFEM for incompressible flows. *SIAM J. Numer. Anal.*, 40:2171–2194, 2003.

[Tém66] R. Témam. Sur l'approximation des solutions des équations de Navier-Stokes. *C. R. Acad. Sci. Paris Sér. I Math.*, 216:219–221, 1966.

[Tém68] R. Témam. Une méthode d'approximation de la solutions des équations de Navier-Stokes. *Bull. Soc. Math. France*, 98:115–152, 1968.

[Tém79] R. Témam. *Navier-Stokes equations. Theory and numerical analysis*. North-Holland, Amsterdam, 1979.

[Tos02] A. Toselli. hp-discontinuous Galerkin approximations for the Stokes problem. *Math. Models Methods Appl. Sci.*, 12:1565–1616, 2002.

[Var99] R. S. Varga. *Matrix Iterative Analysis*. Series in Computational Mathematics. Springer, 2 edition, 1999.

Index